Author:
Stanisic, Stasja

Title:
On The Quantum Information Of Photons

Entanglement And Distinguishability In Linear Optics

On the Quantum Information of Photons: Entanglement and Distinguishability in Linear Optics

By

STASJA STANISIC

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Science.

APRIL 2020

Word count: forty three thousand

ABSTRACT

Research into the design and development of quantum devices has gained significant traction in recent years. As classical technologies are pushed ever closer to the limit of their capabilities, the need for the realisation of quantum devices increases. Through a combination of promising theoretical results and engineering advances, realisation of quantum technologies is becoming possible. A platform used for development of many of these quantum technologies is linear optics. Unfortunately, there are still significant engineering challenges to overcome in order to build relevant devices, from metrology to linear optical quantum computing (LOQC). Luckily, we can alleviate the pressure on developing engineering solutions, by examining and improving the theory behind some of the major challenges.

Firstly, optimizing the generation of entanglement necessary for the smallest building block of a measurement-based quantum computer, the contending LOQC platform, would help with scaling of resources. We offer numerical evidence of optimality of the current scheme for generation of Bell pairs. We also find limits to entanglement generation more generally between two subsets of modes with no encoding present, using only linear optical components. Beyond the implications for engineering of entanglement generation, these general results raise interesting foundational questions. Secondly, we present a framework for examining the distinguishability of particles. We take a quantum information inspired approach by giving a bipartite model where distinguishability can arise as correlation with an environment. This offers a new formalism for distinguishable states and uncovers intriguing observations of the underlying dynamics of bosonic behaviour. We use this model to generalize Hong-Ou-Mandel interference as unambiguous state discrimination, giving analytical and numerical evidence for optimal interferometers with a small number of photons. This demonstrates the utility with which standard quantum information tools can be applied within this new framework. We conclude with a discussion of some observations and directions for possible future research.

Words: 291

To Jason
for giving me strength
and to Vanja
for giving me hope

ACKNOWLEDGEMENTS

First, I would like to say thank you to the most patient and determined person in my life, Jason. Thank you for being patient with me while I was doing the research to write this thesis, while I was panicking about writing this thesis, and while I was finally writing this thesis. And everywhere else in between. And thank you to Vanja for making our home a little bigger and a little brighter, and just being the most bubbly and curious ittle monkey ever. My life would be that much emptier without the two of you.

Next, I have to say big thank you to my supervisors and collaborators. None of this would be possible without the enthusiasm and guidance of Peter Turner. The long chats about finer details of representation theory and quantum optics will be sorely missed as I am not sure anyone will be as passionate as you on the topic. Noah Linden and Ashley Montanaro, for the opportunities they have given me, and for enduring being bombarded by weekly numerical data. Sister in arms, Alex Moylett, for sharing the office, the laughs, and the frustration. Tony Short for the APMs and the advice. Chris, Laura, Lidia, and Miriam for the board game evenings. And of course the rest of the Quantum Information group, for sharing exciting papers at group meetings and seminars. No acknowledgment would be complete without mentioning the original comrades of this wacky CDT experiment, Alasdair, Euan, Janna, Jeremy, Mack, Matt, Sam, SMS, and Stefan. One could not have wanted for better companions over the last five years, I can only hope our paths keep crossing. Also to the rest of the CDT team that brought us together, Andrea, Lin, Renuka, Clarence, Chris, and Mark. Phil and Jake for their guidance during the Project B and offering me the opportunity to break things in their lab. Nicola, for the shared hot chocolates and cakes. But also to the rest of the CQP and QETLabs over the years, admin, students, and postdocs. It has been amazing being under a single roof with such a diverse crowd (personally and professionally). To the British Federation of Women Graduates for intellectually stimulating meetings and for bestowing upon me such a prestigious award.

More than a little help has come from my incredible family. Mami i tati hvala za više stvari nego što bi stale ovde. Ali više od sve podrške i strpljenja (i hrane, piča, bejbisitovanja, društva), hvala što ste me pustili da budem ko želim i ko mi je bilo neophodno da budem (većinom vremena). Hvala Staši za njegovo bezgranično razumevanje i pored svih godina sporadičnog kontakta. I Gali što se brine za njega i pravi mu društvo. Hvala i tetka Ani i teča Koletu za njihovo konstano bodrenje i gostoprimstvo. Nemoguće je ne spomeniti

DECLARATION

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

SIGNED: .... DATE: ....

STATEMENT OF WORK

This thesis contains my previously published work verbatim and with modification, as well as work that has been previously used for my Annual Progress Monitoring reports and course assessments in the School of Physics, University of Bristol during the course of the Ph.D.

Chapter 1 contains extracts from an essay written by me on the topic of:

[1] S. Stanisic, "Universal quantum computation by linear optics." unpublished, 2015

All content of this chapter is my own.

Chapter 2 is largely based on the paper:

[2] S. Stanisic, N. Linden, A. Montanaro, and P. S. Turner, "Generating entanglement with linear optics," *Physical Review A*, vol. 96, no. 4, p. 043861, 2017

The work was performed under the supervision and in collaboration with Noah Linden, Ashley Montanaro, and Peter S. Turner, with the majority of work and writing being my own. The data associated with Chapter 2 is available for download at the University of Bristol data repository, data.bris [3].

Parts of Chapter 3 and most of Chapter 4 are based on:

[4] S. Stanisic and P. S. Turner, "Discriminating distinguishability," *Physical Review A*, vol. 98, no. 4, p. 043839, 2018

The work was performed under the supervision and in collaboration with Peter S. Turner, with the majority of work and writing being my own. The data associated with Chapter 4 is available for download at the University of Bristol data repository, data.bris [5]. The code used for Chapter 3 and Chapter 4 was done in collaboration with Alexandra E. Moylett.

The following papers are based on work that was also carried out during the course of this Ph.D., but are excluded from this thesis due to the difference in scope.

[6] J. Adcock, E. Allen, M. Day, S. Frick, J. Hinchliff, M. Johnson, S. Morley-Short, S. Pallister, A. Price, and S. Stanisic, "Advances in quantum machine learning," *arXiv:1512.02900*, 2015

[7] P. Sibson, J. E. Kennard, S. Stanisic, C. Erven, J. L. O'Brien, and M. G. Thompson, "Integrated silicon photonics for high-speed quantum key distribution," *Optica*, vol. 4, no. 2, p. 172, 2017

CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Symbols

$\cong$ Basis change relation, page 65

$\hat{a}_j$ Annihilation operator on mode j, page 8

$\hat{a}_j^\dagger$ Creation operator on mode j, page 8

$\hat{a}_{sl}^\dagger$ Creation operator for composite bosons, page 85

$B(\theta, \phi), B(\theta)$ Beamsplitter, page 11

$\mathrm{BS}_{50:50}$ Balanced beamsplitter, page 11

$\mathbb{C}^d$ d-dimensional complex vector space, page 69

$\mathbb{C}^{d_{(\lambda)}}$ Carrier space of λ irrep of symmetric group, page 69

$\mathbb{C}^{d_{\{\lambda\}}}$ Carrier space of the irrep λ of unitary group, page 71

$(\mathbb{C}^d)^{\otimes N}$ N-tensor space of d-dimensional complex vector space, page 69

$\mathfrak{D}_{j_1 j_2 \ldots j_N}^{12 \ldots N}$ Slater determinant, page 78

d Number of modes/dimension of the system, page 9

$d_{(\lambda)}$ Dimension of the λ irrep of the symmetric group, page 67

$d_{\{\lambda\}}$ Dimension of irrep λ of the unitary group, page 71

d_{L} Dimension of the Label degree of freedom, page 84

d_{S} Dimension of the System degree of freedom, page 84

$\det$ Determinant, page 15

E_k POVM element, page 108

f_λ Number of standard Young tableaux, page 67

$\mathcal{H}^N$ Hilbert space of N particles, page 7

$\mathcal{H}_{\mathrm{SL}}$ Composite Hilbert space of System and Label degrees of freedom, page 84

$\mathrm{imm}_\lambda(M)$ Immanant of matrix M for irrep λ, page 127

$K_{\lambda,\underline{n}}$ Kostka number, page 67

$\mathcal{L}_{jk}$ Lowering operator from state $|j\rangle$ to state $|k\rangle$, page 71

$M_{\underline{n}}$ Measurement operator for pattern $\underline{n}$, page 87

$m_{(\lambda)}$ Multiplicity of the λ irrep of symmetric group, page 69

$m_{\{\lambda\}}$ Multiplicity of irrep λ of the unitary group, page 71

$\hat{n}_j^\dagger$ Number operator on mode j, page 8

N Number of particles, page 9

N_X Number of particles on a subsystem X, page 40

$\mathbb{N}_0^d$ Set of non-negative integers, page 70

$P(\omega)$ Phaseshifter, page 11

$P_{\underline{n}}$ Probability of pattern $\underline{n}$, page 88

$P_{\underline{n}}^\lambda$ Probability of pattern $\underline{n}$ within irrep λ, page 88

$P_{\underline{n}}^{\text{classical}}$ Scattering probability of classical particles, page 17

P_{T} Normalized Young symmetrizer, page 69

perm   Permanent, page 14

$\text{Par}(N)$ Set of all partitions of a number N, page 65

$\text{Par}(N, d)$ Set of all partition of a number N with at most d parts, page 65

QFT_d Quantum Fourier Transform of size d, page 12

ρ Density matrix, page 6

ρ_{d} The reduced System state of the completely distinguishable state, page 96

ρ_{i} The reduced System state of the completely indistinguishable state, page 93

ρ_{s} The reduced System state of the singly distinguishable state, page 94

$\rho_{\bar{\text{i}}}$ A state with no support in the symmetric subspace, page 111

$\mathcal{R}_{kj}$ Raising operator from state $|k\rangle$ to state $|j\rangle$, page 71

σ Permutation in S_N, page 14

$\mathcal{S}$ Representation of symmetric group, page 69

S von Neumann entropy, page 40

$\text{SSTab}(\lambda)$ Set of all semi-standard Young tableaux of shape λ and weight $\underline{n}$, page 66

$\text{SSTab}(\lambda, \underline{n})$ Set of all semi-standard Young tableaux of shape and weight $\underline{n}$, page 66

$\text{STab}(\lambda)$ Set of all standard Young tableaux of shape λ, page 66

$\text{ST}(\lambda)$ Semi-standard Young tableau of shape and weight λ where the row j is filled with number j, page 66

$s(T)$ Young symmetrizer of tableau T, page 68

SU(d)   Special unitary group, page 17

S$_N$ Symmetric group of all permutations $\{1, 2, \ldots, N\}$, page 14

$\mathcal{T}$ Representation of group G, page 64

$t_\lambda(d)$ Number of semi-standard Young tableaux, page 67

Tab(λ)   Set of all Young tableaux of shape λ, page 66

T(λ, j)   Semi-standard Young tableau of shape λ and with its content indexed by j, page 66

Tr   Trace function, page 6

$\mathcal{U}$ Representation of unitary $U \in$ U(d), page 9

U Element of U(d), usually transformation of mode operators, page 9

$\chi(g)$ Characteristic of g in a group G, page 65

$|\lambda, p, \underline{n}, r\rangle$ Schur-Weyl basis notation, λ is the irrep of the unitary and symmetric groups, p is the basis state of the symmetric irrep, $\underline{n}$ is the occupation number of the basis state of the unitary irrep, and r is its inner multiplicity, page 74

$|\lambda, q, p\rangle$ Schur-Weyl basis notation, λ is the irrep of the unitary and symmetric groups, q is the basis state of the unitary irrep, and p is the basis state of the symmetric irrep, page 73

$|n_1, ..., n_d\rangle$ Fock state, page 7

$|\psi_{\mathrm{d}}\rangle$ Completely distinguishable pure state, page 88

$|\psi_{\mathrm{i}}\rangle$ Completely indistinguishable pure state, page 88

$|\mathrm{vac}\rangle$ Vacuum, page 8

Abbreviations

BBS   Balanced beamsplitter, page 11

BCH   Baker-Campbell-Hausdorff formula, page 10

BFGS   Broyden–Fletcher–Goldfarb–Shanno algorithm, page 46

BS   Beamsplitter, page 11

CNOT   Controlled not gate, page 18

CZ   Controlled Z gate, page 18

GHZ   Greenberger-Horne-Zeilinger states, page 2

H   Hadmard gate, page 18

HOM   Hong-Ou-Mandel effect, page 2

KLM   Knill-Laflamme-Milburn scheme, page 5

LOQC   Linear optical quantum computation, page 2

MBQC   Measurement-based quantum computation, page 5

MCM Maximum Confidence Measurements, page 108

MED Minimum Error Discrimination, page 108

MZI   Mach-Zehnder interferometer, page 12

NS   Non-linear sign shift gate, page 22

NV   Nitrogen-vacancy, page 29

PBS   Polarizing beamsplitter, page 11

PNRD Photon number resolving detector, page 3

POVM Positive operator-value measure, page 108

PS   Phaseshifter, page 11

QFT   Quantum Fourier Transform, page 12

QKD Quantum key distribution, page 1

SPDC Spontaneous parametric down conversion, page 28

UD   Unambiguous Discrimination, page 108

INTRODUCTION

Decades in the making, the quantum revolution finally seems imminent. Theoretical results are demonstrating both the achievability and potential of quantum devices, and advances in engineering are making building them ever more possible, but most importantly, as we reach the limits of what classical technologies can offer, the need for breaking into the quantum domain becomes ever more urgent [8]. The interest spans from foundational physics research and practically driven engineering companies, to end users in companies with reams of data or need for greater and greater precision.

In quantum computing great strides have been made: IBM Q System One which is a fully packaged quantum computer containing a 50-qubit chip and more expected; Google have released a 53-qubit chip with evidence of achieving quantum supremacy, but have also produced a 72-qubit chip; IonQ claims a system with up to 160 qubits (79 of which are computational qubits); Rigetti have plans for a 128-qubit chip, and there are efforts globally to expand on this throughout academic groups and other start-ups. Quantum computing promises a wealth of potential applications including, but not limited to, prediction of new superconducting materials, new medicines, breaking cryptographic security, and even helping with seemingly routine, but very complex tasks such as better air traffic scheduling [9]. In quantum metrology, big projects such as LIGO have announced plans for quantum enhanced gravitational wave detection, and smaller groups and start-ups are working towards bringing devices such as better gravity sensors out of the labs and into production. Similarly, in quantum communications, quantum key distribution (QKD) networks are already being tested in the UK, USA, China, and elsewhere.

Optical systems are an important player in all the aforementioned fields [10]. Like any platform, there are numerous challenges with photonics based devices of varying degrees of importance. To start with, an effect central to many quantum devices is particle interference. However, the famous Hong-Ou-Mandel (HOM) effect demonstrates that boson distinguishability is detrimental to quantum interference [11], and particle distinguishability remains significant engineering challenge for quantum technologies, photonics in particular. If two photons significantly differ in their properties, they are "distinguishable" and therefore will not achieve the necessary degree of interference. Current single photon sources do not have enough control over these properties to create uniform photons. The lack of uniformity of photons produces errors in the system; this is a major problem for

technologies like quantum computing where scaling up might require millions of photons causing these errors to compound.

Nevertheless, new schemes for linear optical quantum computing (LOQC) are making it possible for a universal quantum computer to be imagined on a photonics based platform [12, 13]. A change in paradigm from a circuit based [14] to a measurement-based model [15, 16] and then finally a ballistic percolation based model [17], not only dramatically improves the utilisation of resources, but also improves the system's resilience to loss, and (mostly) reduces the engineering challenges to construction of an efficient source of small entangled states, such as Bell pairs and 3-photon Greenberger-Horne-Zeilinger (GHZ) states [18]. Integration of quantum optical devices on-chip with known technologies [19–21], allowing for millions of linear optical elements on a single chip, promises the ability to scale up in a straight-forward manner [22]. However, LOQC challenges still remain: effects of photons loss, efficient generation of small cluster states, photon number resolving detection, reliable single photon sources, and errors due to distinguishability, to name only a few [21, 23].This thesis focuses on attempting to improve our understanding of two of these issues: entanglement generation and photon distinguishability [18, 24].

The leading LOQC scheme is based on highly-entangled large cluster states. Building these cluster states requires entangled resources such as Bell pairs or 3-photon GHZ states. The qubits in LOQC platforms are commonly realised with dual-rail encoding, where the position of a photon in one of the two modes defines its state (usually polarization). In this type of encoding, generating entanglement with linear optics alone is non-deterministic, with probabilities lower than 20 %. The current leading proposal for building LOQC relies on a source of deterministic 3-photon GHZ states [25]. To produce these states with a probability near unity, 42 Bell pairs are needed. To then get a renormalised qubit in a Raussendorf lattice, thousands of these deterministic GHZ states might be required [18]. When we take into consideration that these Bell pairs are non-deterministic, we can understand the scale of this problem and why an optimal way to generate Bell pairs is needed. The issue of optimal generation of entanglement is therefore tackled in this thesis.

The other challenge addressed here is understanding the distinguishability of photons. Most commonly, distinguishability occurs due to the source of single photons in a linear optical system. While there have been many approaches [26–35], it remains unclear as to what is the "best" model. At minimum however, the one factor that is indisputably detrimental to linear optics is pairwise distinguishability (as demonstrated by the HOM effect [11]), as this leads to the loss of quantum interference and results in the behaviour of particles becoming the same as that of classical particles. Ideally we would have a model for distinguishability that not only allows us to better understand the behaviour of particles, but also allows for the straightforward development of tools to combat distinguishability, through detection, filtering, or postselection. To date, approaches to modelling it have been varied, but finding one which is both sufficiently general and useful for the interpretation of experimental observation remains a challenge. We take a quantum information inspired approach to modelling distinguishability, by separating a particle's Hilbert space into

degrees of freedom that we control and those we do not. This gives a bipartite model where distinguishability can arise as correlation with an environment consisting of unobserved degrees of freedom, and takes into consideration that for the majority of the time we do not know how the photons differ from one another.

Assuming ideal conditions of lossless linear optics and perfect photon number resolving detectors (PNRDs), this thesis examines generation of entanglement and modelling distinguishability of particles in a way that allows the application of quantum information tools on linear optical systems. Firstly, in Chapter 1, we give a review of quantum information and linear optics concepts and results.

This is followed by Chapter 2, in which we present some novel results on the generation of entanglement in linear optics. We start by examining the previous work on entanglement generation in dual-rail encoding, followed by a numerical search for interferometers which can produce "event-ready" Bell pairs with higher success probabilities than that of the currently known setups (Sections 2.1 and 2.3). Finding no improvement on the current schemes, we then turn towards the comparison of entanglement generation in dual-rail encoding and what is known as "mode entanglement" (Section 2.4). We show that mode entanglement produces high rates of entanglement even for a random choice of unitaries, significantly higher than the corresponding dual-rail setup. We examine in further detail how well mode entanglement can perform, as it may offer some explanation for the dual-rail results (Section 2.5), and we uncover some fairly peculiar behaviour. For example, examining the entanglement between two single output modes we find that in the case of photons bunched in a single input mode the entanglement is unbounded, however, coincident photon input over many modes is bounded by a constant. In the more general situation where we examine entanglement between sets of modes, we find two bounds due to the dimensionality of the Hilbert spaces of the mode sets, and linearity of the mode transformation. Optimizing the amount of entanglement generated numerically, we can clearly see these two bounds, but we also see the existence of another unexplained bound. We finish by listing optimal interferometers for entanglement generation in those setups where we know the optimum has been achieved. See Table 2.1 for a summary of results.

In Chapter 3 we turn towards modelling the distinguishability of particles using group representation theory quite generally, with the practical aim of applying this framework to bosons. We then dive into understanding the necessary representation theory tools of the symmetric and unitary groups (Section 3.1). We focus on constructing a Schur-Weyl basis which offers us the most insight on the occupation numbers of bosons, as this offers us a neat way to track distinguishability later on. We then look at how this basis allows for a natural introduction of distinguishability into the description of the states, and how it immediately offers simplifications due to unitary-unitary duality (Section 3.2). We then take this new basis, and through expressing some of the most interesting states in it reveal some compelling features (Sections 3.3 and 5.1). To show how all these elements come together, and to highlight the complex relationship between permanents, determinants, immanants, and representation of the $U(d)$, we offer two examples of states of three particles

and their scattering amplitudes and probabilities in Sections 3.4.1 and 3.4.

After developing the theory needed to use quantum information tools on bosonic states in Chapter 3, we then demonstrate how easily these can be carried over and offer a way to generalize the seminal HOM effect. We start by reviewing the current quantum state discrimination tools available and how they translate into the picture of lossless linear optics with ideal photon number resolving detectors (Section 4.1 and Section 4.2). In the Schur-Weyl basis, almost immediately we find a general bound on the probability of successful discrimination of distinguishable and indistinguishable states based solely on their description in this new basis (Section 4.3). Turning towards states of photons in two modes, we start by examining how the HOM can be expressed as an unambiguous discrimination problem of distinguishable and indistinguishable states, and then demonstrate the formalism for three photons in two modes (Section 4.4). Next, increasing the number of modes to three, we show a mixture of numerical and analytical results towards optimal discrimination (Section 4.5). Finally using a mix of numerical and analytical methods we look at discrimination of four or more photons, finding some novel interferometers that are effective at discrimination of the specific distinguishable states under consideration (Section 4.6). See Table 4.1 for a summary of results and relevant interferometers.

We finish this work in Chapter 5 with a discussion of closely related results in the field on which the results presented in this work could cast a new light. Of particular interest are the relation of immanants to elements of the scattering matrix, in the light of results such as BosonSampling and the immanants fitting well with the framework based around representation theory. We examine suppression laws in more detail, a closely related idea to the discrimination of distinguishabilty which is the focus of Chapter 4. We also leave an important but possibly philosophical discussion on the definition of distinguishability of particles to this Chapter.

QUANTUM INFORMATION WITH LINEAR OPTICS

1.1 Overview

To delve deeper into the topics of entanglement and distinguishability in linear optics, we will first cover some preliminary concepts, definitions, and results. In Section 1.2 we mention the fundamental quantum information terminology and notation as used in this thesis. In Section 1.3 we show how these fundamental concepts are carried over into linear optics and go into the necessary details for understanding the material within this thesis. First we introduce states of bosonic (and fermionic systems). We then turn towards the evolution of these states and the most common linear optical elements. We briefly talk about the current standard for measurements in linear optical system, finishing this Section with discussion on the probabilities that are to be expected from those measurements and their connection to the permanents of matrices. The fundamental concepts are then followed by Section 1.4 which combines all of these elements into LOQC platform. This section starts by reminding the reader of the general criteria and components of a quantum computer. This is followed by a description of possible encodings for LOQC with different challenges that need to be overcome, which subsequently motivate different schemes for building such a computer. We then discuss three proposed schemes for building an LOQC: Knill-Laflamme-Milburn (KLM), measurement-based quantum computation (MBQC), and ballistic percolation based QC, which are presented in the historical order, but also from most theoretical (KLM) to most implementable from an engineering perspective (ballistic percolation based QC). In Section 1.4.6, we take a look at a special type of a non-universal quantum computer – a boson sampler. Its importance is found in its complexity (in the Computer Science sense of the word), but also its simplicity (from the Engineering perspective). Finally, we finish with a Section 1.5 covering the current literature on distinguishability of particles, focusing on bosons. The theory and history presented in these sections should invoke ample motivation to understand, research, and expand the knowledge of both entanglement and distinguishability and its effects in linear optical systems, both for fundamental and engineering purposes.

1.2   Quantum information

Here we have a brief review of terminology and meanings of some of fundamental quantum information concepts. The smallest unit of classical information is a *bit*, whose two possible states are usually denoted by "0" or "1". In a similar manner, to discuss quantum information, we can define a *qubit* whose possible states are $|0\rangle$ or $|1\rangle$. In more physical terms, qubits describe the states of a two-dimensional quantum system. In general, a *quantum state* is a normalized vector $|\psi\rangle$ in d-dimensional complex Hilbert space describing the state of a quantum system. The conjugate transpose of this state is denoted $\langle\psi|$. These quantum units of information in dimension d are referred to as *qudits* (in the case of a qubit $d = 2$). Unlike classical information, quantum information can be found in a state of *superposition*, that is a linear combination of states. For qubits, we can have any normalized vector $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ where $|\alpha|^2 + |\beta|^2 = 1$ and $\alpha, \beta \in \mathbb{C}$. We can interpret these amplitudes as the probabilities of finding the quantum state $|\psi\rangle$ in basis state $|0\rangle$ ($|1\rangle$) with probability $|\alpha|^2$ ($|\beta|^2$). We can immediately see how the space of the smallest unit of quantum information could be fairly rich and more complex than that of a bit.

So far we have only mentioned a single quantum system with state space $\mathbb{C}^d$, however to describe a composite system, we will need to consider multiple single particle states together. For this we use the tensor product of single particle states, $|\Psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, or just $|\psi_A\rangle |\psi_B\rangle$ for short. From this composite state we can take a *partial trace* which removes the information on one of the component subsystems. Given a basis $\{|1\rangle, |2\rangle, \ldots, |d\rangle\}$ for the system "A" we are tracing over, then $\text{Tr}_A(|\Psi\rangle) = \sum_{k=1}^{d} \langle k | \Psi\rangle \langle \Psi | k\rangle$. Multiparticle states are called *product* states, if they can be described as a tensor product of states of the component systems. Otherwise, the states are called *entangled*. An example of this is one of the famous Bell states, $\frac{1}{\sqrt{2}} (|0\rangle |0\rangle + |1\rangle |1\rangle)$, which cannot be written as a product $|\psi_A\rangle |\psi_B\rangle$.

The quantum states we have described so far are called *pure* states. We also define *mixed* states, given by *density operators* or *density matrices*, which describe an ensemble of pure states $\{p_j, |\psi_j\rangle\}$ as $\rho = \sum_j p_j |\psi_j\rangle \langle\psi_j|$. The density operator always has a trace of one and is positive. For a composite system, we have $\rho = \rho_1 \otimes \rho_2 \otimes \ldots \rho_N$.

To learn about the quantum state of a system, we perform (projective or von Neumann) *measurement operators*, $\{M_m\}$, where m is an outcome of the measurement and the probability to get that outcome is defined as $P_m = \langle\psi|M_m^\dagger M_m|\psi\rangle$. The following four properties hold:

- Hermitian: $M_m = M_m^\dagger$ for all outcomes m.

- Positive: $\langle\psi|M_m|\psi\rangle \geq 0$ for all states $|\psi\rangle$ and all outcomes m.

- Complete: $\sum_m M_m^\dagger M_m = \mathbb{1}$.

- Orthonormal: $M_m M_{m'} = \delta_{mm'} M_m$.

When applied to a mixed state, the probability of outcome m is $P_m = \text{Tr}[M_m^\dagger M_m \rho]$. Finally,

closed systems evolve according to a unitary transformation $\mathcal{U}$, such that $\left|\psi'\right\rangle = \mathcal{U}\left|\psi\right\rangle$. It follows that mixed states evolve as $\rho' = \mathcal{U}\rho\mathcal{U}^\dagger$.

We will see how all of these standard concepts find a place in linear optics.

1.3 Linear optics

We now examine how the standard purely mathematical concepts from Quantum Information introduced in the previous section correspond to an actual physical system.

1.3.1 States

The standard quantum information description, detailed in Section 1.2 is useful when thinking about information, however it does not offer easy manipulation of quantum states in real physical systems. This "particle" or "first quantized" picture implicitly labels particles, implying that a state such as $\left|12\right\rangle$ is distinct from the state $\left|21\right\rangle$ even for identical particles. However, if the particles are truly identical, there should be no difference in the physics they demonstrate based on their ordering. Instead of thinking about a first particle in state "1" and second particle in state "2", we should be thinking there is one particle in state "1" and one particle in state "2". The "second quantizated" picture introduces notation for this, where instead of tracking the states particles are in, we track occupations of those states. We discuss this and similar problems further in Chapter 3.

Let ψ be a state in the first quantized picture of N qudits with basis,

$$\psi = \psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_N \tag{1.1}$$

where the k-th particle is in the state $\left|j_k\right\rangle \in \{1, 2, \ldots, d\}$. This would not be a valid state for identical bosons, whose many-particle state has to be symmetric, therefore the corresponding boson state would be

$$\psi_B = \sum_{\sigma \in \mathrm{S}_N} \psi_{\sigma^{-1}(1)} \otimes \psi_{\sigma^{-1}(2)} \otimes \cdots \otimes \psi_{\sigma^{-1}(N)}, \tag{1.2}$$

where σ is an element of S_N, the symmetric group of all permutations of $\{1, 2, \ldots, N\}$. We can then specify this state by referring to occupations of the single particle states (the reasons why this is unique can be seen in Section 3.1.3). For each N we can construct a Hilbert space, $\mathcal{H}^N$, for the symmetric N-particle wavefunctions. Then *Fock space* is a direct-sum of all N-particle Hilbert spaces. The basis states of this space are called *Fock states* and they represent the occupation of the states in the N-particle wavefunction. We denote a Fock state as $\left|n_1, ..., n_d\right\rangle$, where n_j gives us the number of particles in state j and d is the total number of states. The total number of particles is $N := \sum_{j=1}^{d} n_j$. Sometimes the notation $\left|n_j\right\rangle_j$ will be used, marking the exact occupation of the j-th state. Throughout this thesis unless if otherwise specified, the number of modes will be denoted as d, the number of photons as N, and state occupations as $\underline{n} = (n_1, n_2, \ldots, n_d)$.

We identify the states of particles with harmonic oscillator modes, and the change in their occupation with the excitation of those modes. For photons, this is sensible as the energy of the quantized electromagnetic field can be written in the same form as the energy of an harmonic oscillator

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \tag{1.3}$$

A Fock state of a single mode is then an eigenstate of H, and in a way similar to that in the theory of oscillators, we have raising and lowering operators, with a ground state – we have *creation* and *annihilation* operators, with vacuum, acting upon Fock space states in the following fashion

$$\hat{a}_j^\dagger |n_j\rangle_j = \sqrt{n_j + 1} |n_j + 1\rangle_j \tag{1.4}$$

$$\hat{a}_j |n_j\rangle_j = \sqrt{n_j} |n_j - 1\rangle_j \tag{1.5}$$

and $[\hat{a}_j, \hat{a}_k^\dagger] = \delta_{jk}$ is also true for all $j, k \in \{1, ..., d\}$. The number operator is

$$\hat{n}_j |n_j\rangle_j = \hat{a}_j^\dagger \hat{a}_j |n_j\rangle_j = n_j |n_j\rangle_j. \tag{1.6}$$

If no particles are in any of the states, we call this *vacuum*, and denote it as $|\text{vac}\rangle := |0\rangle^{\otimes n}$. Vacuum on a single mode is denoted as $|0\rangle$ or $|0\rangle_j$. By acting with raising operators on the vacuum, we can get any Fock state of a single mode,

$$|n_j\rangle_j = \frac{(\hat{a}_j^\dagger)^{n_j}}{\sqrt{n_j!}} |\text{vac}\rangle, \forall j \in \{1, ..., d\}. \tag{1.7}$$

Similarly,

$$|n_1, ..., n_d\rangle = \prod_{j=1}^{d} \frac{(\hat{a}_j^\dagger)^{n_j}}{\sqrt{n_j!}} |\text{vac}\rangle. \tag{1.8}$$

The proof is straightforward from above definitions. Given a list of numbers such as $\underline{n} = (n_1, n_2, \ldots, n_d)$ we define

$$\underline{n}! = \prod_{j=1}^{d} n_j! \tag{1.9}$$

as the *list factorial*. Then we can write

$$|n_1, ..., n_d\rangle = \frac{1}{\sqrt{\underline{n}!}} \prod_{j=1}^{d} (\hat{a}_j^\dagger)^{n_j} |\text{vac}\rangle. \tag{1.10}$$

Beyond using Fock states, we also use the *mode assignment list*

$$a(\underline{n}) = \oplus_{j=1}^{d} \oplus_{k=1}^{n_j} j. \tag{1.11}$$

That is $a(\underline{n}) = (\overbrace{1, 1, \ldots, 1}^{n_1}, \ldots, \overbrace{d, d, \ldots, d}^{n_d})$.

Thinking about how we can place N identical balls into d boxes gives us a combinatorial

way to calculate the dimension of the Hilbert space of N photons in d modes. The dimension is then

$$\binom{d + N - 1}{N}.$$

(1.12)

Certain Fock states are common in the literature and will be referred to in this thesis often, so they have their own notation. The *coincident* (also, unbunched or collision-free) state is a Fock state where n_k is either 0 or 1 for all k. The coincident state where $N = d$, that is $n_k = 1$ for all k, will have its occupation denoted as $\underline{n} = \underline{1}$. Similarly, a state that has all the photons found in a single mode and vacuum in rest will be referred to as *completely bunched*. When the photons are bunched in the first mode it will commonly be shortened to $\underline{n} = (N, \underline{0})$.

1.3.2 Interferometers

Optical elements are defined by the effective interaction Hamiltonian of the medium H [36]. The unitary that evolves the state of the system $\mathcal{U} |n_1, ..., n_j\rangle$ is given as $\mathcal{U} = e^{iH}$.

A unitary *transformation on optical modes*, U, is an isomorphism from the space of input operators to output operators. *Linear optical* elements are such that the mode transformation under evolution $\mathcal{U}$ can be described by matrices U and V, which transform the modes linearly, that is, $\hat{a}_j^\dagger \rightarrow \sum_k U_{kj} \hat{a}_k^\dagger + V_{kj} \hat{a}_k$. Linear optical elements are called *passive* if the energy of the incoming photons is conserved which, in turn, implies that the number of photons is conserved thus $V = 0$ (from the previous definition). In the case of passive linear optics, the interaction Hamiltonian is bilinear in the creation and annihilation operators and is of the form $\hat{H} = \sum_{jk} h_{jk} \hat{a}_j^\dagger \hat{a}_k$ where the annihilation operator comes second by convention[36].

From now on, any reference to linear optics will implicitly mean passive linear optics. More often than not, in this thesis (but also in general in LOQC) we focus on the mode transformations, and refer to them as *interferometers*, *scattering matrices*, or *transfer matrices*, and we talk about probabilities for certain output states as *scattering probabilities*. Unless stated otherwise, when a unitary U is discussed, we will be talking about this mode transformation, as opposed to the unitary that maps the input states to output states (they act on different spaces, the former acting within the operator space while the latter acts within the state Hilbert space). Therefore, when describing an optical element, this unitary transformation U with its action,

$$U : \hat{a}_j^\dagger \rightarrow \sum_k U_{kj} \hat{a}_k^\dagger.$$

(1.13)

will be how we describe the interactions in the system. Similarly we will use, $\mathcal{U}$, the representation of the mode transformation U on the multimode Fock space.

Lemma 1.3.1. *For passive linear optics we have that* $\mathcal{U} |\text{vac}\rangle = |\text{vac}\rangle$.

Proof. We will want to Taylor expand the equation $\mathcal{U} |\text{vac}\rangle = e^{i\hat{H}} |\text{vac}\rangle = (1 + i\hat{H} + \frac{1}{2!}(i\hat{H})^2 + ...) |\text{vac}\rangle$ first. Since $\hat{H} = \sum_{jk} h_{jk} \hat{a}_j^\dagger \hat{a}_k$, each term in the Taylor expansion except for the

first one will be a polynomial of creation and annihilation operators. The i-th term in the Taylor expansion is the polynomial $\hat{H}^i$, and from here notice that raising $\hat{H}$ to the i-th power will still leave an annihilation operator as the right-most operator for each term of the i-th polynomial. We have mentioned that $\hat{a}_j \left|0\right\rangle_j = 0$ for any j, then all the terms in the i-th polynomial are 0 and the polynomial is actually 0 when acting on vacuum. Then the terms except for the first one in the Taylor expansion will be 0 and thus $\mathcal{U} \left|\text{vac}\right\rangle = \left|\text{vac}\right\rangle$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Lemma 1.3.2. *The modes transform as* $\hat{a}_j^\dagger \to \mathcal{U}\hat{a}_j^\dagger\mathcal{U}^\dagger$.

Proof. Let some state have a single photon in the mode i. Write this state as $\hat{a}_i^\dagger \left|\text{vac}\right\rangle$. Then its evolution under some unitary operator U for a passive linear optical element is given as $\mathcal{U}\hat{a}_i^\dagger \left|\text{vac}\right\rangle = \mathcal{U}\hat{a}_i^\dagger\mathcal{U}^\dagger\mathcal{U}\left|\text{vac}\right\rangle = \mathcal{U}\hat{a}_i^\dagger\mathcal{U}^\dagger\left|\text{vac}\right\rangle$ as $\mathcal{U}$ is unitary and $\mathcal{U}\left|\text{vac}\right\rangle = \left|\text{vac}\right\rangle$. This shows that the state modes transform as $\hat{a}_j^\dagger \to \mathcal{U}\hat{a}_j^\dagger\mathcal{U}^\dagger$. $\qquad\qquad\qquad\qquad$ $\square$

According to the Baker-Campbell-Hausdorff (BCH) formula,

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = B + [A, B] + \frac{1}{2!}[A, [A, B]] + ... \qquad (1.14)$$

where $\hat{A}$ and $\hat{B}$ are some operators. Given a linear optical element defined through its Hamiltonian, we can find what its mode transformation U by using BCH expansion on $\mathcal{U}\hat{a}_i^\dagger\mathcal{U}^\dagger = e^{iH}\hat{a}_i^\dagger e^{-iH} = \sum_k U_{kj}\hat{a}_k^\dagger$. In most common linear optical elements these tidy away to nice equations and we will see their definitions now.

As a consequence, for linear optical elements, transformations on modes, U, are unitary (so the first definition introduced in this subsection is well defined). Moreover, for any given mode transformation unitary U, there is a way to construct this unitary using only beamsplitters and phaseshifters as shown by Reck et al [37].

Therefore, on a multimode state, we can also describe the new state using just U.

Lemma 1.3.3. *Given an input Fock state* $\left|\psi_{\text{in}}\right\rangle = \left|n_1, n_2, \ldots, n_d\right\rangle$ *with* N *photons and* d *modes, the action of* $\mathcal{U}$ *can be described as*

$$\mathcal{U}\left|\psi_{\text{in}}\right\rangle = \prod_{k=1}^{d} \frac{\left(\sum_{j=1}^{d} U_{jk}\hat{a}_j^\dagger\right)^{n_k}}{\sqrt{n_k!}} \left|\text{vac}\right\rangle \qquad (1.15)$$

Proof.

$$\mathcal{U}\left|\psi_{\text{in}}\right\rangle = \mathcal{U}\prod_{k=1}^{d} \frac{\left(\hat{a}_k^\dagger\right)^{n_k}}{\sqrt{n_k!}} \left|\text{vac}\right\rangle = \mathcal{U}\prod_{k=1}^{d} \frac{\left(\hat{a}_k^\dagger\mathcal{U}^\dagger\mathcal{U}\right)^{n_k}}{\sqrt{n_k!}} \left|\text{vac}\right\rangle \qquad (1.16)$$

$$= \prod_{k=1}^{d} \frac{\left(\mathcal{U}\hat{a}_k^\dagger\mathcal{U}^\dagger\right)^{n_k}}{\sqrt{n_k!}}\mathcal{U}\left|\text{vac}\right\rangle = \prod_{k=1}^{d} \frac{\left(\sum_{j=1}^{d} U_{jk}\hat{a}_j^\dagger\right)^{n_k}}{\sqrt{n_k!}} \left|\text{vac}\right\rangle \qquad (1.17)$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Common optical elements

Here we define some common optical elements, such as, phaseshifters, beamsplitters, polarizing beamsplitters, and Mach-Zehnder interferometers. The former two we will use often throughout the thesis and we will often treat them in a more general (mathematical) sense without much regard for the physical meaning of the modes they act on (beyond them being bosonic modes). The latter two are mentioned due to their significance in certain areas of linear optical experiments. We also define another $d \times d$ unitary matrix which can be viewed as a generalization of a balanced beamsplitter, where $|U_{jk}| = \frac{1}{\sqrt{d}}$ for all j and k.

A *phaseshifter* P_ω (PS) is a one mode passive linear optical element with Hamiltonian $H_P(\omega) = \omega \hat{a}^\dagger \hat{a}$. Its unitary matrix is $P(\omega) = e^{i\omega}$. *Beamsplitter* $B(\theta, \phi)$ (BS) is a two mode passive linear optical element generated by the Hamiltonian $H_B(\theta, \phi) = \theta e^{i\phi} \hat{a}^\dagger \hat{b} + \theta e^{-i\phi} \hat{b}^\dagger \hat{a}$. Its unitary matrix is then

$$B(\theta, \phi) = \begin{pmatrix} \cos(\theta) & -e^{i\phi}\sin(\theta) \\ e^{-i\phi}\sin(\theta) & \cos(\theta) \end{pmatrix}, \tag{1.18}$$

acting on modes $\begin{pmatrix} \hat{a}^\dagger \\ \hat{b}^\dagger \end{pmatrix}$. The modes can be path modes, in which case we commonly think about the beamsplitter as a semi-reflective mirror, but they could also be polarization modes, in which case we can implement it using birefringent material called waveplates (commonly quarter and half-wave plates).

In this work we will often drop the phase parameter and only use a single parameter beamsplitter, $B(\theta) := B(\theta, 0)$, that is

$$B(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{1.19}$$

and this should be clear from the text by reference to a single parameter. When we mention beamsplitters in the rest of the thesis, we refer to the single parameter beamsplitter unless stated otherwise.

A commonly used beamsplitter is the "50:50" beamsplitter, where the transmission and reflection are equal. There are a few different representations of this beamsplitter which are equivalent up to a global phase ($\theta \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$ would all give different types of this beamsplitter in the Equation 1.18). The one[*] we will refer to will be

$$\mathrm{BS}_{50:50} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{1.20}$$

A *Polarizing beamsplitter* (PBS) that separates horizontal and vertical polarization is

---

[*]equivalent with the "50:50" beamsplitters found from Equation 1.18

a four mode passive linear optical element with following matrix

$$
\text{PBS} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}
\tag{1.21}
$$

acting on modes $\begin{pmatrix} \hat{a}_H^\dagger \\ \hat{b}_H^\dagger \\ \hat{a}_V^\dagger \\ \hat{b}_V^\dagger \end{pmatrix}$, where $\hat{a}_H^\dagger, \hat{a}_V^\dagger$ ($\hat{b}_H^\dagger, \hat{b}_V^\dagger$) are two possible path-polarization modes in the path mode "a" ("b").

A *Mach-Zehnder interferometer* (MZI) consists of two "50:50" beamsplitters (BS$_{50:50}$) and a phase shifter in the first mode between the beamsplitters (see Figure 1.5).

$$
\text{MZI}(\theta) = \text{BS}_{50:50} P(2\theta) \text{BS}_{50:50} = e^{i\theta} \begin{pmatrix} \cos(\theta) & i\sin(\theta) \\ i\sin(\theta) & \cos(\theta) \end{pmatrix}
\tag{1.22}
$$

The MZI is equivalent to a single parameter beamsplitter. By adding another variable phaseshifter, we can make a two parameter beamsplitter, this setup is also commonly referred to as an MZI when it comes to photonic chips [19] (the importance of this setup can be found in Section 1.4.2).

Finally, in general we can define balanced interferometers on d modes such that $|U_{jk}| = 1/\sqrt{d}$ for all $j, k \in \{1, 2, \ldots, d\}$. The best known unitary that satisfies this is the *Quantum Fourier Transform* (QFT), defined as

$$
\text{QFT}_d = \frac{1}{\sqrt{d}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega^1 & \cdots & \omega^{d-1} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{d-1} & \cdots & \omega^{(d-1)(d-1)} \end{bmatrix}
\tag{1.23}
$$

where $\omega = \exp \frac{2\pi i}{d}$. This matrix is sometimes referred to as a Discrete Fourier transform (DFT), Schur matrix, or Bell multiport beam splitter.

"Reck" Scheme

We can construct any $U \in \text{U}(d)$ on d modes using only phaseshifters and (single-parameter) beamsplitters (see Figure 1.1) with what is known as a "Reck" scheme in optics [37]. As shown in Figure 1.1, such a scheme can be viewed as $d-1$ layers, indexed by j, each with j phaseshifters and beamsplitters, followed by a final phase shift on each mode. Therefore we see that there is a total of $d(d-1)/2$ of beamsplitters and $d(d+1)/2$ phaseshifters needed for any implementation of this unitary. This gives a total of d^2 parameters, which is the number of parameters needed for a unitary from $\text{U}(d)$ and therefore this parametrization
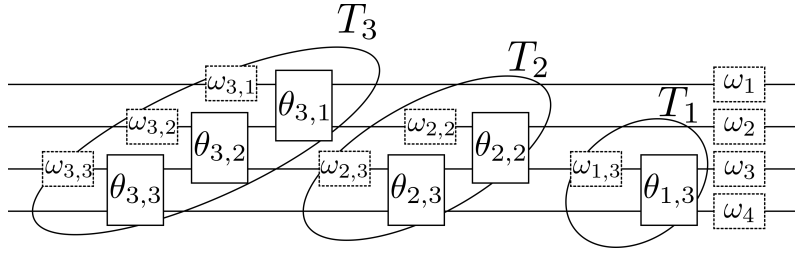
Figure 1.1: Example of a Reck scheme parametrising an arbitrary unitary transformation on four modes ($d_S = 4$), grouped into 'layers' T_j. Each one- (phaseshifter) and two-mode (beamsplitter) subtransformation contributes one real parameter. Only the phaseshifters situated between beamsplitters ($\omega_{1,3}, \omega_{2,2}, \omega_{2,3}$) contribute to our problem.

is optimal.

A layer T_j consists of phaseshifters $P(\omega_{j,k})$ where k denotes the mode it acts on in layer j and beamsplitters $B(\theta_{j,k})$, where k and $k+1$ are the modes it acts on in layer j. The j-the layer can then be described as

$$T_j(\underline{\theta}_j, \underline{\omega}_j) = \prod_k^{M-1} B(\theta_{j,k}) \prod_{k=j}^{M-1} P(\omega_{j,k}). \tag{1.24}$$

The interferometer is then

$$U(\underline{\theta}, \underline{\omega}) = \prod_{k=1}^{M} P(\omega_k) \prod_{j=1}^{M-1} T_j(\underline{\theta}_j, \underline{\omega}_j). \tag{1.25}$$

Notice however, the number of optical elements a photon interacts with is different depending on which mode the photon enters.
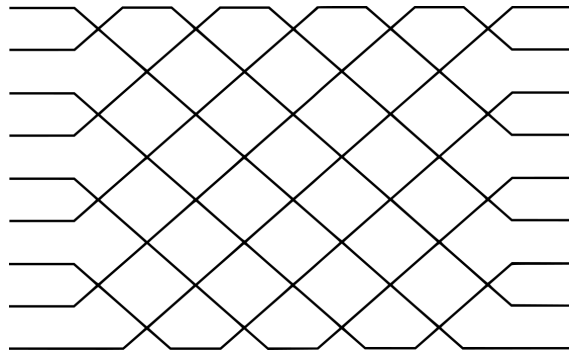
Figure 1.2: A rectangular grid of two parameter MZI, parametrising an arbitrary unitary transformation on $d = 9$ modes. Image taken from [38], copyright(2016) by the OSA under the terms of CC-BY licence.

There are other ways to break down a $d \times d$ interferometer into smaller optical elements. For example in 2016, Clements et al. suggested a factorization with more symmetry in its design and therefore is smaller and has more consistent optical depth [38] (see Figure 1.2). As we see in Figure 1.2 the idea behind their scheme is a rectangular structure of (two-parameter) beamsplitters. The longest path in the Reck scheme would consist of $2d - 3$

beamsplitters (a photon going from the bottom mode to the top and back to bottom), however in this scheme the longest path is d.

Recently however, a way to factorize any U of dimension $d \times d$ into two unitaries of dimensions $(d-1) \times (d-1)$ connected by a beamsplitter was reported [39], which can then be recursively applied to construct different parameterization that might suit the problem in question better (not necessarily made of just beamsplitters and phaseshifters). While in this thesis, when a parameterization of unitary is needed, the Reck scheme has been used, these alternative parameterizations are of interest for future work. For example, in Section 3.4.3, we mention using numerical optimization or manual inspection for small numbers of photons, we can see how certain scattering probabilities are linked. However how to prove this is not clear, but this might be an artefact of the choice of parameterization. It is possible that using some more uniform parameterization such as that of Clements et al. mentioned above would make it more apparent what the connections between different scattering probabilities are.

It is also interesting to mention that given any non-trivial beamsplitter, it will be possible to generate any transformation on $d \geq 3$ modes [40] (so every non-trivial beamsplitter is universal for quantum optics, where by non-trivial it is meant that the beamsplitter does something more than just change the phase of a mode or a simple mode permutation).

1.3.3 Measurements

Single photon detectors are commonly talked about in optical quantum information literature [24]. More than just single photon detectors, photon number resolving detectors (PNRD) are key components in LOQC [14], quantum metrology [41], entanglement distribution [42], and QKD [43] amongst others. There has been significant effort directed towards creating a high quality, photon number resolving detector [24], although commonly, current setups utilise multiplex photon "bucket" detectors to mimic the work of number resolving detectors. With respect to these aforementioned efforts and the significance of them in the literature, it is not unreasonable to assume access to one. Moreover the detectors we will model will have no loss, dark counts, or other errors. They are ideal PNRDs which enact a projection onto the Fock space in the case of ideal photons.

Given an occupation, $\underline{n}$, we would then model this photon counting measurement as

$$M_{\underline{n}} = |\underline{n}\rangle \langle \underline{n}| . \tag{1.26}$$

We define the scattering probabilities $P_{\underline{n}}$ related to a measurement operator $M_{\underline{n}}$ using the Born rule. Given a pure input state $|\psi\rangle$, the probability related to output occupation $\underline{n}$ is then defined as

$$P_{\underline{n}} = |\langle \underline{n} \mid \psi \rangle|^2 . \tag{1.27}$$

1.3.4 Permanents and scattering probabilities

We now define a functional on a square matrix $M = (M_{jk})$ of size $m \times m$ denoted the *permanent*,

$$\text{perm}(M) = \sum_{\sigma \in \text{S}_m} \prod_{j=1}^{m} M_{j,\sigma(j)} \tag{1.28}$$

where σ is an element of S_m, the symmetric group of all permutations of $\{1, 2, \ldots, m\}$. We can see how the permanent is related to a determinant of the matrix which is defined as

$$\det(M) = \sum_{\sigma \in \text{S}_m} \text{sgn}(\sigma) \prod_{j=1}^{m} M_{j,\sigma j}, \tag{1.29}$$

where $\text{sgn}(\sigma)$, the sign of a permutation, is one, if the permutation is even (has an even number of transpositions), or minus one if the permutation is odd (has an odd number of transpositions). However, unlike a determinant which is multiplicative, that is

$$\det(M_1 M_2) = \det(M_1) \det(M_2) \tag{1.30}$$

which can then be used to compute the determinant efficiently [44], this property does not hold for permanents in general.

Given an occupation $\underline{s}$ of the input state, and $\underline{t}$ of the output state, we construct a new matrix U_{st} from U in two steps. First, define the matrix U_{s} consisting of s_j copies of the j-th column of U for all $j \in \{1, \cdots, d\}$. Next, construct the matrix U_{st} by using t_j copies of the j-th row of U_{s} for all $j \in \{1, \cdots, d\}$. Notice that if $s_j \leq 1$ and $t_k \leq 1$ for all j and k, the matrix U_{st} is just a submatrix of U; otherwise it is a submatrix with repeated rows and columns, and is often bigger than the original matrix. The scattering amplitudes, then, are connected to the permanents of this matrix [44, 45]. The scattering amplitude for bosons given an input state $|\underline{s}\rangle$ and output state $|\underline{t}\rangle$ is

$$\langle \underline{t} | \mathcal{U} | \underline{s} \rangle = \frac{\text{perm}(U_{\text{st}})}{\sqrt{\underline{s}!\, \underline{t}!}}, \tag{1.31}$$

where $\mathcal{U}$ is the representation of the mode transformation U on the multimode Fock space as mentioned earlier.

Therefore, the scattering probability is

$$P_{\underline{t}} = |\langle \underline{t} | \mathcal{U} | \underline{s} \rangle|^2 = \frac{|\text{perm}(U_{\text{st}})|^2}{\underline{s}!\, \underline{t}!}. \tag{1.32}$$

Classical particles

We can similarly compute the scattering probability of a "classical state", that is a state which has no quantum interference occurring, or a case with non-identical bosons. To understand the difference between the two cases let us examine the scattering probabilities of two "quantum" and two "classical" particles arriving at the balanced beamsplitter that
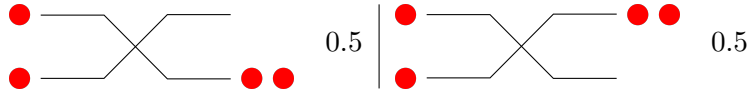
Figure 1.3: Scattering outcomes for quantum particles on a $BS_{50:50}$ interferometer. Two identical bosons bunch after entering the two separate arms of a balanced beamsplitter. The resulting outcome is either two particles in the first mode with probability of 0.5 or two particles in the second mode with probability of 0.5.
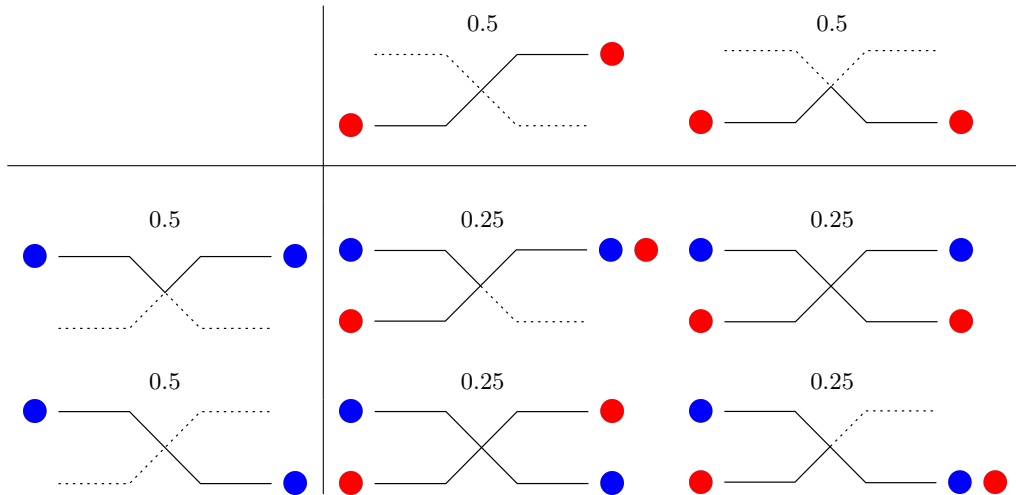
Figure 1.4: Scattering outcomes for classical particles on a $BS_{50:50}$ interferometer. Two distinguishable bosons scatter like classical balls. The row index of the table shows different paths the particle from the first mode can take with the corresponding probabilities. The column index of the table shows different paths the particle from the second mode can take with the corresponding probabilities. The total probability is then calculated through the familiar classical means of multiplying the probabilities of the two independent events together. The outcomes can then be: two particles found in the first mode with probability of 0.25 in the top left corner; two particles found in the second mode with probability of 0.25 in the bottom right corner; one particle in each of the modes with a total probability of 0.5 corresponding to the sum of the two outcomes from the top right and bottom left corners.

was defined in 1.20. Let two identical bosons impinge on a beamsplitter in separate paths (see Figure 1.3). They then transform as

$$\hat{a}_1^\dagger \hat{a}_2^\dagger \xrightarrow{\text{BS}_{50:50}} \frac{1}{2}(\hat{a}_1^\dagger + \hat{a}_2^\dagger)(\hat{a}_1^\dagger - \hat{a}_2^\dagger) = \frac{1}{2}((\hat{a}_1^\dagger)^2 - \hat{a}_1^\dagger \hat{a}_2^\dagger + \hat{a}_2^\dagger \hat{a}_1^\dagger - (\hat{a}_2^\dagger)^2) = \frac{1}{2}((\hat{a}_1^\dagger)^2 - (\hat{a}_2^\dagger)^2). \quad (1.33)$$

This corresponds to a scattering probability of 0.5 for finding both particles in the top mode and a 0.5 for finding both particles in bottom mode (after applying appropriate normalization). This is the well known effect of bosonic bunching, this and related phenomena are the topic of Chapter 3 and Chapter 4.

However, in the case of two distinguishable bosons, we have to keep track of which boson is scattering where through the beamsplitter – the fact that nature does this is what we mean by "distinguishable". We will do this by using two different raising operators.

They transform as

$$\hat{a}_1^\dagger \hat{b}_2^\dagger \xrightarrow{\text{BS}_{50:50}} \frac{1}{2}(\hat{a}_1^\dagger + \hat{a}_2^\dagger)(\hat{b}_1^\dagger - \hat{b}_2^\dagger) = \frac{1}{2}(\hat{a}_1^\dagger \hat{b}_1^\dagger - \hat{a}_1^\dagger \hat{b}_2^\dagger + \hat{a}_2^\dagger \hat{b}_1^\dagger - \hat{a}_2^\dagger \hat{b}_2^\dagger). \tag{1.34}$$

Therefore, there is a chance of 0.25 of finding the two particles in the top or bottom mode, but also a 0.5 chance of finding two particles in two different modes. This matches the probability we would expect if we had two classical states, for example two balls, each scattering with probability of a half (see Figure 1.4).

This holds true more generally as there is no quantum interference in the case of distinguishable particles. Therefore, instead of a complex matrix U describing the scattering, it is sufficient to describe it using a real matrix, these statistics can be understood through classical combinatorial means of N balls instead. Given these N balls, ordered in starting positions such that there is s_k of them in the k-th position (corresponding to occupation $\underline{s}$), we find the probability of finding t_j of them in the j-th position (corresponding to occupation $\underline{j}$) is [46]

$$P_{\underline{t}}^{\text{classical}} = \frac{1}{\underline{s}!\,\underline{t}!} \sum_{\sigma \in S_N} \prod_{j=1}^{N} |U_{a_j(\underline{s}), a_{\sigma(j)}}|^2. \tag{1.35}$$

From the definition of the permanents and the matrix U_{st}, we see the following claim stands.

Lemma 1.3.4. *The scattering probability for distinguishable bosons starting in occupation $\underline{s}$ and ending in occupation $\underline{t}$ is*

$$P_{\underline{t}}^{\text{classical}} = \frac{\text{perm}(|U_{\text{st}}|^2)}{\underline{s}!\,\underline{t}!}, \tag{1.36}$$

where $|M|$ for matrix M denotes the element-wise absolute value.

1.4 Linear optical quantum information

While linear optics has many important usages, we focus further on the theory of linear optical computing, as it demonstrates the importance of entanglement and distinguishability in a straight-forward manner.

1.4.1 Quantum computing

With the development of quantum algorithms, which can solve certain groups of problems faster [47] [48] [49], research into quantum computation has intensified. This was further strengthened by the discovery that quantum computers can be error-correctable [50] [51]. Over the years, many different potential quantum computing platforms have been investigated such as: ion and atom traps, nuclear magnetic resonance, superconducting systems, quantum dots and optical platforms [12]. Due to restrictions inherent in bosonic systems, it was believed that a universal computer built exclusively using linear optics was not possible until, in 2001, Knill, Laflamme and Milburn [14] realized that measurements

on parts of the circuit can be used to evoke non-linearity and still deliver scalability. This changed the global quantum computation landscape making linear optical quantum computing a viable option.

In 2000, DiVincenzo [52] laid out the following criteria for quantum computing:

- Qubits need to be well defined

- Qubit specific measurements can be carried out

- Initialization to a simple pure state such as $|00...0\rangle_{\mathrm{L}}$

- Universal set of quantum gates has to be implementable

- Long coherence times

With regards to optical architectures, photons lend themselves well due to various degrees of freedom that can be used to represent a qubit. Further than this, they have low levels of decoherence and interaction, thus long coherence times are achievable. Some types of qubit encoding are discussed in Section 1.4.2. Initialization to a simple, pure state is usually done using single photon sources and measurements are carried out using single or number resolving photodetectors. The initialization and detection problems are shared between all optical architectures. A universal set of gates (as defined in 1.4.1) is the greatest hurdle for photonics, and the biggest difference in architectures rests on the scalability of these gates.

Even in the "worst case" scenario where an optical quantum computer is not possible, photons will still, most likely, be incorporated into the future of quantum computing as information carriers.

Definition 1.4.1. A set of quantum gates is universal for quantum computation if any unitary operation can be approximated to arbitrary accuracy by a quantum circuit involving only those gates.

We now introduce three gates which together are universal for quantum computation. The first one is CNOT gate which is a two qubit "entangling" gate given in matrix representation as

$$U_{\mathrm{CNOT}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{1.37}$$

Hadamard gate is a single qubit gate which in matrix representation is

$$U_{\mathrm{H}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{1.38}$$

Finally a T gate is a single qubit gate which in matrix representation is

$$U_{\mathrm{T}} = \begin{pmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{pmatrix}. \tag{1.39}$$

The gate set $\{\text{CNOT}, \text{H}, \text{T}\}$ is (approximately) universal for quantum computation. We also introduce another controlled gate, CZ,

$$U_{\text{CZ}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.40}$$

Further, as CNOT can be obtained by applying HCZH, $\{\text{CZ}, \text{H}, \text{T}\}$ is also universal.

Before we can talk about how these gates can be implemented and achieve the needed parts for universal quantum computer using linear optics, we will first talk about how to actually encode the so-far abstract idea of a qubit in a physical optical platform.

1.4.2   Qubit encoding

Different types of qubit encoding can be used in LOQC due to photons having several degrees of freedom. The most commonly used are polarization and spatial. Usually, only one degree of freedom is chosen to encode a qubit, but sometimes two qubits can be encoded on one photon by mixing two (or even more) degrees of freedom. Some of the types of encoding used are single-rail, dual-rail, mixed polarization and spatial encoding, parity encoding and redundant encoding.

Single-rail encoding

In *single-rail encoding*, a photon being present in a path mode is considered to be the logical $|1\rangle$, while vacuum is considered to be $|0\rangle$

$$|0\rangle_{\text{L}} = |0\rangle_{\text{Fock}}, \quad |1\rangle_{\text{L}} = |1\rangle_{\text{Fock}}. \tag{1.41}$$

We can generate entangled states with respect to single-rail encoding by using a 50:50 beamsplitter. Take two qubits such that

$$|10\rangle_{\text{L}} = |10\rangle_{\text{Fock}} = \hat{a}_1^\dagger |\text{vac}\rangle \xrightarrow{\text{BS}_{50:50}} \frac{1}{\sqrt{2}}(\hat{a}_1^\dagger + \hat{a}_2^\dagger) |\text{vac}\rangle \tag{1.42}$$

$$= \frac{1}{\sqrt{2}}(|10\rangle_{\text{Fock}} + |01\rangle_{\text{Fock}}) = \frac{1}{\sqrt{2}}(|10\rangle_{\text{L}} + |01\rangle_{\text{L}}). \tag{1.43}$$

But, in this encoding single-qubit operations are complex [53]. For example, take X rotation, that is

$$|0\rangle_{\text{L}} \rightarrow |1\rangle_{\text{L}}, \quad |1\rangle_{\text{L}} \rightarrow |0\rangle_{\text{L}}. \tag{1.44}$$

Using the single qubit encoding, this can not be done by a number-preserving (passive) linear optical element, since in the first case a photon is created and in the second case a photon is lost. Ancilla states would be needed, and in case we need to also measure some

of them, the operation becomes non-deterministic. More specifically, there is a way to implement an arbitrary phase rotation deterministically, but it is not as straightforward as it would be in dual-rail encoding and an Hadamard gate is so-far only implementable non-deterministically.

Dual-rail encoding
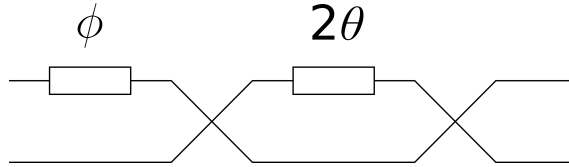
Figure 1.5: Two balanced beamsplitters and two phase-shifters are commonly used for state preparation in integrated photonics. They are commonly depicted as a single unit as they give a general two parameter beamsplitter. We have mentioned in the text that any U(d) interferometer can then be built up using this most basic unit.

In the *dual-rail encoding*, two modes are employed to represent a qubit,

$$|0\rangle_{\rm L} = |10\rangle_{\rm Fock}, \quad |1\rangle_{\rm L} = |01\rangle_{\rm Fock}. \tag{1.45}$$

These two modes can represent polarization or spatial modes. In the case of polarization, the qubits are also written as $|0\rangle_{\rm L} = |H\rangle$ and $|1\rangle_{\rm L} = |V\rangle$. In this encoding, single qubit operations are easy to implement using phaseshifters and beamsplitters (see Figure 1.5 for MZI setup that is typical in integrated optical setups). For example, given a qubit in starting state $|0\rangle_{\rm L}$, it transforms as

$$P(\omega)\mathrm{MZI}\,|10\rangle_{\rm Fock} = e^{i\theta}P(\omega)(\cos(\theta)\,|10\rangle_{\rm Fock} + i\sin(\theta)\,|01\rangle_{\rm Fock}) \tag{1.46}$$

$$= e^{i\theta}(e^{i\omega}\cos(\theta)\,|10\rangle_{\rm Fock} + i\sin(\theta)\,|01\rangle_{\rm Fock}) \tag{1.47}$$

$$= e^{i\theta}(e^{i\omega}\cos(\theta)\,|0\rangle_{\rm L} + i\sin(\theta)\,|1\rangle_{\rm L}). \tag{1.48}$$

On the other hand entanglement is not straightforward.

Lemma 1.4.1. *In a passive linear optical setup using dual-rail encoding, ancillas, and heralding, it is not possible to create a Bell state using a two single photon input.*

Proof. Assume we have initial state $|00\rangle_{\rm L}$ and we want to generate a Bell state $|00\rangle_{\rm L} + |11\rangle_{\rm L}$ (as previously mentioned, we get single qubit gates for free in dual-rail encoding, so we focus just on this input and output without a loss of generality). We will also assume that the first qubit is described by modes 1 and 2 and the second qubit is described by modes 3 and 4. Then all the possible states we can reach from the input state $|00\rangle_{\rm L}$ are described as

$$|00\rangle_{\rm L} = |1010\rangle_{\rm Fock} = \hat{a}_1^\dagger \hat{a}_3^\dagger\,|{\rm vac}\rangle \to (\sum_{j=1}^4 U_{j1}\hat{a}_j^\dagger)(\sum_{k=1}^4 U_{k3}\hat{a}_k^\dagger)\,|{\rm vac}\rangle = \sum_{jk} U_{j1}U_{k3}\hat{a}_j^\dagger \hat{a}_k^\dagger\,|{\rm vac}\rangle. \tag{1.49}$$

An example of the output state we are looking for is $|00\rangle_{\text{L}} + |11\rangle_{\text{L}} = |1010\rangle_{\text{Fock}} + |0101\rangle_{\text{Fock}} = \hat{a}_1^\dagger \hat{a}_3^\dagger + \hat{a}_2^\dagger \hat{a}_4^\dagger |\text{vac}\rangle$. We can see that the input state produces only separable expressions in terms of creation operators, while the output state has entangled creation operators, so the latter cannot be reached. $\qquad\square$

Notice that we are requiring the modes on which we are creating a Bell state to be dual-rail encoded. That means we require one photon per two modes of a single logical qubit to be present, and the four modes in which we are generating a Bell state to be found in this dual-rail basis. By heralding, in this lemma, but also throughout the thesis, we consider only postselection on the ancillary modes and not on the target modes. The lemma above is in contrast to postselected Bell state which is generated using only two photons in four modes that demonstrated a violation of the Bell inequality by picking to include only correlations that match those expected of a Bell state (so the postselection is occurring on the target modes) [54, 55]. It was later shown by Popescu et al. [56] through a generalization of the Bell inequality to include the results that are thrown away, that indeed this does violate the inequality in a satisfactory way. However, we are interested in creating entangled states that do not have to be destroyed at the point of postselection to give access to the necessary correlations (instead we are focused on "event-ready" entanglement, where a buffer of these entangled states can be created, ready to be used with no further postselection or measurements needed). We do also allow heralding on vacuum. The inputs are considered to be Fock states (so for example no weakly entangled state that is to be distilled to a Bell state), and generally throughout the thesis we want to start from the simplest type of states such as Fock states.

Moreover, we will see in Proposition 2.3.1 that not only can no Bell state but no entangled state can be generated under similar condition as in this Lemma using three photons. This would then imply that no entangled state can be generated using two photons either, as if it were possible, then it would also be possible with three photons by using the same passive linear optical network as for two photons and using identity on the third photon (postselecting on it).

Similar to single-rail and Hadamard, in dual-rail encoding, entanglement is so far only implementable non-deterministically (we will discuss this further in Section 2.1).

Dual-rail encoding of qubits is generally preferred over single-rail encoding because both of the dual-rail logical qubits are marked by the presence of a photon as opposed to the absence and presence. From a practical perspective the latter is more error-prone and the former can give easier error-detection.

Parity encoding error-correction

In parity encoding, a qubit is described by an equal superposition of states which can be split into "odd" and "even" [53]. For example, with two photons that carry polarization

information we can have

$$|0\rangle_{\mathrm{L}} = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle), \quad |1\rangle_L = \frac{1}{\sqrt{2}}(|HV\rangle + |VH\rangle). \tag{1.50}$$

This type of encoding is useful for error-correction. Consider a state

$$\alpha\,|0\rangle_{\mathrm{L}} + \beta\,|1\rangle_{\mathrm{L}} = (\alpha\,|HH\rangle + \beta\,|HV\rangle) + (\alpha\,|VV\rangle + \beta\,|VH\rangle), \tag{1.51}$$

with a probabilistic gate failing on this state with the failure case resulting in the measurement of the first qubit. Measurement of one qubit with result $|H\rangle$ would leave the other qubit in state $\alpha\,|H\rangle + \beta\,|V\rangle$ and similarly result $|V\rangle$ would leave the other qubit in the state $\alpha\,|V\rangle + \beta\,|H\rangle$. The second state can be easily corrected since we know it has occurred. The resulting state then preserves the information, reducing the original state by one qubit. As we will see in Section 1.4.3, linear optical quantum computing depends on probablistic gates, and doing a simple trick such as parity encoding can protect the information being computed and significantly improve the overall success probability. It can further improve the resources for LOQC when combined with a special set of parity gates called fusion gates (see Section 1.4.4).

The parity encoding can be further expanded onto N physical qubits, so that

$$|0\rangle_{\mathrm{L}} = \frac{1}{\sqrt{2}}((|H\rangle + |V\rangle)^{\otimes N} + (|H\rangle - |V\rangle)^{\otimes N}) \tag{1.52}$$

$$|1\rangle_{\mathrm{L}} = \frac{1}{\sqrt{2}}((|H\rangle + |V\rangle)^{\otimes N} - (|H\rangle - |V\rangle)^{\otimes N}). \tag{1.53}$$

We can now take this another step further and parity encode the new logical qubits into a logical qubit, so that $|0\rangle_{\mathrm{L}}^{(2)} = \frac{1}{\sqrt{2}}(|00\rangle_{\mathrm{L}} + |11\rangle_{\mathrm{L}})$ and $|1\rangle_{\mathrm{L}}^{(2)} = \frac{1}{\sqrt{2}}(|01\rangle_{\mathrm{L}} + |10\rangle_{\mathrm{L}})$, and so on, thus making the loss of a logical qubit, also error correctable.

Further encoding ideas

As an example of mixed encoding polarization and path mode, two qubits can be encoded in a single photon, with the polarization information representing a qubit and path information representing a second qubit. Redundant encoding is somewhat similar to parity encoding, and for polarization, for example, logical qubits are $|0\rangle_{\mathrm{L}} = |H\rangle^{\otimes N}$ and $|1\rangle_{\mathrm{L}} = |V\rangle^{\otimes N}$ [16].

As well as the above, there are much more complex ways of encoding qubits for the purposes of fault-tolerance, including surfaces codes [57] and the Raussendorf lattice [58]. The closer we get to realising a quantum computer the more important the topics of qubit encoding and fault-tolerance become.

1.4.3 Knill-Laflamme-Milburn scheme

The KLM scheme showed for the first time that is possible to build a theoretically scalable LOQC by adding the non-linearity needed to carry out entanglement of state through measurement. Previous to this, it was believed that a computer based on optics would have to use non-linear optical elements such as Zeno gates and Kerr nonlinearities [36]. In no known element have these nonlinearities been strong enough [12], thus this solution remains impractical.

Another attempt at making a LOQC possible before the KLM scheme that should be mentioned is the Adami-Cerf-Kwiat scheme [59]. They use 2^{n-1} paths to represent n qubits, and the qubits are path and polarization encoded. While this computer can perform the needed gates for universality, it is not scalable and further to this there is a lack of "non-locality" which raises significant concerns. Further, specific algorithms have been shown to work such as Shor's algorithm [60] and Grover's search [61], but none of these are a demonstration of a universal LOQC.

In KLM, dual-rail qubit encoding is used except when the teleportation protocol is implemented, in which case the encoding is changed to single-rail. The state preparation consists of using the single photon source to prepare a photon in one of the two modes. To complete the DiVincenzo criteria, a universal gate set is required. As mentioned before (Section 1.4.2), in the dual-rail encoding, single qubit gates are easy to implement. On the other hand, we also need a two-qubit entangling gate, such as CNOT or CZ. KLM proposes the idea of "non-deterministic quantum computation" which introduces gates that work some of the time, with a known probability. It is known when they succeeded by the measurement result.

The first component they add is a non-linear sign shift gate. A *non-linear sign shift gate* is a gate that takes photons in input mode 1 to output mode 1 as follows

$$\mathrm{NS}_x : \alpha_0 \ket{0}_1 + \alpha_1 \ket{1}_1 + \alpha_2 \ket{2}_1 \xrightarrow{\mathrm{NS}_x} \alpha_0 \ket{0}_1 + \alpha_1 \ket{1}_1 + x\alpha_2 \ket{2}_1, \qquad (1.54)$$

where x is the phase shift applied. We will be interested in $\mathrm{NS} := \mathrm{NS}_{-1}$ which takes the input to $\alpha_0 \ket{0}_1 + \alpha_1 \ket{1}_1 - \alpha_2 \ket{2}_1$. They described a way to perform NS_{-1} gate with probability of 0.5. Two of these gates can then be combined with the HOM effect to give a CZ gate. Then the CZ gate is constructed with probability of $\frac{1}{16}$.

They also make use of something called the "teleportation trick" introduced by Gottesman and Chuang [62], increasing the probability of a successful CZ gate to $\frac{1}{4}$. This protocol uses the properties of the Clifford group to move the entanglement gate from entangling input qubits to applying CNOT or CZ on the ancilla states instead. Then the entanglement can now be teleported through to the input qubits that were intended to get entangled. This trick allows the probabilistic part of the computation (the entangling gate) to be done "off-line", preparing many entangled ancillas beforehand, ready to be used when the resource is need as part of the main computation. Benefits of this are two-fold: one is

that we can move the probabilistic part of the computation to be pre-computed and use only successful ancillary states, making the gate deterministic and increasing the overall success of computation, the other is that we can make the computation be non-destructive. In linear optics though the Bell measurement which is need for the teleportation trick is non-deterministic as well, so at best, the probability of successful gate teleportation will be limited by the number of Bell measurements needed to be carried out. Finally, they combine this trick with clever ancilla states to make the gate be near-deterministic at the expense of the size of the ancilla state, proving efficiency in the sense of polynomial resources.

They mention various ways in which the failure of a gate or a Bell measurement can be used to error correct. There are also codes which would stop from losing information every time a gate was applied, thus the gates can actually be tried more times without the need of making them near deterministic [14]. After this landmark paper, a few simplifications of the NS gate were proposed, either by reducing the number of ancillas or beamsplitters, or with higher success probability, such as for example an implementation of a CZ gate with success rate $\frac{2}{27}$ [63]. Another result that should be mentioned is from 2006, by Spedalieri et al [64] in which the teleported qubits are dual-rail polarization qubits.

Already in KLM we see the importance of generating Bell states, which can be produced on-demand or offline, as they will be consumed as a resource to teleport through the entanglement into the circuit when needed. However, the number of Bell pairs required per logical CZ gate is on the order of 10^6 [18]. While the KLM scheme proved it was scalable from the perspective of information theory, from an implementation perspective, the resource overhead is large and the level of control is too high, due to the scheme being a very large interferometer over which the photons need to be phase controlled. Further, the depth of the quantum circuit is too much considering the exponential loss of photons, ideally it should not be more than $O(1)$ and in KLM it is poly(n). These challenges did not just require better engineering, they required a change of computational model.

1.4.4 Measurement-based quantum computing

The search for new and improved computational models continued, most prominently in 2001, Raussendorf and Briegel introduced one-way computing or MBQC using cluster states [15]. Cluster state (graph state) is, possibly a very large, pure state represented as a lattice (graph) structure where vertices represent qubits and the edges their entanglement with their neighbours (see Figure 1.7). While cluster states are sometimes used interchangeably with graph states, by the definitons of Raussendorf [65], graph states are a generalization of cluster states which are restricted to connected subsets of a simple d-dimensional lattice. The computing is then done through measurements on the nodes of these pre-generated lattices (see Figure 1.6). In 2003, Yoran and Reznik improved the KLM scheme with ideas similar to the cluster state model [66]. They used the polarization and path qubit encoding mentioned earlier. Creating long entangled chains and then using teleportation like in KLM, they transfer the computation from that chain to the state of
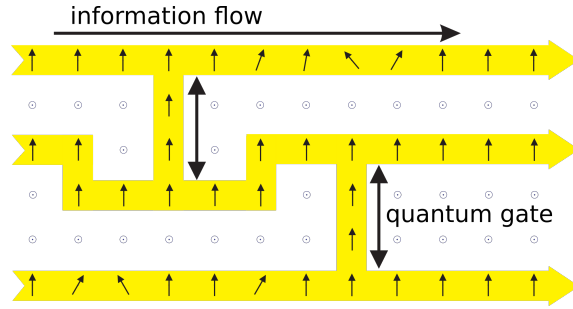
Figure 1.6: In a measurement-based quantum computer, gates are implemented through suitable measurements on the nodes of the lattice. A single qubit (and the evolution of it) is represented by one row of the lattice (with the leftmost node being the input state and the rightmost node being the resulting state). Multi-qubit states are then represented by multiple rows entangled together, and two qubits gates are then implemented by measuring multiple rows at the same time. Image taken from [15], copyright(2011) by the APS.
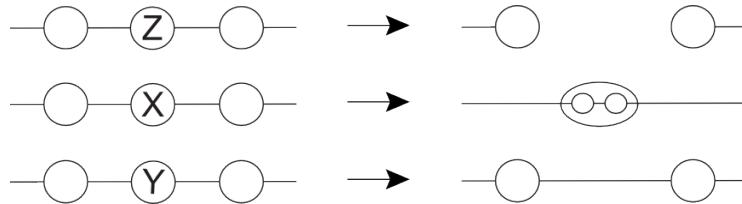
Figure 1.7: The image on the left shows segments of one dimensional cluster states. The nodes connected with lines are entangled together. The image on the right shows the result after application of the suitable Pauli measurement. Image taken from [16], copyright(2005) by the APS.

the system. To add a link between two qubits in their model, they did not need a CZ gate to be near-deterministic, but just succeed with a probability higher than $\frac{1}{2}$, so the expected value of the change in number of qubits in a chain would be positive. This improved on the number of consumed Bell pairs to the order of 10^5 per logical CZ gate [18].

This was followed by a measurement-based computing scheme by Nielsen in 2004 with large lattices of entangled states and computation carried out through measurement on those lattices [67]. Taking the KLM scheme implementation of CZ gate, the paper presents a cluster state equivalent and thus shows that MBQC is universal. He introduces the idea of using failure cases of gates to save resources, which was another major breakthrough. He also describes the idea of microclusters, located at the end of our larger cluster state. These microclusters are small and have a lot of links, thus further protecting from failure affecting the rest of the cluster. This implementation finally removes the need for physical depth photons go through present in KLM [67]. The number of Bell pairs needed for a logical CZ gate is therefore further reduced to order of 10^4.

The Browne-Rudolph scheme introduces two fusion gates, so called because they can fuse clusters together [16]. Essentially, they are parity gates, but Browne and Rudolph have an innovative way of applying them and give insight on how they can be used for cluster computing, from resources being used more efficiently through control of failure modes, and simpler experimental setup. They also show that one dimensional cluster

states are not enough for universal quantum computation. To create vertical links between 1D clusters another type of fusion gate is introduced. In the failure case of type I fusion gate, Z-measurement is performed which is not desirable as it would destroy big clusters (Figure 1.7). The gates make a major improvement on the resources needed for computation with an average of 52 Bell pairs needed to add second dimension to a cluster (which would be an equivalent of entangling two qubits in the circuit based scheme). The fusion gates are now one of the most important building blocks for generating other entangled states, as well as for building very large lattices that MBQC can be carried out on.

1.4.5   Ballistic percolation based architectures

Both the KLM and MBQC presented so far rely heavily on active switching. The newest LOQC schemes are passive, so called "ballistic" schemes which use percolation theory to build clusters. This new idea started with Kieling's paper [17], which shows that there is a significant amount of long-range entanglement in a non-ideal cluster, as long as there are enough links (successfully entangled states) within the cluster. Percolation theory is then applied to determine what the necessary threshold for the number of links is. This proposal really opened up the field of LOQC to the very real possibility of building a quantum computer [13], as now we can create large cluster lattices "on-the-go", which are measured and corrected on one side, whilst being built up on the other. While the original proposal depended on a source of GHZ state of five and more qubits, it was recently shown that a source of three qubit entangled states are enough and also they significantly improve the number of Bell pairs needed to build such architectures. Again, due to the change in the computational model, it is hard to compare the resources of the new schemes to those of the old ones, especially as the new schemes seem to require significantly more Bell pairs [18], however the ballistic schemes have significantly greater robustness and are practically more feasible.

From the discussion on LOQC we can see that entanglement, whether in the form of entangling gates or Bell pair sources, is a major focal point of the work and one of key engineering challenges. However, universal quantum computing is not the only goal. For a universal quantum computer we will need many physical qubits, due to the need for error correction. In the mean time though, it might be possible to do computations on a smaller number of qubits or with smaller number of particles to demonstrate a computation which is not possible on a classical computer. We will describe one interesting contender for this on the linear optics based platform called boson sampling, which is good example of the power of quantum interference for computation.

1.4.6   Boson sampling

As the goal of a universal quantum computer through linear optics remains ever distant due to extensive engineering challenges, an analogue quantum computer offers the chance of demonstrating the advantages of quantum over classical computing. Let us recall from Section 1.3.4, that scattering probabilities of linear optical interferometers can be computed

using permanents of matrices. It turns out that permanents of a matrix are hard to compute (more specifically they are in a #P, which is as at least as hard as the famous NP class), so it is natural to ask the question as to whether or not this can give photonics based platforms some kind of computational advantage, leading to the problem of BOSONSAMPLING.

There are a few different algorithms for computing permanents. The naïve way consists of carrying out the permanent calculation as written, that is for each of the permutations, finding the product over the matrix elements and then summing over all the permutations. The naïve approach requires $m!\,m$ arithmetic operations for a square matrix M of size $m \times m$, and quickly becomes intractable on a classical computer.

A more commonly used method is the so-called "Ryser formula". Let $\mathcal{M}_k$ be the set of all matrices obtained by deleting k columns from matrix M. Let Σ_k then be

$$\Sigma_k = \sum_{\mu \in \mathcal{M}_k} \prod_{j=1}^{m} \sum_{l=1}^{m-k} \mu_{jl}, \tag{1.55}$$

where (μ_{jl}) is some matrix which had k columns removed from the original matrix M. Then the Ryser formula gives

$$\mathrm{perm}(M) = \sum_{k=1}^{m-1} (-1)^k \Sigma_k. \tag{1.56}$$

This formula works based on the inclusion-exclusion principle, significantly diminishing the number operations needed to $O(2^m m^2)$. There are a few other methods for computing permanents [44, 68–71], which offer a speed-up but leave the problem in #P, as well as approximations of permanents [44, 72] that can be evaluated in polynomial time but that work on a specific subset of cases. Aaronson and Arkhipov show that not only is the computation of the permanent in #P, but so is the approximation of the permanent within a constant factor.

However, permanents naturally show up as scattering probabilities of bosons [44, 45] (see Section 1.3.4). Even though the permanents arise naturally as scattering probabilities, this does not automatically give us an algorithm to calculate a permanent of a matrix. This is due to the problem of extraction of "classical" information from a quantum computer, such as a permanent – in some situations it can require an exponential number of samples to get the result to the required accuracy. In Section 1.3.4 we also talked about how the probabilities of a classical state can also be described by a permanent, so one might wonder how this can be reconciled with the Church-Turing thesis. In the case of a classical state, unlike the complex matrix in the case of indistinguishable bosons, the matrix we are dealing with is a nonnegative real matrix, therefore there is an approximation of the permanent in probabilistic polynomial time, so there is no controversy here. Instead, Aaronson and Arkhipov turn towards sampling problems, drawing a sample from a distribution that is close to the actual bosonic distribution.

Let the input state be that of N photons, in d modes where $d \gg N^2$. In the original

paper the claims are proven for d of the order $N^5 \log^2 N$, however in practice $O(N^2)$ is commonly used. The input is also taken to be that of coincident bosons, that is if $|\underline{s}\rangle$ is the input state, $s_j = 0$ or 1 for all $j \leq d$. As there is $d \gg N^2$ modes, the output is also likely to be coincident (or "collision-free" as the authors refer to it) due to the boson birthday paradox [73]. Now taking a Haar random matrix U in d modes as above, the $N \times N$ submatrices of U are approximately Gaussian. On these submatrices the authors use the "Permanent of Gaussian" conjecture, which states that the estimation of the permanent of a Gaussian matrix is also #P hard. To bridge the gap to the scattering probabilities, they make another conjecture of "Permanent Anti-Concentration Conjecture", which stops the Gaussian matrices from being concentrated around zero.

Since the paper introducing the boson sampling problem, there have been many experiments attempting boson sampling on a small number of photons (from bulk optics, integrated chips, to scattershot setups). Originally it was believed that with 20 to 30 bosons, it would be possible to demonstrate quantum advantage, with some of the lowest estimates being as low as 7 bosons in 50 modes[74]. The recent experimental records are five photons in nine modes in a standard linear optical setup [75], seven photons with two lost photons in 16 modes [76], or six sources in 13 modes in a scattershot setup [77]. However, more recently, a paper by Neville et al. [70] has demonstrated that scale of the system needed for quantum advantage is more in the range of 50 bosons or possibly even more. Using statistical modelling tools, they demonstrate boson sampling of 30 bosons carried out on a departmental server, and predict a 50 boson case could be calculated in approximately 10 days on a supercomputer.

Considering that we need a significant number of modes more than bosons (on the order of N^2), the feasibility of boson sampling for near-term demonstration of quantum advantage has been somewhat diminished. However, it has been shown that there can be some loss of photons where boson sampling will still work (although less than $O(N)$ leftover photons make boson sampling classically simulatable [71, 78]). It is also quite likely that far less modes can be used, possibly even on the scale of $O(N)$, but the same tools cannot be applied to prove BOSONSAMPLING still remains classically intractable. It also needs to be proven that BOSONSAMPLING remains hard outside of collision-free subspace as well, allowing for the reduction of the number of modes. Recently Gaussian boson sampling has been explored in more detail, as it turns out to also be hard [79] and allows for a few different applications [80]. For the traditional boson sampling though, beyond significance in terms of demonstrating quantum advantage, a possible application is in terms of verification for large scale implementations when classical simulation becomes unfeasible.

To bring boson sampling closer to a working demonstration of quantum advantage and to allow it to be useful as a verification technique, a better understanding of how photon loss, distinguishability and similar errors affect boson sampling is needed. The framework presented in Chapter 3 can be used for modelling both loss and noise in boson sampling [81]. There, the authors develop a scheme for classical simulation of boson sampling by treating

noise and loss as partial distinguishability. This scheme, while not asymptotically best in terms of run-time, offers substantial improvement in runtime for near-term experiments of 50 to 100 bosons.

For a more detailed look at boson sampling, refer to a recent review [80].

1.5 Introduction of errors through distinguishability

In either of the two computing models (universal and boson sampler) introduced above, errors are a major challenge[18, 82]. While certain techniques that might defend against e.g. loss such as encoding (Section 1.4.2), and engineering better systems helps [18, 83], there are other theoretical paths that can be followed as well. We will focus on one such path, which is the modelling of distinguishability. Distinguishability in itself is a problem which can introduce potentially fatal errors, as we can see on the example of boson sampler with significant amount of distinguishability, the computationally hard problem gets reduced to a much easier one [34, 81]. Moreover, modelling distinguishability in first quantization as we do in Chapter 3 also allows us to model further errors such as loss through the tracing out of qudits [81, 84]. Here we will review major results behind distinguishability modelling, detection, and some more general multiparticle interference results. We start with the seminal HOM effect that has two major parts to it, one is the signature of multiparticle interference and the other is characterization of the distinguishability of the particles. The literature that then follows on both multiparticle interferometry and distinguishability of bosons is highly linked, and we will mention the ideas behind generalizations of both.

1.5.1 Hong-Ou-Mandel experiment
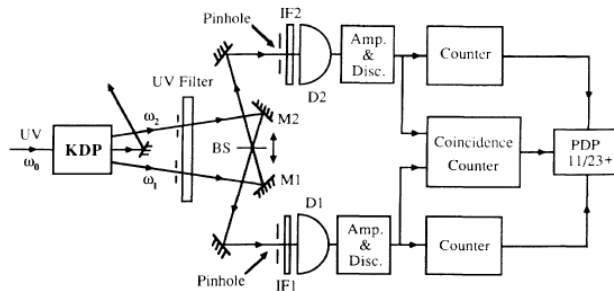
Figure 1.8: The original setup from the Hong, Ou, Mandel paper. The pump is passed thorough a crystal, with a small number of pump photons being split into signal and idler photons which then impinge on the beamsplitter (following the law of conservation of energy and momentum). By moving the beamsplitter the path length, and therefore the timing, of the two photons changes, introducing distinguishability ("which-way" information) into the system. The two photons are then detected and the coincidences (both detectors clicking) are counted. Image taken from [11], copyright(1987) by the APS.

While the relevance of particle distinguishability was theoretically well documented,

and even some initial experiments established the presence of interference that cannot be described through classical theory, it is not until the seminal HOM experiment in 1987 that the physical and practical significance was demonstrated [11]. Therefore we start with a description of the experimental setup and results reported in the HOM paper.

In this bulk optics experiment they used a single spontaneous parametric down-conversion (SPDC) source, generating a signal and idler photon (see Figure 1.8). This is a probabilistic process commonly used to generate single photons exploiting second-order nonlinearity of a material. While SPDC is used in bulk optics (using crystals such as KDP, KTP, BBO, $LiNbO_3$, KDP) and waveguides, optical fibres do not allow the second-order nonlinearity, so a process called spontaneous four-wave mixing (SFWM) is used instead. In the case of SFWM, third-order nonlinearity is exploited and the process has been demonstrated in a variety of fibres (dispresion shifted fibre, photonic crystal fibres, birefringent single-mode fibre, silicon-on-insulator, ring resonators) [21, 85].

Two photons create using an SPDC source were then impinged on a 50:50 beamsplitter, and the rate of coincidences detected. However, the beamsplitter was mounted on a moveable stage, with which they could offset from the symmetric position of the two paths, by lengthening the path of one of the photons. In turn, this would change the arrival timing of one of the generated photons. This timing difference allowed for the length of the photon wave packet to be determined, and offered the ability to measure the difference in arrival of 50 fs with accuracy of 1 fs. Controlling this timing difference allowed for the control of the distinguishability of the photons.
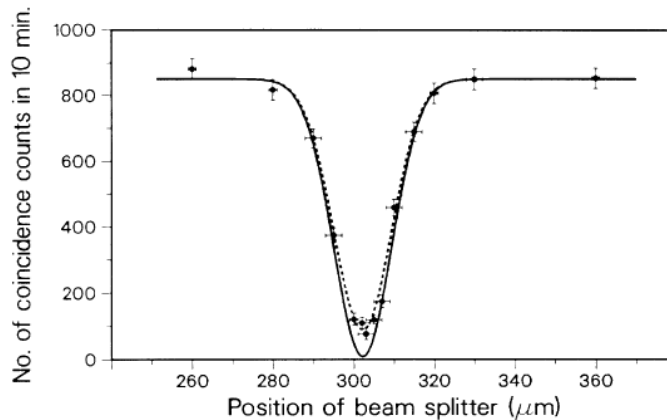
Figure 1.9:   The results of the experiment whose setup is seen in Figure 1.8. We can notice the dip in the coincidence count between the two detectors which are detecting signal and idler photons. The position of beamsplitter has the effect of changing the arrival timing of two photons, with the bottom of the dip representing nearly perfect alignment in the timing. The dip is predicted by quantum but not classical statistics. The full line represents the theoretical values, while the dashed line is the fit to the experimental data. Image taken from [11], copyright(1987) by the APS.

The photons were generated such that they had the same frequency and polarisation, and the timing could be adjusted, therefore producing the "HOM dip", which can be seen in Figure 1.9. While in the classical state where coincidence should still be detected in

half of the situations, for two indistinguishable bosons the coincidence measurement drops to zero. For the theoretical treatment of results presented here, we refer to Section 3.2.3 (and also to Section 1.3.4 which compares quantum to classical particles). The idler and signal photons created from the SPDC source respect the law of conservation of energy and momentum. However, as long as the quantities are conserved, the actual wavelength and momentum each has can be varied. Moreover, it is the energy of the original pump photon that is conserved, which will also come from a certain probability distribution depending on the type of pump laser being used. If we want signal and idler photon to be indistinguishable, then we have to carefully select them amongst those that are generated (by filtering for wavelength, polarisation, ands spatial modes). If we want to use signal photons from two SPDC sources (or even from the same source), we have to filter the spectrum of those photons to make sure they are indistinguishable (lowering the success rate of this process further), but also the timing can be a problem. This is commonly fixed by using a pulsed laser beam, where the pulses are then used as a synchronization method. There are many variables upon which the visibility of the HOM effect depends, which led to new SPDC sources of different types and new demonstrations of the HOM effect using different techniques.

Recent HOM experiments

Since the HOM experiment, other experiments have either repeated these results or demonstrated the same effect with other platforms and techniques. Shortly after, a slightly different setup from that of the original HOM experiment confirmed the same effect [86], and again with a setup using an MZI [87].

A series of experiments were done as stepping stones towards two particle interference from two independent photon sources: same pump passing through same SPDC crystal twice [88]; two photons from different pulses of a mode-locked laser through the same SPDC [89]; two photons generated in one quantum dot [90]; two separate SPDC crystals using the same laser [91]; and finally two independent sources [92].

Still using an SPDC source, Walbron et al. [93] demonstrated the HOM effect with two photons in a multimodal setup. An atomic HOM experiment was achieved using bosonic atoms from a Bose-Einstein condensate of Helium-4 atoms [94], and again using tunneling and optical tweezers [95].Quantum dots as a source have been demonstrated in a variety of setups as capable of producing two photons that give a HOM effect [90, 96–99]. Finally two-photon interference from a set of other interesting sources has also been achieved such as: independently trapped atoms [100]; two ions from different traps [101]; single molecule [102]; two separate nitrogen-vacancy centres in diamond [103]; quantum memory [104]; surface plasmon polaritons [105]; microwave frequency independent source [106].

The fermion two particle interference effect has also been investigated, such as two independent electron sources [107], fermionic atoms in atomic Fermi gas [108], neutrons [109], and collision of electrons [110]. For a review of intereference experiments we direct the reader to Pan et al. [111].

1.5.2 Three photon literature

Considering the important position the HOM effect has had on the development and analysis of optics experiments, there was a natural desire for its generalization. The apparent next step towards generalization of the HOM effect, and a more comprehensive understanding of distinguishability and its effects, was the addition of another mode and/or photon. As the HOM effect involved a balanced beamsplitter, the first interferometer to be considered of significance was a balanced tritter. In [112], the authors investigate the coincident counts of two photons impinging on a tritter and demonstrate the presence of the dip between pairs of ports. Then in [113], the author theoretically investigates the HOM effect for three photons on a tritter, showing that the effect is not quite the same, for example in the case of indistinguishable photons the output coincidences do not cancel out. Instead the author proposes slight tweaking that allows for a more analogous behaviour to HOM interference. Soon after, the group theory of SU(3) is studied in more detail to extend the interference effects found in beamsplitter to that of a suitable tritter [114].

It is not until Tan et al. [115] and de Guise et al. [116] though that the group theoretical approach offers a more extensive model of three photon scattering probabilities using the group theory of SU(3). They find scattering probabilities of the coincident subspace and its relation to the immanants of the given SU(3) matrix [115]. Almost simultaneously to this, Spagnolo et al. reported this three photon interference, experimentally achieved in a tritter [117]. They characterize the device using the two photon HOM effect and then, they use coincident input with delays on different ports making various sets of the photons distinguishable, comparing experimental results to theoretical expectation. Additionally, Metcalf et al. demonstrated the first on-chip three photon non-classical interference [118]. Another tool developed for multiparticle interferometery was reported by Chaboyer et al [119]. They made a three path counterpart of a Mach-Zehnder interferometer, however they did not conduct any three photon distinguishability experiments, but solely focused on characterization using two photon HOM effect. Mahrlein et al. also conducted a theoretical study of integrated devices which demonstrate interesting three photon quantum interference effects [120].

More recently, an experiment using a tritter as the interferometer, spatial modes for the external degree of freedom, and timing along with polarization for internal degrees of freedom, demonstrated that pairwise distinguishability is not enough to explain the behaviour of distinguishable bosons (contrary to the previous experiments which assumed so) [28]. Instead, a collective phase of three photon interference named a *triad phase* needs to also be measured to characterize and describe the multi-particle interference and distinguishability fully. An explanation of how the triad phase arises in a group theory focused picture, using immanants, can be found in [121]. In a three photon Franson type interferometer using energy-time entangled photons, Agne et al. also demonstrated genuine three photon interference, with no two-photon interference effects [122]. This is another paper that raises the importance of multiparticle interference effects beyond only pairwise. Another interesting paper shows three photon interference in the time domain using an

atomic medium [123].

However, while the above mentioned works focused on examining the properties of the balanced tritter (commonly referred just as the tritter), it was not in any way proven that the tritter would actually be optimal generalization for the HOM effect on three photons (using the suitable figure of merit for each of the generalizations) nor would it necessarily provide a way to generalize for any number of photons.

1.5.3   Generalization of the Hong-Ou-Mandel effect

Given coincident input of a QFT interferometer (see Section 1.3.2), Lim and Beige showed that when the number of particles N is even, the coincident output is always suppressed [124]. Similarly, given an input state with an equal number of photons impinging on both input ports of a balanced beamsplitter, all the output states where there is bunching an odd number of particles is suppressed [125, 126].

Ou shows a relation between temporal distinguishability of N photons and quantum interference [127], and then continuing with the same formalism and setup, the author examines a way to characterize temporal distinguishability through interference of an N photon state with a single photon state [128]. However, interest in this field does not rise significantly until boson sampling triggers a more general interest in multiparticle interferometry.

In [46], the author introduces a second, internal degree of freedom to model distinguishability of particles. Even for as few as two modes with many particles, the complex world of parital distingusihability can be demonstrated [46, 129]. They notice that the area in between completely indistinguishable and completely distinguishable particles is not just a simple extrapolation demonstrating the non-monotonicity. This is followed up by an experimental demonstration using a setup of four photons in two modes [129] and analyzed for a proposed four photons in four modes scenario [130]. The orthonormal basis for this second degree of freedom is found through Gram-Schmidt orthonormalization, which then allows the scattering probability to be expressed even for partially distinguishable pure states and summing over all the possible internal degree of freedom configurations giving this chosen scattering pattern in the external degree of freedom. While the basis constructed here specifically follows the Gram-Schmidt procedure starting with the temporal component of the photon in the first mode, then finding orthogonal projection of the second mode, etc., the idea where the internal degree of freedom characterizes distinguishability can be generalised further [33, 131]. The scattering probabilities can then be suitably expressed in any basis, allowing us to treat the internal degree of freedom as just another mode (at least for pure states).

It was noticed soon after, that much in the same regard that the permanents are related to the scattering probabilities of bosons (see Section 1.3.4) and determinants are related to the scattering probabilities of fermions (see Section 1.3.4), immanants, a generalization of permanents and determinants, are related to the scattering probabilities of distinguishable particles for the coincident subspace of three photons [115, 116] and more photons [31]

(see Section 5.1). While this offers a beautiful way to interpolate the gray area of partial distinguishaility, it was soon pointed out that this is not necessarily a scalable way of expressing partial distinguishability [132], at least not without further work.

The formalism we have seen so far is focused on modelling partial distinguishability using pure-states, which is not enough to treat partial distinguishability fully [133]. An important issue here is the lack of the ability to model mixed states appropriately due to the way the basis of the second degree of freedom is orthonormalized. Instead in [134] and [132], Shchesnovich focuses on density-matrix based input states with possible distinguishability up to one photon per mode, allowing for mixed initial states. Shchesnovich introduces the idea of the measuring of a certain output pattern through an operator that projects on the suitable space without differentiating the internal states. The scattering probability is then just the expectation value. The author also points out that some of the results acquired by establishing a relationship between an immanant of the matrix and scattering probabilities, can also be restated as results on the permanent and a matrix slightly different than in the usual indistinguishable photons approach (as in Section 1.3.4).

Tichy then takes these ideas, introduces a *distinguishability matrix*, and multi-dimensional permanents, to allow for efficient modelling of partial distinguishability. The distinguishability matrix contains information on the pair-wise distinguishability of all the $(N-1)N/2$ pairs from a set of N particles. A few of the papers mentioned here introduce matrices to quantify distinguishability [32, 33, 116, 132, 135], but they can all be related to each other in the case of pure states [133]. The projector Shchesnovich introduced to model measurements applied to this framework results in a multi-dimensional tensor permanent constructed using the distinguishability matrix (permanent over N^3-dimensional 3-tensor). This allows the standard indistinguishable calculations such as Ryser's algorithm to be carried out even in the partially distinguishable case. Tichy also links the permanent of the distinguishability matrix to a few important questions such as the deviation of the scattering probability of partially distinguishable to the completely indistinguishable case. The bounds on the deviation are then tightened in [136].

We have mentioned an approach showing the relation of immanants to partial distinguishable states which relies on the representation theory of SU(N). Further work has been developed around this topic, for example, generalizing the types of input and output states for which the immanants can be used to describe scattering amplitudes [137]. In a related paper, [138], Dhand et al. present an algorithm to construct the functions needed for the calculation of the immanants. More recently, Khalid et al. prove that the coincidence rate of partially distinguishable states can always be expressed in terms of immanants [139], generalizing [31, 115, 116]

In a similar vein, the introduction of the relevance of the triad phase we discussed in Section 1.5.2, Shchesnovich and Bezerra [140], show that quantum interference of N distinguishable bosons in pure states have $(N-1)(N-2)/2$ independent triad phases. Moreover, for four or more particles, to allow for genuine ($N \geq 4$)-particle interference, the interference of k particles where $3 \leq k \leq N-1$ is prohibited. Their results again suggest

that more than pairwise correlations are essential for the full interference landscape.

In other interesting results that are not exactly related to modelling distinguishability, but that do deepen our understanding of boson sampling with partially distinguishable photons, it is worth mentioning [34]. This research uses the previously mentioned formalism of Tichy with multidimensional permanents and a distinguishability matrix [133] to demonstrate an efficient algorithm for boson sampling with partial distinguishability. An adequate amount of distinguishability is needed to allow the boson sampler with partial distinguishability to be broken down into a combination of smaller interferometers that act as standard boson samplers. However, if the partial pairwise distinguishability is decreasing fast as the number of added photons increases, the complexity of boson sampling remains intact.

Subsequently, Brod et al. developed a simple distinguishability witness based on pairwise distinguishability[141]. The idea behind it is to take a single photon and apply a QFT to it and a $N - 2$ vacuum ancillas. Then, each of the remaining $N - 1$ photons interacts with one of the output modes of the QFT on a balanced beamsplitter, allowing for pairwise comparison of distinguishability. This was followed by an experiment in which they use this method to show four photon distinguishability [142]. The paper also uses the result from Galvao and Brod, where certain pairwise overlaps allow bounds for the remaining unknown overlaps [143].

Lastly, an interesting paper to mention which does not tackle the problem of modelling distingushability directly, but is related to the approach we are exploring here is [144]. In this paper the authors look at the behaviour and scattering probabilities of particles they name "immanons". These particles, unlike bosons with symmetric exchange symmetry, and fermions with anti-symmetric exchange symmetry, can have mixed exchange symmetry. This type of behaviour can be simulated by distinguishable bosons or fermions and will be discussed later.

For more results on the topic of boson sampling, we refer to the recent review [80]. The tutorials from 2014 by Tichy and 2019 by Walschaers are also good introductions on the topic of multi-particle interference [26, 79]

1.5.4 Suppression laws

As already mentioned, the topics of distinguishability and multiparticle interferometry are highly related. Here we discuss an area of research which does not concern modeling distinguishability, but it does give insights into behaviour of distinguishable particles. Suppression laws are concerned with suppression events that occur due to quantum interference [145][146, and references therein]. There are other related topics that explore multi-particle interference, its signatures, as well as ways for it to verify or validate boson sampling [79]. However, we give a brief description here as the suppression laws are connected to the research presented in Chapter 4 and we will talk about it more as future work prospect in Chapter 5.

The suppression laws compare two states of interest. The first state is the "classical"

state, that is the state of completely distinguishable photons (see Section 3.3.3) in some input configuration, most commonly a Fock state. The other state is the "quantum" state, that is the state of completely indistinguishable photons (see Section 3.3.1), in the same input configuration. Then, the scattering probabilities of the quantum and classical states after some fixed choice of interferometer, U, are compared. Most of the papers mentioned here do not just examine indistinguishable bosons versus completely distinguishable particle, but also indistinguishable fermions.

While the formulation of suppression laws was first defined by Tichy et al. in 2010 [145], some earlier papers have provided a basis for this idea. When it comes to three photons in three modes and suppressing coincident events, some of the earliest work is done by Campos in 2000 [113]. Then, in 2005, Lim and Beige look at QFT in the more general setting of N photons in $N = d$ ports and find that the coincident events are suppressed for all even d [124]. However, the 2010 work [145] and the follow up work from 2012 by Tichy et al. [147] are viewed as really defining the concept.

The interferometers investigated have some symmetric properties to them: QFT [145, 147], Sylvester interferometer [148], Jx matrix[149, 150], hypercube matrix [151], interferometers constructed using permutation operators which leave the input state invariant [146, 152]. This is not surprising as the scattering probabilities of the completely indistinguishable state are computationally hard, and thus seemingly natural that there are not many analytical results that can be evoked. These papers do not just restrict to interferometers with symmetric properties, but commonly the input states examined will have symmetric properties as well.

While in this thesis we do not dig any deeper into suppression laws, we can link the idea of unambiguous discrimination of distinguishable states which is the topic of Chapter 4 with this prior work done on suppression. We will then see in Chapter 5 how this could allow for more general input state and matrices to be examined.

These breakthroughs in linear optical quantum computing (both universal and boson sampling) have made photonics an appealing contender for the quantum computing revolution. However, significant engineering challenges remain. We can see that interference plays a substantial role on any optical platform, therefore theoretical investigations of entanglement generation and effects of distinguishability are not only well motivated, but can offer major improvements. In this thesis we focus on finding novel ways of understanding both restrictions on generation of entanglement and behaviour of distinguishable bosons. The building blocks of an LOQC are small entangled states, so we investigate optimal ways of generating such states, specifically Bell states, but we also look for more general bounds on entanglement generation. Similarly, we saw that scattering probabilities of classical particles is so different from that of quantum particles, that sampling from its distribution is in two very different complexity classes. The effect of distinguishability of quantum particles is not just an introduction of errors into the system, but it can completely invalidate the computing speedup it offers. New ways to model distinguishability that open routes to investigate these effects, or allows for filtering or detection of the magnitude of

these effects are essential. Here we present a framework which grants such a new approach, and we immediately apply it to the problem of discrimination of distinguishable states, offering a possible generalization of the famous Hong-Ou-Mandel effect.

GENERATING ENTANGLEMENT WITH LINEAR OPTICS

Research into quantum technologies has gained significant momentum in the last several years, with applications ranging across metrology, communications, security, simulation and computation [10, 53, 153, 154]. As we saw in Chapter 1, one of the important resources lying behind many of these advances is quantum entanglement [155, 156]. Long before it was a potential technological resource, entanglement was studied as one of the phenomena lying at the foundations of quantum mechanics [157–159]. That there exist non-classical correlations between physical systems is now well established, while how best to generate, verify, and quantify such entangled states in practice is an ongoing field of activity. What is practical in any given situation depends on the physical platform under consideration; here we will be interested in the generation of entanglement using linear optics and postselection.

As we saw in Section 1.3, in linear optics we study collections of optical modes, modelled as harmonic oscillators whose excitations correspond to photons. Interactions are restricted to Hamiltonians that leave the total number of photons fixed, giving rise to unitary transformations on modes (interferometers), as well as possible measurement and postselection of quantum states. This realization introduces an interesting set of constraints on the entanglement problem. Most work to date focuses on either single- or dual-rail encoding of photons into two-dimensional qubits (see Section 1.4.2), and then applying the usual approaches to quantum computation such as the circuit model (see Section 1.4.3) or measurement-based schemes (see Section 1.4.4). Gates are carried out via ancilla modes and photon detection measurements [53]. The dual-rail encoding is the commonly accepted standard for quantum computation with linear optics, and allows us to discuss entanglement in terms of standard concepts such as Bell and GHZ states [19, 25, 53]. However, the requirement of postselection means generation of such states is nondeterministic, and the probability of success is often low; for example, the best known Bell state generation scheme has success probability of $1/4$ [160] and if the postprocessing technique known as procrustean distillation is not allowed, then the probability drops to 0.1875 [161]. When we consider the number of Bell states needed to construct two-dimensional cluster states [25], the requirements can be quite daunting, though promising proposals exist (see Section 1.4.5) [13, 18].

This motivates the study of entanglement generation in linear optics more generally; in particular, it is natural to consider entanglement between two subsets of modes, foregoing encoding altogether. While this is currently not the preferred way of generating entanglement, any bounds that can be found present fundamental limits on linear optical architectures, as well as for other quantum information processing tasks such as boson sampling [162]. A different perspective on this issue, which considers bosonic entanglement in terms of observables, can be found in (e.g.) [163, 164].

In this Chapter we will consider two main themes regarding bipartite entanglement in linear optics; that where the parts are encoded qubits, and that where they are collections of modes. Section 2.1 establishes what the known results on generation of entanglement are and their relevance. Section 2.2 introduces the setup and notation used throughout. Section 2.3 examines qubit entanglement within the standard linear optical dual-rail encoding. When we speak of dual-rail encoding, we mean qubit states that are post-selected such that there is exactly one photon in each pair of modes. In Section 2.4 we compare qubit and mode entanglement, including an investigation of the expected average entanglement over uniformly (Haar) distributed interferometers. In Section 2.5 we shift our focus to mode entanglement, considering bipartite systems made from two sets of optical modes, Alice and Bob, with a fixed total number of photons.

2.1 Previous entanglement results in linear optics

As we saw in Section 1.4, entanglement is fundamental for LOQC, whether in terms of entangling gates in the standard model, or Bell pairs building up cluster states in the MBQC approach. It is not surprising then that there is a body of research looking into optimization of CZ and CNOT gates, postselected and "event-ready" Bell pairs. For an extensive review of the resources needed for various LOQC schemes, we refer the reader to [25]. Gimeno-Segovia also covers important results on Bell measurements, which are closely related to fusion gates, and therefore related to the problem of scaling up entanglement. A thorough description of fusion gates and their inputs and outputs can be found in [25, 53]. Also Gimeno-Segovia in [25] talks about boosted fusion gates, which improves the probability of success of the original gates.

In this Chapter we are interested only in generation of heralded Bell pairs, as well as a more general generation of mode entanglement, with no encoding. While there are other possible approaches for generation of entanglement (such as postselected entanglement), this will not be talked about here (instead we refer an interested reader to [165]).

2.1.1 Entangling gates

The KLM discovery of feasibility of LOQC with the addition of postselection started the research into optimization of LOQC gates. When it comes to optimization of these gates we are usually interested in optimizing the simplest linear optical elements, so we assume access only to photon sources, PNRD, and static linear optical network of any

size with any ancillas, but no feedforward (no switching or teleportation). As we mention in Section 1.4.3, KLM showed how to create an NS gate using three beamsplitters, one phase shifter, two extra ancilla modes, and one ancilla photon. Two of these gates can be combined succeeding with probability of $\frac{1}{4}$ giving a CZ gate with probability of $\frac{1}{16}$. In the paper they also demonstrate how to boost it through teleportation up to unity. However, this boosting requires feedforward, so the probability they demonstrated for the NS gate is then $\frac{1}{4}$ in terms of pure linear optical elements. A further paper shows a CZ gate, with probability of $\frac{2}{27}$, using two extra ancilla photons [63] It also shows that the CZ gate cannot be implemented with probability of 1.

The first papers on the optimization of gates set out to find what the limits for making a CZ gate are. Knill shows that the success probability of an NS gate is, at most a half, and of the CZ gate $\frac{3}{4}$ [166]. Next, Scheel and Lutkenhaus show that for certain setups, a NS gate is upper bounded by a success probability of $\frac{1}{4}$ [167], introducing a new approach for these proofs – breaking up the interferometer as per Reck et al. [37] and focusing on the single beamsplitter, at which, the target mode is interacting with the ancillas. Soon after, Eisert proved that indeed a postselected NS gate without feedforward is bounded by $\frac{1}{4}$ and therefore the KLM scheme is optimal as far as NS gate goes [168]. Shortly after, Scheel and Audenaeret were able to show that a non-linear sign gate more generally scales as $\frac{1}{N^2}$ for a target state with $N + 1$ photons [169]. In the paper by Uskov et al. [170], the authors provide numerical evidence for optimality of a CZ gate with success probability of $\frac{2}{27}$. The methods used to prove some of the results in this Chapter draw inspiration from these papers. Finally we mention a paper which describes the theory of polynomial relaxation mixed with a relaxation of the unitary constraint to find upper bounds for a range of problems, although this method is not used in the work presented here [171].

2.1.2 Bell pairs

In the MBQC paradigm, we are interested in connecting many Bell pairs and small cluster states together, instead of entangling gates. While we can use the entangling gates to generate Bell states, we might be able to get better probabilities by focusing on generation of Bell states only. The current way of generation of "event-ready" Bell pairs in dual-rail encoding is based on the work of Zhang et al. [161] and is indeed better than the current way to carry out a CZ gate. The setup of the scheme can be seen in Figure 2.1. The starting state has four photons in four distinct modes. There is a total of eight available modes (these could for example be four spatial modes along with the two polarization modes). As the state being generated is a Bell state in dual-rail encoding, then the output state requires four modes with two photons between them. Therefore, four modes can have detectors on them.To generate a Bell state, we have to remain in the qubit subspace, which means we want to measure exactly two photons. There are 10 possible candidate patterns (of detecting two photons in four modes) for which we might have made a Bell state.

It turns out that four out of those 10 outcomes for the circuit given in Figure 2.1 give a Bell state with probability 1/8. Two more also give a Bell state with probability of
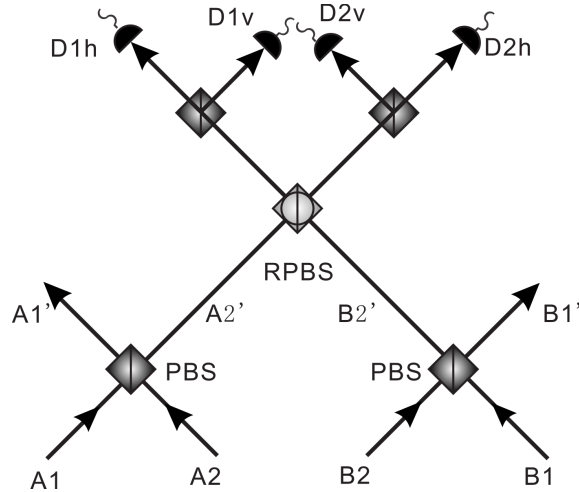
Figure 2.1:   Linear optical scheme for generating "event-ready" Bell pairs. The photons enter through modes "A1", "A2", "B1", "B2" with diagonal polarization. A Bell state is found at outputs "A1' " and "B1' " whenever there is a coincidence detected between two pairs of detectors ("D1h", "D1v", "D2v", "D2h"). Image taken from [161], copyright(2008) by the APS.

1/16 but require correction, some of the modes need to be swapped returning the state to the qubit subspace. Combining all of these probabilities together, we see that the total probability of generating a heralded Bell state with this scheme of four photons and eight modes is 3/16.

However, the scheme can actually do even better. It was shown in the paper by Joo et al. [160] that the failure modes could also be corrected to Bell pairs through procrustean distillation.   These failure outcomes look like unbalanced Bell states, therefore some entanglement is present. The idea behind the distillation is to balance the terms in the superposition through application of a beamsplitter that reduces the coefficient on one of the terms. As this filtering is probabilistic, the heralded Bell states than have smaller probability than that of original outcomes, giving a total probability for a corrected Bell state of 1/16. Combined with the other six outcomes, this gives a total probability of 1/4. However, notice that this will not be a purely linear optical scheme, as we will require a switch which for failure modes diverts the states to the correction circuit.

Another scheme offers a way to generate Bell states with four photons and in six modes [19]. The probability of acquiring a Bell state with this scheme is then $\frac{2}{27}$, matching that of a CZ gate, which is lower than the previously presented scheme.

2.2   Setup and notation

Figure 2.2 introduces the generic linear optical setup and notation used throughout the Chapter. The interferometer has d input modes and d output modes. The mode transformation describing this (photon number preserving) interferometer is an $d \times d$ unitary matrix $U \in \mathrm{U}(d)$ (see Section 1.3.2). The top d_I input modes contain N input photons, while the bottom d_V modes are ancilla vacua. The representation of U carried
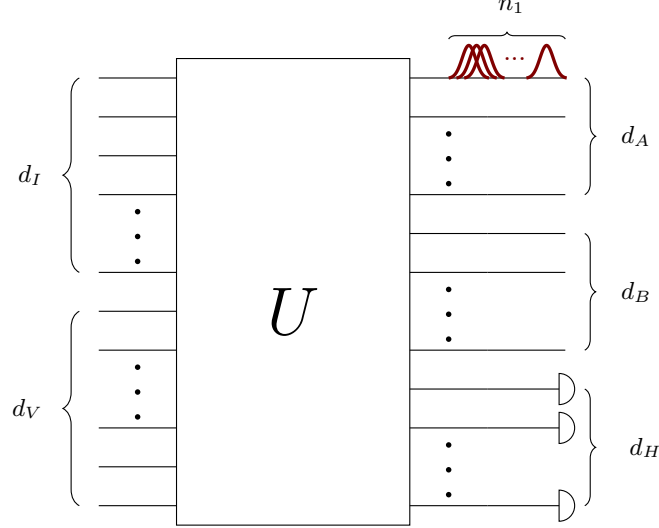
Figure 2.2: The generic setup used throughout this Chapter; see text for an explanation of the notation.

by the d mode Hilbert space in the number state (Fock) basis is denoted $\mathcal{U}$. The top d_A output modes belong to one party, Alice, the middle d_B modes belong to Bob, and the bottom d_H modes – Harold – these are measured using photon counting detectors. Harold's detection pattern is labelled

$$\underline{h} = (n_{d_S+1}, \cdots, n_d) \tag{2.1}$$

where n_i gives the photon number of output mode i, and $d_S = d_A + d_B$ is the number of modes in the "system", i.e. modes that do not belong to Harold and are therefore unmeasured. If $N_H = \sum_{k=d_S+1}^{d} n_k = ||\underline{h}||_1$ total photons have been detected, the number of photons left in the system is $N_S = N - N_H = N_A + N_B$. The Hilbert space of subsystem X (a subset of modes), given that it contains exactly N_X photons, is denoted $\mathcal{H}_X^{N_X}$.

Let the input to the interferometer be a Fock state

$$|\psi_{\text{in}}\rangle = |n_1, n_2, \cdots, n_{d_I}, \underbrace{0, \cdots, 0}_{d_V}\rangle = \prod_{k=1}^{d_I} \frac{(\hat{a}_k^\dagger)^{n_k}}{\sqrt{n_k!}} |\text{vac}\rangle, \tag{2.2}$$

where $|\text{vac}\rangle = |0\rangle^{\otimes N}$. As we saw in Section 1.3.2, the input transforms according to

$$\mathcal{U} |\psi_{\text{in}}\rangle = \prod_{k=1}^{d_I} \frac{1}{\sqrt{n_k!}} \left(\sum_{j=1}^{d} \hat{a}_j^\dagger U_{jk} \right)^{n_k} |\text{vac}\rangle, \tag{2.3}$$

where U_{jk} are the matrix elements of the mode transformation U and $\mathcal{U}$ is the representation of U on the multimode Fock space.

When $d_H > 0$, the ideal number resolving detectors will register a detection pattern $\underline{h} = (n_{d_S+1}, \cdots, n_d)$ of N_H photons. The output will be the post-measurement state

consisting of $N_S = n - N_H$ photons remaining in the system modes $1, \cdots, d_S$, given by

$$|\psi_S(\underline{h}, U)\rangle = \frac{\langle \underline{h}|\mathcal{U}|\psi_{\text{in}}\rangle}{\| \langle \underline{h}|\mathcal{U}|\psi_{\text{in}}\rangle \|}. \tag{2.4}$$

Note that this is a pure state on the system $S = AB$, because $|\underline{h}\rangle$ only has support on subsystem H. We will denote the unnormalized output by $|\widetilde{\psi}_S(\underline{h}, U)\rangle = \langle \underline{h}|\mathcal{U}|\psi_{\text{in}}\rangle$. The Hilbert space of the system is

$$\mathcal{H}_S^{N_S} = \bigoplus_{N_A=0}^{N_S} \mathcal{H}_A^{N_A} \otimes \mathcal{H}_B^{N_B}, \tag{2.5}$$

where $N_B = N_S - N_A$ is the number of photons in Bob's subsystem. We are interested in entanglement with respect to this tensor product structure. As mentioned in Section 1.3.1, the dimension of the Hilbert space of N photons in d modes is $\binom{d+N-1}{N}$, and so

$$\dim \mathcal{H}_S^{N_S} = \sum_{N_A=0}^{N_S} \binom{d_A + N_A - 1}{N_A} \binom{d_B + N_B - 1}{N_B} \tag{2.6}$$

$$= \binom{d_S + N_S - 1}{N_S} \tag{2.7}$$

as $d_S = d_A + d_B$ and $N_S = N_A + N_B$. The totality of states available to Alice can be thought of as the Hilbert space $\bigoplus_{N_A=0}^{N_S} \mathcal{H}_A^{N_A}$, and we may index its Fock basis as $\{|\underline{a}\rangle_A : \underline{a} = (n_1, n_2, \cdots, n_{d_A}), \|\underline{a}\|_1 = N_A\}$. Similarly for Bob. Expanding the output in this basis, we have

$$|\widetilde{\psi}_S(\underline{h}, U)\rangle = \sum_{\underline{a}, \underline{b}} \widetilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U) |\underline{a}\rangle_A \otimes |\underline{b}\rangle_B. \tag{2.8}$$

The coefficients $\widetilde{C}$ are related to permanents of the matrix U, as we saw in Section 1.3.4. Here we have input Fock state $|\psi_{\text{in}}\rangle = |\underline{n}, \underline{0}\rangle$ and an output Fock state $|\underline{abh}\rangle$, then per Section 1.3.4

$$\langle \underline{abh}|\mathcal{U}|\psi_{\text{in}}\rangle = \frac{\text{perm}(U_{\underline{n}, \underline{abh}})}{\sqrt{\underline{n}!\,\underline{a}!\,\underline{b}!\,\underline{h}!}}. \tag{2.9}$$

Therefore, $\widetilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U) = \langle \underline{abh}|\mathcal{U}|\psi_{\text{in}}\rangle$ and the probability of detecting pattern $\underline{h}$ is $P_H(\underline{h}, U) = \sum_{\underline{a}, \underline{b}} |\widetilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U)|^2$. Defining $C_{\underline{a}, \underline{b}} = \widetilde{C}_{\underline{a}, \underline{b}}/\sqrt{P_H(\underline{h}, U)}$, the normalized state can be written as $|\psi_S(\underline{h}, U)\rangle = \sum_{\underline{a}, \underline{b}} C_{\underline{a}, \underline{b}}(\underline{h}, U) |\underline{a}\rangle_A |\underline{b}\rangle_B$. Notice that this slightly different probability than that defined in Section 1.3.3, as it gives the probability of part of a state being found in a certain Fock state.

We denote relevant coefficients in mode assignment notation (see Section 1.3.1) by γ, which are related to the above mentioned permanent as

$$\gamma_{a(\underline{n})}(\underline{h}, U) = \frac{\widetilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U)}{\underline{n}!}. \tag{2.10}$$

These are the coefficients of the output states as expressed in terms of the creation operators

assuming unbunched input to the interferometer, see Eq.(2.12).

Equation (2.8) provides a decomposition we can use to bound the entanglement. However, the fact that the total number of photons in the system, N_S, is preserved implies that not all conceivable bipartite basis states $|\underline{a}\rangle_A \otimes |\underline{b}\rangle_B$ are available, so the system should not simply be viewed as the tensor product of two qudits i.e. Eq.(2.7) is not simply the product of $\dim \mathcal{H}_A$ and $\dim \mathcal{H}_B$. In particular, this means that states that are maximally entangled in the usual sense do not exist. For example, Alice can have many states with N_S photons, but there is only one possible Bob state to which they can be correlated, namely the vacuum (see Section 2.5.2).

The entanglement measure that will be used is the von Neumann entropy; given a pure state $|\psi_S(\underline{h}, U)\rangle$, its density matrix is defined $\rho_{AB}(\underline{h}, U) = |\psi_S(\underline{h}, U)\rangle \langle \psi_S(\underline{h}, U)|$, and its reduced density matrices on subsystems are the marginals $\rho_A(\underline{h}, U) = \mathrm{Tr}_B[\rho_{AB}(\underline{h}, U)]$ and $\rho_B(\underline{h}, U) = \mathrm{Tr}_A[\rho_{AB}(\underline{h}, U)]$. The von Neumann entropy is then

$$S(\rho_A(\underline{h}, U)) = -\mathrm{Tr}[\rho_A(\underline{h}, U) \cdot \log \rho_A(\underline{h}, U)] = -\sum_a \lambda_a \cdot \log \lambda_a \qquad (2.11)$$

where $\{\lambda_a\}_a$ are the non-zero eigenvalues of the reduced state. Unless stated otherwise, logarithms will be assumed to be base 2. Finally, we will use ebits as the unit of bipartite entanglement where 1 ebit corresponds to the von Neumann entropy of a Bell state.

2.3 Qubit entanglement

In this section we will be considering the dual-rail encoding of two qubits that was introduced in Section 1.4.2. This means that $d_A = d_B = 2$ and states are postselected so that subsystems A and B have exactly one photon each, $N_A = N_B = 1$; all the other states are discarded. Despite the full Hilbert space of the system being of dimension 10 (see Eq. (2.13)), these constraints limit the space of permissible states to $\dim \mathcal{H}_A = \dim \mathcal{H}_B = 2$, encoding two qubits. To entangle photons in this encoding using only passive linear optics, the use of ancillas and postselection is necessary [53], so $d_H > 0$.

2.3.1 Generating Bell states with three photons is impossible
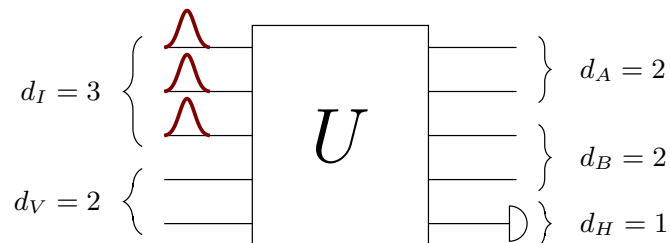

Figure 2.3: The setup used in Section 2.3.1, with $d_I = N = 3$, $d_V = 2$, $d_A = d_B = 2$, and $d_H = 1$. We show that no such setup can create an entangled state in dual-rail qubit encoding with any non-zero probability. On the other hand, with 4 input photons it is possible to create a Bell state with probability of 1/4 [161].

In Section 1.4.2, we mention the familiar result which states that generating a Bell state in dual-rail encoding with just two photons is impossible [53, 172]. Kieling observed it is also impossible with three photons, using an algebraic geometry approach to the problem [172]. Here we offer an explicit proof that not only is it impossible with three photons, it is only possible to create product states.

Proposition 2.3.1. *In a passive linear optical setup using dual-rail encoding, ancillas and heralding, it is not possible to create an entangled state using 3 photon input.*

Proof. First, let us consider the case where there are five modes ($d = 5$); four system modes ($d_A + d_B = 4$) and one ancilla ($d_H = 1$), as illustrated in Figure 2.3. Let the input be three unbunched photons ($N = d_I = 3$). Dual-rail encoding has a total of two photons in a valid qubit state output ($N_S = 2$), implying here that one photon is detected ($N_H = 1$). As there is only one measurement ancilla, the only possible measurement pattern is $\underline{h} = (1)$ (one photon in the fifth mode).

As discussed in Section 1.3.4, the amplitudes are related to the permanents of the matrix U:

$$\gamma_{kj}((1), U) = \begin{cases} \frac{1}{2} \sum_{\sigma \in S_3} U_{k,\sigma(1)} U_{k,\sigma(2)} U_{5,\sigma(3)}, & k = j \\ \sum_{\sigma \in S_3} U_{k,\sigma(1)} U_{j,\sigma(2)} U_{5,\sigma(3)} & k \neq j \end{cases} \tag{2.12}$$

defined $\forall k, j \in \{1, 2, 3, 4\}$. The unnormalized state following detection is

$$\begin{aligned} |\widetilde{\psi}((1), U)\rangle = {}& \sqrt{2}\gamma_{11} |2000\rangle + \sqrt{2}\gamma_{22} |0200\rangle \\ & + \sqrt{2}\gamma_{33} |0020\rangle + \sqrt{2}\gamma_{44} |0002\rangle \\ & + \gamma_{13} |1010\rangle + \gamma_{24} |0101\rangle \\ & + \gamma_{12} |1100\rangle + \gamma_{34} |0011\rangle \\ & + \gamma_{14} |1001\rangle + \gamma_{23} |0110\rangle , \end{aligned} \tag{2.13}$$

occurring with probability $P_H((1), U) = 2 \sum_{k=1}^{4} |\gamma_{kk}|^2 + \sum_{\substack{k,j=1 \\ k \neq j}}^{4} |\gamma_{kj}|^2$.

In dual-rail encoding it is possible to do any local unitary deterministically by adding beamsplitters and phase shifters to each of the qubits [53]. Thus it suffices to show that it is not possible to create any state of the form $\alpha |0\rangle_A |0\rangle_B + \beta |1\rangle_A |1\rangle_B$ where $|\alpha|^2 + |\beta|^2 = 1$ and $\alpha \neq 0$, $\beta \neq 0$, because any entangled pure state can be transformed into one of this form by local unitary operations. The coefficients must therefore satisfy

$$\gamma_{11} = \gamma_{22} = \gamma_{33} = \gamma_{44} = 0, \tag{2.14}$$

$$\gamma_{12} = \gamma_{14} = \gamma_{23} = \gamma_{34} = 0 \quad \text{and} \tag{2.15}$$

$$|\gamma_{13}| = \alpha\sqrt{p}, |\gamma_{24}| = \beta\sqrt{p}, \tag{2.16}$$

where $p = P_H((1), U)$, the probability of one photon being detected in the last mode. We will now try to find a unitary U that satisfies these constraints. Define $K_k := U_{k2} U_{53} + U_{k3} U_{52}, \forall k \in \{1, \ldots, 4\}$.

First, let us consider the case where at least one of U_{51}, U_{52} and U_{53} is 0. Without loss of generality (wlog) we can label modes so that $U_{51} = 0$, because we can swap A for B and mode 1 for 2 without affecting entanglement. Then the equations in (2.12) can be rewritten as $\gamma_{kk} = U_{k1}K_k$ and $\gamma_{kj} = U_{k1}K_j + U_{j1}K_k$ for $k \neq j$. Since $\gamma_{11} = U_{11}K_1 = 0$ and $\gamma_{13} = U_{11}K_3 + U_{31}K_1 \neq 0$, then one and only one of U_{11} or K_1 can be equal to 0. First, assume that $U_{11} = 0$. Since $K_1 \neq 0$, from the constraints $\gamma_{12} = U_{21}K_1 = \gamma_{14} = U_{41}K_1 = 0$ and $\gamma_{24} = U_{21}K_4 + U_{41}K_2 \neq 0$, we see that there is no solution. Similarly, if $K_1 = 0$, then $U_{11} \neq 0$ and the constraints $\gamma_{12} = U_{11}K_2 = \gamma_{14} = U_{11}K_4 = 0$ and $\gamma_{24} = U_{21}K_4 + U_{41}K_2 \neq 0$ again results in no solution. Therefore there is no solution for which at least one of U_{51}, U_{52}, U_{53} is zero.

Next we assume $K_k \neq 0 \ \forall k \in \{1, \dots, 4\}$, with $U_{51}U_{52}U_{53} \neq 0$. Then solving for U_{k1} from $\gamma_{kk} = 0$ we get $U_{k1} = -U_{k2}U_{k3}U_{51}/K_k, \forall k \in \{1, \dots, 4\}$. Substituting this into the expression for γ_{kj} we get

$$\gamma_{kj} = \frac{U_{51}U_{52}U_{53}(U_{k2}U_{j3} - U_{j2}U_{k3})^2}{K_kK_j}, \tag{2.17}$$

for all $k, j \in \{1, \dots, 4\}, k \neq j$. The only way $\gamma_{12} = \gamma_{23} = 0$, is if $U_{12}U_{23} = U_{22}U_{13}$ and $U_{22}U_{33} = U_{32}U_{23}$. If $U_{22}U_{23} \neq 0$ then $U_{12}U_{33} = U_{13}U_{32}$, which means $\gamma_{13} = 0$ also, thus cannot be a solution. If only one of U_{22} or U_{33} is zero, assume $U_{2j} = 0$ where j is 2 or 3. But then $U_{1j} = U_{3j} = 0$ and again $\gamma_{13} = 0$. If both are zero, then $\gamma_{24} = 0$. Therefore, there is no solution with $K_k \neq 0 \ \forall k \in \{1, \dots, 4\}$.

Lastly, assume that at least one of the $K_k = 0$ and that $U_{51}U_{52}U_{53} \neq 0$; wlog, $K_1 = 0$. Then $U_{12} = -U_{13}U_{52}/U_{53}$ combined with the constraint $\gamma_{11} = U_{12}U_{13}U_{51} = 0$ means $U_{12} = U_{13} = 0$. This gives $\gamma_{1j} = U_{11}K_j, \forall j \in \{1, \dots, 4\}$. Since $\gamma_{12} = \gamma_{14} = 0$ and $\gamma_{13} \neq 0$, then $U_{11} \neq 0$, while $K_2 = K_4 = 0$. However, this implies $U_{22} = U_{23} = 0$ by a similar argument, further implying that $\gamma_{24} = 0$ and hence there is no solution.

We see that under no conditions is there a solution to the given equations where $\alpha \neq 0$ and $\beta \neq 0$.

This proves the claim for 5 modes. To see that it is true for any number of vacuum ancillas, notice that as long as there are no photons added, Eqs. (2.12) do not change other than the mode number 5 being replaced with the new detection ancilla. Each new case therefore gives rise to the same constraints implied by Eqs. (2.16), with a lack of solution in the same way. Thus, vacuum ancillas can only increase the probability of creating a state if that probability was nonzero in the first place.

Finally, if we allow inputs other than completely unbunched, Eqs. (2.12) become even more restrictive. For example, if there were two photons in input mode 1 and one photon in input mode 2, then the matrix elements U_{i3}, U_{3i} would not appear in Eqs. (2.12), serving only to make the constraints harder to satisfy.

$\square$

Corollary 2.3.2. *In a passive linear optical setup using dual-rail encoding, ancillas and heralding, it is not possible to create a Bell state using 3 photon input.*
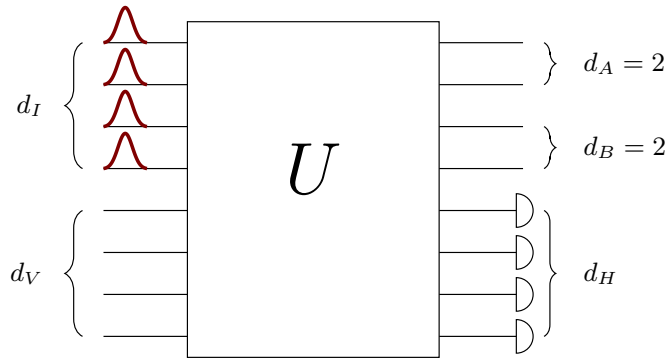
2.3.2 Optimal Bell state generation

Figure 2.4: The setup used in Section 2.3.2 with four photons in eight modes; $d_I = N = 4$, $d_V = 4$, $d_A = d_B = 2$, $d_H = 4$. We give extensive numerical evidence for optimal Bell state generation using this setup when looking for specific Bell states as output.

The previous section showed that Bell state generation with non-zero success probability requires at least four photons. In Section 2.1 we mention two schemes which accomplish this task using four photons and six [19] or eight [160, 161] modes, with success probabilities of 2/27 and 1/4 respectively.

We performed a numerical search for a linear optical Bell state generator that gives a higher success probability. We used a gradient descent based optimization algorithm over $d = 8$ unitaries with $N = 4$ photon input. Numerical optimization was carried out in Python, using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm from the SciPy library [173]. This algorithm finds local minima so it needs to be run many times with different seed unitaries, which were randomly selected according to the Haar measure.

The cost function we consider is based on the overlap with the desired Bell states. We allow for six different Bell states, which in the Fock basis after measurement correspond to $|B_{1,2}\rangle = (|1010\rangle \pm |0101\rangle)/\sqrt{2}$, $|B_{3,4}\rangle = (|1001\rangle \pm |0110\rangle)/\sqrt{2}$ and $|B_{5,6}\rangle = (|1100\rangle \pm |0011\rangle)/\sqrt{2}$, where the latter can be corrected to the usual dual-rail qubit encoding using a switch [161]. After detecting measurement pattern $\underline{h}$, the overlap between each of these states with the post-selected state is calculated. We found that raising the overlap to the exponent 10 optimized the numerical efficacy, penalizing states far from a Bell state heavily. Multiplying by the probability of detection gives the target cost function to be minimized, $f(\underline{h}, U) = -\sum_{\underline{h}} P_H(\underline{h}, U) \sum_{k=1}^{6} |\langle B_k | \psi(\underline{h}, U)\rangle|^{10}$.

Figure 2.5 shows the results of this minimization. The optimal known scheme, when evaluated for this cost function, gives a value of approximately -0.1875. It produces one of these 6 Bell states with probability 1/32 for 6 out of the 10 possible measurement patterns [161]. We can see from the figure that the minimum achieved by the numerical optimization over $53,000$ trials is also approximately -0.1875, thus giving solutions which are equivalent to the known scheme in terms of this parameter. While not a proof, this numerical evidence strongly suggests that the known scheme is optimal for generating the above set of Bell states. Other cost functions were also attempted, as well as other optimization libraries, but all gave the same results as the technique above.
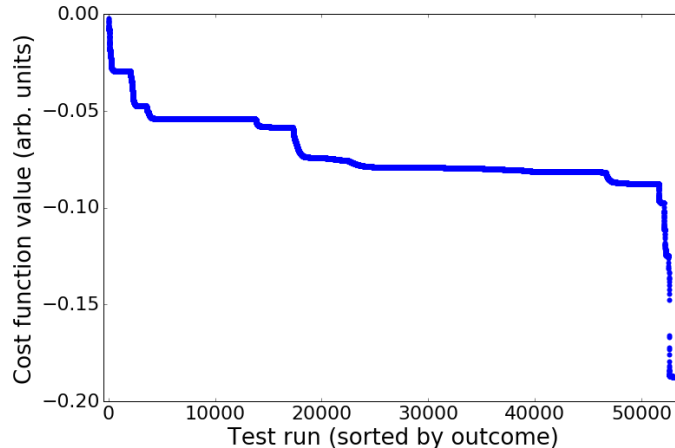
Figure 2.5: Results of optimization looking for interferometers that generate Bell states with highest probability. The minimum found of ≈ -0.1875 is exactly bounded by the values of cost function for the known U_{Bell} interferometer as described in the text. Out of $53,000$ test runs, 0.75 % of minima found were within 0.001 of the minimum corresponding to U_{Bell}, i.e. within numerical error. Besides the trials depicted in this graph, the cost function was also optimized with other parameters given to the optimizing algorithm as well as over the space of orthogonal matrices. Thus the number of test runs for which a better solution could not be found is more than $150,000$.

We also investigated the case of non-orthogonal Bell states; for example, allowing $|00\rangle + |11\rangle$ as well as $|00\rangle + i|11\rangle$ as target states. The possibility of both of these states being generated from the same U for different measurement patterns was explored by running similar numerical optimizations rewarding such situations. We found no such unitary, which is an interesting result in itself.

Though the complexity of the problem grows quickly, we also looked at how the situation changes with higher numbers of input photons and modes. We numerically optimized over $n = 5$, $d = 10$ using a similar algorithm and no improved solutions were found over 5000 runs. Similarly, we checked $n = 6$, $d = 12$ over 1000 runs and here as well there was no improvement over the -0.1875 result for $n = 4$, $d = 8$.

2.4 Random unitaries

In this section, we move from the dual-rail qubit encoding of Section 2.3 to mode entanglement in Section 2.5. First, we look at how much mode entanglement can be generated with random elements of the unitary group, which we can then use to compare with the dual-rail encoding. We do so by setting Alice and Bob's number of modes to 2, and numerically computing the average amount of entanglement over measurement patterns. Notice that this is different from the setting in Section 2.3, where we aimed to get a maximally entangled Bell state with the highest possible probability. Here and in the rest of this work we will study this average entanglement, namely

$$\langle S(U)\rangle_H = \sum_{\underline{h}} P_H(\underline{h}, U) S(\rho_A(\underline{h}, U)). \tag{2.18}$$

48

The expectation over the unitary group (for fixed d and N) is then $\langle S \rangle_{H,\mathrm{U}} = \int_{\mathrm{U}(d)} dU \, \langle S(U) \rangle_H$, where dU is the normalized Haar measure.
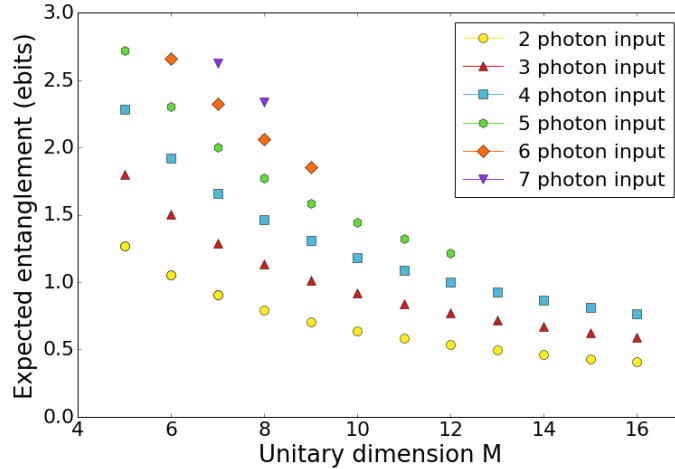
Figure 2.6: The expectation, over the unitary group, of the average, over measurement patterns, mode entanglement versus the number of modes d, for various numbers of unbunched input photons. $d_A = d_B = 2$, and if the number of photons N is smaller than d, vacuum input modes are added. The number of heralding detectors is $d_H = d - d_A - d_B$. The entanglement for a single unitary U is averaged over all measurement patterns, and subsequently averaged over 100,000 randomly Haar-sampled unitaries U. Colours and symbols represent different number of input photons, with $2 \le N \le 7$.

Figure 2.6 shows the numerical results. We notice that often the average is higher than 1 ebit, which is the maximum we can achieve in dual-rail qubit encoding. Adding input photons for the same d increases the average entanglement, while adding vacuum ancillas decreases it. We see that the average entanglement of $N+1$ photons in $d+1$ modes can be lower than that for d and N (see $N = d = 5$ and 6). That is, we do not expect more average entanglement by adding a photon at the cost of adding another mode. Further, we note that even with 2 photons, there is more average entanglement generated than in the optimal Bell state generator with 4 photons. We explore this in more detail for a better comparison.

In the usual Bell state generation scenario discussed in Section 2.3.2, if the measurement outcome indicates that the output state is outside of the qubit subspace, the output is discarded. Here we include the entanglement of the discarded states in accordance with Eq. (2.18). We compare the optimal Bell state generator to random unitaries with the same parameters; $d_A = d_B = 2$, $N = 4$ and $d = 8$.

In Fig. 2.7 we see the results of the comparison. Firstly, in Section 2.3.2 we saw that the probability of getting a Bell state for a state correctable with a single switch is 3/16 [161]. A Bell state gives a single ebit, and if all the other states are discarded, the average entanglement would be 0.1875 ebits. If all the outputs from this unitary were counted towards average entanglement as discussed in the previous paragraph (where Equation 2.18 is utilized), the entanglement obtained is marked on the Figure 2.7 as U_{Bell}. As we
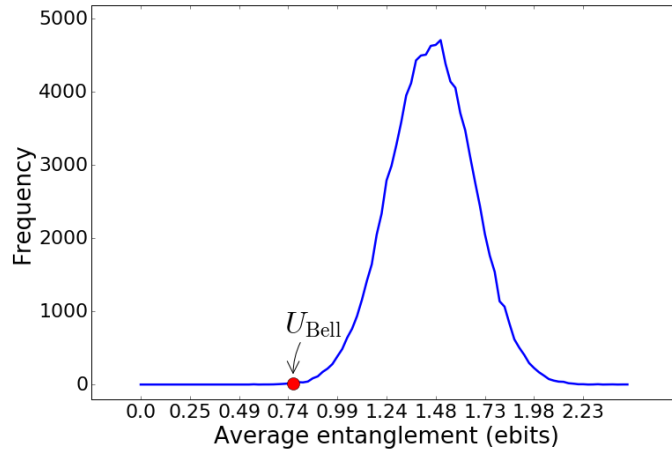
Figure 2.7: Numerical evaluation of $\langle S(U) \rangle_H$ for $100,000$ unitaries U chosen using the Haar measure in the case $d_A = d_B = 2$, $d_H = 4$, and $n = 4$ unbunched input photons. Average entanglement for a given U was calculated according to Equation (2.18) and then binned in one of 100 bins with a minimum of 0 and maximum obtained in the samples. The red dot marks the value of average entanglement that the Bell generating unitary from Section 2.3.2 can give, denoted as U_{Bell}, if all of its output states were used.

can see from the graphs, U_{Bell} gives a markedly lower amount of entanglement than what could be generated on average with a random unitary on the same number of modes.

2.5 Mode entanglement

The previous section shows that, on average, random unitaries give significantly more mode entanglement than dual-rail encoding. We therefore turn our attention to the investigation of mode rather than qubit entanglement as defined in Section 2.2.

Equation (2.5) states that the total system Hilbert space is a direct sum of Hilbert subspaces such that the sum of Alice and Bob's photon numbers is N_S, the number of photons left after heralding. Let $\rho_{AB} = |\psi_S(\underline{h}, U)\rangle \langle \psi_S(\underline{h}, U)|$ as in Eq. (2.4). Alice's reduced density matrix is

$$\rho_A(\underline{h}, U) = \text{Tr}_B[\rho_{AB}(\underline{h}, U)] \tag{2.19}$$

$$= \sum_{\underline{b}''} \langle \underline{b}''| \left(\sum_{\underline{a}, \underline{b}, \underline{a}', \underline{b}'} C_{\underline{ab}} \overline{C}_{\underline{a}'\underline{b}'} |\underline{ab}\rangle \langle \underline{a}'\underline{b}'| \right) |\underline{b}''\rangle$$

$$= \sum_{\underline{a}, \underline{a}'} \left(\sum_{\underline{b}} C_{\underline{ab}} \overline{C}_{\underline{a}'\underline{b}} \right) |\underline{a}\rangle \langle \underline{a}'|, \tag{2.20}$$

where only the terms with $\|\underline{a}\|_1 = \|\underline{a}'\|_1$ are non-zero, because $\|\underline{b}\|_1 = \|\underline{b}'\|_1 = \|\underline{b}''\|_1$ and $n_S = \|\underline{a}\|_1 + \|\underline{b}\|_1 = \|\underline{a}'\|_1 + \|\underline{b}'\|_1$. Therefore, there exists a Fock basis ordering in which Alice's reduced state is block diagonal, which allows us to derive a bound on the entanglement (see Section 2.5.2). In the case that Alice has a single mode, this implies her

state is diagonal in Fock basis. The total number of orthogonal states available to Alice is

$$\dim(\mathcal{H}_A^{N_S}) = \sum_{N_A=0}^{N_S} \binom{d_A + N_A - 1}{N_A} = \binom{d_A + N_S}{N_S}. \tag{2.21}$$

In Section 2.5.1, we find entanglement bounds when Alice only has one mode. The bound depends on the input state; if the input photons are bunched in a single mode, entanglement is unbounded as the number of photons increases. Surprisingly, if the input is unbunched, we find a constant bound independent of the number of Bob's modes and independent of the number of photons. More general bounds can be found, though they are also more loose. In Section 2.5.2 we give the bound on entanglement due to the block diagonal structure of Alice's reduced density matrix in Fock basis. In Section 2.5.2 we give a bound which is a consequence of the linearity of the mode transformations. Unlike in Sec. 2.5.1, neither of these bounds depend on the unitarity of the mode transformations, which we expect should affect the amount of entanglement that can be achieved. In Section 2.5.3 we conjecture a general unitarity bound based on numerical evidence.

2.5.1 Entanglement when Alice has a single mode

Entanglement for bunched input can be unbounded

First, we show that mode entanglement is unbounded if we are not restricted to unbunched input.
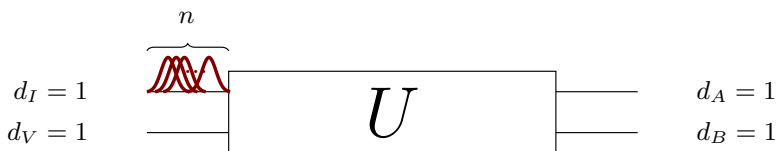
Figure 2.8: The setup used in Section 2.5.1, where we consider only $d = 2$ modes. The input consists of all N photons bunched in the top mode; $d_I = d_V = d_A = d_B = 1$, $d_H = 0$. We prove that in this setup maximal entanglement grows as $\log N$. The special case where U is a balanced beamsplitter was analyzed in [174].

Proposition 2.5.1. *Let the input into a $d = 2$ interferometer consist of N photons bunched in a single mode (see Figure 2.8). Then the entanglement across the two output modes is at most $O(\log N)$ ebits, which is achieved when U is a balanced beamsplitter.*

Proof. Parameterize the $d = 2$ unitary matrix U acting on Alice and Bob's single mode Hilbert spaces as

$$U = \begin{bmatrix} c & d \\ -d^* & c^* \end{bmatrix}, \tag{2.22}$$

where $|c|^2 + |d|^2 = 1$. The output state is

$$
\begin{aligned}
|N0\rangle &= \left(\hat{a}_1^\dagger\right)^N / \sqrt{N!} \, |0\rangle \\
&\mapsto \left(c\hat{a}_1^\dagger - d^*\hat{a}_2^\dagger\right)^N / \sqrt{N!} \, |0\rangle \\
&= \frac{1}{\sqrt{N!}} \sum_{k=0}^N \binom{N}{k} (c\hat{a}_1^\dagger)^k (-d^*\hat{a}_2^\dagger)^{N-k} |0\rangle \\
&= \frac{1}{\sqrt{N!}} \sum_{k=0}^N \binom{N}{k} c^k (-d^*)^{N-k} \sqrt{k!}\sqrt{(N-k)!} \, |k\rangle \, |N-k\rangle \\
&= \sum_{k=0}^N \sqrt{\binom{N}{k}} \, c^k (-d^*)^{N-k} |k\rangle \, |N-k\rangle .
\end{aligned}
\tag{2.23}
$$

When Alice has only one mode, her reduced density matrix is diagonal in the Fock basis, so we can find the spectrum of her state directly from the above equations:

$$
\lambda_k = \binom{N}{k} (|c|^2)^k (|d|^2)^{N-k} = \binom{N}{k} (|c|^2)^k (1 - |c|^2)^{N-k}.
\tag{2.24}
$$

This is a binomial distribution with a 'success' probability of $p = |c|^2$. The entropy of the binomial distribution for a fixed p is $1/2 \log_2 (2\pi eN \cdot p \cdot (1-p)) + O(1/N)^*$. Thus we see that the entanglement bound is $O(\log N)$, where N is the number of photons. The constant prefactor is maximized for $p = |c|^2 = |d|^2 = 1/2$, whence the entropy of Alice's state is $1/2 \log_2 (2\pi eN \cdot 1/2 \cdot (1 - 1/2)) + O(1/N) = 1/2 \log_2 (\pi eN/2) + O(1/N)$. Finally, notice that solutions to Equation (2.22) where $|c|^2 = |d|^2 = 1/2$ are a family of balanced beamsplitters. $\qquad\square$

This is in stark contrast to the situation where the input is unbunched, where we will see in the next section that the entanglement is bounded by a constant.

Entanglement for unbunched input is bounded

We now consider situations where Alice only has one mode, Bob can have many, and we do not use any measurement. The following Lemma will be of use.

Lemma 2.5.2. *Consider inputting a Fock state $|\underline{n}\rangle = |n_1 \ldots n_d\rangle$ into an arbitrary interferometer that has d modes. Let $\nu = \max\{n_1, \ldots, n_d\}$. Then the mean photon number in each output port is bounded by ν [175].*

---

*From, e.g., the de Moivre-Laplace Theorem
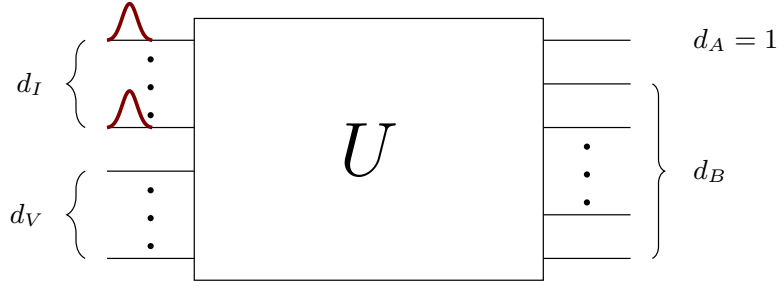
Figure 2.9: The setup used in Section 2.5.1. The input is an unbunched state with $d_I = N$, with $d_V \geq 0$, $d_A = 1$, $d_B \geq 1$ and $d_H = 0$. We prove that entanglement for this setup is bounded by a constant.

Proof. Let $|\underline{n}\rangle$ be an arbitrary Fock state.

$$
\begin{aligned}
\langle \hat{n}_j \rangle &= \langle \underline{n} | \, \mathcal{U}^\dagger \hat{n}_j \mathcal{U} \, | \underline{n} \rangle \\
&= \langle \underline{n} | \, \mathcal{U}^\dagger \hat{a}_j^\dagger \mathcal{U} \mathcal{U}^\dagger \hat{a}_j \mathcal{U} \, | \underline{n} \rangle \\
&= \langle \underline{n} | \left(\sum_{j'} \hat{a}_{j'}^\dagger \overline{U_{jj'}} \right) \left(\sum_{j''} \hat{a}_{j''} U_{jj''} \right) | \underline{n} \rangle \\
&= \sum_{j'} \sum_{j''} \overline{U_{jj'}} U_{jj''} \langle \underline{n} | \, \hat{a}_{j'}^\dagger \hat{a}_{j''} \, | \underline{n} \rangle \\
&= \sum_{j'} \overline{U_{jj'}} U_{jj'} \, n_{j'} \qquad\qquad\qquad\qquad (2.25)
\end{aligned}
$$

If, as hypothesized, $n_j \leq \nu$ for all modes j, then

$$
\langle \hat{n}_j \rangle = \sum_{j'} |U_{j'j}|^2 \, n_{j'} \leq \sum_{j'} |U_{j'j}|^2 \, \nu = \nu. \qquad\qquad (2.26)
$$

$\square$

In the following calculations we shall assume that $N \to \infty$ as any bound on the entropy found for this infinite case would also hold for a finite one with the same set of constraints.

Lemma 2.5.3. *Let $\{p_j\}_{j=0}^\infty$ be a probability distribution subject to the constraint $\sum_j j p_j \leq \nu$. Then the entropy of this distribution is at most $\log\left((1+\nu)^{1+\nu}/\nu^\nu\right)$.*

Proof. The entropy of the probability distribution $\{p_j\}_{j=0}^\infty$ is $S = -\sum_{j=0}^\infty p_j \log p_j$. We maximize this subject to the constraints $\sum_{j=0}^\infty j p_j = \overline{n} \leq \nu$ and $\sum_{j=0}^\infty p_j = 1$ using the method of Lagrange multipliers.

Let the Lagrangian be

$$
L = S + (\lambda_0 + \log e) \left(\sum_{j=0}^\infty p_j - 1 \right) + \lambda_1 \left(\sum_{j=0}^\infty j p_j - \overline{n} \right). \qquad (2.27)
$$

Then $\partial L/\partial p_j = -\log p_j + \lambda_0 + \lambda_1 j$. Setting $\partial L/\partial p_j = 0$ gives $p_j = 2^{\lambda_0 + \lambda_1 j}$. Substituting

the value of p_j into the constraints, we get

$$\sum_{j=0}^{\infty} jp_j = 2^{\lambda_0} 2^{\lambda_1}/(1 - 2^{\lambda_1})^2 = \overline{n} \tag{2.28}$$

$$\sum_{j=0}^{\infty} p_j = 2^{\lambda_0}/(1 - 2^{\lambda_1}) = 1 \tag{2.29}$$

This allows us to solve for λ_0 and λ_1, giving

$$\lambda_0 = \log\left[1/(1 + \overline{n})\right], \ \lambda_1 = \log\left[\overline{n}/(1 + \overline{n})\right]. \tag{2.30}$$

Notice that

$$S = -\sum_j p_j \log p_j = -\sum_j p_j(\lambda_0 + \lambda_1 j) = -\lambda_0 - \lambda_1 \overline{n}$$
$$= \log\left((1 + \overline{n})^{1+\overline{n}}/\overline{n}^{\overline{n}}\right) \tag{2.31}$$

The function above increases monotonically for $\overline{n} \geq 0$ and since $\overline{n} \leq \nu$ we get

$$S \leq \log\left((1 + \nu)^{1+\nu}/\nu^{\nu}\right). \tag{2.32}$$

$\square$

Corollary 2.5.4. *Let $\{p_j\}_{j=0}^{\infty}$ be some probability distribution subject to the constraint $\sum_j jp_j \leq \nu$, $\nu \in [0,1]$. Then the entropy of this distribution is at most 2 ebits.*

Theorem 2.5.5. *Let Alice have one output mode, $d_A = 1$, and Bob have $d_B = k$. Let the input be a single photon in each of the $k + 1$ modes. Then the entanglement between Alice and Bob is bounded by 2 ebits for all k.*

Proof. Alice's reduced density matrix is diagonal in the Fock basis, where each entry $\langle j| \rho_A |j\rangle$ corresponds to the probability that Alice's mode contains j photons. By Lemma 2.5.2, this distribution satisfies the conditions of Corollary 2.5.4. Thus the von Neumann entropy of this state is bounded by 2 for any k, as the bound which holds for $k \to \infty$ also holds for any finite k as well. $\square$

Notice that extra vacuum modes will not increase this limit on the entanglement as the limit is due to the expected number of photons in Alice's mode being at most 1. We see that despite the fact that the dimension of Alice's Hilbert space grows with the number of photons as $N + 1$, and Bob's can be even larger, the maximum entanglement is severely constrained to be less than 2 ebits.

Because we are interested in the average entanglement, the result will hold for heralding as well:

Corollary 2.5.6. *Let Alice have one output mode, $d_A = 1$, while Bob and Harold have $d_B + d_H = k$. Let the input be a single photon in each of the $k + 1$ modes. Then the entanglement between Alice and Bob is bounded by 2 ebits for all k.*

Proof. No LOCC operation can increase the amount of entanglement in the system on average [176]. Therefore, $\langle S(U) \rangle_H = \sum_{\underline{h}} P_H(\underline{h}, U) S(\rho_A(\underline{h}, U)) \leq S(\rho_A(U))$, where $\rho_A(U)$ is Alice's reduced density matrix before any measurement, and by Theorem 2.5.5, $S(\rho_A(U)) \leq 2$ ebits. $\qquad\square$

We can also examine inputs that have different numbers of bunched photons. If the highest number of photons in a single input mode is ν, as per Lemma 2.5.2, the expected number of photons in Alice's mode will then be bounded by ν. Because the function which bounds the entropy, Eq. (2.32), is monotonically increasing, the entropy of Alice's (diagonal) state $(p_0, \ldots, p_n)$ is at most $\log\left((1+\nu)^{1+\nu}/\nu^\nu\right)$ by Lemma 2.5.3. In the extreme case where all the photons are bunched in a single mode, S scales as $O(\log(\nu+1))$, consistent with Proposition 2.5.1.

Entanglement when Bob also has a single mode
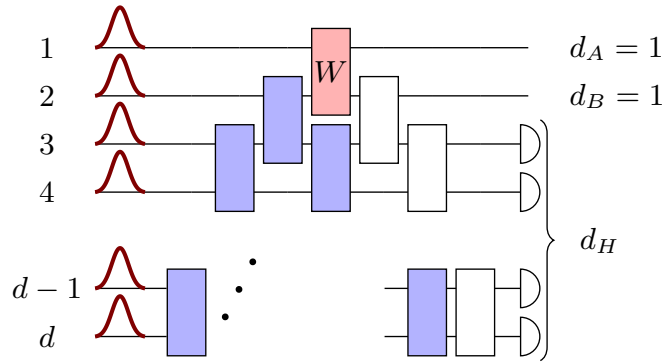
Figure 2.10: The setup used in Section 2.5.1, where $d = N$, $d_A = d_B = 1$, and $d_H \geq 1$. An arbitrary d mode interferometer can be decomposed into $d(d+1)/2$ two-mode interferometers [37, 177]. Note that this also applies to an arbitrary $d-1$ mode sub-interferometer (blue). By focusing on the only component that entangles Alice and Bob (red), we show that the maximum entanglement is the $d = N = 2$ value of $\log 3$ ebits.

In this section we consider a similar setup to the previous section, except now we fix the number of Bob's modes to 1 and assign the rest to Harold. Recall that we are interested in generating the highest amount of entanglement between Alice and Bob on average, thus the probability of detection patterns must be taken into account. More precisely, we are looking for the maximum of $\langle S(U) \rangle_H = \sum_{\underline{h}} P_H(\underline{h}, U) S(\rho_A(\underline{h}, U))$. Some patterns might yield a state with high entanglement, but be very unlikely to occur. In a practical setting we might prefer states that are less entangled but we can generate more consistently.

We first prove a technical lemma that will be useful later.

Lemma 2.5.7. *Given a probability distribution $(p_0, \ldots, p_N)$ such that $\sum_{j=0}^{N} j p_j = 1$, the sum $\sum_{j=0}^{N} p_j \log(j+2)$ is bounded by $\log 3$ which can be achieved by $p_1 = 1$ and $p_k = 0$ for $k \in \{0, 2, 3, \ldots, n\}$.*

Proof. Since $f(x) = \log(x+2)$ is a concave function, by Jensen's inequality $\sum_{j=0}^{N} p_j f(j) \leq f\left(\sum_{j=0}^{N} p_j j\right) = f(1) = \log 3$, which is achieved by substituting $p_1 = 1$ and $p_k = 0$ for

$k \in \{0, 2, 3, \ldots, n\}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Theorem 2.5.8. *Consider an interferometer with $d \geq 3$ modes, where both Alice and Bob have one mode and the other output modes are measured using photon counting detectors. Let the input be the $N = d$ unbunched Fock state. Then the maximal average entanglement that can be created between Alice and Bob is $\log 3$ ebits.*

Proof. First, notice that the average entanglement achievable by an $d = 2$ interferometer can be achieved for $d \geq 2$ by having modes 3 to d transform trivially, since photons in these modes will be detected with unit probability. Thus $\max \langle S(U_d) \rangle_H \geq \max \langle S(U_{d=2}) \rangle_H = \log 3$ ebits, $\forall d \geq 3$. The interferometer given in Section 2.5.4, Eq. (2.43) below achieves this.

Any $U \in \mathrm{U}(d)$ can be decomposed as in Fig. 2.10. Then the bottom left triangle (colored blue in the figure) is a unitary $V \in \mathrm{U}(d-1)$. Since the input is unbunched, Lemma 2.5.2 implies that each output from V has a mean photon number of 1. In particular, Bob's mode before beamsplitter W (red in the figure) will satisfy $\sum_k k q_k = 1$ where k is the number of photons occuring with probability q_k. Since the remaining beamsplitters (white in the figure) act only on Bob and Harold's systems, they have no effect on Alice's reduced state and can therefore be ignored.

Let the probability of detecting pattern $\underline{h}$ be $p_{\underline{h}}$, and the probability of detecting a total of $N_H = \|\underline{h}\|_1$ photons be $p_{N_H} = \sum_{\underline{h}:\|\underline{h}\|_1 = N_H} p_{\underline{h}}$. The average entanglement is

$$
\begin{aligned}
\langle S(U) \rangle_H = \sum_{\underline{h}} p_{\underline{h}} S(\rho_A(\underline{h})) &= \sum_{N_H=0}^{N} p_{N_H} \sum_{\underline{h}:\|\underline{h}\|_1 = N_H} p_{\underline{h}} / p_{N_H} S(\rho_A(\underline{h})) \\
&\leq \sum_{N_H=0}^{Ns} p_{N_H} \sum_{\underline{h}:\|\underline{h}\|_1 = N_H} p_{\underline{h}} / p_{N_H} \log (N - N_H + 1) \\
&= \sum_{N_H=0}^{N-1} p_{N_H} \log (N - N_H + 1),
\end{aligned}
\tag{2.33}
$$

where we've used the fact that the entanglement of $S(\rho_A(\underline{h}))$ is upper bounded by the Schmidt rank $\log(N - N_H + 1)$.

As the photon number found in modes 1 and 2 is set before the beamsplitter W, if N_H photons have been detected, then there were already N_H photons in modes 3 through d. Alice contributes one photon through her mode to their joint system, which implies that Bob must contribute $N - N_H - 1$ photons through mode 2, occurring with probability q_{N-N_H-1}. Therefore $p_{N_H} = q_{N-N_H-1}$ and recall that Bob's probability distribution is constrained by $\sum_{k=0}^{N-1} k q_k = 1$. By Lemma 2.5.7 $\sum_{N_H=0}^{N-1} q_{N-1} \log (N - N_H + 1) = \sum_{j=0}^{N-1} q_j \log (j + 2)$ is maximized for $j = N - N_H - 1 = 1$, that is $q_1 = 1$, yielding $\langle S(U) \rangle_H \leq \log 3$. This also implies that $N_H = N - 2$ photons are detected in the optimal situation. $\qquad$ $\square$

Note that this agrees with the bound in Theorem 2.5.5 found in the previous section, which follows from the entanglement measure property $\langle S(U) \rangle_H = \sum_{\underline{h}} P_H(\underline{h}, U) S(\rho_A(\underline{h}, U)) \leq S(\rho_A(U))$, where $\rho_A(U)$ is Alice's reduced density matrix before any measurement. Here

the maximum entanglement is $\log 3 < 2$ ebits. Moreover, adding more vacuum input modes will not affect this bound, as this would only change Bob's expected number of photons before the beamsplitter W to be at most 1 instead of exactly 1 as per Lemma 2.5.2.

2.5.2   Entanglement when Alice has many modes

In this section we give two bounds on entanglement for more general situations when Alice has more than one mode, based on the Schmidt rank of Alice's reduced state. They are independent of the input state or any interferometer transformation, depending only on the given number of photons and modes; we assume the latter is the same for both Alice and Bob. This generality comes at a price, however, in that the bounds loosen; we will discuss a conjectured tighter bound in the following section.

Dimensionality

By looking solely at the dimensions of Alice and Bob's Hilbert spaces, we can derive an entanglement bound as follows.

Proposition 2.5.9. *Let Alice's and Bob's joint postselected state have a total of N_S photons. Let Alice and Bob have $d_A = d_B$ modes. The Schmidt rank, ω, is at most*

$$\omega = 2\binom{d_A + \frac{N_S-1}{2}}{\frac{N_S-1}{2}} \qquad\qquad N_S \ \text{odd}, \qquad\qquad (2.34)$$

$$\omega = 2\frac{N_S + d_A}{N_S}\binom{d_A + \frac{N_S}{2} - 1}{\frac{N_S}{2} - 1} \qquad\qquad N_S \ \text{even}. \qquad\qquad (2.35)$$

Proof. Let Alice's and Bob's joint state be $|\psi_S(\underline{h}, U)\rangle = \sum_{k,j} C_{kj}(\underline{h}, U) |k\rangle_A \otimes |j\rangle_B$, where we include the possibility of no measurement ($d_H = 0$). The Schmidt decomposition is achieved by a state dependent change of basis such that

$$|\psi_S(\underline{h}, U)\rangle = \sum_{q=1}^{\min(\dim \mathcal{H}_A, \dim \mathcal{H}_B)} \lambda_q |q\rangle_A \otimes |q\rangle_B, \qquad\qquad (2.36)$$

where $\{|q\rangle_{A,B}\}$ are orthonormal bases for A and B, respectively.

Writing this state in terms of Alice and Bob's photon numbers we have $|\psi_S(\underline{h}, U)\rangle = \sum_{N_A=0}^{N_S} |\psi_S^{N_A, N_B}(\underline{h}, U)\rangle$ with $N_B = N_S - N_A$. The overlap $\langle \psi_S^{N_A, N_B}(\underline{h}, U) | \psi_S^{N'_A, N'_B}(\underline{h}, U)\rangle = 0$ for $N_A \neq N'_A$, $N_B \neq N'_B$ as these states belong to different Hilbert subspaces in the direct sum. The reduced density matrix is block diagonal – each block corresponds to a different (N_A, N_B) combination. We may therefore consider each subspace individually, where the maximal Schmidt rank is $\min(\dim \mathcal{H}_A^{N_A}, \dim \mathcal{H}_B^{N_B})$. The total number of Schmidt

coefficients is therefore at most

$$\omega = \sum_{N_A=0}^{N_S} \min\{\dim \mathcal{H}_A^{N_A}, \dim \mathcal{H}_B^{N_B}\} \tag{2.37}$$

$$= \sum_{N_A=0}^{N_S} \min\left\{\binom{d_A + N_A - 1}{N_A}, \binom{d_B + N_B - 1}{N_B}\right\}. \tag{2.38}$$

For $d_A = d_B$ this gives the result. $\qquad\square$

Since the entanglement is given by the number of nonzero Schmidt coefficients, this gives a bound on the entanglement $S \leq \log(\omega)$.

Linearity bound

Here we consider a bound due to the linearity of the interferometer transformations. In the following we do not assume anything about the form of the input Fock state, nor whether measurement occurs or not.

Proposition 2.5.10. *Given an N photon Fock state as input to a d-mode linear optical device, with Alice and Bob having d_A and d_B output modes respectively, the maximal entanglement achievable between Alice and the rest of the modes for any state is bounded by N ebits.*

Proof. Starting with the arbitrary linear optical mode transformation in Eq.(2.3), we can group the sum into Alice's modes and the 'rest':

$$\hat{a}_k^\dagger \mapsto \sum_{j=1}^{d} \hat{a}_j^\dagger U_{jk} = \sum_{j=1}^{d_A} \hat{a}_j^\dagger U_{jk} + \sum_{j=d_A+1}^{d} \hat{a}_j^\dagger U_{jk}$$
$$=: \hat{A}_k(U) + \hat{R}_k(U). \tag{2.39}$$

The degree one polynomials $\hat{A}_k(U)$, $\hat{R}_k(U)$ in the creation operators are not canonical raising operators, because e.g. $[\hat{A}_k(U), \hat{A}_{k'}(U)] \neq \delta_{kk'}$. This means that different monomials in $\{\hat{A}_k(U)\}_k$ do not necessarily give rise to orthogonal states; however, this can only reduce the Schmidt rank of the resulting state.

An arbitrary input Fock state is of the form $\prod_{k=1}^{d}(\hat{a}_k^\dagger)^{n_k}/\sqrt{n_k!} |\text{vac}\rangle$, so that the output state is of the form

$$\prod_{k=1}^{d} \frac{1}{\sqrt{n_k!}}(\hat{A}_k(U) + \hat{R}_k(U))^{n_k} |\text{vac}\rangle \tag{2.40}$$

i.e. it is a product of N terms, not all of which are necessarily different. We can rewrite it as

$$\mathcal{N} \prod_{k=1}^{N}(\hat{A}_{j_k}(U) + \hat{R}_{j_k}(U)) |\text{vac}\rangle, \tag{2.41}$$

where $j_k \in \{1, \ldots, d\}$ and $\mathcal{N}$ is the necessary normalization. The highest number of monomial terms in this product is bounded by 2^N and after tracing out Bob and Harold this also bounds the number of monomial terms that can be in Alice's reduced state. □

Consider a balanced 50:50 beam splitter coupling one of Alice's modes (say k) to one of Bob's modes (say $k + d_A$). If Alice's mode contained one input photon and Bob's none, we get 1 ebit of entanglement. Proposition 2.5.10 tells us we can only get up to N ebits using N photons, so as long as $N \leq d_A = d_B$, a beamsplitter coupling mode k with mode $k + d_A$ for $k = 1$ through N in this way would give us a state that achieves the bound.

The dimensionality (Section 2.5.2) and linearity bounds above hold for all d and all N. We can find numerically the photon number $N_L(d_A) \in \mathbb{N}$, which depends on the number of Alice's modes. For a given d_A it represents the number of photons up to which the linearity bound is smaller than the dimensionality bound. For $N > N_L(d)$, the dimensionality bound is a tighter limit on the entanglement (see Figure 2.11).

2.5.3   Hints of another bound

In this section we explore a potential bound that is motivated by numerical evidence (see Figure 2.11). While adding more photons to the interferometer increases the size of Alice's and Bob's Hilbert spaces, and according to the results from the previous section should allow for higher amounts of entanglement, we see that this is not what happens in general (assuming the number of modes that Alice and Bob have are fixed). Based on the analytical results from Section 2.5.1 and the numerical evidence for all cases up to $N = 7$ photons and $d_A = d_B = 3$ modes, we make a conjecture that there is another bound which seems to arise from the unitarity of the mode transformation.

Conjecture 2.5.11. *For N unbunched photons input into an interferometer with d_A Alice and d_B Bob output modes with $N > d_A + d_B$, the average amount of entanglement, obtained over Harold's measurements, is bounded above by the maximal average amount of entanglement achieved when $N = d_A + d_B$.*

We provide numerical evidence supporting this "unitarity bound" for various numbers of input photons and modes. We assume that the input states are unbunched, ancillas and measurement are allowed, and Alice and Bob have the same number of modes; $d_I = N$, $d_V \geq 0$, $d_A = d_B \geq 1$ and $d_H \geq 0$.

Propositions 2.5.9 and 2.5.10 provide tight entanglement bounds when all input photons are kept in the system, i.e. when there is no detection. We know that it is possible to postselect states that exceed these bounds, but because we are interested in average entanglement these cases must be weighted with their heralding probabilities. Our findings are consistent with a generic trade off between these two quantities, leading to a bounded average entanglement.

Figure 2.11 shows the results of numerical optimization of the average entanglement given by Eq. (2.18) for various numbers of input photons and modes. We can see how the linearity and dimensionality bounds of Sec. 2.5.2 are indeed limiting the entanglement.
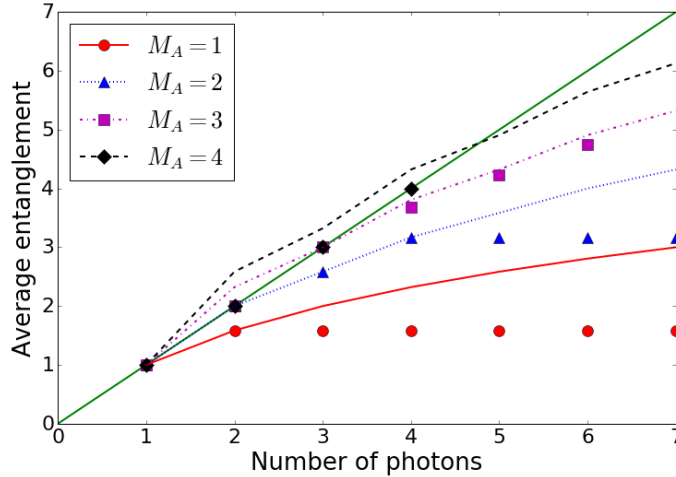
Figure 2.11: Plot of the maximum average entanglement found through numerical optimization, along with the dimensionality and linearity bounds for $d_A = d_B$. The input are unbunched states. If $N > d_A + d_B$, the remaining $d_H = N - d_A - d_B$ modes contain detectors. The green solid (straight) line is the linearity bound. Other lines are dimensionality bounds for the value of d_A whose dots have the corresponding colour. We can see that the values of N_L for specific d_As are: $N_L(1) = 1, N_L(2) = 2, N_L(3) = 3, N_L(4) = 4$. The markers show values found through numerical optimization.

We also see the appearance of what looks like a third bound, seemingly when the number of photons is larger than the total number of modes in the system $(d_A + d_B)$. This new behaviour is not captured by the bounds we have obtained and we conjecture that it is due to the unitarity of the interferometric transformation. This leads to the hypothesis that the maximum possible average entanglement, in situations with unbunched input and Alice and Bob have the same number of modes, can be reached using a $(d_A + d_B)$-mode interferometer with $d_A + d_B$ photons.

2.5.4 Optimal interferometers

Finally, in this section we report some of the explicit interferometers (unitaries) that produce the optimal entanglement found for small number of modes.

In the case of $d_A = d_B = 1$ and a single photon $N = 1$, the well known balanced 50:50 beamsplitter is optimal,

$$BS_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \tag{2.42}$$

This is familiar, as in single-rail encoding it creates a Bell state. Let $\theta = \frac{1}{2} \arccos(1/\sqrt{3})$. When we input two photons, $N = 2$, with one in each mode, the unitary

$$BS_2 = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{2.43}$$

produces a state with $\log 3$ ebits of entanglement. Conjecture 2.5.11 says that for all higher photon numbers, $\log 3$ will still be the maximum, achieved by using BS_2 between any pair

of Alice and Bob's modes and identity on all the others (they are just routed straight to the detectors).

For $d_A = d_B = 2$, we have that all the optimal interferometers are actually combinations of BS_1 and BS_2. An example for $d = N = 4$ is:

$$
\begin{bmatrix}
\cos\theta & 0 & 0 & \sin\theta \\
0 & \cos\theta & \sin\theta & 0 \\
0 & -\sin\theta & \cos\theta & 0 \\
-\sin\theta & 0 & 0 & \cos\theta
\end{bmatrix}, \tag{2.44}
$$

where as before $\theta = \frac{1}{2}\arccos\left(1/\sqrt{3}\right)$. This interferometer corresponds to a BS_2 beamsplitter between modes 1 and 4 and another BS_2 beamsplitter between modes 2 and 3, giving $\log 9 \approx 3.17$ ebits of entanglement. When $N = 3$, the optimal value of $\log 6 \approx 2.58$ ebits is achieved by using BS_2 on modes 1 and 4 and BS_1 on modes 2 and 3. For $N = 2$, the maximum of 2 ebits is achieved by two BS_1 beamsplitters, similar to $N = 4$ case. Finally, for $N = 1$ we just use a single BS_1 to achieve 1 ebit.

2.6 Conclusion

Bound (ebits)	Parameters	Input state	Page
$O(\log N)$	$d_A = 1, d_B = 1, d_H = 0$	Bunched	51
2	$d_A = 1, d_B \geq 1, d_H = 0$	Unbunched	54
$\log 3$	$d_A = d_B = 1, d_H \geq 1$	Unbunched	56
$\log\left(2\binom{d_A + \frac{N-1}{2}}{d_A}\right)$	$d_A = d_B$, N odd	Any state in the fixed N subspace of Fock space	57
$\log\left(2\frac{N+d_A}{N}\binom{d_A + \frac{N}{2}-1}{d_A}\right)$	$d_A = d_B$, N even	Any state in the fixed N subspace of Fock space	57
N	$d_A = d_B$	Fock state	58

Table 2.1: Entanglement bounds proven in this Chapter. The notation is as defined in Section 2.2 (see Figure 2.2).

This Chapter tackles the generation of entanglement in purely linear optical systems – an important question for linear optical quantum computing, be it universal (see Section 1.4) or near-term (see Section 1.4.6). We considered entanglement found in dual-rail encoding of qubits, but also the mode entanglement with no encoding (where the excitation number of a mode is understood as a qudit, although this is not a proper qudit encoding as we cannot just simply represent a whole qudit hilbert space).

For the dual-rail encoding, previous research has demonstrated a no-go result for two photons and interferometers which generate on-demand Bell states for four photons. We expand upon these results by showing a no-go result for three photons with any number

of modes. We also demonstrate strong numerical results that the existing four photon Bell state generator (with a success probability of 0.1875) is optimal in the case of eight modes. This does not eliminate the possibility of there being more efficient Bell state generators with four photons and more modes than eight, and the numerical methods used could be used with higher number of modes (up to 12 has been tested). The numerical optimization was also attempted for five photons in ten modes, and six photons in 12 modes, with no immediate improvement of success probability. However, these tests were not as comprehensive as four photons in eight modes so it is not possible to make any definitive claims here. The numerical optimization was not the type that guarantees a global minimum, and further work could consist of developing a new, different method which does give a bound to this minimum.

Another interesting development is the lack of non-orthogonal Bell states being generated by the optimal circuits found through numerical optimization. This gives a possible direction of how to prove the bound on the entanglement generation in this case, starting by demonstrating that only certain types of Bell states can be generated at the same time.

We then looked into the average entanglement found in random circuits with no encoding present. For the same setup as the above mentioned Bell state generator (that of four photons in eight modes, two mode Alice, two mode Bob and two photons used for heralding), we find that most random circuits produce more ebits than that of the Bell state generator (see Figure 2.7). If entanglement retains an important role in LOQC as it does due to MBQC being the leading scheme, we might wonder if there are more efficient ways of encoding the information and generating these highly entangled states. We therefore carry on investigating generation of entanglement in this more general case, where these is no encoding as this will give ultimate linear optical bounds of how well we can do.

We see two types of behaviour. In the case of bunched photon input and single mode Alice, we find the entanglement can grow as $\log N$ where N is the number of photons. On the other hand, looking at the case of at most a single photon per input mode (as in, for example, boson sampling [162], see Section 1.4.6), a single mode Alice and no measurement, the entanglement is bounded above by 2 ebits regardless of how many photons are present. This is a surprising result as in both cases the size of Alice's Hilbert space is growing with the number of photons (and the size of Bob's system is not even taken into consideration) yet in the latter case there is a strict constraint. If we also restrict Bob to a single mode and furnish the remaining modes with number resolving detectors, the expected entanglement drops further, and is bounded by $\log 3$ ebits.

To understand these limitations of linear optical interferometers and their manipulation of photons we turn towards some more general results. We find provable universal bounds on the mode entanglement stemming from the dimensionality of the bipartite Fock states involved, and from the linearity of the optical transformation. Finally, we conjecture a third bound due to limitations of unitarity which extends the previously mentioned constant bound in the case of Alice having a single mode to multi-mode Alice, and we provide

numerical evidence for this conjecture.

Further work could look into proving this conjecture, which we speculate is due to unitarity constraints which have not been previously taken into consideration. Reflecting back on how these bounds behave in the restricted case of dual-rail qubit encoding would be another interesting further direction. It might be possible to extract some ideas on bounds of entanglement generation in specific encodings. The maximum mode entanglement is summarized in Table 2.1.

While we saw that the generation of single photon states as well as their entanglement is challenging, the more naturally arising Gaussian states can be easier to generate and manipulate, with the elements used to build up optical circuit similar to one we already discussed: beamsplitters, phaseshifters, with addition of squeezers [178].Gaussian states have already been researched significantly in terms of entanglement generation (and also more generally in terms of quantum information). A few results exist to help determine whether certain bipartite Gaussian state is separable (for example when Alice and Bob have one mode each, similar to our scenario presented in Figure 2.8, or when Alice has one mode, and Bob has many, similar to Figure 2.9). It is also known that if the input state is a classical state (a coherent state), even in multi-mode settings with any number of beamsplitters, the final state remains separable. A useful entanglement resource is two mode squeezed state, which can be produced using one-mode squeezed state and a beamsplitter. It is also possible to distill many weakly entangled Gaussian state to gain a a highly entangled one, although it does require non-Gaussian operations. As the entanglement of the pure Gaussian states is also evaluated using von Neumann entropy, and the setup used for bipartite entanglement is similar to the one in this Chapter, it would be interesting future work to see the how Gaussian input states could be compared to the bounds found in the Table 2.1 (as the work here was focused on states with fixed particle number only).

).

particle number only).

).

LINEAR OPTICS IN FIRST QUANTIZATION

Interference lies at the heart of quantum mechanics, and thus its promise of fundamental advantages over non-quantum technologies, with far-reaching ramifications in communication, metrology, simulation and computation. The nemesis of quantum interference is distinguishability, with the HOM effect [11] being a prototypical example. Recent advances in scaling linear optics for universal quantum computation [13, 14, 19], and the race to demonstrate quantum computational 'supremacy' via analog computations that sample the scattering amplitudes of multipartite states [44, 70, 75, 77, 179–183], highlight the need for a thorough understanding of distinguishability in multimode quantum interference [26–35].

Rather than the usual second quantized approach, we can gain insight by bringing quantum information concepts to bear in first quantization [84, 184–186]. Distinguishability can then be modelled, for example, as entanglement between controlled and uncontrolled degrees of freedom of individual particles, with loss of interference being caused by the decoherence that results when the uncontrolled Hilbert space is marginalized. This can be formalized by observing that bosonic (and fermionic) Fock states of two degrees of freedom can have natural Schmidt decompositions, corresponding to so called unitary-unitary duality in many-body physics [187].

In principle the formalism accommodates any number of particles and modes, and we show how this generalises for multimode quantum interference, taking a representation theoretic approach; this complements a number of generalizations in the literature [93, 113, 124, 141, 145, 148, 188–190].

The first part of this Chapter is an introduction to the representation theory necessary for understanding, constructing, and applying the basis which will be used to model distinguishability in this thesis (Section 3.1). It first introduces the relevant elements from the theory of the symmetric and the unitary groups (see Section 3.1.1). Next, it introduces the Schur-Weyl duality and unitary transform, the workhorse of the framework, along with relevant notation in Section 3.1.2. This is followed by Section 3.1.3, which applies all the previously introduced concepts onto particles (more specifically bosons), and Section 3.1.4 then shows how to construct the relevant Schur-Weyl basis. The section that follows, Section 3.2 then offers a model of distinguishability using yet another duality,

this time between unitaries acting on the particle states. It starts off by describing the behaviour of bosons with two degrees of freedom, then gives a new model for measurements in linear optics. The seminal HOM effect in this new framework is examined, finishing the section by using the formalism to compute number of parameters to represent any distinguishable state. In Section 3.3, we present some basic states and their scattering probabilities, describing them in this new framework. The Chapter finishes with an example of calculations for three bosons in two and three modes, bringing together all the previously introduced tools. We finish this section with an example of calculations for three bosons in two modes, bringing together all the previously introduced tools. Then the seminal HOM effect in this new framework is examined, finishing the section by presenting different situations we need to think about to define distinguishability.

3.1 Representation theory and the Schur-Weyl basis

In this section we introduce the necessary theory to get from the symmetric and unitary groups, to the Schur-Weyl basis of particles indexed by Young tableaux. There are no novel results in this Section, most of the concepts introduced are standard group and representation theory results. The notation introduced in Section 3.1.2 based on the occupation numbers, while not novel, to the best of our knowledge has not been used in such a form before. The example given in Section 3.4.1 is also an analysis of states that have been looked at many times before, but is given here in a form demonstrating more explicitly and clearly the links between the irreps, the scattering matrix construction made famous by the boson sampling paper, and immanants (see Section 5.1), than in other literature on the topic.

3.1.1 Representation theory of the symmetric and unitary groups

We will first recall some basic definitions and results from the representation theory of the symmetric and unitary groups. Further details on the material presented in this section can be found in standard and historical group and representation theory texts [191–195].

   A *group homomorphism* is a structure preserving map from group G_1 to group G_2, that is $\mathcal{T}: G_1 \to G_2$, $\mathcal{T}(a)\mathcal{T}(b) = \mathcal{T}(ab)$, $\forall a, b \in G_1$. The *general linear group* of a vector space V, $\mathrm{GL}(V)$, is the group of all invertible linear maps of V to itself, equipped with function composition as the group operation (i.e. the set of all *automorphisms* of V). A *representation* $\mathcal{T}$ of a group G is a homomorphism $\mathcal{T}: G \to \mathrm{GL}(V)$. More commonly, if $\dim V = N$ over some field K, we can pick a basis for V, and the representation becomes a homomorphism $\mathcal{T}: G \to \mathrm{GL}(N, K)$, where $\mathrm{GL}(N, K)$ is the group of invertible $N \times N$ matrices over the field K. The dimension of such a representation is taken to be N. It is also common to identify the representation by referring to its *carrier space* (instead of specifying $\mathcal{T}$ we would specify V). Given two representations $\mathcal{T}_1$ and $\mathcal{T}_2$ and their carrier spaces V_1 and V_2, respectively, if there is a unitary change of basis, U, such that $U\mathcal{T}_1(g)U^\dagger = \mathcal{T}_2(g)$ for all $g \in G$ then the two representations are said to be equivalent,

$V_1 \cong V_2$.

We call a subspace $W \subseteq V$ *invariant* under a representation $\mathcal{T}$ if the image is the subspace itself, that is $W = \{w \mid \mathcal{T}(g)w \in W, \forall g \in G\}$. An *irreducible representation*, or *irrep*, $\mathcal{T}$ has no proper invariant subspaces (subspaces other than itself and 0). If we can split a representation $\mathcal{T}$ into a direct sum of irreducible representations, then it is *fully reducible*. These irreducible representations need not be distinct. The irreps reveal the existence of more "natural" bases for describing the group action. A representation $\mathcal{T}$ is called *unitary* if $\mathcal{T}^{\dagger}(g) = \mathcal{T}(g^{-1})$, $\forall g \in G$. Also, if the carrier space is a finite dimensional vector space, then the representation is finite dimensional. One can prove that every finite dimensional unitary representation is fully reducible. The symmetric group, S_N, and unitary group, $U(d)$, that we will introduce later on, both have irreps that are fully reducible.

Next, we define characters and conjugacy classes whose importance will be made clear when we discuss irreps of the symmetric group. The *characteristic* of g in a group G under representation $\mathcal{T}$ is $\chi(g) = \mathrm{Tr}\,[\mathcal{T}(g)]$. The *conjugacy classes* are then formed by defining an equivalence relation on the characteristics of group elements $a \sim b$ if $\chi(a) = \chi(b)$. The number of distinct irreducible representations of a finite group G is equal to the number of conjugacy classes of G. Each of these classes will be denoted by a characteristic $\chi_k = \chi(g)$ for any g belonging to the same conjugacy class k. The *character* χ of a representation $\mathcal{T}$ of a group G is defined as the set of all the characteristics from different conjugacy classes. Further, each of the irreps of $\mathcal{T}$ has its own character χ and is uniquely determined by it.

Young diagrams and Young tableaux

A *partition* of an integer N, denoted $\lambda \vdash N$, is a sequence of numbers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > 0$ and $\sum_k \lambda_i = N$. The set of all partitions of a number N will be denoted as $\mathrm{Par}(N)$ and the set of all partitions of a number N with at most d parts will be denoted as $\mathrm{Par}(N, d)$. Partitions of an integer λ are in one-to-one correspondence with *Young diagrams* of N boxes. Given a partition $(\lambda_1, \ldots, \lambda_k)$, its corresponding Young diagram will have λ_i boxes in the row i, and as $\lambda_1 \geq \ldots \geq \lambda_k$ the number of rows will therefore be k and maximum number of boxes in a row will be λ_1. The partition λ will be used interchangeably with the Young diagram that corresponds to it. If we want to be specific we are referring to a Young diagram of partition λ for clarity, it will be referred to as Y^{λ}. The λ will also be referred to as the *shape* of the Young diagram. The conjugate of a Young diagram is a Young diagram in which columns are exchanged for rows. Some examples of Young diagrams and conjugates are given below.

$$(3,2,2,1) \in \mathrm{Par}(8,4), \qquad (3,2,2,1) \equiv \;\young(~~~,~~,~~,~)\;, \qquad \young(~~~,~~,~~,~)^{*} = \young(~~~~,~~~,~)\;,$$

$$(4,2,1) \in \mathrm{Par}(8,4), \qquad (4,2,1) \;\equiv\; \young(~~~~,~~,~)\;, \qquad \young(~~~~,~~,~)^{*} = \young(~~~,~~,~,~)\;.$$

A *Young tableau* is a Young diagram filled by the numbers from 1 to N (no repetition and in any ordering). Let $\text{Tab}(\lambda)$ be the set of all Young tableaux of shape λ. A *standard Young tableau* is a filling of the Young diagram boxes with numbers 1 through N such that the numbers strictly increase in each row and each column (again no repetition of numbers, but there is a specific ordering). We also single out the standard Young tableau $T(\lambda)$ where the numbers are placed into the boxes consecutively down the columns from left to right. Let $\text{STab}(\lambda)$ be the set of all standard Young tableaux of shape λ. As an example for partition $\lambda = (3, 2)$ we have

$$\text{STab}((3,2)) = \left\{ \begin{smallmatrix}1&2&3\\4&5\end{smallmatrix}, \begin{smallmatrix}1&2&4\\3&5\end{smallmatrix}, \begin{smallmatrix}1&2&5\\3&4\end{smallmatrix}, \begin{smallmatrix}1&3&4\\2&5\end{smallmatrix}, \begin{smallmatrix}1&3&5\\2&4\end{smallmatrix} \right\},$$
$$\text{T}((3,2)) = \begin{smallmatrix}1&3&5\\2&4\end{smallmatrix}.$$

A *semi-standard Young tableau* is a filling of the Young diagram boxes with numbers 1 through d such that the numbers are weakly increasing in each row and strictly increase in each column (repetition is allowed, but there is a specific ordering). Let $\text{SSTab}(\lambda)$ be the set of all semi-standard Young tableaux of shape λ. As an example of some of the semi-standard tableaux we have for shape $\lambda = (3, 2)$

$$\text{SSTab}((3,2)) \supset \left\{ \begin{smallmatrix}1&1&2\\2&3\end{smallmatrix}, \begin{smallmatrix}1&1&3\\2&2\end{smallmatrix}, \begin{smallmatrix}1&1&1\\2&2\end{smallmatrix}, \begin{smallmatrix}1&2&2\\3&4\end{smallmatrix}, \begin{smallmatrix}1&1&1\\5&5\end{smallmatrix} \right\}.$$

The numbers in the semi-standard Young tableaux are called the *content* of the Young tableaux, but we will also refer to them as *weights* for reasons that will become apparent later. A tableau can then be referred to by its shape, λ, but also its weight, $\underline{n}$ (often denoted μ but $\underline{n}$ serves our subsequent notation better). The content $\underline{n} = (n_1, n_2, ..., n_N)$ tells us that the j-th number appears n_j times in the tableau. Let $\text{SSTab}(\lambda, \underline{n})$ be the set of all semi-standard Young tableaux of shape λ and weight $\underline{n}$. For example for partition $\lambda = (3, 2)$, we can now add the description of some content, $\underline{n} = (2, 2, 1)$ which gives the following semi-standard tableaux only:

$$\text{SSTab}((3,2),(2,2,1)) = \left\{ \begin{smallmatrix}1&1&2\\2&3\end{smallmatrix}, \begin{smallmatrix}1&1&3\\2&2\end{smallmatrix} \right\}.$$

The content $\underline{n} = (2, 2, 1)$ indicates that the diagram should be filled with number 1 twice, number 2 twice, and number 3 once.

Let the semi-standard Young tableau where the row j is filled with number j be referred to as $\text{ST}(\lambda)$ (here $\underline{n} = \lambda$). This semi-standard Young tableaux of weight $\underline{n} = \lambda$ is unique.

Looking at a more complicated example of the above mentioned tableaux, for partition $\lambda = (4, 2, 1, 1)$ we have

$$\lambda = (4,2,1,1), \qquad \text{T}(\lambda) = \begin{smallmatrix}1&5&7&8\\2&6\\3\\4\end{smallmatrix}, \qquad \text{ST}(\lambda) = \begin{smallmatrix}1&1&1&1\\2&2\\3\\4\end{smallmatrix}.$$

Now taking content to be $\underline{n} = (2, 3, 2, 1)$, we get the semi-standard tableaux set

$$\mathrm{SSTab}(\lambda, \underline{n}) = \left\{ \begin{array}{c}\begin{array}{|c|c|c|c|}\hline 1 & 1 & 2 & 2 \\\hline 2 & 3 \\\cline{1-2} 3 \\\cline{1-1} 4 \\\cline{1-1}\end{array} , \begin{array}{|c|c|c|c|}\hline 1 & 1 & 2 & 3 \\\hline 2 & 2 \\\cline{1-2} 3 \\\cline{1-1} 4 \\\cline{1-1}\end{array}\end{array} \right\}.$$

A semi-standard Young tableau of partition λ will be referred to as $\mathrm{T}^{\lambda, j}$ where j indexes the content. There is no ordering to these tableaux, the content is mapped to numbers for the sake of labelling. In the case of standard Young tableaux we could impose a partial ordering, but we do not need this here. We see that the standard Young tableaux are a subset of the semi-standard ones (with weight $\underline{n} = \underline{1} = (1, 1, \dots, 1)$, and as long as $d \geq N$), therefore they are also referred to in the same way.

Let all the boxes in the Young diagram be marked as b_{jk} where j marks the row the box is in and k marks the column the box is in. The labelling of boxes starts from 0. The *hook length* of the box b_{jk}, $h(j, k)$, is the number of boxes to the right of it in the row j plus the number of boxes down from it in the column k including the box itself. We highlight the following two results.

Lemma 3.1.1 (Hook formula). *The number of standard Young tableaux corresponding to Young diagram Y^λ, for any $\lambda \in \mathrm{Par}(N, d)$, is given as*

$$f^\lambda = \frac{N!}{\prod_{j,k} h(j, k)} \tag{3.1}$$

where $0 \leq j < d$ and $0 \leq k < \lambda_j$.

We will see later that this also gives us the dimension of the λ irrep of the symmetric group, $\dim (\mathcal{S}^\lambda) =: d_{(\lambda)}$. Similarly, we can also count how many semi-standard Young tableaux there are using the hook content formula (which is a special case of Weyl's dimension formula).

Lemma 3.1.2 (Hook content formula). *The number of semi-standard Young tableaux corresponding to Young diagram Y^λ, for any $\mathrm{Par}(N, d)$, of all weights $\underline{n}$ in $\mathrm{Par}(N, d)$, is*

$$t^\lambda(d) = \prod_{j,k} \frac{d - j + k}{h(j, k)} \tag{3.2}$$

where $0 \leq j < d$ and $0 \leq k < \lambda_j$.

Another number that is worth mentioning is the *Kostka number* or *Kostka coefficient*, $K_{\lambda \underline{n}}$ which specifies the number of semi-standard tableaux of shape λ and content $\underline{n}$. While there is no general formula, there are a few properties worth mentioning.

- $K_{\lambda \lambda} = 1$ for any partition λ (the only possibly filling is $\mathrm{ST}(\lambda)$).

- $K_{\lambda \underline{1}}$ is just equal to the number of standard tableaux of shape λ, that is f^λ.

- If $\underline{n}$ and $\underline{n}'$ have the same entries, just in different order (for example $\underline{n} = (3, 1, 2)$ and $\underline{n} = (1, 3, 2)$), their Kostka numbers are the same, $K_{\lambda \underline{n}} = K_{\lambda \underline{n}'}$.

For example, we have the following dimensions and Kostka numbers in the case of partition $(2,1) \in \text{Par}(3)$

$$\lambda = (2,1), \qquad f^\lambda = 2, \qquad \text{STab}(\lambda) = \left\{ \begin{array}{|c|c|}\hline 1 & 2 \\\hline 3 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 3 \\\hline 2 \\\cline{1-1}\end{array} \right\},$$

$$d = 3, \qquad t^\lambda(d) = 8, \qquad \text{SSTab}(\lambda) = \left\{ \begin{array}{|c|c|}\hline 1 & 1 \\\hline 2 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 1 \\\hline 3 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 2 \\\hline 2 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 3 \\\hline 3 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 2 & 2 \\\hline 3 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 2 & 3 \\\hline 3 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 2 \\\hline 3 \\\cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 3 \\\hline 2 \\\cline{1-1}\end{array} \right\},$$

$$K_{(2,1)(2,1,0)} = 1 = K_{(2,1)(2,0,1)}, \qquad K_{(2,1)(1,1,1)} = 2 = f^\lambda.$$

Representation theory of the symmetric group

We recall some important properties of the representations of the symmetric group using the above definitions and results.

The group of all $N!$ permutations of N objects is the *symmetric group* S_N. Any permutation $\sigma \in \text{S}_N$ can be written as a product of cycles. All permutations that have the same cycle structure (same number of cycles of the same order) are in the same conjugacy class. Cycle structures can be specified by the partitions of N (which we recall are interchangeable with Young diagrams), therefore the conjugacy classes of the symmetric group S_N can be labelled by Young diagrams as well. Further, S_N is a finite group, consequently the number of irreducible representation of S_N is the number of conjugacy classes, that is the number of Young diagrams.

We focus on two permutation subgroups based on how they affect some chosen tableau T. The action of some permutation $\sigma \in \text{S}_N$ on a tableau T is that of exchanging the numbers in the boxes, e.g., $(12)\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\\hline 4 & 5 \\\cline{1-2}\end{array} = \begin{array}{|c|c|c|}\hline 2 & 1 & 3 \\\hline 4 & 5 \\\cline{1-2}\end{array}$. The *row group* of T and the *column group* of T are defined as

$$R(T) = \{\sigma \in \text{S}_N \mid \sigma \text{ preserves the rows of } T\}, \tag{3.3}$$

$$C(T) = \{\sigma \in \text{S}_N \mid \sigma \text{ preserves the columns of } T\}. \tag{3.4}$$

That is, the row (column) group contains permutations which permutes the row (column) entries within a row For example,

$$\begin{array}{|c|c|}\hline 1 & 2 \\\hline 4 & 5 \\\hline\end{array} = \left\{ \begin{array}{|c|c|}\hline 1 & 2 \\\hline 4 & 5 \\\hline\end{array}, \begin{array}{|c|c|}\hline 2 & 1 \\\hline 4 & 5 \\\hline\end{array}, \begin{array}{|c|c|}\hline 1 & 2 \\\hline 5 & 4 \\\hline\end{array}, \begin{array}{|c|c|}\hline 2 & 1 \\\hline 5 & 4 \\\hline\end{array} \right\}$$

is a set of all tableaux where permutations of the numbers in the two rows keep the entries in a row invariant.

$$R\left(\begin{array}{|c|c|}\hline 1 & 2 \\\hline 4 & 5 \\\hline\end{array}\right) = \{e, (12), (45), (12)(45)\}$$

is the corresponding row group of $\begin{array}{|c|c|}\hline 1 & 2 \\\hline 4 & 5 \\\hline\end{array}$. Then the *row symmetrizer* and the *column skew*

symmetrizer are defined as

$$r(T) = \sum_{r \in R(T)} r \tag{3.5}$$

$$c(T) = \sum_{r \in C(T)} \text{sgn}(c)c \tag{3.6}$$

in the group algebra of S_N, where $\text{sgn}(c)$ can be defined as the parity of the permutation c (-1 for odd number of transpositions, 1 for even number of transpositions in the decomposition of the permutation). We define the *Young symmetrizer* as

$$s(T) = c(T)r(T). \tag{3.7}$$

Finally, the normalized Young symmetrizer for a standard tableau T of shape λ is

$$P_T = \frac{f^\lambda}{N!} s(T) \tag{3.8}$$

which is a minimal projection of the group algebra of S_N for the shape λ as there are f^λ standard tableaux of shape λ. It is known that the sum of the squares of the dimensions of inequivalent irreps is equal to the order of the group, and in case of symmetric group the dimension of an irrep λ is equal to number of standard tableaux of shape λ, $\sum_\lambda (f^\lambda)^2 = N!$. For a finite group, there is a one-to-one relation of irreps to minimal projectors of the group algebra.

Let $\mathbb{C}^d$ be a d-dimensional complex vector space with some basis $\{\psi_1, ..., \psi_d\}$ and $(\mathbb{C}^d)^{\otimes N}$ be the N-fold tensor product. The natural representation $\mathcal{S}$ of S_N on the space $(\mathbb{C}^d)^{\otimes N}$ is that of permutation matrices, permuting the tensor factors. That is for $\sigma \in S_N$, the representation $\mathcal{S}(\sigma)$ acts on N-tensors as

$$\mathcal{S}(\sigma)\psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_N = \psi_{\sigma^{-1}(1)} \otimes \psi_{\sigma^{-1}(2)} \otimes \cdots \otimes \psi_{\sigma^{-1}(N)}. \tag{3.9}$$

As mentioned before, $\mathcal{S}$ is fully reducible, thus we can express $\mathcal{S}$ in terms of the disjoint irreps $\mathcal{S}^\lambda$ using a basis change, that is

$$\mathcal{S}(\sigma) \cong \bigoplus_{\lambda \in \text{Par}(N,d)} \mathcal{S}^\lambda(\sigma) \otimes \mathbb{1}_{m_{(\lambda)}} \tag{3.10}$$

$$(\mathbb{C}^d)^{\otimes N} \cong \bigoplus_{\lambda \in \text{Par}(N,d)} \mathbb{C}^{d_{(\lambda)}} \otimes \mathbb{C}^{m_{(\lambda)}}. \tag{3.11}$$

where $m_{(\lambda)}$ is the multiplicity of the λ irrep. In general, we will denote the carrier space of the irreducible representations λ of S_N as $\mathbb{C}^{d_{(\lambda)}}$ and its dimension as $d_{(\lambda)}$.

For a tableau $T \in \text{Tab}(\lambda)$ we can construct a corresponding tensor, $\Psi_T = \prod_k \psi_{r_k}$ where r_k is the row in which the number k occurs. Then $\mathbb{C}^{d_{(\lambda)}} = \text{span}\{\Psi_T \mid T \in \text{Tab}(\lambda)\}$. It can be proven that the tensors constructed using the Young symmetrizers corresponding

to the standard tableaux of shape λ, $\{s(T)\Psi_T \mid T \in STab(\lambda)\}$, give a basis for the carrier space $\mathbb{C}^{d(\lambda)}$. Therefore, not only do the Young diagrams label the irreps of the symmetric group, the dimension of the irrep is equal to the number of corresponding standard Young tableaux, and the standard tableaux label basis states of the irrep (which is why the dimension of an irrep λ is equal to f^λ as mentioned before).

While there are more interesting results in the representation theory of the symmetric group, we will introduce them as they are needed later in the text.

Representation theory of the unitary group

We will also need to understand some properties of the unitary group to carry on with our discussion. An element U of the unitary group $U(d)$ is a $d \times d$ unitary matrix U with an adjoint matrix $U^\dagger$ such that $UU^\dagger = \mathbb{1}_d$. Let the unitary group act on the previously defined N-tensors as

$$\mathcal{U}(U)\psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_N = U\psi_1 \otimes U\psi_2 \otimes \cdots \otimes U\psi_N. \qquad (3.12)$$

To understand the irreducible representations of the unitary group in detail we would have to cover Lie groups and Lie algebras, which is out of the scope of this thesis. Instead, we will just state some of the relevant results and refer the reader to a standard group theory text [193, 195].

The maximal torus of $U(d)$ is the Abelian subgroup of all the $d \times d$ diagonal matrices. The one-dimensional irreducible representations of this torus are labelled by a set of weights $\underline{n} = (n_1, n_2, \ldots, n_d)$ (there is some choice to what these weights are, and we will be using non-negative integers so $\underline{n} \in \mathbb{N}_0^d$, and we will often drop the trailing zeros). The weights are commonly marked with μ, and they split the $U(d)$ carrier space into *weight space* $\mathbb{C}^{\underline{n}}$ (sometimes referred to as μ-weight space when the weights are denoted by μ). Weight spaces of distinct weights are orthogonal.

We can impose a lexicographical ordering to these weights (given $\underline{n} = (n_1, ..., n_d)$ and $\underline{n}' = (n'_1, ..., n'_d)$, $\underline{n} > \underline{n}'$ if for the first j where n_j and n'_j are different, $n_j > n'_j$). Notice that this is the weight description of an N-tensor that is in lexicographical order, not the description in terms of basis states. It can be proven that for every irrep of $U(d)$, there is a unique highest weight occurring with a multiplicity of one. Therefore the irreducible representations of $U(d)$ can be completely specified by their highest weights, with $\mu = \lambda$. The highest weight determines a highest weight state in the carrier space of irrep $\{\lambda\}$ of $U(d)$, and this state is annihilated by an appropriate set of raising operators (see Figure 3.1). The lowering operators act on states to achieve the following change in weight

$$\mathcal{L}_{jk} : (n_1 \ldots n_j \ldots n_k \ldots n_d) \to (n_1 \ldots n_j - 1 \ldots n_k + 1 \ldots n_d). \qquad (3.13)$$

As the ordering of the weights is lexicographical, j is always a smaller integer than k. The corresponding lowered state is considered to be "lower" than the initial state, and some
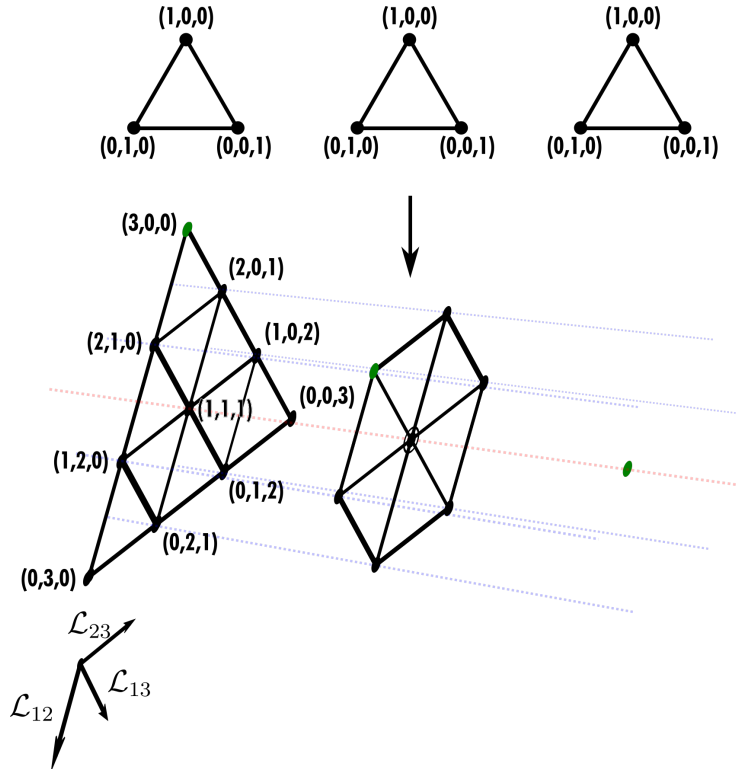
Figure 3.1: The first row represents the weight diagrams of three copies of U(3), that is single particle basis states with the corresponding weights $((1,0,0), (0,1,0),$ and $(0,0,1))$. Dots represent the weights and lines represent lowering and raising operator links between the weights. The space then decomposes into $(3,0,0)$ irrep (the large triangle on the left), two copies of $(2,1,0)$ irrep (hexagon in the middle), and $(1,1,1)$ irrep (single dot on the right), labelled by the highest weights. The highest weight state is marked in each by a green dot (top of the triangle, top left dot of the hexagon). Circles denote the number of extra copies of that weight, refereed to as "inner multiplicites" in this thesis (only the middle dot of the hexagon has an inner multiplicity). The dimension of the weight space of an irrep is given by the number of dots and circles. The bottom arrow represent lowering operators and how they act on the weight space given in the previous two rows.

normalization is required. Similarly the raising operator does the reverse

$$\mathcal{R}_{kj} : (n_1 \ldots n_j \ldots n_k \ldots n_N) \to (n_1 \ldots n_j + 1 \ldots n_k - 1 \ldots n_N). \qquad (3.14)$$

When the raising operator is applied to the highest weight it is annihilated, and similarly for lowest weights and lowering operators. The Kostka coefficient $K_{\lambda \underline{n}}$ gives us the multiplicity of the weight $\underline{n}$ in the irrep of U(d) with highest weight λ.

The lowering and raising operators and the weights present in an irrep can be found visually using weight diagrams (construction of which can be found in standard representation theory texts and will not be further discussed here). In Figure 3.1, we see an example of the irreps of three copies of U(3). The irreps of U(3) are partitions of 3, so $\lambda \in \{(3), (2,1), (1,1,1)\}$. We then have highest lexicographical weight in the symmetric irrep (3) to be the $\underline{n} = (3,0,0)$ weight. Applying the lowering operator corresponding to

$\mathcal{L}_{12}$ on the Figure 3.1, we see that the weight $\underline{n} = (3, 0, 0)$ is taken to the weight $\underline{n} = (2, 1, 0)$. Applying it again, we get the weight $\underline{n} = (1, 2, 0)$ and finally $\underline{n} = (0, 3, 0)$, which is the bottom left point in the triangle in Figure 3.1. The triangle represents the weights of the symmetric irrep. Similarly, we can apply the other lowering operator, $\mathcal{L}_{13}$ to get to the bottom right weight, $\underline{n} = (0, 0, 3)$. In the irrep $\lambda = (2, 1)$ (represented by hexagon in the Figure), the highest lexiographical weight is now $\underline{n} = (2, 1, 0)$. Finally, the irrep $\lambda = (1, 1, 1)$ has a single weight (represented by single dot in the Figure), which is $\underline{n} = (1, 1, 1)$. The diagrams help visualize the connection of different weights through the irreps (for example the absence of the $(3, 0, 0)$ weight in the $(2, 1)$ irrep), as well as the connection of the weights within the irrep using the lowering/raising operators. It also allows us to find the points of multiplicites, such as $\underline{n} = (1, 1, 1)$ in the $(2, 1)$ irrep.

The unitary group is fully reducible so we can express $\mathcal{U}$, using a basis change, as follows

$$\mathcal{U} \cong \bigoplus_{\lambda} \mathcal{U}^{\lambda} \otimes \mathbb{1}_{m_{\{\lambda\}}} \tag{3.15}$$

$$(\mathbb{C}^d)^{\otimes N} \cong \bigoplus_{\lambda} \mathbb{C}^{d_{\{\lambda\}}} \otimes \mathbb{C}^{m_{\{\lambda\}}} \tag{3.16}$$

where $m_{\{\lambda\}}$ is the multiplicity of the $\{\lambda\}$ irrep, the carrier space of the irreducible representations λ of U(d) is $\mathbb{C}^{d_{\{\lambda\}}}$, its dimension is $d_{\{\lambda\}}$, and we will often refer to $\mathcal{U}(U)$ as $U^{\otimes N}$, as well as shorten $\mathcal{U}^{\lambda}(U)$ to U^{λ} We can see that $\mathcal{U}^{(1)}(U) = U$. The next section we will specify what exactly these highest weights λ are.

Summary

To summarize, any representation of a finite or compact group can be written as a direct sum of irreps. The symmetric group can be split into irreps which can be labelled by Young diagrams, and the standard tableaux give a basis for the irrep. The unitary group can also be split into irreps which are labelled by highest weights. Next section shows how these are in $1 - 1$ correspondence. We have formulas that calculate the dimensions and multiplicities of these irreps for both the symmetric and unitary groups. Next we will determine the basis changes that implement these decompositions.

3.1.2 Schur-Weyl duality

In this section we put together all the previously mentioned definitions to produce a useful decomposition of a N-tensor space. It can be shown that the unitary and symmetric group commute with each other on $(\mathbb{C}^d)^{\otimes N}$. This allows for a nice relation between the irreducible representations of the two groups.

Theorem 3.1.3 (Schur-Weyl). *The Hilbert space of N-tensors can be decomposed as*

$$(\mathbb{C}^d)^{\otimes N} \cong \bigoplus_{\lambda \in \mathrm{Par}(N, d)} \mathbb{C}^{\{\lambda\}} \otimes \mathbb{C}^{(\lambda)}, \tag{3.17}$$

where λ is a partition of the integer N with d parts, $\mathbb{C}^{\{\lambda\}}$ carries irrep λ of the group of unitary transformations, $U(d)$, $\mathbb{C}^{(\lambda)}$ carries irrep λ of the group of permutations, S_N, and $\cong$ signifies that the left and right hand sides are related by a change of basis (a Schur-Weyl transform).

This implies the existence of a basis in which the symmetric and unitary actions are decomposed simultaneously. Therefore in the decompositions given in Eqs. (3.10) and (3.15), the multiplicity $m_{(\lambda)} = d_{\{\lambda\}}$ and similarly, $m_{\{\lambda\}} = d_{(\lambda)}$. For proofs of this theorem and the existence of the following choice of decomposition refer to standard representation theory literature [193, 195]. Notice that there is a dependence on the dimension of the original space, d. We have actually suppressed the dependence of the $\mathbb{C}^{\{\lambda\}}$ space on this dimension, which will become clear later with examples (this space is indexed by semi-standard tableaux, and if $d < N$, then not all the possible partitions of N can be included).

We see that not only the irreducible representations of S_N are labelled by Young diagrams, but so are those of $U(d)$. Examining again the Young symmetrizer, we can find that the normalized Young symmetrizer for a tableau of shape λ projects the N-tensor space onto an irrep with highest weight λ. Consequently, we will take the highest weight of the irrep marked with Young diagram λ to be $\underline{n} = \lambda$, and we will see in the following Section 3.1.4 why this makes sense. The theorem does not specify how to construct this basis, but this makes sense, as there is more than one "Schur-Weyl" basis. We will introduce a specific construction in Section 3.1.4.

We will now describe the following three equivalent notations for the Schur-Weyl basis state on the next pages. The first notation, which we will refer to as Young tableaux notation, will describe each of the basis states using a pair of Young tableaux. This notation will be useful for understanding how the Schur-Weyl basis is constructed. While it allows for easy understanding of symmetries and occupations of the states, it becomes cumbersome pretty quickly as the number of particles grows. Thus we introduce two other notations to handle general cases. The second notation, which we will refer to as $|\lambda, q, p\rangle$ notation, is a fairly standard type of notation for the Schur-Weyl basis, referring to the states using the irreps and the basis states within those irreps. The final notation we will use is a new type of notation which we will refer to as $|\lambda, p, \underline{n}, r\rangle$. It is similar to the $|\lambda, q, p\rangle$ notation however in the Schur-Weyl basis we construct, the occupation numbers also identify a basis state within an irrep up to inner multiplicities, and this notation takes advantage of that. This notation is very useful for linear optics, serving a purpose of an extension of the concept of Fock state for distinguishable bosons, allowing us to think about input and output occupation numbers and their scattering probabilities with ease in a complex setting of distinguishability. All of the notations have a shortened version that is taken advantage of to reduce clutter when it can be used unambiguously.

Young tableaux notation

We can appoint the standard Young tableaux of shape λ to index the λ irreps of S_N, and the semi-standard Young tableaux of shape λ to index the corresponding λ irreps of U_d. More specifically, each basis state can then be described by a pair of Young tableaux, $\left| \left(T_{\text{ind}}^{\lambda,p}, T_{\text{states}}^{\lambda,q} \right) \right\rangle$ where λ denotes the irrep (and thus defines the Young diagram), $T_{\text{ind}}^{\lambda,p}$ is the standard tableau labelling the state of the symmetric group irrep and $T_{\text{states}}^{\lambda,q}$ is the semi-standard tableau labelling the state of the unitary group irrep (both in some ordering we assign). In Section 3.1.3 we will discuss the Clebsch-Gordan coefficients and connection to Schur-Weyl in further detail. However, here we will assume some standard familiarity with it and proceed to give an example of the new notation being introduced in this case of two qubits, giving the triplet and the singlet states.

$$|(\boxed{1\,2}, \boxed{1\,1})\rangle = |J = 1, M = 1\rangle \quad = |\uparrow\uparrow\rangle, \tag{3.18}$$

$$|(\boxed{1\,2}, \boxed{1\,2})\rangle = |J = 1, M = 0\rangle \quad = 1/\sqrt{2}\,|\uparrow\downarrow\rangle + 1/\sqrt{2}\,|\downarrow\uparrow\rangle, \tag{3.19}$$

$$|(\boxed{1\,2}, \boxed{2\,2})\rangle = |J = 1, M = -1\rangle = |\downarrow\downarrow\rangle, \tag{3.20}$$

$$\left| \left(\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}, \boxed{\begin{smallmatrix}1\\2\end{smallmatrix}} \right) \right\rangle \quad = |J = 0, M = 0\rangle \quad = 1/\sqrt{2}\,|\uparrow\downarrow\rangle - 1/\sqrt{2}\,|\downarrow\uparrow\rangle, \tag{3.21}$$

with the new notation found on the left-hand side, and the spin-up and spin-down (computational) basis on the right-hand side.

This notation allows us to see how the permutations act on the indices (taking us between the different standard tableau of partition λ), and similarly how the unitaries act on the states (taking us between the different semi-standard tableaux). For example in the case of thre qutrits, from the notation $\left| \left(\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}, \boxed{\begin{smallmatrix}1&1\\2\end{smallmatrix}} \right) \right\rangle$, we can see that we are talking about a basis state in irrep $\lambda = (2, 1)$, and the tensor is built out of two basis states "1" and one basis state "2". Another reason why this notation is useful can be found in Section 3.1.4 when the basis is being constructed.

For some of the partitions λ and some weights $\underline{n}$, there will be more than one filling of a Young diagram. An example is $\lambda = (2, 1)$, $\underline{n} = (1, 1, 1)$, where both $\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}$ and $\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}$ are an acceptable filling. This would lead to both an "outer" multiplicity that we talked about earlier (more than one standard Young tableaux can label the particle indices, e.g. $\left| \left(\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}, T_{\text{states}}^{\lambda,q} \right) \right\rangle$ and $\left| \left(\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}, T_{\text{states}}^{\lambda,q} \right) \right\rangle$), as well as an "inner"' multiplicity, where there is more than one state of the irrep λ or U(d) with the given weight (e.g. $\left| \left(T_{\text{ind}}^{\lambda,p}, \boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}} \right) \right\rangle$ and $\left| \left(T_{\text{ind}}^{\lambda,p}, \boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}} \right) \right\rangle$)). Refer to Section 3.4.2 for examples of states of the 3-tensor space on $\mathbb{C}^3$.

$|\lambda, q, p\rangle$ notation

We will also make use of notation mentioned in [196]. Here the decomposition is denoted $|\lambda q p\rangle$ where λ labels the irrep of both the unitary and the symmetric groups simultaneously, $q = 1, 2, \ldots, d_{\{\lambda\}}$ indexes a basis of the unitary irrep, and $p = 1, 2, \ldots, d_{(\lambda)}$ indexes a basis of the symmetric irrep. There is an implied dependence of q on λ and d, and p on λ. This notation is similar to the previous one, with p being used instead of the standard Young

tableaux $\mathrm{T}_{\mathrm{ind}}^{\lambda,p}$ and q instead of the semi-standard Young tableaux $\mathrm{T}_{\mathrm{states}}^{\lambda,q}$. The numbering of p when identified to a previous state follows lexicographical ordering of the corresponding weights (up to a weight multiplicity). Referring back to the triplet and singlet example, the notation is as follows,

$$|(2),1,1\rangle \quad = |J=1,M=1\rangle \quad = |\uparrow\uparrow\rangle\,, \tag{3.22}$$

$$|(2),2,1\rangle \quad = |J=1,M=0\rangle \quad = 1/\sqrt{2}\,|\uparrow\downarrow\rangle + 1/\sqrt{2}\,|\downarrow\uparrow\rangle\,, \tag{3.23}$$

$$|(2),3,1\rangle \quad = |J=1,M=-1\rangle = |\downarrow\downarrow\rangle\,, \tag{3.24}$$

$$|(1,1),1,1\rangle = |J=0,M=0\rangle \quad = 1/\sqrt{2}\,|\uparrow\downarrow\rangle - 1/\sqrt{2}\,|\downarrow\uparrow\rangle\,, \tag{3.25}$$

with the new notation found on the left-hand side, first value denoting the irrep λ, second value is the state in the unitary irrep, third value is the state in the symmetric group irrep.

Refer to Section 3.4.2 for examples of states of the 3-tensor space on $\mathbb{C}^3$. When it is apparent from the setup (after we set our basis to be fixed in Section 3.1.4), we will also combine this and the Young tableau notation, shortening it to just the semi-standard tableaux with a subscript p denoting the basis of the symmetric irrep. For the triplet and singlet, notation would then just be simplified to

$$\left|\boxed{\begin{smallmatrix}1&1\end{smallmatrix}}_1\right\rangle = \left|\left(\boxed{\begin{smallmatrix}1&2\end{smallmatrix}},\boxed{\begin{smallmatrix}1&1\end{smallmatrix}}\right)\right\rangle \tag{3.26}$$

$$\left|\boxed{\begin{smallmatrix}1&2\end{smallmatrix}}_1\right\rangle = \left|\left(\boxed{\begin{smallmatrix}1&2\end{smallmatrix}},\boxed{\begin{smallmatrix}1&2\end{smallmatrix}}\right)\right\rangle \tag{3.27}$$

$$\left|\boxed{\begin{smallmatrix}2&2\end{smallmatrix}}_1\right\rangle = \left|\left(\boxed{\begin{smallmatrix}1&2\end{smallmatrix}},\boxed{\begin{smallmatrix}2&2\end{smallmatrix}}\right)\right\rangle \tag{3.28}$$

$$\left|\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}_1\right\rangle = \left|\left(\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}},\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right)\right\rangle. \tag{3.29}$$

In the case of three qutrits where we do have outer multiplicities, we will find states such as

$$\left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_1\right\rangle = \left|\left(\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}},\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}\right)\right\rangle,\ \left|\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_1\right\rangle = \left|\left(\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}},\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}\right)\right\rangle,\ \left|\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_2\right\rangle = \left|\left(\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}},\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}\right)\right\rangle, \tag{3.30}$$

where the left-hand side is the shortened combined notation, and the right hand side is the Young tableaux notation.

$|\lambda,p,\underline{n},r\rangle$ notation

In the Schur-Weyl basis we choose, weights (which we will identify with occupations of a state) play a major notational role. Therefore, we introduce a new notation, $|\lambda p \underline{n} r\rangle$. This notation is similar to that of $|\lambda q p\rangle$, except there is a slight reordering of arguments, and also instead of q which is in $\{1,2,\ldots,d_\lambda\}$, we map it on a tuple of occupation $\underline{n}$ and an 'inner' multiplicity $r \in \{1,2,\ldots,K_{\lambda,\underline{n}}\}$, which accounts for the fact that there can be more than one orthogonal state with the same weight in a unitary irrep λ) [*]. As we are focusing on the action of the unitary group, p will be referred to as an 'outer' multiplicity accounting for the fact that the same unitary irrep λ can occur more than once. The irrep

---

[*]Moreover, we can actually assign a more substantial role to r, which is for it to specify the subgroup chain depending on the construction of our basis, but we will not do that as we do not need that information.

dependence of p, $\underline{n}$ and r has again been suppressed to prevent clutter (as well as the dependence of $\underline{n}$ and r on d, and dependence of r on $\underline{n}$). Finally, the triplet and singlet example is as follows in this notation,

$$|(2), 1, (2,0), 1\rangle \quad = |J = 1, M = 1\rangle \quad = |\uparrow\uparrow\rangle \,, \tag{3.31}$$

$$|(2), 1, (1,1), 1\rangle \quad = |J = 1, M = 0\rangle \quad = 1/\sqrt{2}\,|\uparrow\downarrow\rangle + 1/\sqrt{2}\,|\downarrow\uparrow\rangle \,, \tag{3.32}$$

$$|(2), 1, (2,0), 1\rangle \quad = |J = 1, M = -1\rangle = |\downarrow\downarrow\rangle \,, \tag{3.33}$$

$$|(1,1), 1, (1,1), 1\rangle = |J = 0, M = 0\rangle \quad = 1/\sqrt{2}\,|\uparrow\downarrow\rangle - 1/\sqrt{2}\,|\downarrow\uparrow\rangle \,, \tag{3.34}$$

with the new notation found on the left-hand side, first value denoting the irrep λ, second value is the state in the symmetric group irrep, third value is the occupation numbers of the state, and the fourth value is the inner multiplicity as defined.

We will often shorten the notation such that $|\lambda p \underline{n}\rangle := |\lambda, p, \underline{n}, r = 1\rangle$, $|\lambda \underline{n} r\rangle :=$ $|\lambda, p = 1, \underline{n}, r\rangle$, $|\lambda \underline{n}\rangle := |\lambda, p = 1, \underline{n}, r = 1\rangle$, reducing clutter when the multiplicity is trivial; since λ and $\underline{n}$ are vectors while p and r are scalars there should be no ambiguity. Finally, the triplet and singlet example look like

$$|(2), (2,0)\rangle = |(2), 1, (2,0), 1\rangle \tag{3.35}$$

$$|(2), (1,1)\rangle = |(2), 1, (1,1), 1\rangle \tag{3.36}$$

$$|(2), (0,2)\rangle = |(2), 1, (0,2), 1\rangle \tag{3.37}$$

$$|(1,1), (1,1)\rangle = |(2), 1, (1,1), 1\rangle \,. \tag{3.38}$$

in this notation. We will also sometimes write $\underline{n}$ and r in a grouped manner, such as $(\underline{n}, r)$ or just replace it with the equivalent q. As before, refer to Section 3.4.2 for examples of states of the 3-tensor space on $\mathbb{C}^3$.

Schur-Weyl transform

The basis change needed to take the N-tensor basis state to the $|\lambda p \underline{n} r\rangle$ state will be referred to as U_{Sch}, so that

$$|\lambda, p, \underline{n}, r\rangle = \sum_{j_1, j_2, \ldots, j_N} [U_{\mathrm{Sch}}]^{\lambda, p, \underline{n}, r}_{j_1, j_2, \ldots, j_N} |j_1, j_2, \ldots, j_N\rangle \tag{3.39}$$

where $|j_1, j_2, \ldots, j_N\rangle$ has the weight $\underline{n}$. For U(2) this is the familiar Clebsch-Gordan transformation of angular momentum theory. We will usually order the basis based on the irrep λ and the outer multiplicity p. This will then give us a block diagonal form for the action of any $U \in \mathrm{U}(d)$, that is

$$U_{\mathrm{Sch}} U^{\otimes N} U^{\dagger}_{\mathrm{Sch}} = \oplus_\lambda U^\lambda \otimes \mathbb{1}^\lambda . \tag{3.40}$$

where $\operatorname{Tr} \mathbb{1}^\lambda = d_{(\lambda)}$. We can also see from this form that the representation of U in the Schur-Weyl basis is block diagonal,

$$\langle \lambda p \underline{n} r \mid U_{\text{Sch}} U^{\otimes N} U_{\text{Sch}}^\dagger \mid \lambda' p' \underline{n}' r' \rangle = \delta_{\lambda\lambda'} \langle \lambda p \underline{n} r \mid U^\lambda \otimes \mathbb{1}^\lambda \mid \lambda' p' \underline{n}' r' \rangle \qquad (3.41)$$

$$= \delta_{\lambda\lambda'} \delta_{pp'} \langle \lambda p \underline{n} r \mid U^\lambda \mid \lambda' p' \underline{n}' r' \rangle . \qquad (3.42)$$

But also the block diagonals that correspond to the same irrep λ can be made identical as U^λ only depends on the choice of λ so,

$$\langle \lambda p \underline{n} r \mid U_{\text{Sch}} U^{\otimes N} U_{\text{Sch}}^\dagger \mid \lambda p \underline{n}' r' \rangle = \langle \lambda p \underline{n} r \mid U^\lambda \mid \lambda p \underline{n}' r' \rangle \qquad (3.43)$$

$$= \langle \lambda p' \underline{n} r \mid U^\lambda \mid \lambda p' \underline{n}' r' \rangle , \qquad (3.44)$$

that is the two corresponding outer multiplicity matrix elements are equivalent.

The notation also makes clear that unitary actions cannot take a state from irrep λ or outer multiplicity p into a different irrep or outer multiplicity (see Equation 3.142 for example). That is the action of the unitary and symmetric group on this new basis state can be described as

$$U : |\lambda p \underline{n} r\rangle \rightarrow \sum_{\underline{n}',r'} U^\lambda_{(\underline{n}',r'),(\underline{n},r)} |\lambda p \underline{n}' r'\rangle \qquad (3.45)$$

$$\sigma : |\lambda p \underline{n} r\rangle \rightarrow \sum_{p'} S^\lambda_{p'p} |\lambda p' \underline{n} r\rangle \qquad (3.46)$$

where $U^\lambda_{(\underline{n}',r'),(\underline{n},r)}$ $(S^\lambda_{p'p})$ are the suitable matrix elements of the matrix U^λ (S^λ) for the corresponding $\mathrm{U}(d)$ (S_N) irrep. This tells us that a state of certain permutation symmetry cannot change its symmetry by using unitary actions, for example we cannot make a symmetric state be anti-symmetric by applying a unitary.

3.1.3 Schur-Weyl duality and first quantisation

Most of the discussion of the Schur-Weyl duality so far was in abstract mathematical terms, but now we will want to think about how it can be applied to particles and their behaviour.

Assume we have a particle that lives in a Hilbert space of dimension d, $\mathcal{H}_d$. Next assume we have N such particles and denote $\mathcal{H}_d^N = \mathcal{H}_d^{\otimes N}$. From the Schur-Weyl duality, we now know there is a basis change U_{Sch}, that takes the computational (tensor product) basis $|j_1 j_2 \dots j_N\rangle$ to the Schur-Weyl basis $|\lambda q p\rangle$.

Schur-Weyl transformation in quantum physics

For $d = 2$ dimensional systems, commonly in physics we use the $|jm\rangle$ basis where j is the angular momentum quantum number and m denotes the projection on the z-axis, that is j labels the representation and m denotes the eigenvalue of $\hat{J}_z$, the angular momentum operator in the z-direction. The Clebsch-Gordan transform describes how two states of "good" angular momentum, $|j_1 m_1\rangle$ and $|j_2 m_2\rangle$, are coupled together to make a combined

state of "good" angular momentum, that is it gives the coefficients $\langle j_1 m_1 j_2 m_2 \mid JM \rangle$. For example, if coupling two electrons together, the result will be the familiar triplet basis for $J = 1$ (see Equations (3.47), (3.48), and (3.49)), and the singlet basis for $J = 0$ (see Equation (3.50)). This is just a special case of the Schur-Weyl duality, where $j = (\lambda_1 - \lambda_2)/2$ (we can see that this is unique for any set N, as $\lambda_1 + \lambda_2 = N$, and since $d = 2$, there will only ever be two rows allowed in a Young diagram). There are only two partitions in case of $N = 2$, $\boxed{}$ so $\lambda_1 = 2$ and $\lambda_2 = 0$, and $\boxed{}$ so $\lambda_1 = 1$ and $\lambda_2 = 1$. Translated in one of the notations defined previously, we see that

$$|(\boxed{1\,2}, \boxed{1\,1})\rangle = |J = 1, M = 1\rangle \;\; = |\uparrow\uparrow\rangle, \tag{3.47}$$

$$|(\boxed{1\,2}, \boxed{1\,2})\rangle = |J = 1, M = 0\rangle \;\; = 1/\sqrt{2}\,|\uparrow\downarrow\rangle + 1/\sqrt{2}\,|\downarrow\uparrow\rangle, \tag{3.48}$$

$$|(\boxed{1\,2}, \boxed{2\,2})\rangle = |J = 1, M = -1\rangle = |\downarrow\downarrow\rangle, \tag{3.49}$$

$$\left|\left(\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}, \boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right)\right\rangle \;\; = |J = 0, M = 0\rangle \;\; = 1/\sqrt{2}\,|\uparrow\downarrow\rangle - 1/\sqrt{2}\,|\downarrow\uparrow\rangle. \tag{3.50}$$

From the above example we see that Schur-Weyl transform has already had a use in physics for tensor products of $SU(2)$. However, in the more recent decades, it has found its use in quantum information as well. Specifically the symmetric subspace and its properties has found many uses in the situations where we have many copies of the same state [197, 198]. For example, it has allowed insight in state estimation of a mixed state ρ when N copies of it are given [199], or producing a maximally entangled state through entanglement concentration of many copies of unknown pure states [200]. Efficient quantum circuits for the transform have also been explored, by applying the Clebsch-Gordan transform recursively [196]. Here we will use the Schur-Weyl transform and representation theory in general to approach questions of boson sampling and boson distinguishability [81, 186].

Schur-Weyl basis and bosons

In this thesis we will be interested in the behaviour of bosons, which in multi-particle states, always behave in a symmetric manner under the exchange of particle labellings. Thinking about it in Young diagrams and partitions, the Young diagram is always a single row of N boxes where N is the number of particles and the partition is just (N) (e.g. for two particles $\boxed{}$). If we were interested in fermions, there is a very similar argument for most things presented, except the exchange of particles is always antisymmetric, so the corresponding Young diagram is a single column with one box per row and the partition is $(1, 1, \ldots, 1)$ (e.g. for two particles $\boxed{}$).

We used the notion of weights to label states forming the basis of the carrier spaces of the representations of the unitary group. These weights can be associated with occupation notation for particles. Given a state in computational basis $|i_1 i_2 \ldots i_N\rangle$ we can express the occupation of that state as $\underline{n} = (n_1, n_2, \ldots, n_d)$ where $n_k = \sum_{j=1}^{N} \delta_{i_j k}$. For example, considering two modes, if we have two qubits in state "1" and three qubits in state "2", this is can be given as weight $(2, 3)$. Five qubits in state "1" would then give occupation

$(5, 0)$. That is we have a mapping

$$|11111\rangle \rightarrow (5, 0), \tag{3.51}$$

$$|11222\rangle \rightarrow (2, 3). \tag{3.52}$$

Notice however, that this is not a one-to-one mapping, as both $|11222\rangle$ and $|12122\rangle$ have the same occupation of $(2, 3)$. For that, we would have to introduce another argument, just like in the case of weights, where we have the "inner multiplicity", r.

The coincident state of particles will be denoted with occupation number $\underline{1} = (1, 1, \ldots, 1)$, with exactly one particle in each mode (corresponding in first quantisation to one qudit in each basis state). Since the weight of a weight vector is equivalent to occupation of a state, and the weight splits the carrier space into disjoint weight-spaces, we can see that the occupations will also split the space into disjoint subspaces. While weights (referred to as "types" as well, see [196]) are well studied in representation theory, and types have been used to understand certain properties in QI theory [201], the link between weight and occupation numbers, especially in regards to bosons, has not previously been utilized.

Let d be the number of modes. Let N be the number of bosons. Then we can model bosons in the first quantized picture (particle picture, see Section 1.2) as N qudits. As these particles are bosons, their states only have support in the symmetric subspace, $\mathrm{Sym}\left((\mathbb{C}^d)^{\otimes N}\right)$. If we were instead talking about fermions, the setup would be similar, except their support would be only on $\mathrm{Alt}\left((\mathbb{C}^d)^{\otimes N}\right)$.

We will use weights and occupations interchangeably. Moreover, the raising and lowering operators for these weights are then related to bosonic annihilation and creation operators. That is

$$\mathcal{L}_{jk} = \hat{a}_k^\dagger \hat{a}_j \tag{3.53}$$

$$\mathcal{R}_{kj} = \hat{a}_j^\dagger \hat{a}_k \tag{3.54}$$

where $j < k$. As the symmetric subspace has multiplicity of one, the weight (that is occupation) uniquely identifies the state. The occupation notation is familiar to those that work in quantum optics as this is exactly what Fock states are (see Section 1.3.1). The interferometer which acts on the bosonic modes is then a unitary from the $\mathrm{U}(d)$ group acting as defined in Eq. (3.12).

As mentioned, fermions and bosons have simple antisymmetric and symmetric exchange symmetry (respectively). We can also use the above formalism for particles with other exchange symmetries, such as fermions (antisymmetric), but also purely theoretical particles "immanons" (mixed symmetry) introduced recently by Tichy and Mølmer [144].We can see these mixed exchange symmetries more generally even for particles whose total exchange symmetry is normally symmetric or antisymmetric (in the same way that fermionic behaviour can be reproduced by certain states of bosons). Most of the work following this section will actually be spent in defining what distinguishable bosons are, their properties, and how they are related to this mixed exchange symmetry which occurs naturally after

introducing a second degree of freedom to bosons.

3.1.4 Constructing the Schur-Weyl basis

We now describe one possible way to construct the representation basis for N qudits . From Schur-Weyl duality we know that the space decomposes as $(\mathbb{C}^d)^{\otimes N} \cong \bigoplus_\lambda \mathbb{C}^{\{\lambda\}} \otimes \mathbb{C}^{(\lambda)}$. The first step is to construct all of the relevant partitions of the number N into at most d parts (if $N = d$ then all possible partitions are present). We will then find the highest weight for each of the corresponding unitary group irreps. On the highest weight states we can use the necessary permutation action on particle labels to find the other highest weights for that irrep, and the lowering operators to find the other basis states in the irrep. Finally, we make sure that the states form a basis by orthonormalizing them using Gram-Schmidt procedure. As there is more than one Schur-Weyl basis, we have some freedom of choice on how to construct it, what the highest weight state is, and how we orthonormalize.

Highest weight

We start by constructing our chosen highest weight state. As seen previously, every irrep $\{\lambda\}$ of U(d) can be assigned a highest weight state, which is annihilated by an appropriate set of raising operators that are realised in terms of the bosonic creators and annihilators. For N particles the ordering of the weight can be understood by identifying weights with a number in base $N + 1$. For example, given two qutrits, we would have the weights ordering $200 > 110 > 101 > 020 > 011 > 002$ (as would be expected for numbers in base 3). Given four qubits, the weights are ordered as $40 > 31 > 22 > 13 > 04$ (as would be expected for numbers in base 5). This is the lexicographical ordering we have mentioned previously in Section 3.1.1. For an irrep λ we set the highest state of that irrep to be $\mathrm{ST}(\lambda)$ which we recall has weight $\underline{n}$ and is unique (see Section 3.1.1).

The other freedom we have when choosing the highest weight state is connected to the state of particle indices. We take this to be a standard Young tableau which we are familiar with, $\mathrm{T}(\lambda)$. Here are some examples of highest weight states, $|(\mathrm{T}(\lambda), \mathrm{ST}(\lambda))\rangle$:

$$N = 3,\ d = 2,\ \lambda = (2,1):\qquad \left| \left(\young(13,2)\ ,\ \young(11,2) \right) \right\rangle$$

$$N = 5,\ d = 2,\ \lambda = (3,2):\qquad \left| \left(\young(135,24)\ ,\ \young(111,22) \right) \right\rangle$$

$$N = 8,\ d = 4,\ \lambda = (3,2,2,1):\qquad \left| \left(\young(158,26,37,4)\ ,\ \young(111,22,33,4) \right) \right\rangle .$$

The highest weight state can then be expressed in terms of single particle (qudit) states using products of Slater determinants. Let ψ_i describe the wavefunction of the j-th particle.

Then the *Slater determinant* is defined as

$$
\mathfrak{D}^{12\ldots N}_{j_1 j_2 \ldots j_N} =
\begin{vmatrix}
\psi_1(j_1) & \psi_1(j_2) & \ldots & \psi_1(j_N) \\
\psi_2(j_1) & \psi_2(j_2) & \ldots & \psi_2(j_N) \\
\ldots & \ldots & \ldots & \ldots \\
\psi_N(j_1) & \psi_N(j_2) & \ldots & \psi_N(j_N)
\end{vmatrix}
\tag{3.55}
$$

To get the state we find the product of the Slater determinants of the columns of the tableaux. An example for a mix of labellings and states of particles,

The set of highest weight states

The highest weight state was set by combining two specifically chosen Young tableaux together. However, we could have chosen a different standard Young tableaux for the labelling of the particle indices, which give us an equally valid highest state of weight λ (we do have to be careful about the choice of Young tableau we use to construct the states so we do not end up with an overcomplete basis; this is why we are limited by standard Young tableaux, which we know form a basis of the symmetric group irrep). This set of outer multiplicities is handled by utilising the dual S_N action to permute a highest weight state in order to find corresponding highest weights for the multiple copies of irrep $\{\lambda\}$. Again, the number of such linearly independent highest weight states is known, namely $d_{(\lambda)}$. Therefore, we use permutations on the particle labels on the state $|(\mathrm{T}(\lambda), \mathrm{ST}(\lambda))\rangle$, that is

$$
\sigma \, |(\mathrm{T}(\lambda), \mathrm{ST}(\lambda))\rangle = |(\sigma \mathrm{T}(\lambda), \mathrm{ST}(\lambda))\rangle = \left| \left(\mathrm{T}^{\lambda,p}_{\mathrm{ind}}, \mathrm{ST}(\lambda) \right) \right\rangle
\tag{3.56}
$$

where $\sigma \in S_N$ is a permutation. This will give us the highest weight states set,

$$
\left\{ \left| \left(T^{\lambda,p}_{\mathrm{ind}}, \mathrm{ST}(\lambda) \right) \right\rangle \mid p \in \{1, \ldots, d_{(\lambda)}\} \right\}.
\tag{3.57}
$$

This set of states might not be orthonormal, so the Gram-Schmidt procedure is then used to orthonormalize the set.

Set of states with lower weights

In much the same way as is done for U(2) in angular momentum theory, we then apply corresponding lowering operators to find a set of states that span the unitary group irrep λ (size of which is known, namely $d_{\{\lambda\}}$). In the single particle basis, these operators act by taking a particle in state $|j\rangle$ to state $|k\rangle$, where k is a "lower" state based on the

lexicographical weight ordering. That is these operators act as follows in the single particle basis

$$\mathcal{L}_{kj} = \sum_{k=l}^{N} \left(I^{\otimes l-1} \otimes |k\rangle \langle j| \otimes I^{\otimes N-l} \right). \tag{3.58}$$

These operators are then applied to the highest weight states,

$$\mathcal{L}_{kj} \left| \left(T_{\text{ind}}^{\lambda,s}, SA_{\text{states}}(\lambda) \right) \right\rangle = \left| \left(T_{\text{ind}}^{\lambda,s}, T_{\text{states}}^{\lambda,t} \right) \right\rangle \tag{3.59}$$

where $SA_{\text{states}}(\lambda)$ was of weight λ, and $T_{\text{states}}^{\lambda',t}$ has weight $\lambda' = (\lambda_1 \ldots \lambda_{j-1} \ldots \lambda_{k+1} \ldots \lambda_N)$. We now keep applying these operators until we find the right number of orthogonal states (this might require orthnormalizing the states again). Notice that it will not necessarily be enough to apply the operator just on the highest weight, but recursively on the new states we get until we have a complete basis. For a given irrep λ multiplicity p this will create the unitary irrep set of states $\left\{ \left| \left(T_{\text{ind}}^{\lambda,p}, T_{\text{states}}^{\lambda,q} \right) \right\rangle \mid q \in \{1, \ldots, d_{\{\lambda\}}\} \right\}$. We summarize the above in a pseudoalgorithm.

3.2 Representation theoretic approach to distinguishability

In this section we take the previously introduced Schur-Weyl basis, add a second degree of freedom to each particle to represent distinguishing information, and connect the result to what is known as unitary-unitary duality. These components together combine to give a new framework for modelling distinguishability of particles. The other novel concept in this Section is the measurement model based on occupation numbers of all states (including mixed symmetry states).

3.2.1 Composite bosons

In Section 1.3.1 we discuss that bosons have symmetric exchange interaction and this is reflected in Fock States. In Section 3.1.3 we then talk about how we can reconcile the standard linear optics notation of Section 1.3 with the representation theoretic approach we focus on in this thesis. We also mention that exchange interaction of fermions can incite a sign change. Bar that difference, most of the theory and calculations we do here can also be carried out in a similar fashion for fermions, but our interest is in bosons, so we will maintain that focus. Before we delve deeper into distinguishability, we will introduce another representation theory tool that we will need frequently. We will look at some specific examples with no notion of distinguishability for now, but just focusing on the mathematical implications.

Two photons with two degrees of freedom

Assume we have two photons ($N = 2$) in two degrees of freedom, denoting the first degree of freedom as "S(ystem)" and second as "L(abel)" (the notation will make sense later, but no further meaning is attributed to it here). For example, we can manipulate the photon's

---

Algorithm 1 Constructing Schur-Weyl basis

1: **function** SCHURWEYL(d, N)
2:     partitions ← Partition(N, d)   ▷ Partitions of N up to length d
3:     loweringOperators ← LoweringOperators(N,d)
4:     permOperators ← PermutationOperators(N,d)
5:     **for** λ in partitions **do**
6:         $d_{(\lambda)}$ ← HookFormula(λ) (see Eq. (3.1))
7:         $d_{\{\lambda\}}$ ← HookContentFormula(λ) (see Eq. (3.2))
8:         columns ← ST(λ)   ▷ Construct highest weight semi-standard tableaux
9:         highestState ← identity
10:         **for** column in columns **do**   ▷ Construct highest state for irrep λ
11:             columnState ← SlaterDeterminant(column)
12:             highestState ← highestState ⊗ columnState
13:         **end for**
14:         highestStates ← GenerateStates($d_{(\lambda)}$, permOperators, highestState)   ▷ Find highest state basis for $\mathbb{C}^{d_{\{\lambda\}}}$
15:         **for** state in highestStates **do**
16:             state ← GenerateStates($d_{\{\lambda\}}$, loweringOperators, state)   ▷ Find basis for $\mathbb{C}^{d_{(\lambda)}}$
17:             add states to allStates
18:         **end for**
19:     **end for**
20: **end function**
21: **function** GENERATESTATES(d, operators, startState)
22:     empty stateQueue
23:     queue startState to stateQueue
24:     add startState to genStates
25:     **while** number of states in genStates $< d$ **do**
26:         currentState ← Dequeue(stateQueue)
27:         **for** operator in operators **do**
28:             opState ← operator acting on currentState
29:             **if** opState not in genStates **then**
30:                 add opState to genStates
31:                 queue opState to stateQueue
32:             **end if**
33:         **end for**
34:     **end while**
35:     genStates ← GramSchmidt(genStates)
36:     **return** genStates
37: **end function**

---

path modes and let there be two of them (this is S, and the dimension is $d_S = 2$); we can also manipulate the photon polarization (this is L, and the dimension is $d_L = 2$). The two photons can then be in the path mode "1" or path mode "2", and also can be vertically ("V") or horizontally polarized ("H"). Therefore each photon has two degrees of freedom. The particles are ququarts, and there are four modes that a photon can be found in, 1H, 1V, 2H, and 2V. We could denote the set of creation operators acting on these states by the four modes

$$\left\{ \hat{a}_1^\dagger, \hat{a}_2^\dagger, \hat{a}_3^\dagger, \hat{a}_4^\dagger \right\}, \tag{3.60}$$

but we could also denote them by the action on the two degrees of freedom, that is

$$\left\{ \hat{a}_{1H}^\dagger, \hat{a}_{1V}^\dagger, \hat{a}_{2H}^\dagger, \hat{a}_{2V}^\dagger \right\}, \tag{3.61}$$

respectively. The Fock basis is the usual two photons in four modes basis, and the single particle basis is

$$\left\{ |1_H\rangle, |1_V\rangle, |2_H\rangle, |2_V\rangle \right\}, \tag{3.62}$$

or equally when we rearrange the tensor product structure of the two degrees of freedom,

$$\left\{ |1\rangle |H\rangle, |1\rangle |V\rangle, |2\rangle |H\rangle, |2\rangle |V\rangle \right\}, \tag{3.63}$$

arriving at a state in the S(ystem) $\otimes$ L(abel) basis. As we mentioned in Section 3.1.2, writing states in terms of Young tableaux can be more compact for small N and d, more specifically dropping the Young tableau labelling the state of indices, but instead just writing the U(d) basis state with the subscript p denoting the outer multiplicity (see Eq. 3.30). Moreover, as there are two degrees of freedom here, we will actually have both of the degrees of freedom described in this notation.

Recall that in the case of two qubits, the unitary irrep basis is the standard triplet singlet basis, that is

$$\left\| \boxed{1 \, 1} \right\rangle = |11\rangle \tag{3.64}$$

$$\left\| \boxed{1 \, 2} \right\rangle = \frac{1}{\sqrt{2}} (|12\rangle + |21\rangle) \tag{3.65}$$

$$\left\| \boxed{2 \, 2} \right\rangle = |11\rangle \tag{3.66}$$

$$\left\| \boxed{\begin{smallmatrix}1\\2\end{smallmatrix}} \right\rangle = \frac{1}{\sqrt{2}} (|12\rangle - |21\rangle) \tag{3.67}$$

Example 3.2.1. Start with a state $\hat{a}_{1H}^\dagger \hat{a}_{1H}^\dagger |\text{vac}\rangle$, that is two photons both in first path mode and horizontally polarized. In first quantization we can write it as

$$\hat{a}_{1H}^\dagger \hat{a}_{1H}^\dagger |\text{vac}\rangle = |1_H\rangle_1 |1_H\rangle_2, \tag{3.68}$$

where the subscripts 1 and 2 are being used as (fictitious) particle labels that get permuted.

We can now rearrange to a tensor of the two degrees of freedom,

$$\hat{a}_{1H}^\dagger \hat{a}_{1H}^\dagger \left|\text{vac}\right\rangle = \left|1_H\right\rangle_1 \left|1_H\right\rangle_2 = \left|11\right\rangle_\text{S} \left|HH\right\rangle_\text{L}. \tag{3.69}$$

This makes it clear how the state is written in the Schur-Weyl basis:

$$\hat{a}_{1H}^\dagger \hat{a}_{1H}^\dagger \left|\text{vac}\right\rangle = \left|11\right\rangle \left|HH\right\rangle = \left|\boxed{1\,1}\right\rangle \left|\boxed{H\,H}\right\rangle. \tag{3.70}$$

We can do this for any state in a similar fashion, rearranging it into two degrees of freedom followed by a transformation to the Schur-Weyl basis.

Example 3.2.2. Another state to take a look at could be $\hat{a}_{1H}^\dagger \hat{a}_{2H}^\dagger \left|\text{vac}\right\rangle$, with photons in different path modes.

$$\hat{a}_{1H}^\dagger \hat{a}_{2H}^\dagger \left|\text{vac}\right\rangle = \frac{1}{\sqrt{2}} \left(\left|1_H 2_H\right\rangle + \left|2_H 1_H\right\rangle\right) \tag{3.71}$$

$$= \frac{1}{\sqrt{2}} \left(\left|12\right\rangle \left|HH\right\rangle + \left|21\right\rangle \left|HH\right\rangle\right) \tag{3.72}$$

$$= \frac{1}{\sqrt{2}} \left(\left|12\right\rangle + \left|21\right\rangle\right) \left|HH\right\rangle \tag{3.73}$$

$$= \left|\boxed{1\,2}\right\rangle \left|\boxed{H\,H}\right\rangle. \tag{3.74}$$

Example 3.2.3. We could also have photons in different path and polarization modes, $\hat{a}_{1H}^\dagger \hat{a}_{2V}^\dagger \left|\text{vac}\right\rangle$.

$$\hat{a}_{1H}^\dagger \hat{a}_{2V}^\dagger \left|\text{vac}\right\rangle = \frac{1}{\sqrt{2}} \left(\left|1_H 2_V\right\rangle + \left|2_V 1_H\right\rangle\right) \tag{3.75}$$

$$= \frac{1}{\sqrt{2}} \left(\left|12\right\rangle \left|HV\right\rangle + \left|21\right\rangle \left|VH\right\rangle\right) \tag{3.76}$$

$$= \frac{1}{2\sqrt{2}} \left(\left|12\right\rangle + \left|21\right\rangle\right) \left(\left|HV\right\rangle + \left|VH\right\rangle\right) + \frac{1}{2\sqrt{2}} \left(\left|12\right\rangle - \left|21\right\rangle\right) \left(\left|HV\right\rangle - \left|VH\right\rangle\right) \tag{3.77}$$

$$= \frac{1}{\sqrt{2}} \left(\left|\boxed{1\,2}\right\rangle \left|\boxed{H\,V}\right\rangle + \left|\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right\rangle \left|\boxed{\begin{smallmatrix}H\\V\end{smallmatrix}}\right\rangle\right). \tag{3.78}$$

Possibly the most interesting example is the last one. We can see that the states in the two degrees of freedom are entangled. If a particle is in the path mode "1", then the polarization of that particle is "H", and if a particle is in the path mode "2", then the polarization of that particle will be "V".

Another curious thing about this last example is how the two degrees of freedom couple across the irreps. When the first degree of freedom is in irrep $\lambda = (2)$, so is the second ($\left|\boxed{1\,2}\right\rangle \left|\boxed{H\,V}\right\rangle$). When the first degree of freedom is in irrep $\lambda = (1^2)$, so is the second ($\left|\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right\rangle \left|\boxed{\begin{smallmatrix}H\\V\end{smallmatrix}}\right\rangle$). This is also true in the first two examples, but since there is no anti-symmetric component, it is not as clear that this is something that might be a general rule. The last

example leads us to wonder if for the full multiboson wavefunction to be totally symmetric, the two degrees of freedom have to be coupled in this very specific way, with both degrees of freedom found in the same irrep.

Indeed, this is another result that comes out of representation theory. There exists a generalization of the Schur-Weyl duality to groups called dual reductive pairs. The one we will be interested in is the so called unitary-unitary duality, which is between two unitary groups.

Unitary-unitary duality

Assume we have a particle with a tensor product $\mathcal{H}_{SL} = \mathcal{H}_S \otimes \mathcal{H}_L$ over two degrees of freedom of dimension d_S and d_L, respectively. Further, assume this particle is a boson, so the possible states are only the ones from the fully symmetric subspace of the total space $\mathcal{H}_{SL}$, $\mathrm{Sym}(\mathcal{H}_{SL})$. Then the following theorem considers the behaviour of multipartite Hilbert space $\mathrm{Sym}(\mathcal{H}_{SL})$.

Theorem 3.2.4 (Unitary-unitary duality, adapted from Rowe et al. [187])**.** *The groups* $\mathrm{U}(d_S)$ *and* $\mathrm{U}(d_L)$ *have dual representations on the fully symmetric subspace of* $\mathrm{U}(d_S d_L)$*, decomposing as*

$$\mathrm{Sym}\left((\mathbb{C}^{d_S} \otimes \mathbb{C}^{d_L})^{\otimes N}\right) \cong \bigoplus_{\lambda} \mathbb{C}^{\{\lambda\}_S} \otimes \mathbb{C}^{\{\lambda\}_L}, \tag{3.79}$$

where the sum extends over all λ *from* $\mathrm{Par}(N, \min(d_S, d_L))$*,* $\mathrm{Sym}\left((\mathbb{C}^{d_S} \otimes \mathbb{C}^{d_L})^{\otimes N}\right)$ *is the carrier space of the symmetric irrep of* $\mathrm{U}(d_S d_L)$*,* $\mathbb{C}^{\{\lambda\}_S}$ *is the carrier space of the* λ *irrep of* $\mathrm{U}(d_S)$*,* $\mathbb{C}^{\{\lambda\}_L}$ *is the carrier space of the* λ *irrep of* $\mathrm{U}(d_L)$*.*

An analogous result holds for fermions and the antisymmetric subspace, although here one couples irreps with their conjugate Young diagrams.

It follows that given a fully symmetric state of N particles with two degrees of freedom, we can always describe this multiparticle state using coupled mixed symmetry states on the separate degrees of freedom. The totally symmetric pure System-Label states are of the form

$$\sum_{\lambda q q'} \psi_{\lambda q q'} |\lambda q q'\rangle_{SL}, \tag{3.80}$$

where we define

$$|\lambda q q'\rangle_{SL} := \frac{1}{\sqrt{d_{(\lambda)}}} \sum_{p=1}^{d_{(\lambda)}} |\lambda q p\rangle_S |\lambda q' p\rangle_L \tag{3.81}$$

with equal coupling coefficients independent of p [185, 195], and p, λ label the irreps and q, q' basis states of the two unitary groups $(\mathrm{U}(d_S), \mathrm{U}(d_L))$ and the symmetric group (S_N). As discussed in Section 3.1.2, we can replace q with pairs $(\underline{n}, r)$ in all of these expressions.

Photons with two degrees of freedom notation

The discussion of the two photon examples shows us some new notation needed to simplify talking about states of photons with two degrees of freedom. Photon creators will be written as

$$\hat{a}_{sl}^{\dagger}, \quad s \in \{1, 2, \ldots, d_{\mathrm{S}}\}, l \in \{1, 2, \ldots, d_{\mathrm{L}}\} \tag{3.82}$$

where the index s refers to what "System" mode is being raised, and index l refers to the "Label" mode being raised. For example, $\hat{a}_{sl}^{\dagger} = \int \mathrm{d}\omega\, f_l(\omega) \hat{a}_s^{\dagger}(\omega)$ where f_l is a spectral envelope function indexed by l for a photon created in System mode s.

Another notation we introduce is that of Fock arrays. From the doubly indexed photon annihilation and creation operators, we can define two dimensional arrays that correspond to Fock states. The rows correspond to the System modes and the columns correspond to Label modes. In the case of the usual Fock states, we have the following action

$$\hat{a}_j^{\dagger} |n_1, n_2, \ldots, n_j, \ldots, n_d\rangle = \sqrt{n_j + 1} |n_1, n_2, \ldots, n_j + 1, \ldots, n_d\rangle, \tag{3.83}$$

where d is the number of modes (see Section 1.3). Then similarly,

$$\hat{a}_{sl}^{\dagger} \left| \begin{matrix} n_{11} & \cdots & n_{1l} & \cdots & n_{1d_{\mathrm{L}}} \\ n_{21} & \cdots & n_{2l} & \cdots & n_{2d_{\mathrm{L}}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ n_{s1} & \cdots & n_{sl} & \cdots & n_{sd_{\mathrm{L}}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ n_{d_{\mathrm{S}}1} & \cdots & n_{d_{\mathrm{S}}l} & \cdots & n_{d_{\mathrm{S}}d_{\mathrm{L}}} \end{matrix} \right\rangle = \sqrt{n_{sl} + 1} \left| \begin{matrix} n_{11} & \cdots & n_{1l} & \cdots & n_{1d_{\mathrm{L}}} \\ n_{21} & \cdots & n_{2l} & \cdots & n_{2d_{\mathrm{L}}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ n_{s1} & \cdots & n_{sl}+1 & \cdots & n_{sd_{\mathrm{L}}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ n_{d_{\mathrm{S}}1} & \cdots & n_{d_{\mathrm{S}}l} & \cdots & n_{d_{\mathrm{S}}d_{\mathrm{L}}} \end{matrix} \right\rangle. \tag{3.84}$$

These operators behave just as usual operators on their two separate sets of modes. We retrieve the usual one dimensional Fock state on sl modes by flattening the above Fock array. Also, if there is only a single degree of freedom, the Fock array just reduces to the regular Fock state. We can also retrieve the Fock state of the System modes only or Label modes only by finding the "Fock marginals". The System Fock marginal is then

$$|\underline{r}\rangle = \left| \begin{matrix} r_1 \\ r_2 \\ \vdots \\ r_{d_{\mathrm{S}}} \end{matrix} \right\rangle, \quad r_s = \sum_{l=1}^{d_{\mathrm{L}}} n_{sl}, \tag{3.85}$$

and the presentation of this Fock marginal as a column is purely a visual aid, it can be treated as a regular Fock state. Similarly, the Label Fock marginal is then

$$|\underline{c}\rangle = |c_1, c_2, \ldots, c_{d_{\mathrm{L}}}\rangle, \quad c_l = \sum_{s=1}^{d_{\mathrm{S}}} n_{sl}. \tag{3.86}$$

Commonly when we refer to the occupation numbers of a state which has System and Label degree of freedom, it is the System occupation numbers that we are interested in. We will commonly just use the standard notation for this, i.e. $\underline{n}$, to mean $\underline{r}$.

Interferometer action

We introduce an interferometer which acts only upon the System, corresponding to a unitary transformation on the d_S System modes,

$$\hat{a}_{sl}^\dagger \mapsto \sum_t \hat{a}_{tl}^\dagger U_{ts} \tag{3.87}$$

leaving the label modes untouched. Here U is a $d_\mathrm{S} \times d_\mathrm{S}$ unitary matrix corresponding to a multi-port interferometer, the same matrix from previous discussions in Section 1.3.2, while the interferometer acts trivially upon the Label modes, corresponding to the $d_\mathrm{L} \times d_\mathrm{L}$ identity transfer matrix $\mathbb{1}$. For a suitable choice of ordering of the $d_\mathrm{S}d_\mathrm{L}$ possible creators, the full $d_\mathrm{S}d_\mathrm{L} \times d_\mathrm{S}d_\mathrm{L}$ transfer matrix U_{SL} acts on all $d_\mathrm{S}d_\mathrm{L}$ modes by

$$U_\mathrm{SL} = U \otimes \mathbb{1}, \tag{3.88}$$

where $U \in \mathrm{U}(d_\mathrm{S})$, $\mathbb{1} \in \mathrm{U}(d_\mathrm{L})$. It is tempting to interpret the tensor product in Eq. (3.88) as that between the System and the Label. A quantum information theoretic approach to distinguishability would then ignore (trace out) the Label, arriving at reduced states on the System where all the nontrivial transformations and measurements occur. However, this matrix acts on the space of operators, not on the state space which is a tensor product of $d_\mathrm{S}d_\mathrm{L}$ System-Label harmonic oscillators in the second quantized model.

Applying the unitary-unitary duality theorem, we see that such an interferometer U acts on states in irrep λ according to the irreducible matrix representation U^λ

$$U : |\lambda q q'\rangle_\mathrm{SL} \mapsto \sum_{q''} |\lambda q'' q'\rangle_\mathrm{SL}\, U_{q''q}^\lambda \tag{3.89}$$

leaving q' unchanged as it does not act on the second degree of freedom.

This grants us a description of an input state changing to a new output state in a simpler fashion. Originally we would have had to construct a transfer matrix for the whole space. In this new particle basis, the unitary is block diagonal by irreps (as seen in Section 3.1.2), so we need to compute a set of smaller matrices (some of which are identical), and we might only need a small submatrix to express the parts of the states that we are interested in.

3.2.2   Measurements

The current way to model photon counting detectors makes the assumption that what are being detected is Fock states, and that is what we do in Section 1.3.3. Given an occupation of a state to be $\underline{n}$ we would then model this photon counting measurement as

$$|\underline{n}\rangle \langle \underline{n}|_\mathrm{Fock} = |(N), \underline{n}\rangle \langle (N), \underline{n}| \tag{3.90}$$

However, this might not be the best model, especially without knowing further details about the detector. For example, in the previous section where we had two photons with two degrees of freedom, one being path and one polarization, we might have a PNRD that is polarization insensitive. While the full state was symmetric, we saw in Example 3.2.3 that when broken onto the different degrees of freedom, amplitude in other irreps of the unitary group might occur.

Assume we have a PNRD that only acts on the System degree of freedom. Whatever else is happening in the Label degree of freedom is inaccessible to such a detector. Let $\underline{n}$ be the Fock marginal of the System degree of freedom. Then indeed, this detector will not accurately be modeled by

$$|\underline{n}\rangle \langle\underline{n}|_{\text{Fock}}, \tag{3.91}$$

as this will only account for the amplitudes of the state in the symmetric subspace, yet we saw in Eq. (3.67) that there is also an anti-symmetric state with the same occupation numbers. We will see in the following Section 3.2.3 how these measurement statistics play out.

Therefore, our photon number resolving detectors should really be modeled such that they detect the occupation of the symmetric but also all the other mixed symmetry states as well. That is, it should be modeled as

$$M_{\underline{n}} = \sum_{\lambda pr} |\lambda p\underline{n}r\rangle \langle\lambda p\underline{n}r|, \tag{3.92}$$

a projection onto the $\underline{n}$-weight subspace.

We will also use the notation where we include the interferometer in our definition of a measurement, yielding POVM elements parameterized by unitaries,

$$M_{\underline{n}}(U) = \left(\oplus_\lambda U^\lambda \otimes \mathbb{1}^\lambda\right)^\dagger M_{\underline{n}} \left(\oplus_{\lambda'} U^{\lambda'} \otimes \mathbb{1}^{\lambda'}\right), \tag{3.93}$$

where $\mathbb{1}^\lambda$ corresponds to the irrep of the identity permutation in accordance with Eq. (3.40), (note that we omit the identity when it is only one dimensional, e.g. Eq. (4.25)).

We define the scattering probabilities $P_{\underline{n}}$ related to a measurement operator $M_{\underline{n}}$ using the Born rule. Given an input state $|\psi\rangle$, the probability related to output occupation $\underline{n}$ is then defined as

$$P_{\underline{n}} = \sum_{\lambda,p,r} |\langle\lambda, p, \underline{n}, r \mid \psi\rangle|^2. \tag{3.94}$$

We also define

$$P_{\underline{n}}^\lambda = \sum_{p,r} |\langle\lambda, p, \underline{n}, r \mid \psi\rangle|^2 \tag{3.95}$$

which is the scattering amplitude in the given irrep.

3.2.3 Hong-Ou-Mandel effect in Schur-Weyl basis

Here we take a look at how the setup with two degrees of freedom in the Schur-Weyl basis from Section 3.2.1 along with measurements from Section 3.2.2 come together applied in the HOM scenario. We will come back to the HOM example a few more times throughout this thesis to explain new features, so here we focus only on the basis change, its implications, and the detection probabilities.

In the HOM scenario we have two photons that can be in one of two spatial System modes where we will label "top" with 1 and "bottom" with 2, and two temporal Label modes, where we will label "early" with a and "late" with b. We will look at two states that are relevant for the HOM scenario.

The first state we will name "completely distinguishable", (we will argue later on why this name is suitable),

$$|\psi_{\rm d}\rangle = \hat{a}_{1a}^\dagger \hat{a}_{2b}^\dagger |{\rm vac}\rangle = \left|\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right\rangle, \tag{3.96}$$

with an early photon in the top arm and a late one in the bottom. The other state we are going to name "completely indistinguishable",

$$|\psi_{\rm i}\rangle = \hat{a}_{1a}^\dagger \hat{a}_{2a}^\dagger |{\rm vac}\rangle = \left|\begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix}\right\rangle \tag{3.97}$$

where both photons are early.

Following the same procedure as in Section 3.2.1 we arrive at the following states in the triplet-singlet basis,

$$|\psi_{\rm i}\rangle = \left|\begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix}\right\rangle \tag{3.98}$$

$$= \frac{1}{\sqrt{2}} \left(|12\rangle_{\rm S} + |21\rangle_{\rm S}\right) |aa\rangle_{\rm L}, \tag{3.99}$$

$$= |\boxed{1\,2}\rangle_{\rm S} |\boxed{a\,a}\rangle_{\rm L} \tag{3.100}$$

and

$$|\psi_{\rm d}\rangle = \left|\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right\rangle \tag{3.101}$$

$$= \frac{1}{\sqrt{2}} |12\rangle_{\rm S} |ab\rangle_{\rm L} + \frac{1}{\sqrt{2}} |21\rangle_{\rm S} |ba\rangle_{\rm L} \tag{3.102}$$

$$= \frac{1}{\sqrt{2}} \left(|\boxed{1\,2}\rangle_{\rm S} |\boxed{a\,b}\rangle_{\rm L} + \left|\boxed{\begin{matrix}1\\2\end{matrix}}\right\rangle_{\rm S} \left|\boxed{\begin{matrix}a\\b\end{matrix}}\right\rangle_{\rm L}\right). \tag{3.103}$$

We assume that the interferometer acts trivially upon the Label modes (the photons remain early or late), corresponding to the 2×2 identity transfer matrix $\mathbb{1}$, as mentioned in the interferometer discussion in Section 3.2.1. Let the interferometer we are using be

that of a balanced beamsplitter that is,

$$\text{BS}_{50:50} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}. \tag{3.104}$$

Order the triplet-singlet basis as $\left\{ \left|\boxed{1\,1}\right\rangle, \left|\boxed{1\,2}\right\rangle, \left|\boxed{2\,2}\right\rangle, \left|\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right\rangle \right\}$, Then in this ordering of the basis we have (after applying the Schur-Weyl transform)

$$U^{\otimes 2} \cong U^{\boxed{}} \oplus U^{\boxed{}} = \begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} & \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} & \\ & & & -1 \end{bmatrix}. \tag{3.105}$$

It follows that the indistinguishable state transforms as

$$|\psi_i\rangle = \left|\boxed{1\,2}\right\rangle \left|\boxed{a\,a}\right\rangle \xrightarrow{\text{BS}_{50:50}} 1/\sqrt{2} \left(\left|\boxed{1\,1}\right\rangle + \left|\boxed{2\,2}\right\rangle \right) \left|\boxed{a\,a}\right\rangle. \tag{3.106}$$

The scattering probabilities after the transformation are then

$$P_{(1,1)}(U) = 0, \quad P_{(2,0)}(U) = 0.5, \quad P_{(0,2)}(U) = 0.5. \tag{3.107}$$

where the probabilities and the notation have been defined in Section 3.2.2. The distinguishable state transforms as

$$|\psi_d\rangle = 1/\sqrt{2} \left(\left|\boxed{1\,2}\right\rangle \left|\boxed{a\,b}\right\rangle + \left|\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right\rangle \left|\boxed{\begin{smallmatrix}a\\b\end{smallmatrix}}\right\rangle \right) \xrightarrow{\text{BS}_{50:50}} \tag{3.108}$$

$$1/2 \left(\left|\boxed{1\,1}\right\rangle + \left|\boxed{2\,2}\right\rangle \right) \left|\boxed{a\,b}\right\rangle - 1/\sqrt{2} \left|\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right\rangle \left|\boxed{\begin{smallmatrix}a\\b\end{smallmatrix}}\right\rangle. \tag{3.109}$$

This gives rise to the following scattering probabilities,

$$P_{(1,1)}(U) = 0.5, \quad P_{(2,0)}(U) = 0.25, \quad P_{(0,2)}(U) = 0.25. \tag{3.110}$$

In this notation we easily see why it is that the distinguishable state produces different statistics to the indistinguishable. It contains an antisymmetric part, $\left|\boxed{\begin{smallmatrix}1\\2\end{smallmatrix}}\right\rangle \left|\boxed{\begin{smallmatrix}a\\b\end{smallmatrix}}\right\rangle$ which is the only basis state in the irrep (1^2), leaving no other option of state change except a global phase shift. There is no unitary that will take this antisymmetric state with occupation $(1,1)$ to any other occupation. Therefore $P_{(1,1)}(U) = P^{\boxed{}}_{(1,1)}(U) + P^{\boxed{}}_{(1,1)}(U) = P^{\boxed{}}_{(1,1)}(U) + 0.5 \geq 0.5$ for all choices of U for the input state $|\psi_d\rangle$.

3.2.4 Parameter counting

We can use the formalism to compute the number of parameters that describe an arbitrary partially distinguishable collection of N particles in N (or more generally d) modes. Because of the maximal entanglement over p in Eq. (3.81), when we trace out the Label of an arbitrary totally symmetric state in Eq. (3.80), the resulting mixed state has identical blocks

for each copy of λ (the number of identical blocks being equal to the outer multiplicity). Thus the most general mixed state is described by a single (Hermitian) block for each irrep. Because the number of real parameters in a d-dimensional Hermitian matrix is d^2, we have for an arbitrary partially distinguishable mixed state of N bosons in N modes (subtracting one for normalisation) $\sum_\lambda d_{\{\lambda\}}{}^2 - 1 = \binom{N^2+N-1}{N} - 1$ real parameters. If we restrict to coincident input, the number of states is given by the number of *standard* Young tableaux, $d_{(\lambda)}$. This is because coincidence implies each single particle state is different, and so semistandard tableau become standard; in this case we have $\sum_\lambda d_{(\lambda)}{}^2 - 1 = N! - 1$ real parameters. This number decreases significantly if pure Label states are assumed. A pure state in d dimensions has $2(d-1)$ real parameters, and every pure state added to a set can add at most one parameter beyond those required to describe its projection onto up to $d-1$ dimensional space spanned by the states that came before it (namely the 'angle' it makes with this subspace). Thus there are $\sum_{d=2}^{N}(2(d-2)+1) = (N-1)^2$ real parameters in this case, which agrees with previous analyses [28, 121] but is far fewer than the general case. Note that all of these quantities are of course larger than $\binom{N}{2}$, the number of pairwise distinguishabilities classical intuition might lead one to believe are necessary to measure [185].

3.3 States of interest

For a discussion on what distinguishability means in case of bosons and how this formalism can give a possibly fresh view to distinguishable states and our understanding of them see Section 5.3. The work that follows on the other hand is based around "non-controversial" distinguishable states – the states where each particle is found in its own system mode and its own label mode.

First, let us summarize what the definitions of the previous two Sections mean for distinguishability. By applications of Schur-Weyl and unitary-unitary duality we arrive at a description of bosons with two degrees of freedom that can model partial distinguishability. The total state of bosons then decomposes as seen in Eqs. (3.80) and (3.81),

$$\sum_{\lambda q q'} \psi_{\lambda q q'} \, |\lambda q q'\rangle_{\text{SL}}, \quad |\lambda q q'\rangle_{\text{SL}} := \frac{1}{\sqrt{d_{(\lambda)}}} \sum_{p=1}^{d_{(\lambda)}} |\lambda q p\rangle_{\text{S}} \, |\lambda q' p\rangle_{\text{L}} \qquad (3.111)$$

with the action of interferometer $U \in \text{U}(d_{\text{S}})$ as in Eq. (3.89),

$$U : |\lambda q q'\rangle_{\text{SL}} \mapsto \sum_{q''} |\lambda q'' q'\rangle_{\text{SL}} \, U^\lambda_{q''q}. \qquad (3.112)$$

Just as with a single degree of freedom, the space of second quantized $d_{\text{S}} \times d_{\text{L}}$ Fock arrays can be put into one-to-one correspondence with first quantized totally symmetric states by the procedure exemplified in Eqs. (3.98 - 3.99). Thus we can write an arbitrary partially distinguishable state, which is an element of the totally symmetric subspace of $(\mathbb{C}^{d_{\text{S}}} \otimes \mathbb{C}^{d_{\text{L}}})^{\otimes N}$, in a basis of first quantized states given by Eq. (3.81). We may now trace

out the Label to arrive at mixed states describing any partially distinguishable state of N photons in d_S modes.

Example 3.3.1. Recall Eqs. (3.99, 3.102) describing two states in the HOM setup,

$$|\psi_\mathrm{i}\rangle = \left|\begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix}\right\rangle = |\boxed{1\,2}\rangle_\mathrm{S}\,|\boxed{a\,a}\rangle_\mathrm{L}\,, \tag{3.113}$$

$$|\psi_\mathrm{d}\rangle = \left|\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right\rangle = \frac{1}{\sqrt{2}}\,|\boxed{1\,2}\rangle_\mathrm{S}\,|\boxed{a\,b}\rangle_\mathrm{L} + \frac{1}{\sqrt{2}}\left|\boxed{\begin{matrix}1\\2\end{matrix}}\right\rangle_\mathrm{S}\left|\boxed{\begin{matrix}a\\b\end{matrix}}\right\rangle_\mathrm{L}\,. \tag{3.114}$$

We now see clearly that in this case the Schur-Weyl bases provide a Schmidt decomposition of the Fock arrays, and that the completely distinguishable state has nonzero amplitude outside the totally symmetric irrep.

Tracing out the Label degree of freedom, we arrive at the reduced density matrices that describe the state of the System. Another feature of the Schur-Weyl basis is that these states will be block diagonal, each block corresponding to an irrep λ. Thus, ordering our triplet-singlet basis as before, $\left\{|\boxed{1\,1}\rangle\,,|\boxed{1\,2}\rangle\,,|\boxed{2\,2}\rangle\,,\left|\boxed{\begin{matrix}1\\2\end{matrix}}\right\rangle\right\}$, we have

$$\rho_\mathrm{i} = \mathrm{Tr}_\mathrm{L}\left[|\psi_\mathrm{i}\rangle\langle\psi_\mathrm{i}|\right] = \begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 0 \end{bmatrix}, \quad \rho_\mathrm{d} = \mathrm{Tr}_\mathrm{L}\left[|\psi_\mathrm{d}\rangle\langle\psi_\mathrm{d}|\right] = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 1 \end{bmatrix}. \tag{3.115}$$

This simplification of the states and their projections will be useful in Chapter 4 when we apply discrimination tools to generalize the HOM effect. However, first, we will generalize these distinguishable and indistinguishable states.

We will focus our attention on three types of N-photon states: completely indistinguishable (Section 3.3.1), singly distinguishable (Section 3.3.2), and completely distinguishable (Section 3.3.3), described below. We are not considering loss (where entire qudits would be traced out), so N will be fixed throughout. Situations with mixed System-Label states and/or partial distinguishability can be written in terms of the basis of Eq. (3.80) [202], or equivalently that of Fock arrays. We give examples of this generality with partial distinguishability later in Chapter 4 for two photons in two modes in Sec. 4.4.1, and of mixed System-Label states for three photons in three modes in Sec. 4.5.2. Otherwise we will restrict ourselves to the case where the total System-Label state is pure, corresponding to a source that produces states that are always (in)distinguishable in exactly the same way; generalization is, in principle, straightforward. So far we have not discussed what the size of the System and Label subspaces should be.

In practice the space of Label states available to a photon is as large as that of the uncontrolled degrees of freedom; in general it is infinite dimensional many times over. However, in any given N particle experiment with N fixed, in order to model the distinguishability we need only consider the subspace spanned by the Label states, which can be at most N dimensional. In other words, the most distinguishable N photons can be

is for each of the N Labels to be in an orthogonal state, and so we consider only $d_{\mathrm{L}} \leq N$.

As far as the size of the System, we could consider any size. However, certain states will be more interesting than others. To see this, consider first two photons that are in the same System mode. In first quantisation this means each photon is in the same System state, implying the two photon System state is symmetric, and so in order to maintain total symmetry – or by unitary-unitary duality – the state of the Label must also be symmetric (like in Equations. (5.3.5) and (5.48)). This means that states in the antisymmetric irrep are not available, and so restricts the combined state to a subspace of those allowed in Eq. (3.80). This argument extends to any number of photons, where the situation corresponds to any Fock array that is not full rank N. In order to consider the full set of states available and arbitrary distinguishability, we will have to consider input states that have a single photon in each System mode, giving $d_{\mathrm{S}} \geq N$.

Therefore, we see that the richest distinguishability dynamics will occur in the case with $d_{\mathrm{S}} = d_{\mathrm{L}} = N$, with coincident System photons. If we can find and prove interesting behaviour of the most complex states, one would hope that similar logic would apply for simpler states. The completely indistinguishable state is well motivated as this is the typical true identical boson state (or sometimes in the literature referred as "quantum" state). Similarly, the completely distinguishable state is the typical "classical" state, in which all the particles are completely different from each other. We add another state which is not commonly examined, a "singly distinguishable" state, where only one particle is different than the rest, arguably the most common type of error if we are trying to generate many identical particles.

3.3.1 Completely indistinguishable states

A completely indistinguishable state is one in which every photons' Label state is the same, e.g.

$$\hat{a}_{11}^{\dagger}\hat{a}_{21}^{\dagger}\cdots\hat{a}_{N1}^{\dagger}\left|\mathrm{vac}\right\rangle = \left|\begin{smallmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 1 & 0 & \cdots & 0 \end{smallmatrix}\right\rangle, \tag{3.116}$$

where the System part of the state is that of coincident photons, and we have included $N-1$ redundant zero columns in the Fock array so we can easily compare with the other states in this section.

Proposition 3.3.2. *The reduced System state of the completely indistinguishable state is*

$$\rho_{\mathrm{i}} = \left|(N), \underline{1}\right\rangle \left\langle(N), \underline{1}\right|. \tag{3.117}$$

Proof. In the first quantized picture then we see

$$\left| \begin{smallmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 1 & 0 & \cdots & 0 \end{smallmatrix} \right\rangle = \mathrm{Sym} \left(|11\rangle \, |21\rangle \cdots |N1\rangle \right) \tag{3.118}$$

$$= \frac{1}{\sqrt{N!}} \sum_{\sigma \in \mathrm{S}_N} |\sigma(1)\sigma(2)\dots\sigma(N)\rangle_{\mathrm{S}} \, |11\dots1\rangle_{\mathrm{L}} \tag{3.119}$$

$$= \frac{1}{\sqrt{N!}} \sum_{\sigma \in \mathrm{S}_N} |\sigma(1)\sigma(2)\dots\sigma(N)\rangle_{\mathrm{S}} \, |(N),(N,\underline{0})\rangle_{\mathrm{L}}, \tag{3.120}$$

ordering our Label modes such that the occupied one is first, and with the understanding that the list of zeroes, $\underline{0}$, is as long as it needs to be, in this case $N-1$.

To find what the System part of the state should be, we first notice that in the Eq. (3.80) λ can only be (N), as the Label state lies fully in the symmetric Label subspace, (N). Also from the above equation we can see that q' can be replaced with the Label occupation $(N,\underline{0})$ (changing the notation from that of $|\lambda q p\rangle$ to the equivalent notation $|\lambda p \underline{n} r\rangle$ where q is broken into two components). Since the symmetric irrep of S_N is one dimensional, $d_{(N)} = 1$, and Schur-Weyl duality tells us that the corresponding unitary irrep is always outer multiplicity free, the Eq.(3.81) simplifies to a single term. Moreover, (N) is also inner multiplicity free, (there is only one way to symmetrise a product of single particle states) and here we can see that the occupation in the computational basis is $\underline{1}$, therefore in Eq. (3.80) we can replace q with the System occupation $\underline{1}$. The total state in Eq. (3.80) therefore becomes

$$\hat{a}_{11}^\dagger \hat{a}_{21}^\dagger \cdots \hat{a}_{N1}^\dagger |\mathrm{vac}\rangle = |(N),\underline{1}\rangle_{\mathrm{S}} \, |(N),(N,\underline{0})\rangle_{\mathrm{L}}, \tag{3.121}$$

where we have suppressed the trivial multiplicities. We see that this is always a product state, with no correlation between the System and Label, as expected for completely indistinguishable particles. The reduced System state is

$$\rho_{\mathrm{i}} = \mathrm{Tr}_{\mathrm{L}} \big[|(N),\underline{1}\rangle \, |(N),(N,\underline{0})\rangle \, \langle(N),\underline{1}| \, \langle(N),(N,\underline{0})| \big] \tag{3.122}$$

$$= |(N),\underline{1}\rangle \, \langle(N),\underline{1}|, \tag{3.123}$$

supported on the one dimensional intersection of the symmetric System subspace given by (N), with the coincident subspace defined by the System occupation number $\underline{1}$. $\square$

3.3.2   Singly distinguishable states

The next state we consider is one where a single photon has become distinguishable from the rest; assuming all efforts are being made to produce the completely indistinguishable state, this is the most likely error to occur. Ordering our modes so that the 'bad' photon

is in System mode N and Label mode 2 for ease of writing, we have

$$
\hat{a}_{11}^{\dagger}\hat{a}_{21}^{\dagger}\cdots\hat{a}_{N2}^{\dagger}\left|\text{vac}\right\rangle = \left|\begin{smallmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 1 & \cdots & 0 \end{smallmatrix}\right\rangle \tag{3.124}
$$

Proposition 3.3.3. *The reduced System state of the singly distinguishable state is*

$$
\rho_{\text{s}} = \frac{1}{N}\left|(N),\underline{1}\right\rangle\left\langle(N),\underline{1}\right| + \frac{1}{N}\sum_{p=1}^{N-1}\left|(N-1,1),p,\underline{1},1\right\rangle\left\langle(N-1,1),p,\underline{1},1\right|. \tag{3.125}
$$

Proof. The singly distinguishable state in the computational basis is

$$
\hat{a}_{11}^{\dagger}\hat{a}_{21}^{\dagger}\cdots\hat{a}_{N2}^{\dagger}\left|\text{vac}\right\rangle = \left|\begin{smallmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 1 & \cdots & 0 \end{smallmatrix}\right\rangle \tag{3.126}
$$

$$
= \text{Sym}\left(\left|11\right\rangle\left|21\right\rangle\cdots\left|(N-1)1\right\rangle\left|N2\right\rangle\right), \tag{3.127}
$$

One observes there are $N!$ permutations of the N distinct System indices. On the other hand, only two distinct Label modes are involved, so there are only N single particle states available to the Label degree of freedom, namely those with the jth photon in Label mode 2 and the rest in Label mode 1. Denote these states $\left|2_j\right\rangle_{\text{L}}$, i.e.

$$
\left|2_j\right\rangle_{\text{L}} := \left|1\right\rangle_{\text{L}}^{\otimes j-1}\left|2\right\rangle_{\text{L}}\left|1\right\rangle_{\text{L}}^{\otimes N-j} \tag{3.128}
$$

Such a Label state will be perfectly correlated to all System states with the jth photon in mode N; for each j we can factor these $(N-1)!$ System states off, and denote the resulting normalised state $\left|N_j\right\rangle_{\text{S}}$,

$$
\left|N_j\right\rangle_{\text{S}} := \frac{1}{\sqrt{(N-1)!}}\sum_{\sigma\in\text{S}_{N-1}}\bigotimes_{k=1}^{j-1}\left|\sigma(k)\right\rangle\left|N\right\rangle\bigotimes_{k=j+1}^{N}\left|\sigma(k)\right\rangle \tag{3.129}
$$

Thus in the System-Label basis, the singly distinguishable state can be written as

$$
\text{Sym}\left(\left|11\right\rangle\left|21\right\rangle\cdots\left|N2\right\rangle\right) = \frac{1}{\sqrt{N}}\sum_{j=1}^{N}\left|N_j\right\rangle_{\text{S}}\left|2_j\right\rangle_{\text{L}}, \tag{3.130}
$$

e.g. Eq. (3.102). These sets of states are orthonormal, therefore we recognise this as an entangled state with Schmidt coefficients $1/\sqrt{N}$.

Now consider Schur-Weyl transforming this state into the form of Eq. (3.80). Because there are only two distinct Label modes involved, the only Label irreps that can occur are those whose Young diagrams have two or fewer rows. Moreover, because only a single photon is 'bad', the only two rowed diagram allowed is that with a single box in the second row. Thus the Label state is supported only by irreps $\lambda = (N)$ and $(N-1,1)$. By unitary-unitary duality, the System is therefore also supported only on these two irreps. The totally symmetric irrep (N) is always both inner and outer multiplicity free; for irrep

$(N-1, 1)$, the outer multiplicity is $d_{((N-1,1))} = N - 1$. It remains only to work out the inner multiplicities for irrep $(N-1, 1)$. The System and Label have occupations $\underline{1}$ and $(N-1, 1, \underline{0})$ respectively (the marginals of the Fock array). There is only one Young tableau of shape $(N-1, 1)$ consistent with occupation $(N-1, 1, \underline{0})$, (that with the 2 in the second row box), so the Label states are inner multiplicity free. The System occupation $\underline{1}$ is consistent with $N - 1$ Young tableaux of shape $(N-1, 1)$, (all those without a 1 in the second row box), and so the System inner multiplicity is $N - 1$. Inserting these observations into Eq. (3.80) (and replacing q with the suitable $\underline{n}, r$ pair instead), the Schur-Weyl transformed state is

$$
\begin{aligned}
&\psi_{(N),\underline{1},1,(N-1,1,\underline{0}),1} \left| (N), 1, \underline{1}, 1 \right\rangle_{\mathrm{S}} \left| (N), 1, (N-1, 1, \underline{0}), 1 \right\rangle_{\mathrm{L}} \\
&+ \sum_{r=1}^{N-1} \frac{\psi_{(N-1,1),\underline{1},r,(N-1,1,\underline{0}),1}}{\sqrt{N-1}} \sum_{p=1}^{N-1} \left| (N-1,1), p, \underline{1}, r \right\rangle_{\mathrm{S}} \left| (N-1,1), p, (N-1,1,\underline{0}), 1 \right\rangle_{\mathrm{L}}.
\end{aligned}
$$

(3.131)

where we are yet to work out the coefficients. We can factor the second term and redefine coefficients to yield another Schmidt decomposition:

$$
\begin{aligned}
&\psi_{(N)} \left| (N), 1, \underline{1}, 1 \right\rangle_{\mathrm{S}} \left| (N), 1, (N-1, 1, \underline{0}), 1 \right\rangle_{\mathrm{L}} \\
&+ \frac{\psi_{(N-1,1)}}{\sqrt{N-1}} \sum_{p=1}^{N-1} \left(\sum_{r=1}^{N-1} \phi_r \left| (N-1,1), p, \underline{1}, r \right\rangle_{\mathrm{S}} \right) \left| (N-1,1), p, (N-1,1,\underline{0}), 1 \right\rangle_{\mathrm{L}},
\end{aligned}
$$

(3.132)

where $\psi_{(N-1,1),\underline{1},r,(N-1,1,\underline{0}),1} = \psi_{(N-1,1)} \phi_r$. Because the Schur-Weyl transformations yielding Eq. (3.80) are performed independently on each degree of freedom, the System-Label entanglement cannot be changed. From Eq. (3.130) we know that the Schmidt coefficients are all $1/\sqrt{N}$, so we must have $\psi_{(N)} = 1/\sqrt{N}$ and $\psi_{(N-1,1)} = \sqrt{(N-1)/N}$. The amplitudes ϕ_r do not affect this entanglement at all – they depend on how one chooses to orthonormalise multiplicities in the Schur-Weyl transform, and encode the fact that we chose the 'bad' photon to be in System mode N. We can always choose $r = 1$ to correspond to this specific situation, and then use the subgroup of $\mathrm{U}(d_{\mathrm{S}})$ that permutes System modes to find the states corresponding to the 'bad' photon being in any other mode.

Making this choice and tracing out the Label in Eq. (3.132) yields the singly distinguishable reduced state (now suppressing trivial multiplicities)

$$
\begin{aligned}
\rho_{\mathrm{s}} = & \frac{1}{N} \left| (N), \underline{1} \right\rangle \left\langle (N), \underline{1} \right| \\
& + \frac{1}{N} \sum_{p=1}^{N-1} \left| (N-1,1), p, \underline{1}, 1 \right\rangle \left\langle (N-1,1), p, \underline{1}, 1 \right|.
\end{aligned}
$$

(3.133)

We see that this is mixed over N dimensions of the coincident subspace, overlapping the symmetric and 'almost symmetric' $(N-1, 1)$ irreps. $\qquad \square$

3.3.3 Completely distinguishable state

A completely distinguishable state has each particle in a distinct Label mode, paired with a unique System mode. We can choose to order the modes such that the corresponding Fock array is diagonal, cf. Eq. (3.96), for example,

$$
\hat{a}_{11}^{\dagger}\hat{a}_{22}^{\dagger}\cdots\hat{a}_{NN}^{\dagger}\left|\mathrm{vac}\right\rangle = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{vmatrix}\Big\rangle . \tag{3.134}
$$

Proposition 3.3.4. *The reduced System state of the completely distinguishable state is*

$$
\rho_{\mathrm{d}} = \frac{1}{N!}\sum_{\lambda p r}\left|\lambda, p, \underline{1}, r\right\rangle\left\langle\lambda, p, \underline{1}, r\right| . \tag{3.135}
$$

Proof. Generalising the symmetrisation procedure of Eqs. (3.98 - 3.99) to N particles, one finds that all $N!$ possible terms will occur in the single particle picture (with System and Label state coupled to each other), and they will each occur once, that is

$$
\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{vmatrix}\Big\rangle = \mathrm{Sym}\left(\left|11\right\rangle\left|22\right\rangle\cdots\left|NN\right\rangle\right) \tag{3.136}
$$

$$
= \frac{1}{\sqrt{N!}}\sum_{\sigma\in\mathrm{S}_N}\left|\sigma(1)\sigma(2)\ldots\sigma(N)\right\rangle\left|\sigma(1)\sigma(2)\ldots\sigma(N)\right\rangle . \tag{3.137}
$$

The unique pairing of System and Label modes manifests as maximal entanglement between the System and Label single particle states in the coincident subspace. Recall that in our choice of Schur-Weyl basis, the occupation numbers are preserved. As above, because the Schur-Weyl transformations yielding Eq. (3.80) are local and unitary on the coincident subspace, the System-Label entanglement on this subspace is preserved. This means that the transformed state must also be maximally entangled with the same Schmidt rank. Thus

$$
\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{vmatrix}\Big\rangle = \frac{1}{\sqrt{N!}}\sum_{\lambda p r}\left|\lambda, p, \underline{1}, r\right\rangle_{\mathrm{S}}\left|\lambda, p, \underline{1}, r\right\rangle_{\mathrm{L}} , \tag{3.138}
$$

with the sum running over all allowed values of irrep, outer, and inner multiplicities. The completely distinguishable reduced System state is therefore

$$
\rho_{\mathrm{d}} = \frac{1}{N!}\sum_{\lambda p r}\left|\lambda, p, \underline{1}, r\right\rangle\left\langle\lambda, p, \underline{1}, r\right| \tag{3.139}
$$

$$
= \frac{1}{N!}\left|(N), \underline{1}\right\rangle\left\langle(N), \underline{1}\right| + \frac{1}{N!}\sum_{\lambda\neq(N), p, r}\left|\lambda, p, \underline{1}, r\right\rangle\left\langle\lambda, p, \underline{1}, r\right| , \tag{3.140}
$$

which is completely mixed over the $N!$ dimensional coincident subspace.

$\square$

3.3.4 General states

Although we have focused on single and complete distinguishability, the formalism admits arbitrary states. Consider, for example, Fock arrays with a single excitation in each System mode and an arbitrary Label occupation, call it $\underline{n}_{\mathrm{L}}$. Applying the Schur-Weyl transform and focusing on the symmetric irrep (N), where the support is one dimensional, we see that the reduced system state will be of the form

$$\frac{n_{\mathrm{L}}!}{N!} |(N), \underline{1}\rangle \langle (N), \underline{1}| + \left(1 - \frac{n_{\mathrm{L}}!}{N!}\right) \rho_{\bar{\mathrm{i}}}. \tag{3.141}$$

The exact form of such states could be found by reasoning as in previous sections.

3.3.5 Scattering probabilities for the states of interest

Let us look at what the probability of a specific measurement pattern $\underline{n}$ being detected at the output of an arbitrary interferometer U is for the states of interest, starting with the completely distinguishable state. From Eqs. (3.140) and (3.93),

$$
\begin{aligned}
\mathrm{Tr}\left[\rho_{\mathrm{d}} M_{\underline{n}}(U)\right] = \mathrm{Tr} &\left[\left(\frac{1}{N!} \sum_{\lambda, p, r} |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r| \right) (\oplus_{\mu} U^{\mu} \otimes \mathbb{1}^{\mu})^{\dagger} \right. \\
&\left. \left(\sum_{\lambda', p', r'} |\lambda', p', \underline{n}, r'\rangle \langle \lambda', p', \underline{n}, r'| \right) \left(\oplus_{\mu'} U^{\mu'} \otimes \mathbb{1}^{\mu'}\right) \right] \\
= \frac{1}{N!} &\sum_{\lambda, p, r, r'} \mathrm{Tr}\left[|\lambda, p, \underline{n}, r'\rangle \langle \lambda, p, \underline{n}, r'| \left(U^{\lambda} \otimes \mathbb{1}^{\lambda}\right) |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r| \left(U^{\lambda} \otimes \mathbb{1}^{\lambda}\right)^{\dagger}\right] \\
= \frac{1}{N!} &\sum_{\lambda, p, r, r'} |\langle \lambda, p, \underline{n}, r| U^{\lambda} \otimes \mathbb{1}^{\lambda} |\lambda, p, \underline{1}, r'\rangle|^2 \\
= \frac{1}{N!} &\sum_{\lambda, r, r'} d_{(\lambda)} |\langle \lambda, \underline{n}, r| U^{\lambda} |\lambda, \underline{1}, r'\rangle|^2, \tag{3.142}
\end{aligned}
$$

where in the last line we have used the fact that outer multiplicities p give rise to identical copies of unitary irreps to write the probability in terms of irreducible unitary matrix elements. When $r = r' = 1$ these matrix elements are immanants [203] of a matrix $U_{\underline{1}\mathrm{n}}$ whose rows and columns are determined by the input and output occupations of the interferometer given by U [137, 204] as will be discussed in Section 5.1. Moreover, the completely distinguishable case can be interpreted as independent classical particles evolving stochastically [162], leading to the remarkable fact that the sum in Eq. (3.142) can always be written in terms of the permanent of the matrix given by the elementwise square amplitudes of $U_{\underline{1}\mathrm{n}}$. Note that $U_{\underline{11}} = U$.

The calculation for the singly distinguishable and completely indistinguishable state is the same as Eq. (3.142), only with fewer irreps occurring. Recalling from Sec. 3.3.2 that

$d_{((N-1,1))} = N - 1$, Eq. (3.125) gives

$$\text{Tr}\left[\rho_{\text{s}} M_{\underline{n}}(U)\right] = \frac{1}{N} |\langle(N), \underline{n}| U^{(N)} |(N), \underline{1}\rangle|^2 \tag{3.143}$$

$$+ \frac{N-1}{N} \sum_r |\langle(N-1,1), \underline{n}, r| U^{(N-1,1)} |(N-1,1), \underline{1}, 1\rangle|^2, \tag{3.144}$$

where the sum is over all r consistent with $\underline{n}$, and Eq. (3.117) gives

$$\text{Tr}\left[\rho_{\text{i}} M_{\underline{n}}(U)\right] = |\langle(N), \underline{n}| U^{(N)} |(N), \underline{1}\rangle|^2, \tag{3.145}$$

where as mentioned above these matrix elements are expressible in terms of $\text{per}U_{\underline{1}\text{n}}$ [205].

3.4   Example: Three photons

We now work through the theory introduced in the previous Sections on examples with three bosons. First we take a look at three qubits, or three bosons in two modes (this example is complex enough to demonstrate some of the finer points, but not so much that the number of states is too big). As we refer to the three particle Schur-Weyl often, we also write it out in full for three qutrits (the most general three particle basis, as for any $d \geq 3$ the states that appear would just be a repetition of one of the states in the three trits basis with suitable state replacements). We also look at some specific features of scattering probabilities when separated by the irrep that show up in the case of three photons in three modes when using coincident input state.

3.4.1   Schur-Weyl basis for three qubits or three photons in two modes

In the case of three particles we have three Young diagrams, which are ⬜⬜⬜, ⬛, and ⬛. However, we recall that we cannot have more rows than the size of the single particle Hilbert space, which is two in the case of qubits, therefore there are only two Young diagrams we care about, ⬜⬜⬜ and ⬛, that is $\text{Par}(3,2) = \{(3),(2,1)\}$.

Now we use the hook formula, Eq. (3.1), to find that $d_{(3)} = 1$, $d_{(2,1)} = 2$. Therefore, there is an outer multiplicity in the irrep ⬛. From the hook content formula, Eq. (3.2), we find $d_{\{3\}} = 4$, $d_{\{2,1\}} = 2$. In our three qubit picture, the dimension of the space is 2^3. We check that $d_{(3)} \times d_{\{3\}} + d_{(2,1)} \times d_{\{2,1\}} = 1 \times 4 + 2 \times 2 = 8$ as expected. As we are dealing with qubits we can see from Eq. (3.58) there will be only one lowering operator, that is

$$\mathcal{L}_{12} = |2\rangle\langle 1| \otimes I \otimes I + I \otimes |1\rangle\langle 2| \otimes I + I \otimes I \otimes |1\rangle\langle 2|. \tag{3.146}$$

Irrep (3)

We start with partition $\square\square\square$. As mentioned, $d_{(3)} = 1$ and $d_{\{3\}} = 4$. The four states we have in this irrep of $U(2)$ are:

$$S_3 \times U(2)$$

$$\boxed{1\,2\,3} \, , \; \boxed{1\,1\,1}$$

$$\boxed{1\,2\,3} \, , \; \boxed{1\,1\,2}$$

$$\boxed{1\,2\,3} \, , \; \boxed{1\,2\,2}$$

$$\boxed{1\,2\,3} \, , \; \boxed{2\,2\,2} \, .$$

First we find $\mathrm{ST}(\lambda) = \boxed{1\,1\,1}$, and $\mathrm{T}(\lambda) = \boxed{1\,2\,3}$. The highest weight state for partition (3) will have weight $\underline{n} = (3)$ (or that is $(3,0)$ with the trailing zero). We can see the that highest weight state is then $|(\boxed{1\,2\,3}, \boxed{1\,1\,1})\rangle\rangle = \mathfrak{D}_1^1 \mathfrak{D}_2^1 \mathfrak{D}_3^1 = |1\rangle \otimes |1\rangle \otimes |1\rangle$. We apply the lowering operator on the highest weight state to get the other states in this unitary irrep.

$$|(\boxed{1\,2\,3}, \boxed{1\,1\,2})\rangle\rangle \sim \mathcal{L}_{12} |(\boxed{1\,2\,3}, \boxed{1\,1\,1})\rangle\rangle = |2\rangle\langle 1| \otimes I \otimes I \, |111\rangle + I \otimes |2\rangle\langle 1| \otimes I \, |111\rangle \tag{3.147}$$

$$+ \, I \otimes I \otimes |2\rangle\langle 1| \, |111\rangle \tag{3.148}$$

$$= |211\rangle + |121\rangle + |112\rangle \, . \tag{3.149}$$

Then the rest of the states are

$$|(\boxed{1\,2\,3}, \boxed{2\,2\,2})\rangle\rangle \sim \mathcal{L}_{12} |(\boxed{1\,2\,3}, \boxed{1\,2\,2})\rangle\rangle \sim \mathcal{L}_{12} \left(|211\rangle + |121\rangle + |112\rangle \right) \tag{3.150}$$

$$= |221\rangle + |212\rangle + |221\rangle + |122\rangle + |212\rangle + |122\rangle \tag{3.151}$$

$$|(\boxed{1\,2\,3}, \boxed{2\,2\,2})\rangle\rangle \sim \mathcal{L}_{12} |(\boxed{1\,2\,3}, \boxed{1\,2\,2})\rangle\rangle \sim \mathcal{L}_{12} \left(|221\rangle + |212\rangle + |122\rangle \right) \tag{3.152}$$

$$= |222\rangle + |222\rangle + |222\rangle \, . \tag{3.153}$$

As a check, we can see how the raising operator annihilates the highest state. The raising operator is

$$\mathcal{R}_{21} = |1\rangle\langle 2| \otimes I \otimes I + I \otimes |1\rangle\langle 2| \otimes I + I \otimes I \otimes |1\rangle\langle 2| \, . \tag{3.154}$$

Then

$$\mathcal{R}_{21} |(\boxed{1\,2\,3}, \boxed{1\,1\,1})\rangle\rangle = 0. \tag{3.155}$$

Irrep (2,1)

The next partition is $\begin{smallmatrix}\square\square\\\square\end{smallmatrix}$. As mentioned, $d_{(2,1)} = 2$ and $d_{\{2\}} = 2$. So in this case we will have two outer multiplicities and two states in the $U(2)$ irrep $(2,1)$. To find the states within the irrep we will apply the lowering operator as before, however to find all the

highest weights, we will have to apply a permutation operator. These states are

$$S_3 \times U(2)$$

$$\young(13,2) \,,\; \young(11,2)$$

$$\young(12,3) \,,\; \young(11,2)$$

$$\young(13,2) \,,\; \young(12,2)$$

$$\young(12,3) \,,\; \young(12,2) \,.$$

First we find $\mathrm{ST}(\lambda) = \young(11,2)$, and $\mathrm{T}(\lambda) = \young(13,2)$. The highest weight is then

$$\left| \left(\young(13,2), \young(11,2) \right) \right\rangle = \mathfrak{D}_{12}^{12} \mathfrak{D}_3^1 = \frac{1}{\sqrt{2}} \left(|12\rangle_{12} - |21\rangle_{12} \right) |1\rangle_3 = \frac{1}{\sqrt{2}} \left(|121\rangle - |211\rangle \right). \qquad (3.156)$$

Applying the lowering operator, we get

$$\left| \left(\young(13,2), \young(12,2) \right) \right\rangle \sim \mathcal{L}_{12} \left| \left(\young(13,2), \young(11,2) \right) \right\rangle = \frac{1}{\sqrt{2}} \left(|2\rangle \langle 1| \otimes I \otimes I \right) \left(|121\rangle - |211\rangle \right) \qquad (3.157)$$

$$+ \frac{1}{\sqrt{2}} \left(I \otimes |2\rangle \langle 1| \otimes I \right) \left(|121\rangle - |211\rangle \right) \qquad (3.158)$$

$$+ \frac{1}{\sqrt{2}} \left(I \otimes I \otimes |2\rangle \langle 1| \right) \left(|121\rangle - |211\rangle \right) \qquad (3.159)$$

$$= \frac{1}{\sqrt{2}} |221\rangle - \frac{1}{\sqrt{2}} |221\rangle + \frac{1}{\sqrt{2}} \left(|122\rangle - |212\rangle \right) \qquad (3.160)$$

$$= \frac{1}{\sqrt{2}} \left(|122\rangle - |212\rangle \right). \qquad (3.161)$$

Now we permute the particle labels with $(1)(23)$ permutation which corresponds to the group element of this Young diagram so

$$(1)(23) \left| \left(\young(13,2), \young(11,2) \right) \right\rangle = \frac{1}{\sqrt{2}} \left(|12\rangle_{13} - |21\rangle_{13} \right) |1\rangle_2 \qquad (3.162)$$

$$= \frac{1}{\sqrt{2}} \left(|112\rangle - |211\rangle \right) \qquad (3.163)$$

$$= \left| \left(\young(12,3), \young(11,2) \right) \right\rangle. \qquad (3.164)$$

This is the other highest state in the carrier space of the permutation irrep $\yng(2,1)$. We apply the lowering operator one more time to get the last state of this irrep. We also need to make these state orthonormal, as they are not naturally so by construction.

Again we can check that the raising operator annihilates the highest state,

$$\mathcal{R}_{12} \left| \left(\young(13,2), \young(11,2) \right) \right\rangle = \frac{1}{\sqrt{2}} \left(|1\rangle \langle 2| \otimes I \otimes I \right) \left(|121\rangle - |211\rangle \right) \qquad (3.165)$$

$$= \frac{1}{\sqrt{2}} \left(|111\rangle - |111\rangle \right) = 0 \qquad (3.166)$$

Schur-Weyl basis for three qubits

We need to orthonormalize the states we have found above, so we see how there is freedom in choosing what the Schur-Weyl basis will be. The full 3 qubit basis of our choice is (the first notation is the one with Young tableaux, the second notation is the shortened $|\lambda, p, \underline{n}, r\rangle$ notation, see Section 3.1.2), and the third is the computational basis,

$$|(\boxed{1\,2\,3}, \boxed{1\,1\,1})\rangle = |(3), (3,0)\rangle = |111\rangle \tag{3.167}$$

$$|(\boxed{1\,2\,3}, \boxed{1\,1\,2})\rangle = |(3), (2,1)\rangle = 1/\sqrt{3}\,(|112\rangle + |121\rangle + |211\rangle) \tag{3.168}$$

$$|(\boxed{1\,2\,3}, \boxed{1\,2\,2})\rangle = |(3), (1,2)\rangle = 1/\sqrt{3}\,(|122\rangle + |212\rangle + |221\rangle) \tag{3.169}$$

$$|(\boxed{1\,2\,3}, \boxed{2\,2\,2})\rangle = |(3), (0,3)\rangle = |222\rangle \tag{3.170}$$

$$\left|\left(\begin{smallmatrix}1&3\\2&\end{smallmatrix}, \begin{smallmatrix}1&1\\2&\end{smallmatrix}\right)\right\rangle = |(2,1), (2,1)\rangle = 1/\sqrt{2}\,(|121\rangle - |211\rangle) \tag{3.171}$$

$$\left|\left(\begin{smallmatrix}1&3\\2&\end{smallmatrix}, \begin{smallmatrix}1&2\\2&\end{smallmatrix}\right)\right\rangle = |(2,1), (1,2)\rangle = 1/\sqrt{2}\,(|122\rangle - |212\rangle) \tag{3.172}$$

$$\left|\left(\begin{smallmatrix}1&2\\3&\end{smallmatrix}, \begin{smallmatrix}1&1\\2&\end{smallmatrix}\right)\right\rangle = |(2,1), 2, (2,1)\rangle = 1/\sqrt{6}\,(2\,|112\rangle - |121\rangle - |211\rangle) \tag{3.173}$$

$$\left|\left(\begin{smallmatrix}1&2\\3&\end{smallmatrix}, \begin{smallmatrix}1&2\\2&\end{smallmatrix}\right)\right\rangle = |(2,1), 2, (1,2)\rangle = 1/\sqrt{6}\,(-2\,|221\rangle + |122\rangle + |212\rangle). \tag{3.174}$$

The basis change matrix is

$$U_{\mathrm{Sch}} = \begin{array}{r} \langle(\boxed{1\,2\,3}, \boxed{1\,1\,1})| \\ \langle(\boxed{1\,2\,3}, \boxed{1\,1\,2})| \\ \langle(\boxed{1\,2\,3}, \boxed{1\,2\,2})| \\ \langle(\boxed{1\,2\,3}, \boxed{2\,2\,2})| \\ \langle(\begin{smallmatrix}1&3\\2&\end{smallmatrix}, \begin{smallmatrix}1&1\\2&\end{smallmatrix})| \\ \langle(\begin{smallmatrix}1&3\\2&\end{smallmatrix}, \begin{smallmatrix}1&2\\2&\end{smallmatrix})| \\ \langle(\begin{smallmatrix}1&2\\3&\end{smallmatrix}, \begin{smallmatrix}1&1\\2&\end{smallmatrix})| \\ \langle(\begin{smallmatrix}1&2\\3&\end{smallmatrix}, \begin{smallmatrix}1&2\\2&\end{smallmatrix})| \end{array} \begin{array}{cccccccc} |111\rangle & |112\rangle & |121\rangle & |122\rangle & |211\rangle & |212\rangle & |221\rangle & |222\rangle \\ \left[\begin{array}{cccccccc} \sqrt{6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{2} & \sqrt{2} & 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{6} \\ 0 & 0 & \sqrt{3} & 0 & -\sqrt{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3} & 0 & -\sqrt{3} & 0 & 0 \\ 0 & 2 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & -2 & 0 \end{array}\right] \end{array} \frac{1}{\sqrt{6}} \tag{3.175}$$

Now given any number of particles, N, recall from Eq. 3.40 that we can apply this transform as follows,

$$U_{\mathrm{Sch}} \mathcal{U} U_{\mathrm{Sch}}^{\dagger} = \oplus_{\lambda} U^{\lambda} \otimes \mathbb{1}^{\lambda}. \tag{3.176}$$

States that are within the same unitary irrep can be accessed through single particle unitary transformations. To change symmetric group multiplicities, a permutation of the particle labelling is needed.

Another thing that is interesting is seeing what the $\underline{n}$-weight subspace looks like. For example choosing $\underline{n} = (1,2) = \underline{n}'$, we can find the right subspace of the unitary as per

above, $\mathcal{U}(\underline{n})$ to be equal to

$$
\begin{array}{c}
 \\
\langle(\boxed{1\,2\,3},\boxed{1\,2\,2})| \\
\left\langle\left(\substack{\boxed{1\,3}\\\boxed{2}},\substack{\boxed{1\,2}\\\boxed{2}}\right)\right| \\
\left\langle\left(\substack{\boxed{1\,2}\\\boxed{3}},\substack{\boxed{1\,2}\\\boxed{2}}\right)\right|
\end{array}
\begin{array}{c}
|(\boxed{1\,2\,3},\boxed{1\,2\,2})\rangle \qquad \left|\left(\substack{\boxed{1\,3}\\\boxed{2}},\substack{\boxed{1\,2}\\\boxed{2}}\right)\right\rangle \qquad \left|\left(\substack{\boxed{1\,2}\\\boxed{3}},\substack{\boxed{1\,2}\\\boxed{2}}\right)\right\rangle \\
\left[
\begin{array}{ccc}
U_{22}\left(U_{11}U_{22}+2U_{12}U_{21}\right) & 0 & 0 \\
0 & U_{22}\left(U_{11}U_{22}-U_{12}U_{21}\right) & 0 \\
0 & 0 & U_{22}\left(U_{11}U_{22}-U_{12}U_{21}\right)
\end{array}
\right]
\end{array}
\tag{3.177}
$$

We can see that the block diagonal structure is preserved in this subspace.

3.4.2 Schur-Weyl basis for three particles in three modes

Using the techniques exemplified above, we can takle the $N=3$, $d=3$ case. This is the specific choice of Schur-Weyl basis for three particles we refer to throughout the text.

$\lambda=(3):$

$|(3),1,1\rangle = |(3),(3,0,0)\rangle = |(\boxed{1\,2\,3},\boxed{1\,1\,1})\rangle = |111\rangle\,,$

$|(3),2,1\rangle = |(3),(2,1,0)\rangle = |(\boxed{1\,2\,3},\boxed{1\,1\,2})\rangle = \frac{1}{\sqrt{3}}|112\rangle+\frac{1}{\sqrt{3}}|121\rangle+\frac{1}{\sqrt{3}}|211\rangle\,,$

$|(3),3,1\rangle = |(3),(1,2,0)\rangle = |(\boxed{1\,2\,3},\boxed{1\,2\,2})\rangle = \frac{1}{\sqrt{3}}|122\rangle+\frac{1}{\sqrt{3}}|212\rangle+\frac{1}{\sqrt{3}}|221\rangle\,,$

$|(3),4,1\rangle = |(3),(1,1,1)\rangle = |(\boxed{1\,2\,3},\boxed{1\,2\,3})\rangle = \frac{1}{\sqrt{6}}|123\rangle+\frac{1}{\sqrt{6}}|132\rangle+\frac{1}{\sqrt{3}}|213\rangle$

$\qquad\qquad\qquad\qquad\qquad +\frac{1}{\sqrt{6}}|231\rangle+\frac{1}{\sqrt{6}}|321\rangle+\frac{1}{\sqrt{3}}|312\rangle\,,$

$|(3),5,1\rangle = |(3),(1,0,2)\rangle = |(\boxed{1\,2\,3},\boxed{1\,3\,3})\rangle = \frac{1}{\sqrt{3}}|133\rangle+\frac{1}{\sqrt{3}}|313\rangle+\frac{1}{\sqrt{3}}|331\rangle\,,$

$|(3),6,1\rangle = |(3),(2,0,1)\rangle = |(\boxed{1\,2\,3},\boxed{1\,1\,3})\rangle = \frac{1}{\sqrt{3}}|113\rangle+\frac{1}{\sqrt{3}}|131\rangle+\frac{1}{\sqrt{3}}|311\rangle\,,$

$|(3),7,1\rangle = |(3),(0,3,0)\rangle = |(\boxed{1\,2\,3},\boxed{2\,2\,2})\rangle = |222\rangle\,,$

$|(3),8,1\rangle = |(3),(0,2,1)\rangle = |(\boxed{1\,2\,3},\boxed{2\,2\,3})\rangle = \frac{1}{\sqrt{3}}|223\rangle+\frac{1}{\sqrt{3}}|232\rangle+\frac{1}{\sqrt{3}}|322\rangle\,,$

$|(3),9,1\rangle = |(3),(0,1,2)\rangle = |(\boxed{1\,2\,3},\boxed{2\,3\,3})\rangle = \frac{1}{\sqrt{3}}|233\rangle+\frac{1}{\sqrt{3}}|323\rangle+\frac{1}{\sqrt{3}}|332\rangle\,,$

$|(3),10,1\rangle = |(3),(0,0,3)\rangle = |(\boxed{1\,2\,3},\boxed{3\,3\,3})\rangle = |333\rangle$

$\lambda = (2,1), p = 1:$

$$|(2,1),1,1\rangle = |(2,1),(2,1,0)\rangle = \left|\left(\young(13,2),\young(11,2)\right)\right\rangle = \quad \frac{2}{\sqrt{6}}|112\rangle - \frac{1}{\sqrt{6}}|121\rangle - \frac{1}{\sqrt{6}}|211\rangle,$$

$$|(2,1),2,1\rangle = |(2,1),(2,0,1)\rangle = \left|\left(\young(13,2),\young(11,3)\right)\right\rangle = \quad \frac{2}{\sqrt{6}}|113\rangle - \frac{1}{\sqrt{6}}|131\rangle - \frac{1}{\sqrt{6}}|311\rangle,$$

$$|(2,1),3,1\rangle = |(2,1),(1,2,0)\rangle = \left|\left(\young(13,2),\young(12,2)\right)\right\rangle = \quad -\frac{2}{\sqrt{6}}|221\rangle + \frac{1}{\sqrt{6}}|212\rangle + \frac{1}{\sqrt{6}}|122\rangle,$$

$$|(2,1),4,1\rangle = |(2,1),(0,2,1)\rangle = \left|\left(\young(13,2),\young(22,3)\right)\right\rangle = \quad -\frac{2}{\sqrt{6}}|223\rangle + \frac{1}{\sqrt{6}}|232\rangle + \frac{1}{\sqrt{6}}|322\rangle,$$

$$|(2,1),5,1\rangle = |(2,1),(1,0,2)\rangle = \left|\left(\young(13,2),\young(13,3)\right)\right\rangle = \quad -\frac{2}{\sqrt{6}}|331\rangle + \frac{1}{\sqrt{6}}|133\rangle + \frac{1}{\sqrt{6}}|313\rangle,$$

$$|(2,1),6,1\rangle = |(2,1),(0,1,2)\rangle = \left|\left(\young(13,2),\young(23,3)\right)\right\rangle = \quad \frac{2}{\sqrt{6}}|233\rangle - \frac{1}{\sqrt{6}}|323\rangle - \frac{1}{\sqrt{6}}|332\rangle,$$

$$|(2,1),7,1\rangle = |(2,1),(1,1,1)\rangle = \left|\left(\young(13,2),\young(13,2)\right)\right\rangle = \quad -\frac{1}{\sqrt{12}}|123\rangle + \frac{2}{\sqrt{12}}|132\rangle - \frac{1}{\sqrt{12}}|213\rangle$$
$$-\frac{1}{\sqrt{12}}|231\rangle + \frac{2}{\sqrt{12}}|312\rangle - \frac{1}{\sqrt{12}}|321\rangle,$$

$$|(2,1),8,1\rangle = |(2,1),(1,1,1),2\rangle = \left|\left(\young(13,2),\young(12,3)\right)\right\rangle = \quad \frac{1}{\sqrt{4}}|123\rangle + \frac{1}{\sqrt{4}}|213\rangle$$
$$-\frac{1}{\sqrt{4}}|231\rangle - \frac{1}{\sqrt{4}}|321\rangle$$

$\lambda = (2,1), p = 2:$

$$|(2,1),1,2\rangle = |(2,1),2,(2,1,0)\rangle = \left|\left(\young(12,3),\young(11,2)\right)\right\rangle = \quad \frac{1}{\sqrt{2}}|121\rangle - \frac{1}{\sqrt{2}}|211\rangle,$$

$$|(2,1),2,2\rangle = |(2,1),2,(2,0,1)\rangle = \left|\left(\young(12,3),\young(11,3)\right)\right\rangle = \quad \frac{1}{\sqrt{2}}|131\rangle - \frac{1}{\sqrt{2}}|311\rangle,$$

$$|(2,1),3,2\rangle = |(2,1),2,(1,2,0)\rangle = \left|\left(\young(12,3),\young(12,2)\right)\right\rangle = \quad \frac{1}{\sqrt{2}}|122\rangle - \frac{1}{\sqrt{2}}|212\rangle,$$

$$|(2,1),4,2\rangle = |(2,1),2,(0,2,1)\rangle = \left|\left(\young(12,3),\young(22,3)\right)\right\rangle = \quad \frac{1}{\sqrt{2}}|322\rangle - \frac{1}{\sqrt{2}}|232\rangle,$$

$$|(2,1),5,2\rangle = |(2,1),2,(1,0,2)\rangle = \left|\left(\young(12,3),\young(13,3)\right)\right\rangle = \quad \frac{1}{\sqrt{2}}|133\rangle - \frac{1}{\sqrt{2}}|313\rangle,$$

$$|(2,1),6,2\rangle = |(2,1),2,(0,1,2)\rangle = \left|\left(\young(12,3),\young(23,3)\right)\right\rangle = \quad \frac{1}{\sqrt{2}}|323\rangle - \frac{1}{\sqrt{2}}|233\rangle,$$

$$|(2,1),7,2\rangle = |(2,1),2,(1,1,1)\rangle = \left|\left(\young(12,3),\young(13,2)\right)\right\rangle = \quad \frac{1}{\sqrt{4}}|123\rangle - \frac{1}{\sqrt{4}}|213\rangle$$
$$-\frac{1}{\sqrt{4}}|231\rangle + \frac{1}{\sqrt{4}}|321\rangle,$$

$$|(2,1),8,2\rangle = |(2,1),2,(1,1,1),2\rangle = \left|\left(\young(12,3),\young(12,3)\right)\right\rangle = \frac{1}{\sqrt{12}}|123\rangle + \frac{2}{\sqrt{12}}|132\rangle - \frac{1}{\sqrt{12}}|213\rangle$$
$$+\frac{1}{\sqrt{12}}|231\rangle - \frac{2}{\sqrt{12}}|312\rangle - \frac{1}{\sqrt{12}}|321\rangle$$

$$\lambda = (1^3):$$

$$|(1^3), 1, 1\rangle = |(1^3), (1, 1, 1)\rangle = \left| \left(\boxed{\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}}, \boxed{\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}} \right) \right\rangle = \frac{1}{\sqrt{6}} |123\rangle - \frac{1}{\sqrt{6}} |132\rangle - \frac{1}{\sqrt{3}} |213\rangle$$

$$+ \frac{1}{\sqrt{6}} |231\rangle - \frac{1}{\sqrt{6}} |321\rangle + \frac{1}{\sqrt{3}} |312\rangle$$

.

3.4.3 Scattering probabilities of three distinguishable photons in three modes

We find that the scattering probabilities of certain occupations in an irrep are linked with each other. Let the input be the coincident state of three photons in three modes as mentioned before. Assume we are looking for an interferometer U such that the scattering probability of occupation number $\underline{n}' = (2, 1, 0)$ in the irrep $\lambda = (3)$ is zero. We setup a optimization procedure that minimizes this $P^{(3)}_{(2,1,0)}$ (in a similar manner as mentioned in Section 4.2.2). An interesting feature we notice, is that if $P^{(3)}_{(2,1,0)}$ reaches 0, the scattering probability $P^{(3)}_{(0,1,2)}$ also reaches 0. Similarly all the optimization results where $P^{(3)}_{(1,2,0)}$ is minimized to 0, the scattering probability $P^{(3)}_{(1,0,2)}$, again, is also found to be 0. Finally, $P^{(3)}_{(2,0,1)}$ and $P^{(3)}_{(0,2,1)}$ are matched up in a similar fashion.

If we now take a look at the irrep $\lambda = (2, 1)$, the match up between these scattering probabilities is even greater. More specifically, they are the same values

$$P^{(2,1)}_{(2,1,0)} = P^{(2,1)}_{(0,1,2)} \tag{3.178}$$

$$P^{(2,1)}_{(1,2,0)} = P^{(2,1)}_{(1,0,2)} \tag{3.179}$$

$$P^{(2,1)}_{(2,0,1)} = P^{(2,1)}_{(0,2,1)}. \tag{3.180}$$

To see these scattering probabilties match up, it is necessary to first parameterize the scattering matrix.

These are likely side-effects of the way our basis is constructed (the subgroup chain construction of Schur-Weyl basis is not within the scope of this thesis, as we mentioned before, we do not keep track of the subgroups in our states, but this is possibly related). If we look at Figure 3.1, we can see that the number occupations that are matched together can be found by applying the same raising/lowering operator twice. This could possibly be generalized to higher numbers of photons and modes, simplifying the calculation for the scattering probabilities by irrep.

In the following Chapter, we will see how the optimization of these probabilities and their dependence on each other create bounds on the discrimination of the states of interest mentioned here.

3.5   Conclusion

In this Chapter we discuss a fundamental concept of quantum information – distinguishability of particles, focusing mostly on bosons but with straightforward generalization to fermions. We present a somewhat unconventional framework for thinking about distinguishability of bosons using representation theory and first quantization (as mentioned in the literature section there have been previous attempts to understanding the effects of distinguishability using representation theory). Although we are not the first, we do offer a detailed account of how to construct and interpret distinguishability in first quantization using the Schur-Weyl basis. More importantly we use the unitary-unitary duality between the two degrees of freedom (System and Label) which has rarely been exploited and significantly simplifies the construction of distinguishable states, immediately providing new insight (see Section 3.2.3, Section 3.3, and Section 4.3).

We introduce a new notation for Schur-Weyl basis (see Section 3.1.2) which allows us to identify weight and weight spaces, commonly used in standard representation of Lie groups, with occupation numbers and their subspaces, deepening the relationship between these two fields. The weights have been explored previously in quantum information (usually linked with the concept of "types" from information theory). However, we are unaware of any previous investigation of the connection between occupation numbers, especially in linear optical systems, and weights. Further work could possibly be looking deeper into this link, similar to how the theory of types, when expanded to quantum information and linked to representation theory, offered insight into entanglement concentration.

In Section 3.3, we examine some states of interest in this new framework, and then use them further in the following Chapter. We also examine the simplest interesting case in this new framework, that of three photons (from basis construction up to results on scattering probabilities). The case of three photons is further discussed later, in Chapter 5 as it can provide a starting point for further work, both examining in new light existing results and finding new results on the connection of immanants and the scattering probabilities.

DISCRIMINATING DISTINGUISHABILITY

The formalism introduced in Chapter 3 now allows us to apply regular quantum information tools on the distinguishable states of bosons. An example of a pertinent idea from quantum information is state discrimination [206–208] and we will show how this reproduces the well known HOM distinguishability test for two particles. We set up the state discrimination problem in the linear optical framework, assuming we have access to passive transformations (networks of phaseshifters and beamsplitters) and projective measurements via photon number counting detectors (Sec. 4.2.1). This restriction on the allowed measurements yields a highly nontrivial constraint on the mixed state discrimination scenario – this new problem is what we study here. In particular, the optimisation problem that results is nonlinear, as is usually the case in multiphoton interferometry [171], necessitating numerical techniques described in Sec. 4.2.2.

The chapter is structured as follows: in Sec. 4.1 we remind the reader of the discrimination tools available in quantum information; in Sec. 4.2 we consider the linear optical restriction to these discrimination tools; in Sec. 4.3 we investigate general bounds for discriminating (i) a state with a single distinguishable photon from the completely indistinguishable state, and (ii) the completely distinguishable from the completely indistinguishable state; in Sec. 4.4 we revisit HOM effect as a test of distinguishability for arbitrary states of two photons; in Sec. 4.5 we employ analytical and numerical techniques to find the optimal disciminator of three photons; in Sec. 4.6.1 we look at discrimination of singly distinguishable states with higher photon numbers up to $N = 9$; finally in Sec. 4.6.2 we look at the discrimination of completely distinguishable states with higher photon numbers of up to $N = 8$. Beyond the brief review of discrimination tools, the remaining results in this Chapter are novel.

4.1   Quantum state discrimination

A general state discrimination protocol [207, 208] consists of two parties, a source (Alice) and a detector (Bob), who agree on an ensemble of states $\{p_k, \rho_k\}$ to be discriminated. The source draws a random sample from this ensemble according to the distribution $\{p_k\}$

and sends it to the detector, whose task is to identify which state was sent as best as possible. There are different metrics for what "as best as possible" means depending on the problem in question. Before we introduce these different discrimination strategies, we will need to expand our notion of measurement. We talk about measurements in Section 1.2 and measurement in linear optics in standard formalism in Section 1.3.3 and in the representation theoretic formalism in Section 3.2.2. These measurement operators introduced so far are all *projective operators*, satisfying the four properties mentioned in Section 1.2. We now introduce *positive operator-valued measure* or POVM, which is a generalization of the measurement concept. The same properties as for projective operators stand except for the requirement for orthonormality, so

- $E_k = E_k^\dagger$.

- $\langle\psi|E_k|\psi\rangle \geq 0$ for all $|\psi\rangle$.

- $\sum_k E_k = \mathbb{1}$.

It is now up to Bob to find a measurement, given by a set of POVM elements $\{E_k\}$, that discriminates optimally between the states he might receive from Alice. Here we outline three state discrimination strategies: Minimum Error Discrimination (MED), Unambiguous Discrimination (UD), and Maximum Confidence Measurements (MCM).

4.1.1 Minimum Error Discrimination

The strategy in MED is to find a set E_k such that the chance of the wrong state being inferred from an outcome is minimized [209]. We associate each measurement outcome k to a state ρ_k from the ensemble. Then the probability of a state ρ_j with apriori probability p_j being erroneously assigned outcome $k \neq j$ is

$$P_{\text{error},j} = \sum_{k \neq j} \text{Tr}(\rho_j E_k). \tag{4.1}$$

Total error probability is just the sum over these probabilities for all the states in the ensemble, $P_{\text{error}} = \sum_j p_j P_{\text{error},j}$. Alternatively, we can maximise the expected probability of success: $\sum_k p_k \text{Tr}[\rho_k E_k]$.

There is no generalized expression for the optimal choice of POVM. However, there are certain cases where this has been resolved. For example, in the case of ensemble of two states $\{(p_1, \rho_1), (p_2, \rho_2)\}$, the Helstrom bound gives

$$P_{\text{error}} = \frac{1}{2} \left(1 - \text{Tr} |p_2\rho_2 - p_1\rho_1|\right) \tag{4.2}$$

reducing to

$$P_{\text{error}} = \frac{1}{2} \left(1 - \sqrt{1 - 4p_1 p_2 |\langle\psi_1|\psi_2\rangle|^2}\right) \tag{4.3}$$

for pure states [209]. When the priors are equal and states are pure, the choice of the POVM is simple. It is just that of projective operators placed symmetrically around the

two pure states. Other cases where some results are known when it comes to MED are special highly symmetrical cases where square-root measurement is optimal [208, 210].

In general though, the optimal POVM satisfies the following two conditions

$$\sum_k p_k \rho_k E_k - p_k \rho_k \geq 0, \quad \forall k \tag{4.4}$$

$$E_k(p_k \rho_k - p_j \rho_j)E_j = 0, \quad \forall j, k. \tag{4.5}$$

4.1.2 Unambiguous Discrimination

In the previous strategy we had as many outcomes as the states in the ensemble and we allowed for errors to be made. In the UD strategy we would like the outcome k to correspond to the state ρ_k with certainty [211–213]. This requires any outcomes that previously would have been miscategorized due to ambiguity of the measurements, to be grouped into inconclusive outcome which we will denote with 0,

$$E_0 = \mathbb{1} - \sum_{k \neq 0} E_k, \tag{4.6}$$

$$P_0 = \sum_{k \neq 0} p_k \operatorname{Tr}(\rho_k E_k). \tag{4.7}$$

We can see that if $P_0 = 0$ we can discriminate all states perfectly.

In the case of two pure states, it has been show the optimal POVM is a projection on the states orthogonal to the two given states [214]. That is, given $\{(p_1, |\psi_1\rangle), (p_2, |\psi_2\rangle)\}$, take

$$E_1 \sim |\psi_2^\perp\rangle \langle \psi_2^\perp|, \quad E_2 \sim |\psi_1^\perp\rangle \langle \psi_1^\perp|, \tag{4.8}$$

then these are the optimal operators along with $E_0 = \mathbb{1} - E_1 - E_2$. The optimum failure probability is then

$$P_0 = 2\sqrt{p_1 p_2} \cos(\langle \psi_1 | \psi_2 \rangle) \tag{4.9}$$

when it exists (for certain priors this stops being defined). Actually, as the a priori probabilities change, the optimal measurement become more biased towards UD of the state which is more probable.

In general, for any number of pure states, Chefles has shown that the states have to be linearly independent. For further results on unambigiously discriminating three pure states, n linearly independent symmetric states, and some upper and lower bounds, see the reviews [208, 210, 215, 216].

When it comes to mixed state unambiguous discrimination this is an even more open field of research. However, in [217] the authors show that given two mixed states, it is possible to UD them if their support is not exactly the same.

4.1.3   Maximum Confidence Measurements

As we mentioned in the previous section, the unambiguous discrimination of pure states is only possible if all the given states are linearly independent. If they are not, we cannot avoid making errors in identifying some of the states [215]. However, it is possible to generalize unambiguous discrimination, by defining confidence that outcome j indicates state ρ_j was the one sent. This probability is defined as a posteriori probability given outcome j

$$P(\rho_j|j) = \frac{p_j \operatorname{Tr}(\rho_j E_j)}{Tr(\rho E_j)}, \tag{4.10}$$

that can then be maximized. This strategy is called "maximum confidence measurements" (MCM) [218]. For the set $\{E_k\}_k$ to be complete, an inconclusive outcome, E_0, might have to be added again, and in this case the probability of an inconclusive outcome will usually need to be minimized. There are other strategies related to this discrimination tool, for example, perhaps the inconclusive outcome is not allowed, but instead the smallest posterior value is maximized. Another alternative is allowing an inconclusive outcome up to a probability threshold and then maximizing the posteriors.

The MCM can be easily reduced to the UD task by requiring posteriors to be equal to one. This retrieves the conditions on when the task of UD is possible (as mentioned previously, pure states need to be linearly independent or mixed state need to have distinct supports). Similarly, MED can be retrieved by maximizing the posteriors weighted by the priors.

There are other discrimination strategies as well, such as maximising mutual information between the two parties, or state discrimination for repeated measurements, however these remain out of scope for the work presented here.

4.2   Discrimination in linear optics

In the HOM experiment we find an interferometer that maximises the probability of seeing a coincidence for a distinguishable input state, subject to the constraint that it never gives coincidences for an indistinguishable input state (see Section 3.2.3). In Section 4.4.1 it will become clear that an HOM scenario is an instance of an UD problem in linear optics, with the solution being a balanced beamsplitter. We can then generalize it to any number of particles in any number of modes as a UD problem.

A key distinction from general UD is the restricted form of the available POVM elements, which must be projective measurements defined by the interferometer U and the N-photon occupation $\underline{n}$ being detected. In particular, we expect that known optimal measurements for two-state discrimination will not be available in linear optics. When speaking generally about measurements we will use the notation E for POVM elements, while, as defined in Section 3.2.2, $M_{\underline{n}}(U)$ is reserved for photon counts. Because $M_{\underline{n}}(U)$ is a polynomial of degree N in the variables U and $U^\dagger$, this measurement restriction makes

the UD optimisation problem nonlinear.

We will be interested in two problems

- discriminating the completely indistinguishable state, ρ_i (see Section 3.3.1), from the distinguishable state ρ_s (see Section 3.3.2) and

- discriminating the completely indistinguishable state, ρ_i, from the distinguishable state ρ_d (see Section 3.3.3).

From Eqs. (3.117, 3.125, 3.140), we observe that each of these states is of the form

$$\rho = \alpha\rho_i + (1 - \alpha)\rho_{\bar{i}}, \quad \alpha \neq 0, \tag{4.11}$$

where ρ_i is pure, and $\rho_{\bar{i}}$ is diagonal in the Schur-Weyl basis with support outside the symmetric subspace $\lambda = (N)$. From well known results for the discrimination of two mixed states [219], the fact that ρ_i lies within the support of the mixed state to be discriminated means that the optimal measurement is essentially the same for either Minimum Error or Unambiguous Discrimination; one wishes to project onto the support of $\rho_{\bar{i}}$. In particular for UD, the error-free constraint means that we are forced to set

$$E_i = 0, \tag{4.12}$$

and thus the prior probabilities do not affect the optimal choice of measurement operators. This reflects the fact that there is no way to unambiguously discriminate the indistinguishable state ρ_i – we can either conclude that the state was distinguishable by observing an output that is completely suppressed by quantum interference, or fail to conclude anything at all. Our task is therefore to minimise the probability of failure

$$E_0 = \mathbb{1} - E_{s,d}, \tag{4.13}$$

equivalently maximising the probability of unambiguously detecting a singly or completely distinguishable state, respectively.

If our measurements are unrestricted, the best choice of POVM is to project onto the nonsymmetric subspace. This choice is suitable for not only the states $\rho_{s,d}$, but by extension any state to be discriminated from ρ_i. However, as mentioned in Sec. 3.2.2, in practice we only have access to number counting measurements – we will therefore want to approximate this projection as best possible. The approximation will be sensitive to the state we are discriminating: for example, Eqs. (3.125, 3.140) show that ρ_s can be optimally discriminated by projecting onto only the $(N-1, 1)$ irrep, while for ρ_d one wants to project on to all of the nonsymmetric irreps. As we will see, this can lead to different interferometers being optimal for discriminating different distinguishable states. We will parametrise these unitaries with a Reck scheme as described in Section 1.3.2, decomposing an arbitrary U into a sequence of single mode unitaries (phaseshifters) and unitaries that act on neighbouring modes (beamsplitters). As shown in Fig. 1.1, such a scheme can be

viewed as $d_{\mathrm{s}} - 1$ layers, indexed by k, each with k phaseshifters and beamsplitters, followed by a final phase shift on each mode. Because we are only interested in number state inputs and number counting measurements, only the phaseshifters between beamsplitters play a role. Hereafter when we refer to U we will therefore be referring to this smaller interferometer, without the initial and final sets of phaseshifters.

4.2.1 Restriction to linear optical measurements

In order to discriminate distinguishability in linear optics we wish to find the best we can do with regards to a suitable metric with the measurements we have, namely those in Eq. (3.93). There are many ways we can approach this problem. One way would be, given a specific occupation $\underline{n}$, to find U

$$\underset{U}{\operatorname{argmax}} \operatorname{Tr}\left[\rho M_{\underline{n}}(U)\right], \quad \text{such that } \operatorname{Tr}\left[\rho_{\mathrm{i}} M_{\underline{n}}(U)\right] = 0. \tag{4.14}$$

Notice that any $\underline{n}$ that can be made to satisfy $\operatorname{Tr}\left[\rho_{\mathrm{i}} M_{\underline{n}}(U)\right] = 0$ for a suitable U is an unambiguous discriminator. In general, it is possible for multiple occupations to satisfy the UD constraint simultaneously, each contributing to the total probability of successful discrimination.

We therefore consider a different optimization, where we wish to find the subset of all discriminating occupations, call it D, that optimises the success probability simultaneously, for the same choice of U:

$$\text{find } U \text{ and } D \text{ maximising} \quad \sum_{\underline{n} \in D} \operatorname{Tr}\left[\rho M_{\underline{n}}(U)\right] \tag{4.15}$$

$$\text{subject to, for all } \underline{n} \in D, \quad \operatorname{Tr}\left[\rho_{\mathrm{i}} M_{\underline{n}}(U)\right] = 0. \tag{4.16}$$

Note that the quantity we are maximizing gives us the total probability of successful discrimination, which is the sum over all the unambiguously discriminating events in the set of occupations D.

While the former choice of optimisation focuses on giving an optimal interferometer for discrimination given a specific measurement occupation, the latter optimisation focuses on the highest probability of discrimination across all measurement patterns. In general we find that these two problems give different optimal interferometers; here we will focus on the latter optimisation over both U and D.

We observe that not all occupations are useful for unambiguous discrimination. Measurements where all the photons are bunched into a single mode only occur in the symmetric irrep, that is, if $\underline{n} = (0, .., 0, N, 0, ..., 0)$, then $M_{\underline{n}} = |(N), 1, \underline{n}, 1\rangle \langle (N), 1, \underline{n}, 1|$. In this case Eqs. (3.142) and (3.144) are proportional to Eq. (3.145), and since Eq. (4.16) has to be satisfied, they will always give zero. Completely bunched events can therefore never help unambiguously discriminate the indistinguishable state, and we will exclude such events from our searches.

4.2.2 Numerical optimisation approach

We present a mixture of analytical and numerical results. To construct the cost function for our numerical work we took into consideration the following criteria:

- The measurement operator $M_{\underline{n}}$ can only be included in the optimisation if Eq. (4.16) is satisfied

- When previous point is satisfied, $\underline{n}$ is added to a sum being optimised as per Eq. (4.15).

The cost function chosen was

$$C(U) = -\sum_{\underline{n}} \exp\left(-\xi \operatorname{Tr}\left[\rho_i M_{\underline{n}}(U)\right]\right) \operatorname{Tr}\left[\rho M_{\underline{n}}(U)\right], \tag{4.17}$$

where ξ is adjusted (usually depending on the choice of N, and ranging from 2 to 60) to penalise results where $M_{\underline{n}}$ might be added to Eq. (4.15) and optimised without satisfying Eq. (4.16). A high penalty ξ guarantees that the value of $\operatorname{Tr}\left[\rho_i M_{\underline{n}}(U)\right]$ is close to zero before $\operatorname{Tr}\left[\rho M_{\underline{n}}(U)\right]$ is optimised and added to the sum. Combining this with the Eqs. (3.142) and (3.144) we have

$$C_{\mathrm{d}}(U) = \frac{-1}{N!} \sum_{\lambda \neq (N)} d_{(\lambda)} \sum_{\underline{n}} e^{-\xi|\langle (N),\underline{n}|U^{(N)}|(N),\underline{1}\rangle|^2} \sum_{r,r'} |\langle \lambda, \underline{n}, r| U^{\lambda} |\lambda, \underline{1}, r'\rangle|^2, \tag{4.18}$$

$$C_{\mathrm{s}}(U) = \frac{1-N}{N} \sum_{\underline{n},r} e^{-\xi|\langle (N),\underline{n}|U^{(N)}|(N),\underline{1}\rangle|^2} |\langle (N-1,1),\underline{n}, r| U^{(N-1,1)} |(N-1,1),\underline{1}\rangle|^2. \tag{4.19}$$

Python was used to optimise these functions with the scipy library function `basinhopping` using Broyden–Fletcher–Goldfarb–Shanno (BFGS) as the optimisation algorithm. The seeds were generated using numpy random number generation. Though this optimisation function will help us explore the space, it neither guarantees that the minimum is global, nor does it exactly solve the original optimisation problem. This will be problematic with minima that are close together, as for example $\exp\left(-\xi \operatorname{Tr}\left[\rho_i M_{\underline{n}}(U)\right]\right)$ gets closer to 1 for values of $\operatorname{Tr}\left[\rho_i M_{\underline{n}}(U)\right]$ that are close to 0. In some situations this value can be quite high when combined with a high value of $\operatorname{Tr}\left[\rho M_{\underline{n}}(U)\right]$, skewing the results towards a possible non-optimal solution for the original problem. We could avoid this by choosing an appropriately high ξ as a function of the number of occupations $\binom{N+d_{\mathrm{S}}-1}{N}$, however, if too high, $\exp\left(-\xi \operatorname{Tr}\left[\rho_i M_{\underline{n}}(U)\right]\right)$ will behave like a step function, which does not reward transitional values enough. Therefore, we do not make any strong claims of optimality for the interferometers found numerically when they do not saturate the general bounds presented in Sec. 4.3.

4.3 General bounds

Recall from Sec. 4.2.1 the best possible unrestricted discrimination measurement is to project onto the nonsymmetric subspace,

$$E_{\overline{(N)}} = \sum_{\lambda \neq (N), p, \underline{n}, r} |\lambda p \underline{n} r\rangle \langle \lambda p \underline{n} r| . \tag{4.20}$$

Such a POVM element would be equally good for both singly and completely distinguishable states, and indeed any distinguishable state of the form in Eq. (4.11). The success probability of such a measurement is given by

$$
\begin{aligned}
\mathrm{Tr} & \left[\rho \left(\oplus_\lambda U^\lambda \otimes \mathbb{1}^\lambda \right)^\dagger E_{\overline{(N)}} \left(\oplus_\lambda U^\lambda \otimes \mathbb{1}^\lambda \right) \right] \\
&= \mathrm{Tr} \left[(\alpha \rho_\mathrm{i} + (1 - \alpha) \rho_{\bar{\mathrm{i}}}) E_{\overline{(N)}} \right] \\
&= 1 - \alpha \\
&= \begin{cases} 1 - \frac{1}{N} & \text{if } \rho = \rho_\mathrm{s} \\ 1 - \frac{1}{N!} & \text{if } \rho = \rho_\mathrm{d}, \end{cases}
\end{aligned}
\tag{4.21}
$$

where we have used the fact that any projector onto irreps is unitarily invariant. These then are universal upper bounds on the success probability for singly and completely distinguishable states, respectively. However, since we are restricted to photon number counting measurements, we will see that while the first bound is achievable, the second is not in general. We will go through various examples in detail in the following sections.

Moreover from Section 3.3.4, we can find a bound of $1 - \underline{n}_\mathrm{L}!/N!$ on the probability for successfully discriminating a general state from the completely indistinguishable one, and this includes the singly and completely distinguishable cases above.

4.4 Two modes

4.4.1 Two photons in two modes

Recall from Eq. (3.115) that

$$
\rho_\mathrm{i} = \begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 0 \end{bmatrix}, \qquad \rho_\mathrm{s} = \rho_\mathrm{d} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 1 \end{bmatrix}, \tag{4.22}
$$

Observing that there is only one available state which is not symmetric, it is easy to write down an arbitrary *partially* distinguishable System state in this case, since there is

but one parameter:

$$\rho = \alpha \left| \boxed{1\,2} \right\rangle \left\langle \boxed{1\,2} \right| + (1-\alpha) \left| \boxed{\genfrac{}{}{0pt}{}{1}{2}} \right\rangle \left\langle \boxed{\genfrac{}{}{0pt}{}{1}{2}} \right|. \tag{4.23}$$

As discussed in Sec. 4.2.1, only occupations that do not have all the photons bunched in the same mode can be used for meaningful discrimination, in this case leaving only one choice of projector, the coincidence $M_{(1,1)} = \left| \boxed{1\,2} \right\rangle \left\langle \boxed{1\,2} \right| + \left| \boxed{\genfrac{}{}{0pt}{}{1}{2}} \right\rangle \left\langle \boxed{\genfrac{}{}{0pt}{}{1}{2}} \right|$.

A coincidence count occurs when both the top and bottom modes are occupied, defining the coincidence subspace spanned by $\left\{ \left| \boxed{1\,2} \right\rangle, \left| \boxed{\genfrac{}{}{0pt}{}{1}{2}} \right\rangle \right\}$. The projector onto this subspace has matrix representation

$$M_{(1,1)} = \begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 1 \end{bmatrix}, \tag{4.24}$$

where we have used an occupation (one excitation in each of the two System modes) in the subscript.

The matrix elements in the coincident subspace for an arbitrary two mode interferometer with transfer matrix U are

$$U^{\otimes 2} \cong U^{\boxplus} \oplus U^{\boxminus} = \begin{bmatrix} * & * & * & \\ * & \mathrm{per}U & * & \\ * & * & * & \\ & & & \det U \end{bmatrix}, \tag{4.25}$$

where per and det are the matrix permanent and determinant functions, $*$ are matrix elements for events outside the coincident subspace, and we can see an idea on how to fill in the missing values in Section 5.1.

The probability of a coincidence count is given by the Born rule, which from Eqs. (4.22,4.25) is given by

$$P_{(1,1)} = \mathrm{Tr}\left[\left(U^{\boxplus} \oplus U^{\boxminus} \right) \rho \left(U^{\boxplus} \oplus U^{\boxminus} \right)^{\dagger} M_{(1,1)} \right] \tag{4.26}$$

$$= \mathrm{Tr}\left[\left(U^{\boxplus} \rho U^{\boxplus\dagger} + U^{\boxminus} \rho U^{\boxminus\dagger} \right) M_{(1,1)} \right] \tag{4.27}$$

$$= \begin{cases} |\mathrm{per}U|^2 & \text{if } \rho = \rho_{\mathrm{i}} \\ |\alpha||\mathrm{per}U|^2 + |1-\alpha||\det U|^2 & \text{if } \rho = \rho_{\mathrm{d}} \end{cases} \tag{4.28}$$

where we have written $|U|^2$ for the elementwise absolute value squared of a matrix U. It follows that in order to see no coincidences for an indistinguishable state, which has only a triplet component, we need an interferometer whose transfer matrix permanent vanishes. We can recall from Section 3.2.3 what we expect $U^{\boxplus} \oplus U^{\boxminus}$ to look like for the specific choice of U to be balanced beamsplitter. Also, we find that for the balanced beamsplitter,

$P_{(1,1)}(\mathrm{BS}_{50:50}) = P^{\boxminus}_{(1,1)}(\mathrm{BS}_{50:50}) + P^{\boxminus}_{(1,1)}(\mathrm{BS}_{50:50}) = P^{\boxminus}_{(1,1)}(\mathrm{BS}_{50:50}) + 0.5 \geq 0.5$. This is to be expected as $\mathrm{perBS}_{50:50} = 0$, and the results are consistent taking $\alpha = 0.5$.

Rearranging Eq. (4.26) and defining

$$M_{(1,1)}(U) = \left(U^{\boxminus} \oplus U^{\boxminus}\right)^{\dagger} M_{(1,1)} \left(U^{\boxminus} \oplus U^{\boxminus}\right), \tag{4.29}$$

the HOM measurement scenario described above can now be summarised by

$$\text{find } U \text{ maximising} \quad \mathrm{Tr}\left[\rho_{\mathrm{d}} M_{(1,1)}(U)\right] \tag{4.30}$$

$$\text{subject to} \quad \mathrm{Tr}\left[\rho_{\mathrm{i}} M_{(1,1)}(U)\right] = 0. \tag{4.31}$$

We see that the idea of generalizing HOM using UD as described in Section 4.4.1 is well motivated.

First, note that since there is only one antisymmetric state, the antisymmetric irreducible representation of any U has but one matrix element and so the action of any interferometer on this state is trivial (in Eq. (4.25) given by its determinant). Thus the only contribution to the non-symmetric part of Eq. (3.142) is $\left|\left\langle \boxed{\frac{1}{2}} \middle| U^{\boxminus} \middle| \boxed{\frac{1}{2}} \right\rangle\right| = 1$, and there is nothing to maximise in Eq. (4.30). All that is left is to satisfy the constraint, Eq. (4.31). Parametrising U as

$$\begin{bmatrix} e^{i\phi}\cos\theta & e^{i\varphi}\sin\theta \\ -e^{-i\varphi}\sin\theta & e^{-i\phi}\cos\theta \end{bmatrix}. \tag{4.32}$$

one finds that the constraint is then $\mathrm{per}U = \cos^2\theta - \sin^2\theta = \cos 2\theta = 0$, with the family of solutions $\{(\phi, \varphi, \pi/4)\,|\,0 \leq \phi \leq \pi, 0 \leq \varphi \leq \pi\}$. The solutions do not depend on the phases ϕ or φ, as we would expect from the discussion in Sec. 4.2, but only on the choice of the beamsplitter reflectivity, which is balanced as claimed.

We see that not only does unambiguous discrimination return the HOM measurement as was discussed, it is optimal for an arbitrary partially distinguishable two photon state.

4.4.2 Three photons in two modes

As an example of the utility of the formalism, in this subsection we will make a slight digression and consider the simplest nontrivial case with $N(= 3) > d_{\mathrm{S}}(= 2)$. As mentioned in Sec. 3.3, this restricts the kinds of distinguishable states that can occur; we consider situations with two photons in one System mode and the third in the other. The indistinguishable state is $\hat{a}^{\dagger}_{11}\hat{a}^{\dagger}_{11}\hat{a}^{\dagger}_{21}\,|\mathrm{vac}\rangle = \left|{}^2_1\right\rangle$, with reduced state

$$\rho_{\mathrm{i}} = \left|\boxed{1\,|\,1\,|\,2}\right\rangle \left\langle\boxed{1\,|\,1\,|\,2}\right|. \tag{4.33}$$

There are essentially two types of distinguishable state in this situation. The first is $\hat{a}^{\dagger}_{11}\hat{a}^{\dagger}_{11}\hat{a}^{\dagger}_{22}\,|\mathrm{vac}\rangle = \left|{}^2_0\,{}^0_1\right\rangle$, and the second $\hat{a}^{\dagger}_{11}\hat{a}^{\dagger}_{12}\hat{a}^{\dagger}_{21}\,|\mathrm{vac}\rangle = \left|{}^1_1\,{}^1_0\right\rangle$. Other states are equivalent to the above for the reasons discussed in Sec. 3.3.2. Further, the (now incompletely) distinguishable state $\hat{a}^{\dagger}_{11}\hat{a}^{\dagger}_{12}\hat{a}^{\dagger}_{23}\,|\mathrm{vac}\rangle = \left|{}^1_0\,{}^1_0\,{}^0_1\right\rangle$ has a reduced state that is the same as Eq. (4.34), and will therefore have the same discrimination measurement and success

probability. The reduced state for the first case is

$$\rho_{s_1} = \frac{1}{3} \left| \boxed{1\,1\,2} \right\rangle \left\langle \boxed{1\,1\,2} \right| + \frac{1}{3} \left| \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_1 \right\rangle \left\langle \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_1 \right| + \frac{1}{3} \left| \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_2 \right\rangle \left\langle \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_2 \right|, \tag{4.34}$$

while that for the second case is

$$\rho_{s_2} = \frac{4}{6} \left| \boxed{1\,1\,2} \right\rangle \left\langle \boxed{1\,1\,2} \right| + \frac{1}{6} \left| \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_1 \right\rangle \left\langle \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_1 \right| + \frac{1}{6} \left| \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_2 \right\rangle \left\langle \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_2 \right|. \tag{4.35}$$

Note that Eq. (4.15) does not depend on the amplitude of the symmetric part of the state – its contribution has to be zero by Eq. (4.16). It only depends on the nonsymmetric components, and since ρ_{s_1} and ρ_{s_2} are equally weighted across the available nonsymmetric states, the optimal discriminator will be the same. However ρ_{s_2} does have half of the amplitude of ρ_{s_1} in this subspace, which will halve the success probability.

There are four possible occupations to measure, however as mentioned in Sec. 3.3.5 the bunched ones can be disregarded and the optimisation carried out on $M_{(2,1)}$ and $M_{(1,2)}$. We parametrise U again as in Eq. (4.32). For $M_{(2,1)}$ Eq. (4.16) reduces to

$$\left| \left\langle \boxed{1\,1\,2} \right| U^{\boxed{}} \left| \boxed{1\,1\,2} \right\rangle \right| = \left| (\cos\theta + 3\cos 3\theta)/4 \right| = 0. \tag{4.36}$$

Since $0 \le \theta \le \pi$, this equation is true for

$$\theta \in \{\pi/2, \arccos(\sqrt{2/3}), \arccos(-\sqrt{2/3})\}. \tag{4.37}$$

On the other hand, Eq. (4.16) for $M_{(1,2)}$ is

$$\left| \left\langle \boxed{1\,2\,2} \right| U^{\boxed{}} \left| \boxed{1\,1\,2} \right\rangle \right| = \left| (\sin\theta - 3\cos 3\theta)/4 \right|. \tag{4.38}$$

This equation cannot be zero for the above choice of angles that ensure $\left| \left\langle \boxed{1\,1\,2} \right| U^{\boxed{}} \left| \boxed{1\,1\,2} \right\rangle \right| = 0$. Thus, only one of the outcomes can be used to discriminate these states; without loss of generality, we choose to optimise for $M_{(2,1)}$. In this case we want to maximise

$$\mathrm{Tr}\left[\rho_{s_1} M_{(2,1)}(U) \right] = 2 \left| \left\langle \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_1 \right| U^{\boxed{}} \left| \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}_1 \right\rangle \right|^2 = 2\cos^2\theta/3. \tag{4.39}$$

When $\theta = \pi/2$, we get success probability of 0. When $\theta = \pm\arccos(\sqrt{2/3})$, we get success probability of 4/9. Thus an optimal discriminating interferometer is

$$U = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{2} & 1 \\ -1 & \sqrt{2} \end{bmatrix}, \tag{4.40}$$

with success probabilities 4/9 for ρ_{s_1} and 2/9 for ρ_{s_2}.

4.5 Three modes

From now on we will only consider coincident input with $N = d_{\mathrm{S}}$. For three photons in three System modes, the completely indistinguishable reduced state, is from Eq. (3.117),

$$\rho_{\mathrm{i}} = |(3), \underline{1}\rangle \langle (3), \underline{1}| = |\boxed{1\,2\,3}\rangle \langle \boxed{1\,2\,3}|. \tag{4.41}$$

There are now three different singly distinguishable states, depending on which System mode the 'bad' photon is in. In the Schur-Weyl basis (see Sec. 3.4.2) their full System-Label states, as per the discussion in Sec. 3.3.2, are

$$\sqrt{3}\,\hat{a}_{11}^{\dagger}\hat{a}_{21}^{\dagger}\hat{a}_{32}^{\dagger}|\mathrm{vac}\rangle = \sqrt{3}\left|\begin{smallmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{smallmatrix}\right\rangle$$
$$= |\boxed{1\,2\,3}\rangle\,|\boxed{1\,1\,2}\rangle$$
$$+ \left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_1\right\rangle \left|\boxed{\begin{smallmatrix}1&1\\2\end{smallmatrix}}_1\right\rangle$$
$$+ \left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_2\right\rangle \left|\boxed{\begin{smallmatrix}1&1\\2\end{smallmatrix}}_2\right\rangle, \tag{4.42}$$

$$\sqrt{3}\,\hat{a}_{11}^{\dagger}\hat{a}_{22}^{\dagger}\hat{a}_{31}^{\dagger}|\mathrm{vac}\rangle = \sqrt{3}\left|\begin{smallmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{smallmatrix}\right\rangle$$
$$= |\boxed{1\,2\,3}\rangle\,|\boxed{1\,1\,2}\rangle$$
$$- \frac{1}{2}\left(\left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_1\right\rangle + \sqrt{3}\left|\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_1\right\rangle\right)\left|\boxed{\begin{smallmatrix}1&1\\2\end{smallmatrix}}_1\right\rangle$$
$$- \frac{1}{2}\left(\left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_2\right\rangle + \sqrt{3}\left|\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_2\right\rangle\right)\left|\boxed{\begin{smallmatrix}1&1\\2\end{smallmatrix}}_2\right\rangle, \tag{4.43}$$

$$\sqrt{3}\,\hat{a}_{12}^{\dagger}\hat{a}_{21}^{\dagger}\hat{a}_{31}^{\dagger}|\mathrm{vac}\rangle = \sqrt{3}\left|\begin{smallmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{smallmatrix}\right\rangle$$
$$= |\boxed{1\,2\,3}\rangle\,|\boxed{1\,1\,2}\rangle$$
$$- \frac{1}{2}\left(\left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_1\right\rangle - \sqrt{3}\left|\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_1\right\rangle\right)\left|\boxed{\begin{smallmatrix}1&1\\2\end{smallmatrix}}_1\right\rangle$$
$$- \frac{1}{2}\left(\left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_2\right\rangle - \sqrt{3}\left|\boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_2\right\rangle\right)\left|\boxed{\begin{smallmatrix}1&1\\2\end{smallmatrix}}_2\right\rangle. \tag{4.44}$$

While for completely distinguishable states permuting System modes has no effect on the reduced state, here the reduced states will not be invariant. However, because permutations of System modes lie inside the set of allowed operations, (that is, $\mathrm{S}_{d_{\mathrm{S}}} \subset \mathrm{U}(d_{\mathrm{S}})$), if we optimise for one of these states, the resulting interferometer will be easily related to the others by including some mode swapping. Therefore we can focus on one of these states and the success probabilities that we find will be the same for the other two; Eq. (4.42) has the reduced state (cf. Eq. (3.125))

$$\rho_{\mathrm{s}} = \frac{1}{3}|\boxed{1\,2\,3}\rangle \langle \boxed{1\,2\,3}|$$
$$+ \frac{1}{3}\left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_1\right\rangle \left\langle\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_1\right| + \frac{1}{3}\left|\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_2\right\rangle \left\langle\boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_2\right|. \tag{4.45}$$

It is also natural to ask about discrimination of mixtures of the three states in Eqs. (4.42, 4.43, 4.44); we will discuss this in Sec. 4.5.2.

The completely distinguishable state corresponding to $\hat{a}_{11}^{\dagger}\hat{a}_{22}^{\dagger}\hat{a}_{33}^{\dagger}|\mathrm{vac}\rangle = \left|\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right\rangle$ per

Eq. (3.140) is

$$
\begin{aligned}
\rho_{\mathrm{d}} = {} & \frac{1}{6} \left| \boxed{1\,2\,3} \right\rangle \left\langle \boxed{1\,2\,3} \right| + \frac{1}{6} \left| \boxed{\begin{smallmatrix}1\\2\\3\end{smallmatrix}} \right\rangle \left\langle \boxed{\begin{smallmatrix}1\\2\\3\end{smallmatrix}} \right| \\
& + \frac{1}{6} \left| \boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_{1} \right\rangle \left\langle \boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_{1} \right| + \frac{1}{6} \left| \boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_{2} \right\rangle \left\langle \boxed{\begin{smallmatrix}1&3\\2\end{smallmatrix}}_{2} \right| \\
& + \frac{1}{6} \left| \boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_{1} \right\rangle \left\langle \boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_{1} \right| + \frac{1}{6} \left| \boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_{2} \right\rangle \left\langle \boxed{\begin{smallmatrix}1&2\\3\end{smallmatrix}}_{2} \right| .
\end{aligned}
\tag{4.46}
$$

For the following let us define two sets of measurement operators: those with two photons in one mode, $M_{(2,1,0)} = \sum_{\lambda \neq \boxminus, p} |\lambda, p, (2,1,0)\rangle \langle \lambda, p, (2,1,0)|$, $M_{(2,0,1)}$, $M_{(1,0,2)}$, $M_{(1,2,0)}$, $M_{(0,1,2)}$, and $M_{(0,2,1)}$, which we denote $\mathcal{M}_2$; and those with each photon in a different mode, that is $\mathcal{M}_1 \ni M_{(1,1,1)} = \sum_{\lambda, p, r} |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r|$. As discussed in Sec. 3.3.5, the measurements $M_{(3,0,0)} = |\boxed{1\,1\,1}\rangle \langle \boxed{1\,1\,1}|$, $M_{(0,3,0)}$, and $M_{(0,0,3)}$ will not be helpful for discrimination.

4.5.1 Discriminating singly distinguishable states

Let ρ^λ denote the (unnormalized) part of a state supported on the subspace of irrep λ. Notice that ρ_{s} has no support in the antisymmetric subspace, so that $\sum_{\underline{n}} \mathrm{Tr}\left[\rho_{\mathrm{s}}^{\boxplus} M_{\underline{n}}(U) \right] = 2/3$ and $\sum_{\underline{n}} \mathrm{Tr}\left[\rho_{\mathrm{s}}^{\boxminus} M_{\underline{n}}(U) \right] = 0$. This means that for any subset of occupations D and any U for which Eq. (4.16) holds, the success probability will be bounded by $2/3$. It is well known how to saturate this; use a balanced tritter, $U = QFT_3$, and all the occupations from $\mathcal{M}_2$, where QFT_N is defined as

$$
QFT_N = \frac{1}{\sqrt{N}}
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
1 & \omega^1 & \cdots & \omega^{N-1} \\
\vdots & \vdots & & \vdots \\
1 & \omega^{N-1} & \cdots & \omega^{(N-1)(N-1)}
\end{bmatrix}
\tag{4.47}
$$

and $\omega = \exp \frac{2\pi i}{N}$. A parametrisation that realizes a balanced tritter is given in Figure 4.1.
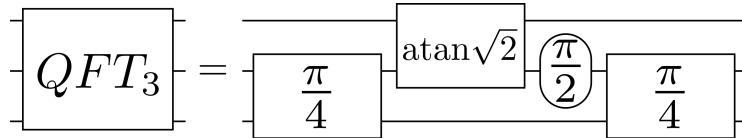
Figure 4.1: The best known interferometer for discriminating completely indistinguishable from distinguishable states of three photons in three modes is QFT_3, with a success probability of $2/3$. Up to phases, it consists of two balanced beamsplitters, one $2:1$ beamsplitter, and one $\pi/2$ phaseshifter.

4.5.2 Discriminating mixed singly distinguishable states

A short digression regarding mixed System-Label states: if we were (uniformly) ignorant about which mode the 'bad' photon was in, we would have an equal mixture of Eqs. (4.42,

4.43, 4.44). The resulting mixed state is

$$
\begin{aligned}
\rho_{\text{sm}} :=& \frac{1}{3} \left| \boxed{1\,2\,3} \right\rangle \left\langle \boxed{1\,2\,3} \right| \\
& + \frac{1}{6} \left| \begin{smallmatrix}\boxed{1\,3}\\\boxed{2}\end{smallmatrix}_1 \right\rangle \left\langle \begin{smallmatrix}\boxed{1\,3}\\\boxed{2}\end{smallmatrix}_1 \right| + \frac{1}{6} \left| \begin{smallmatrix}\boxed{1\,3}\\\boxed{2}\end{smallmatrix}_2 \right\rangle \left\langle \begin{smallmatrix}\boxed{1\,3}\\\boxed{2}\end{smallmatrix}_2 \right| \\
& + \frac{1}{6} \left| \begin{smallmatrix}\boxed{1\,2}\\\boxed{3}\end{smallmatrix}_1 \right\rangle \left\langle \begin{smallmatrix}\boxed{1\,2}\\\boxed{3}\end{smallmatrix}_1 \right| + \frac{1}{6} \left| \begin{smallmatrix}\boxed{1\,2}\\\boxed{3}\end{smallmatrix}_2 \right\rangle \left\langle \begin{smallmatrix}\boxed{1\,2}\\\boxed{3}\end{smallmatrix}_2 \right| .
\end{aligned}
\tag{4.48}
$$

The overlap $\sum_{\underline{n}} \text{Tr}\left[\rho_{\text{sm}}^{\boxplus} M_{\underline{n}}(U) \right] = 2/3$ for $\mathcal{M}_2$ again saturates the bound, and a balanced tritter remains the best choice of interferometer. This can be seen from the symmetry of the QFT which treats a 'bad' photon in any mode essentially the same way, and so should be true for analogous singly distinguishable mixed states for all N, however we will not discuss mixed System-Label states further here.

4.5.3 Discriminating completely distinguishable states

Using the cost function from Eq. (4.18) and a range of penalties $\xi \in \{2, 4, 6, 8, 10\}$ we find that the highest success probability in the completely distinguishable case is $2/3$. The measurement operators are always the full set $\mathcal{M}_2$ with a balanced tritter as a solution, just as in the previous section. However, this does not saturate the bound in Sec. 4.3, which is $5/6$ in the case of three photons. To investigate this further, we try to understand the structure of the state a bit better and use numerical evidence to show that a balanced tritter is likely to be optimal.

From Eq. (4.46) we have $\sum_{\underline{n}} \text{Tr}\left(\rho_{\text{d}}^{\boxplus} M_{\underline{n}}(U) \right) = 2/3$ and $\sum_{\underline{n}} \text{Tr}\left(\rho_{\text{d}}^{\boxminus} M_{\underline{n}}(U) \right) = 1/6$, so that

$$
\sum_{\underline{n}} \text{Tr}\left(\rho_{\text{d}}^{\boxminus\boxminus} M_{\underline{n}}(U) \right) = 5/6,
\tag{4.49}
$$

which is the discrimination bound. Notice that operators from $\mathcal{M}_2$ do not have support on the anti-symmetric subspace. Therefore, if we only pick operators from $\mathcal{M}_2$ as the discriminating operators, and assume they can simultaneously satisfy Eq. (4.16), then

$$
\sum_{\underline{n} \in \mathcal{M}_2} \text{Tr}\left(\rho_{\text{d}} M_{\underline{n}}(U) \right) = \sum_{\underline{n} \in \mathcal{M}_2} \text{Tr}\left(\rho_{\text{d}}^{\boxplus} M_{\underline{n}}(U) \right) \leq 2/3.
\tag{4.50}
$$

This is exactly what happens for the interferometers from our optimisation.

This tells us that if we want the success probability to be larger than $2/3$, the only operator left, $M_{(1,1,1)}$, would have to be included. Our numerical results show that, on the contrary, it is unlikely for any D that includes $M_{(1,1,1)}$ to give a success probability over $1/2$. We do this with a new cost function, much like Eq. (4.18) but modified to force $M_{(1,1,1)}$ to be included:

$$
C_{\text{d},111}(U) = \eta \, \text{Tr}(\rho_{\text{i}} M_{(1,1,1)}(U)) + C_{\text{d}}(U),
\tag{4.51}
$$

where η is a penalty to ensure Eq. (4.16) for $M_{(1,1,1)}$ has to be satisfied, and $C_{\text{d}}(U)$ is

as defined in Eq. (4.17). This penalty is set to $\eta = 10$ making the first term an order of magnitude higher than the second term of Eq. (4.51), where we took $\xi = 6$. As we learned in Sec. 4.2, we can ignore the outside phaseshifters of the standard Reck parametrisation, therefore we are only optimizing over 4 parameters, $\theta_{2,1}, \theta_{2,2}, \theta_{1,2}$, and $\omega_{1,2}$. The lowest value of the cost function found by the optimisation techniques in Sec. 4.2.2 is -0.500426. This corresponds to a success probability of 0.5 in discriminating between the two states, which is lower than the $2/3$ achievable when $M_{(1,1,1)}$ is not included.

While this does not give us definitive proof that no scheme that includes a threefold coincidence can give success probability higher than $2/3$, it does strongly indicate that this should be true. Moreover, with the same optimisation functions we investigated how many of the other operators alongside $M_{(1,1,1)}$ we can pick at the same time, and it seems that the best we can do is to have four from $\mathcal{M}_2$ satisfy Eq. (4.16) simultaneously. However, in all the situations when this occurs, some of the terms in Eq. (4.15) are zero, thus the success probability remains at $1/2$, which can be achieved using just $M_{(1,1,1)}$ and a balanced beamsplitter.

The balanced tritter uses all the measurement operators from $\mathcal{M}_2$, with each contributing $1/9$ to achieve the success probability $2/3$. To draw attention to the difference between optimizing a single operator and multiple operators at once, mentioned in Sec. 4.2.1, we notice that optimizing for one operator from the set $\mathcal{M}_2$ yields a success probability higher than $1/9$ (for some other choice of U). Taking this further, we can search numerically for the single best outcome, with a cost function similar to that of Eq. (4.18), except we now focus only on a single $\underline{n}$, that is $C(\underline{n}, U) = -2 \sum_{\lambda, r, r'} \exp \left(-\xi| \langle (3), \underline{n}|U|(3), \underline{1} \rangle |^2 \right) | \langle \lambda, \underline{n}, r'|U|\lambda, \underline{1}, r \rangle |^2$. We find $M_{(1,1,1)}$ is a clear winner with a total success probability of $1/2$, achievable by a balanced beamsplitter as mentioned above. All of the other operators by themselves only ever give an optimised success probability of $1/8$. Notice that $6 \cdot 1/8 = 3/4 > 2/3$, showing that the strategy that gives us the best success chance with a single operator from $\mathcal{M}_2$ can not be achieved simultaneously by all six of them.

4.6 Four and more modes

4.6.1 Discriminating singly distinguishable states

Using the numerical optimisation described in Sec. 4.2.2, we also examined the discrimination of singly distinguishable states for $N = 4$ and 5 photons. Together with the results for $N = 2$ and 3, we see the optimisation return interferometers equivalent to QFT_N, each giving a success probability $1 - 1/N$, saturating the bound in Sec. 4.3. We have confirmed this behaviour by direct calculation up to $N = 9$.

4.6.2 Discriminating completely distinguishable states

Numerical optimisation for the $N = 4$ and 5 photon completely distinguishable states yields success probabilities of $19/24$ and $31/36$, respectively. Both of these are less than the general bounds of Sec. 4.3, ($23/24$ and $35/36$, respectively), and so we cannot conclude
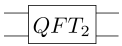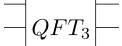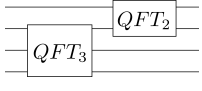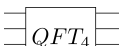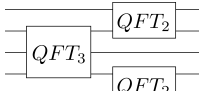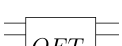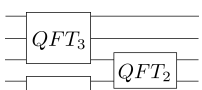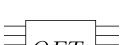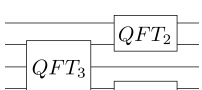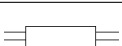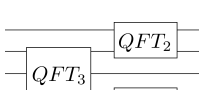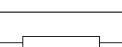
N	U	Singly distinguishable, ρ_s — Success probability				Completely distinguishable, ρ_d — Success probability	
		Bound $(1-1/N)$	Best	Worst	Avg	Bound $(1-1/N!)$	
2	QFT_2	$\frac{1}{2}=0.5000$	$\frac{1}{2}=0.5000$			$\frac{1}{2}=0.5000$	$\frac{1}{2}=0.5000$
3	QFT_3	$\frac{2}{3}\approx0.6666$	$\frac{2}{3}\approx0.6666$			$\frac{5}{6}\approx0.8333$	$\frac{2}{3}\approx0.6666*$
4	QFT_3, QFT_2 network	$\frac{3}{4}=0.7500$	$\frac{25}{36}\approx0.6944$	$\frac{1}{4}=0.2500$	$\frac{7}{12}\approx0.5833$	$\frac{23}{24}\approx0.9583$	$\frac{19}{24}\approx0.7916*$
	QFT_4		$\frac{3}{4}=0.7500$				$\frac{3}{4}=0.7500$
5	QFT_3, QFT_2, QFT_2 network	$\frac{4}{5}=0.8000$	$\frac{8}{15}\approx0.7222$	$\frac{1}{4}=0.2500$	$\frac{13}{18}\approx0.5333$	$\frac{119}{120}\approx0.9917$	$\frac{31}{36}\approx0.8611*$
	QFT_5		$\frac{4}{5}=0.8000$				$\frac{4}{5}=0.8000$
6	QFT_3, QFT_2, QFT_3 network	$\frac{5}{6}\approx0.8333$	$\frac{167}{243}\approx0.6872$	$\frac{167}{243}\approx0.6872$	$\frac{167}{243}\approx0.6872$	$\frac{719}{720}\approx0.9986$	$\frac{671}{729}\approx0.9204$
	QFT_6		$\frac{5}{6}\approx0.8333$				$\frac{65}{72}\approx0.9028$
7	QFT_3, QFT_2, QFT_2, QFT_3 network	$\frac{6}{7}\approx0.8571$	$\frac{695}{972}\approx0.7150$	$\frac{1}{4}=0.2500$	$\frac{361}{567}\approx0.6367$	$\frac{5039}{5040}\approx0.9998$	$\frac{2765}{2916}\approx0.9482$
	QFT_7		$\frac{6}{7}\approx0.8571$				$\frac{6}{7}\approx0.8571$
8	QFT_3, QFT_2, QFT_3, QFT_2 network	$\frac{7}{8}=0.8750$	$\frac{695}{972}\approx0.7150$	$\frac{1}{4}=0.2500$	$\frac{97}{162}\approx0.5988$	$\frac{40319}{40320}\approx1.0000$	$\frac{45095}{46656}\approx0.9665$
	QFT_8		$\frac{7}{8}=0.8750$				$\frac{7}{8}=0.8750$

Table 4.1: The best known interferometers for discrimination of the singly and completely distinguishable states of $N=2$ to 8 photons in N modes. For $N=2$ and 3 the quantum Fourier transform (QFT_N) is optimal for both ρ_s and ρ_d, but for $N\geq4$ the interferometers for each are different; we include all probabilities of success for comparison. For singly distinguishable states, the quantum Fourier transform saturates the bound and so is optimal for each N; due to the QFT's symmetry it does not matter which port the 'bad' photon (see Sec. 3.3.2) is in, however this is not true of the ρ_d interferometers and so we include best, worst and average success probabilities assuming each port is equally likely to be 'bad'. The completely distinguishable state is essentially unique, so there is only one success probability to report; an asterix $*$ indicates extensive numerical optimisation leads us to believe the $N=3,4,5$ cases are optimal despite being far from the bound; it is remarkable that the ρ_d interferometers have constant optical depth (made up of QFT_3s followed by QFT_2s) for each N. Interestingly, the two success probabilities for the QFT are always equal except for $N=6$, the only case in the table that is not a power of a prime (see Discussion). The measurement outcomes that lead to these probabilities are specified in Table 4.2.

they are optimal. We observe that they do both exceed the singly distinguishable bound of $1 - 1/N$, consistent with the intuition that it is easier to discriminate a completely distinguishable state than one that is less distinguishable.

The numerics are sensitive to the penalties used in Eq. (4.18), due to the existence of interferometers with very similar performance. For $N = 4$, a penalty $\xi = 10$ returns an interferometer with success probability $25/32$ that minimises the cost function with a value of -0.839477, while a better interferometer with success probability $19/24$ exists but gives a higher value of -0.836287. Increasing the penalty to 50 yields costs -0.78455 and -0.79277 for these two interferometers respectively, showing that the latter is now the minimum. However, increasing the penalty makes optimisation more difficult, because the landscape flattens and gradients go to zero. For this reason, penalties of 10, 13, 15, 17, 20, 25, 35 and 50 were used for $N = 4$, and 10, 12, 14, 15, 16, 18, 20, 35, and 60 for $N = 5$.

While the complexity of the calculations precluded any further optimisation for $N > 5$, we notice that the best interferometers for $N = 2, 3, 4, 5$ can be composed out of QFT_3 followed by QFT_2s. This suggests a 'recursive' structure for the best discriminating interferometers; for $N = 6, 7, 8$ we tried combinations of QFT_N, QFT_{N-1} and so on, and found that discriminators composed of QFT_3s followed by QFT_2s performed the best. This is remarkable as these are of constant optical depth (the maximum number of beamsplitters and phaseshifters that each photon encounters), independent of N. Indeed, increasing the optical depth beyond this seems to decrease the success probability, which allowed us to limit our search to a manageable number of configurations. These are educated guesses however, and do not rule out the existence of better interferometers that might be found.

Table 4.1 contains a summary of these results. We report the probabilities for the best interferometers found to successfully discriminate ρ_s and ρ_d from ρ_i up to $N = 8$. The measurement outcomes that achieve these probabilities up to $N = 5$ are specified in Table 4.2, where in the interest of saving space we give the occupations that fail (i.e. correspond to the *ambiguous* POVM element E_0) instead of the successful discriminators, because the latter far exceed the former. For comparison, for each interferometer we include success probabilities for both states of interest to be discriminated from the completely indistinguishable state. Note that as discussed above for $N = 3$, although the QFT_N interferometer is optimal for ρ_s no matter which System mode the 'bad' photon is in, this will not be true for interferometers that lack the symmetry of QFT_N. Indeed, the best ρ_d discriminator does not treat each System mode the same way, and so when using such an interferometer to discriminate ρ_s we report best, worst and average success probabilities, assuming each System mode is equally likely to contain the 'bad' photon.

4.7 Conclusion

In this Chapter we provide an application of the framework developed and presented in Chapter 3. By a simple inspection of the states written in the Schur-Weyl basis as per Sections 3.3.1, 3.3.2, and 3.3.3, we present two general upper bounds valid for any

N	ρ_{s}
2	20,02
3	300,030,003
	111
4	4000,0400,0040,0004
	2020,0202
	2101,1210,1012,0121
5	50000,05000,00500,00050,00005
	31001,30110,13100,11030,10301,10013,03011,01310,01103,00131
	22010,21200,20102,20021,12002,10220,02201,02120,01022,00212
	11111

N	ρ_{d}
2	20,02
3	300,030,003
	111
4	4000,0400,0040,0004
	3100,1300,1030,1003,0130,0103
	2011,0211
5	50000,05000,00500,00050,00005
	40010,40001,10040,10004,04010,04001,01040,01004
	31010,31001,13010,13001,10310,10301,10031,10013,01310,01301,01031,01013
	20120,20102,02120,02102

Table 4.2: Measurement occupations corresponding to the ambiguous POVM element E_0 that do *not* discriminate the two states of interest for the numerically optimised interferometers in Table 4.1 ($N = 2, 3, 4, 5$) – these are in general far fewer than the number of successful discriminating occupations, and so easier to list. Recall that for ρ_{s}, the optimal choice of QFT_N does not depend on the mode in which the single distinguishable photon is present, and neither do the occupations. Note that although all of the occupations not listed here satisfy Eq. (4.16), some might have zero probability of occurring and therefore not contribute to discrimination.

photon number when discriminating (i) a state with a single distinguishable photon from the completely indistinguishable state, and (ii) the completely distinguishable from the completely indistinguishable state. This shows how easily a different framework can offer new results.

We then motivated the idea that the HOM test can be generalized using unambigious discrimintation by demonstrating it as a special case of unambigious discrimination when applied to distinguishable states of two photons. We also show that the HOM test is the only test of distinguishability for arbitrary states of two photons, and demonstrate the generality of the formalism by considering three photons in two modes. As a further work goal, this formalism should make it possible to make some more general statements about N photons in two modes (although this is not very practically motivated).

We use a mix of analytical and numerical techniques to argue the optimality of a balanced three mode network (tritter) as a discriminator for both completely distinguishable and singly distinguishable states, and further looking at discrimination of singly distinguishable states with higher photon numbers up to $N = 9$ we show that the quantum Fourier

transform (QFT) saturates the previously found bound, suggesting it is the optimal interferometer for all N. However, in Sec. 4.6.2 we look at the discrimination of completely distinguishable states with higher photon numbers and give examples of the best known interferometers up to $N = 8$, found by observing a pattern emerging from the optimisations, showing that not in all cases QFT is the best choice. Most of these results are summarised in Table 4.1. Although not surprising that the QFT features heavily, the results show that it is not optimal for discriminating completely distinguishable states, motivating the search for optimal discriminating networks for other states of interest.

There are many other state discrimination scenarios we could consider. For example, we could try to unambiguously discriminate ρ_d from ρ_s, two entirely different states, or more than two states. Note that due to the 'nested' structure of our three states of interest (cf. Eq. (4.11)), attempting to find a UD POVM $\{E_i, E_d, E_s, E_0\}$ reduces to only being able discriminate ρ_d from the rest. Another version of discrimination to consider is using bucket (yes/no) instead of number resolving detectors, which are simpler to engineer. While our focus has been on optimizing over all the possible measurement patterns to obtain the highest possible success probability, as mentioned in Sec. 3.2.2 another type of optimisation that can be carried out is choosing a fixed set of patterns and optimizing the interferometer U only. The difference would be that in Eqs. (4.15, 4.16) D would now be fixed, simplifying the problem. As an example, during the preparation of this thesis a closely related paper was released [141], where the authors study a single reference photon input into a QFT_{N-1}, followed by QFT_2 HOM tests on the $N-1$ outputs with the rest of the $N-1$ photons (for a total of N photons in $2N-2$ modes). This is equivalent to a UD procedure where D is fixed as the set of N-fold coincidences. The approach is different and so it is not surprising that it is suboptimal for discrimination, however this interferometer's behaviour is clear for all N.

Finally, we have no doubt that proofs for many of the results here, such as the optimality of QFT_N for discriminating singly distinguishable states, should be possible, but they are left as further work.

DISCUSSION

Throughout out the thesis, we have already highlighted some future work ideas and directions. The new model for distinguishability we introduce in Chapter 3 in particular makes us examine older distinguishability ideas and results, and try to either understand them in different ways or forward them in fresh directions. Here we discuss a few key concepts in a bit more depth. We link the presented distinguishability model back to scattering probabilities of linear optical systems and the permanents (seen in Section 5.1), to find that there is a possibility of some more general relation between immanants and scattering amplitudes of distinguishable particles. Then, we look at suppression laws in more detail and compare them to results from Chapter 4. We finish this thesis by presenting different situations we need to think about to define distinguishability, and their representation and meaning in this framework.

5.1 Permanents, determinants, and immanants

5.1.1 Permanents and determinants

We already mentioned the relation of Fock states and permanents in Section 1.3.4. As a reminder, given an occupation $\underline{n}$ of the input state, and an occupation $\underline{n}'$ of the output state, we construct a new matrix $U_{\mathrm{nn}'}$ from U in two steps. First, define the $d \times N$ matrix U_{n} consisting of n_j copies of the j-th column of U for all $j \in \{1, \cdots, d\}$. Next, construct the $d \times d$ matrix $U_{\mathrm{nn}'}$ by using n'_j copies of the j-th row of U_{n} for all $j \in \{1, \cdots, d\}$.

The scattering probability for bosons is (see Section 1.3.4)

$$\langle \underline{n} | U^{(N)} | \underline{n}' \rangle = \frac{\mathrm{perm}(U_{\mathrm{nn}'})}{\sqrt{\underline{n}!\underline{n}'!}}, \tag{5.1}$$

where $\underline{n}!,\underline{n}'!$ was defined in Eq. (1.9).

Moreover, in the case of fermions, a very similar expression holds. Again we use $\underline{n}$ and $\underline{n}'$ to mark the input and output occupations respectively and as we are dealing with fermions, note that these follow Pauli exclusion principle, that is $n_j \leq 1$ and $n'_k \leq 1$ for all

j and k. Then

$$\langle \underline{n} | U^{(1^N)} | \underline{n}' \rangle = \sum_{\sigma \in S_N} \prod_{j=1}^{N} \text{sgn}(\sigma) U_{\text{nn}'} = \det(U_{\text{nn}'}). \tag{5.2}$$

The permanent and determinant are just special cases of a broader set of matrix functionals called immanants. For a $d \times d$ matrix $M = (M_{jk})$ and character χ_λ, the immanant of the matrix M is then defined as

$$\text{imm}_\lambda(M) = \sum_{\sigma \in S_N} \chi_\lambda(\sigma) M_{1\sigma(1)} M_{2\sigma(2))} \dots M_{d\sigma(d)} \tag{5.3}$$

$$= \sum_{\sigma \in S_N} \chi_\lambda(\sigma) \mathcal{S}(\sigma) \, \text{diag}(M), \tag{5.4}$$

where $\mathcal{S}(\sigma)$ is as defined in Equation (3.9), and $\text{diag}(M)$ is the diagonal matrix $(\delta_{jk} M_{jk})$. Because the $\chi_{(N)}(\sigma) = 1$ and $\chi_{(1^N)}(\sigma) = \text{sgn}\,\sigma$, we see that that $\text{perm} = \text{imm}_{(N)}$ and $\det = \text{imm}_{(1^N)}$.

This connection between bosons and permanents, fermions and determinants, leads us to question whether immanants can tell us something more about the matrix elements occurring in the mixed exchange symmetry space. We have already mentioned "immanons" from [144], and the link to immanants is where these theoretical particles have got their name. However in [144], they focus on scattering probabilities of these immanons, but the link to possible construction of the basis of mixed exchange symmetry is not explored.

In the occupation notation, it might be straightforward to see the relevance of the immanants to the matrix elements of the irreps. The matrix $U_{\text{nn}'}$ is specified by occupation numbers only. The immanants are specified by the irrep λ (i.e. the characters of the irrep λ). Each matrix element in $\mathcal{U}$ is determined by the irrep, occupation number, but also the inner multiplicity (denoting the basis state as $|\lambda p \underline{n} r\rangle$, we know that the matrix elements $\langle \lambda p \underline{n} r | U^\lambda | \lambda p \underline{n}' r' \rangle$ do not depend on p so no need for an extra parameter here). So the back of the envelope check on whether we can combine immanants with the $U_{\text{nn}'}$ to get all the matrix elements of $\mathcal{U}$ shows us that we do not have enough parameters (we need more due to inner multiplicities). When the inner multiplicity is unique though, then it might be possible for the immanant to be linked to the single matrix element U^λ of $U_{\text{nn}'}$. Notice that the matrix $U_{\text{nn}'}$ is actually all we need to know from the original matrix U to see how a $\underline{n}$-weight space transforms into $\underline{n}'$-weight space.

5.1.2 Generalization to immanants

As mentioned the immanons have been explored in [144]. However, previous results linking immanants with representations of $U(d)$ can be found in [204]. Kostant shows a link between immanants of the matrix U and the elements $\langle \lambda, p, \underline{n}, r | U^\lambda | \lambda, p, \underline{n}', r' \rangle$ where $\underline{n}$ and $\underline{n}'$ are occupations from the 0-weight subspace. It is not immediately clear what the 0-weight subspace is in our notation where weights are occupations, however, there is an easy mapping from a standard weight notation of the $(d-1)$-tuple $\mu = (\mu_1, \mu_2, \dots, \mu_{d-1})$ for weights of $U(d)$, to the one we are using. The standard notation μ is then just

$\mu = (n_1 - n_2, n_2 - n_3, \ldots, n_{d-1} - n_d)$. We can see that the 0-weight subspace is the coincident subspace, where $\underline{n} = \underline{1}$. Therefore there is a link between immanants of the matrix U and the matrix elements found in the coincident subspace. According to Kostant

$$\text{imm}_\lambda(U) = \text{Tr}\left[P_\lambda U^\lambda P_\lambda\right] = \sum_r \langle \lambda, p, \underline{n}, r | U^\lambda | \lambda, p, \underline{n}, r \rangle \tag{5.5}$$

where P_λ is the projector to the λ irrep of S_N, $P_\lambda = \sum_\sigma \chi_\lambda(\sigma)\mathcal{S}(\sigma)$, and U^λ is the λ irrep of $U \in U(d)$. The coincident symmetric subspace and the coincident antisymmetric subspace have the inner multiplicity 1, so this just reduces to the Eqs. (1.31) and (5.2). But the Eqs. (1.31) and (5.2) are more general than just the coincident subspace, so there is possibly another generalization.

In [137], the authors show the immanants link to the representation of submatrices of U. Taking $U_{\text{nn}'}$ to be a submatrix of U such that some columns and rows are missing then

$$\text{imm}_\lambda(U_{\text{nn}'}) = \sum_r \langle \lambda, p, \underline{n}, r | U_{\text{nn}'}^\lambda | \lambda, p, \underline{n}, r \rangle. \tag{5.6}$$

That is, the previous result generalizes to submatrices of unitary matrices. While this is more general, it is not of the form we need above, as in general matrices $U_{\text{nn}'}$ have missing rows and columns but can also have rows and columns that show up more than once. In another paper [116], immanants are linked to the irrep matrix elements, however this paper as well with coincidences for three particles. An interesting result in this paper is that on the previous construction of $U_{\text{nn}'}$, they also permute certain rows and columns of this matrix depending on the basis state in the coincident subspace. This could introduce the extra parameter we were mentioning is needed to handle inner multiplicities occurring in irrep λ of an occupation $\underline{n}$.

To summarize, while there is a result linking immanants and the representation matrix elements in the zero-weight subspace, there is no generalized results. There are two things needed to be generalized: one is generalizing to other than zero-weight subspaces with inner multiplicity of one; two is generalizing to other subspaces with more inner multiplicities. We are also left with a question of whether a single element of the matrix can be linked to a single immanant expression (as opposed to the Eq. (5.5) with the trace which possibly arises due to the inner multiplicities and here the idea of permuting rows and columns could come in).

5.1.3 Permanents and immanants as matrix elements for three qubits

In Section 3.4.1 we saw how the Schur-Weyl basis is constructed and what it is for three qubits. Here we see how this basis change affects the unitary applied to the states, leading to matrix elements expressed as immanants. As mentioned before, we build the basis

transformation matrix U_{Sch} and then apply the single particle unitary which is

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}. \tag{5.7}$$

The transformed matrix is then $\oplus U^\lambda \otimes \mathbb{1}^\lambda = U_{\text{Sch}} U^{\otimes 3} U_{\text{Sch}}^{-1}$. The resulting matrix is of the form

$$\begin{bmatrix} \mathcal{U}^{\square\square\square} & 0 & 0 \\ 0 & \mathcal{U}^{\square\!\square} & 0 \\ 0 & 0 & \mathcal{U}^{\square\!\square} \end{bmatrix} \tag{5.8}$$

where we can clearly see the block diagonal matrices belonging to different irreps. The matrix $\mathcal{U}^{\square\square\square}$ is

$$
\begin{array}{c}
 \\
\langle(\boxed{1\,2\,3},\boxed{1\,1\,1})| \\
\langle(\boxed{1\,2\,3},\boxed{1\,1\,2})| \\
\langle(\boxed{1\,2\,3},\boxed{1\,2\,2})| \\
\langle(\boxed{1\,2\,3},\boxed{2\,2\,2})|
\end{array}
\begin{bmatrix}
U_{11}^3 & \sqrt{3}U_{11}^2 U_{12} & \sqrt{3}U_{11}U_{12}^2 & U_{12}^3 \\
\sqrt{3}U_{11}^2 U_{21} & U_{11}\left(2U_{12}U_{21}+U_{11}U_{22}\right) & U_{12}\left(U_{12}U_{21}+2U_{11}U_{22}\right) & \sqrt{3}U_{12}^2 U_{22} \\
\sqrt{3}U_{11}U_{21}^2 & U_{21}\left(U_{12}U_{21}+2U_{11}U_{22}\right) & U_{22}\left(2U_{12}U_{21}+U_{11}U_{22}\right) & \sqrt{3}U_{12}U_{22}^2 \\
U_{21}^3 & \sqrt{3}U_{21}^2 U_{22} & \sqrt{3}U_{21}U_{22}^2 & U_{22}^3
\end{bmatrix}
$$

with column labels $|(\boxed{1\,2\,3},\boxed{1\,1\,1})\rangle$, $|(\boxed{1\,2\,3},\boxed{1\,1\,2})\rangle$, $|(\boxed{1\,2\,3},\boxed{1\,2\,2})\rangle$, $|(\boxed{1\,2\,3},\boxed{2\,2\,2})\rangle$
$$\tag{5.9}$$

and the matrix $\mathcal{U}^{\square\!\square}$ is

$$
\begin{array}{c}
 \\
\langle(\substack{\boxed{1\,3}\\ \boxed{2}},\substack{\boxed{1\,1}\\ \boxed{2}})| \\
\langle(\substack{\boxed{1\,3}\\ \boxed{2}},\substack{\boxed{1\,2}\\ \boxed{2}})|
\end{array}
\begin{bmatrix}
U_{11}\left(U_{11}U_{22}-U_{12}U_{21}\right) & U_{12}\left(U_{11}U_{22}-U_{12}U_{21}\right) \\
U_{21}\left(U_{11}U_{22}-U_{12}U_{21}\right) & U_{22}\left(U_{11}U_{22}-U_{12}U_{21}\right)
\end{bmatrix},
\tag{5.10}
$$

with column labels $\left|\left(\substack{\boxed{1\,3}\\ \boxed{2}},\substack{\boxed{1\,1}\\ \boxed{2}}\right)\right\rangle$ and $\left|\left(\substack{\boxed{1\,3}\\ \boxed{2}},\substack{\boxed{1\,2}\\ \boxed{2}}\right)\right\rangle$

where for $\mathcal{U}^{\square\!\square}$ we give the basis in the outer multiplicity $p = 1$, but the same ordering and matrix apply for $p = 2$.

From what we learned in the Section 5.1 we would expect the first block ($\lambda = (3)$) to consist of permanents of matrices related to U. The following examples shows how to construct the suitable U_{st} matrix and retrieve the right permanent.

Let the occupation of the input state be $\underline{s} = (2,1)$. There are three states that match this occupation, the symmetric state $|(\boxed{1\,2\,3},\boxed{1\,1\,2})\rangle$, the mixed symmetry state $\left|\left(\substack{\boxed{1\,3}\\ \boxed{2}},\substack{\boxed{1\,1}\\ \boxed{2}}\right)\right\rangle$, and the outer multiplicity of this state, $\left|\left(\substack{\boxed{1\,2}\\ \boxed{3}},\substack{\boxed{1\,1}\\ \boxed{2}}\right)\right\rangle$. Construct U_s by repeating the first column of U twice and the second once to get matrix

$$U_{21} = \begin{bmatrix} U_{11} & U_{11} & U_{12} \\ U_{21} & U_{21} & U_{22} \end{bmatrix}. \tag{5.11}$$

Now, let's say the state we want to get out is the output state with occupation $\underline{t} = (1,2)$. Then the matrix $U_{\text{s,t}}$ is built by repeating its row 1 once and row 2 twice from that matrix

U_s. So we get

$$U_{21,12} = \begin{bmatrix} U_{11} & U_{11} & U_{12} \\ U_{21} & U_{21} & U_{22} \\ U_{21} & U_{21} & U_{22} \end{bmatrix}. \tag{5.12}$$

The permanent is then $\mathrm{perm}(U_{21,12}) = 2U_{12}U_{21}^2 + 4U_{11}U_{21}U_{22}$ and this indeed matches the element $\langle(\boxed{1\,2\,3}, \boxed{1\,2\,2})\,|\,\mathcal{U}^{\boxed{}}\,|\,(\boxed{1\,2\,3}, \boxed{1\,1\,2})\rangle$ of Eq. (5.9) (with a factor of two as expected from Eq. (1.31).

In general, we find the matrix $\mathcal{U}^{\boxed{}}$ to be

$$\mathcal{U}^{\boxed{}} = \begin{bmatrix} 6\mathrm{perm}(U_{30,30}) & 2\sqrt{3}\mathrm{perm}(U_{21,30}) & 2\sqrt{3}\mathrm{perm}(U_{12,30}) & 6\mathrm{perm}(U_{30,30}) \\ 2\sqrt{3}\mathrm{perm}(U_{30,21}) & 2\mathrm{perm}(U_{21,21}) & 2\mathrm{perm}(U_{12,21}) & 2\sqrt{3}\mathrm{perm}(U_{21,21}) \\ 2\sqrt{3}\mathrm{perm}(U_{30,12}) & 2\mathrm{perm}(U_{21,12}) & 2\mathrm{perm}(U_{12,12}) & 2\sqrt{3}\mathrm{perm}(U_{12,12}) \\ 6\mathrm{perm}(U_{30,03}) & 2\sqrt{3}\mathrm{perm}(U_{21,03}) & 2\sqrt{3}\mathrm{perm}(U_{12,03}) & 6\mathrm{perm}(U_{03,03}) \end{bmatrix} \tag{5.13}$$

We also find the immanant of the $U_{21,12}$ matrix,

$$\mathrm{imm}_{(2,1)}(U_{21,12}) = 2U_{11}U_{21}U_{22} - U_{11}U_{22}U_{21} - U_{12}U_{21}U_{21} \tag{5.14}$$

$$= U_{11}U_{22}U_{21} - U_{12}U_{21}U_{21} \tag{5.15}$$

which matches $\langle\left(\boxed{\begin{smallmatrix}1&3\\2&\end{smallmatrix}}, \boxed{\begin{smallmatrix}1&2\\2&\end{smallmatrix}}\right)\,|\,\mathcal{U}^{\boxed{}}\,|\,\left(\boxed{\begin{smallmatrix}1&3\\2&\end{smallmatrix}}, \boxed{\begin{smallmatrix}1&1\\2&\end{smallmatrix}}\right)\rangle$.

Finding the rest of the matrices, and applying the immanant, we get the following expressions

$$\mathrm{imm}_{(2,1)}(U_{21,21}) = 2U_{11}U_{11}U_{22} - 2U_{11}U_{12}U_{21} \tag{5.16}$$

$$\mathrm{imm}_{(2,1)}(U_{21,12}) = U_{11}U_{21}U_{22} - U_{12}U_{21}U_{21} \tag{5.17}$$

$$\mathrm{imm}_{(2,1)}(U_{12,21}) = U_{11}U_{12}U_{22} - U_{12}U_{12}U_{21} \tag{5.18}$$

$$\mathrm{imm}_{(2,1)}(U_{12,12}) = 2U_{11}U_{22}U_{22} - 2U_{12}U_{21}U_{22}. \tag{5.19}$$

$$\tag{5.20}$$

This matrix in Eq. (5.10) can then be re-written as

$$\mathcal{U}^{\boxed{}} = \begin{bmatrix} 2\,\mathrm{imm}_{(2,1)}(U_{21,21}) & \mathrm{imm}_{(2,1)}(U_{12,21}) \\ \mathrm{imm}_{(2,1)}(U_{21,12}) & 2\,\mathrm{imm}_{(2,1)}(U_{12,12}) \end{bmatrix}. \tag{5.21}$$

Notice that there is no zero-weight subspace in the case of three qubits, so the immanant statements mentioned in Section 5.1 do not apply.

5.1.4 Matrix elements for inner multiplicities of three qutrits

In Section 3.4.1, we have shown how the framework can be applied to states of three qubits (which as we can see in Section 3.2 is relevant to situations of three photons in two modes when we introduce distinguishability). We also saw what the basis is for three qutrits

in Section 3.4.2. Here, we consider three qutrits, or three photons in three modes, in a coincident state. In Section 3.4.1, since we were working with three qubits, there were no states with inner multiplicities. However, in the case of three qutrits, we do have an irrep with inner multiplicities, the $\lambda = (2,1)$ irrep for occupation $(1,1,1)$. The four states of interest (there are two outer multiplicites as well in this irrep) are then

$$\left| \left(\young(13,2), \young(13,2) \right) \right\rangle, \left| \left(\young(13,2), \young(12,3) \right) \right\rangle, \left| \left(\young(12,3), \young(13,2) \right) \right\rangle, \left| \left(\young(12,3), \young(12,3) \right) \right\rangle. \tag{5.22}$$

The scattering amplitudes of states with no inner multiplicites, such as

$$\left\langle \left(\young(13,2), \young(11,2) \right) \right| \mathcal{U}^{\yng(2,1)} \left| \left(\young(13,2), \young(11,2) \right) \right\rangle, \tag{5.23}$$

or

$$\left\langle \left(\young(13,2), \young(11,2) \right) \right| \mathcal{U}^{\yng(2,1)} \left| \left(\young(13,2), \young(23,3) \right) \right\rangle \tag{5.24}$$

can actually be found to match the immanants in much the same way as in Eq. 5.15.

However, in the case of the states which have inner multiplicities, things are less clear. For example,

$$\left\langle \left(\young(13,2), \young(11,2) \right) \right| \mathcal{U}^{\yng(2,1)} \left| \left(\young(13,2), \young(13,2) \right) \right\rangle = \frac{2U_{11}U_{13}U_{22} - U_{12}(U_{13}U_{21} + U_{11}U_{23})}{\sqrt{2}}. \tag{5.25}$$

We can see that the expression on the right does seem like an immanant of a matrix. Constructing the matrix as given in Section 5.1, we get

$$U_{111,210} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \end{bmatrix}. \tag{5.26}$$

The immanant is then

$$\mathrm{imm}_{(2,1)}(U_{111,210}) = 2U_{11}U_{12}U_{23} - U_{12}U_{13}U_{21} - U_{11}U_{13}U_{22}, \tag{5.27}$$

which is not the same expression as on the right hand side of Equation. 5.25. However, we can see that the matrix (following the ideas from [115, 116] for zero-weight subspace with permutations of rows/columns)

$$\begin{bmatrix} U_{11} & U_{13} & U_{12} \\ U_{11} & U_{13} & U_{12} \\ U_{21} & U_{23} & U_{22} \end{bmatrix} \tag{5.28}$$

has an immanant that matches the expression in Equation 5.25. If we now take a look at the scattering amplitude

$$\left\langle \left(\young(13,2), \young(11,2) \right) \right| \mathcal{U}^{\yng(2,1)} \left| \left(\young(13,2), \young(12,3) \right) \right\rangle = \sqrt{\frac{3}{2}} U_{12}(U_{11}U_{23} - U_{13}U_{21}) \tag{5.29}$$

it is even less clear what matrix gives this expression as an immamant. Like the expressions found in [116] for zero-weight subspace, it is possible that this amplitude can be expressed as a linear combination of immanants but not as an immanant itself. It is also possible that there is a better choice of Schur-Weyl basis which makes the scattering amplitudes appear as immanants. For example we have some freedom in choosing the states when we are orthonormalizing them across the multiplicites. More work is required to formulate a general law that applies well to all the different states.

The work of Kostant though can be easily checked,

$$\left\langle \left(\tfrac{1\,3}{2}, \tfrac{1\,3}{2} \right) \mid \mathcal{U}^{\square} \mid \left(\tfrac{1\,3}{2}, \tfrac{1\,3}{2} \right) \right\rangle + \left\langle \left(\tfrac{1\,3}{2}, \tfrac{1\,2}{3} \right) \mid \mathcal{U}^{\square} \mid \left(\tfrac{1\,3}{2}, \tfrac{1\,2}{3} \right) \right\rangle \qquad (5.30)$$

$$= -U(1,2)U(2,3)U(3,1) - U(1,3)U(2,1)U(3,2) + 2U(1,1)U(2,2)U(3,3) \qquad (5.31)$$

$$= - \mathrm{imm}_{(2,1)}(U_{111,111}) \qquad (5.32)$$

which follows the Eq. 5.5 up to a sign.

5.2 Suppression laws and discrimination of distinguishability

The focus of suppression laws has been on finding formulas that predict suppression of events for either specific unitaries or families of unitaries. The unambiguous discrimination procedure studied in this thesis can be viewed as a search over the space of all interferometers, where the optimality of U is measured by the total probability of completely suppressed occupations. We will now discuss how the two tools can come together. But first let us briefly touch on the matrices and input states that are currently covered by these laws.

As QFT is a matrix that comes up often in Chapter 4 (see Table 4.1), we mention some of the results related to it in a bit more detail here. QFT was first looked into from the suppression law perspective by Tichy et al. in 2010 [145]. In this paper, they find sufficient conditions for suppression in case of coincident input and any output. Interestingly, this condition accounts for all the suppressed outputs in the case for $N = 2, 3, 4, 5, 7, 8, 9, 11, 13$, but not $N = 6, 10, 12, 14$ as noted in the paper (cases that were numerically checked). This seemingly strange behaviour could possibly be explained by the fact that the latter are not powers of prime. We see the same type of behaviour show up for unambigious discrimination in Table 4.1. General suppression effects they observe in the paper are:

- Events with $N - 1$ occupied ports are suppressed for odd N.

- Coincident events are forbidden for even N.

- If N is prime, there are never exactly two occupied ports.

- Output of the type $\underline{r} = (N - 1, 1, 0, ..., 0)$ is strictly suppressed.

Of note, two inputs (outputs) of the QFT matrix related by a cyclic or anticyclic permutation give the same scattering probabilities [145].

The follow-up paper [147] investigates more general periodic inputs and outcomes using a QFT matrix. For this purpose they introduce an *m-periodic* state, $\underline{q}$, of N particles in d modes such that for an integer $m|N$, it consists of $p = N/m$ repetitions of pattern $\underline{k}$ of length m with $\sum_j k_j = N/p$. That is $\underline{q} = (k_1, \ldots, k_m, k_1, \ldots, k_m, \ldots, k_1, \ldots, k_m)$. For example, the coincident input has $p = N$ and $m = 1$. The suppression law they previously found can be expanded to m-periodic initial states.

An interesting observation that they make is that initial states which seem very different but have the same period m show the same suppression. The result is also more general insofar that the number of modes is not restricted to be equal to the number of particles, and in the paper they show the results for $N = 4$ particles in $d = 12$ modes as well. The suppression condition remains only sufficient and there exist suppressed events which are not accounted for. The underlying motivation for suppression laws is identifying signatures of distinguishability based on scattering probabilities of bosons, similar to the HOM effect for two photons. The discrimination work in Chapter 4 is focused on a similar goal, however, unlike here where the suppression laws are not complete due to suppressed events that remain unaccounted, our work finds interferometers that are provably optimal in certain situations, and numerically optimal for others. The two tasks are not completely identical though, however, they are closely linked.

The papers that follow [148–151], look at different interferometers, some even generalizing the type of input states that are looked at [151], or finding necessary and sufficient condition for a specific set of inputs [148]. Most interestingly though, all of the above results are unified and expanded in two papers by Dittel et al. [146, 152]. In the former, the authors show that given a certain initial state which is invariant under some mode-permutation operator $\mathcal{P}$, by constructing unitaries U using the eigendecomposition of the operator $\mathcal{P}$, the product of corresponding eigenvalue distribution generated using the output state can reveal whether that output is "strictly suppressed". In the latter paper [152], they show how to use this unifying law to recreate all the previous supression laws.

Another interesting result in the second paper [152] is that of splitting suppressed events into consequences of single-particle and multi-particle dynamics. The single-particle dynamics that cause suppression of a certain output give a suppressed output in both the distinguishable ("classical") and indistinguishable ("quantum") case. On the other hand, suppression due to multi-particle dynamics is a consequence of quantum interference. Further, they also consider more general states than just Fock state inputs, i.e. arbitrary pure states, entangled states, and partially distinguishable pure states.

There are various ways to rate the success of suppression, but two that come to mind first are either the number of suppressed states or the total probability of suppression. While the former is just a ratio of suppressed states to the total states and possibly easier to find as it does not require to know the suppression probability of a specific state, the second figure of merit is possibly more useful, as the whole idea behind suppression of events is that it allows characterization of "quantumness" of a state. The total probability of suppression is then just a sum over all the suppressed event probabilites. This is

equivalent to our figure of merit in the discrimination paper, that is, it is the same as asking for the unambiguous discrimination between completely distinguishable ("classical") and completely indistinguishable ("quantum") state.

Looking through the inputs and matrices that are covered by the Dittel et al. and calculating the total suppression probability (for small N), we find that none of the matrices offer higher total suppression probability than that of the QFT for the choices of N given in Table 4.1. However, in Chapter 4 our work shows that the QFT is not the optimal choice of interferometer when it comes to total suppression probability. Furthermore, in Chapter 4 we check the success probabilities of the QFT up to $N = 12$, and only the cases $N = 6, 10$ and 12 have higher success probabilities than $1 - 1/N$; these non power of a prime cases are exactly those for which it is known that the suppression law does not account for all the suppressed patterns. An exciting future direction would be exploring the link between the suppression laws and the interferometers we uncover in the work presented here, as it might offer an even more general suppression law. This will probably not be feasible through permanent calculation, as Tichy points out, even in the case of highly symmetric matrices, there are results which show that permanent calculation is not an easy problem. Some more general attack of this problem would be needed, and the distinguishability framework presented here is ideal for taking advantage of symmetries of matrices for calulation of probabilities. This framework, possibly mixed with number theory following Dittel et al., might be a good starting point for future work.

Finally, from the Table 4.1 we are left with a couple of conjectures to be explored:

Conjecture 5.2.1. *The discrimination probability when QFT is used in conjunction with singly distinguishable states as input state is $1 - 1/N$.*

Conjecture 5.2.2. *The discrimination probability when QFT is used with any distinguishable state is at equal to or more than $1 - 1/N$.*

The approaches to proving suppression laws mentioned here might be helpful in proving these conjectures.

5.3 On the definition of distinguishability

As we have seen, photons have many possible modes in which encoding of qubits can be achieved, such as polarization, spectrum, timing, path, etc. However, this also means that when we encode information in, for example, polarization, we also have to control all the other degrees of freedom, otherwise the particles become distinguishable which makes them behave classically [31]. Experimental setups still have no foolproof way of creating many indistinguishable single photons on demand. This is why there have been many studies attempting to analyze the behaviour of bosons which are partially distinguishable [26, 29, 31]. These approaches have been varied, ranging from group theory to analysis (see Section 1.5).

As we mentioned, the total wavefunction for multi-boson states is symmetric. However, we have seen in previous sections how states of mixed symmetry can occur for bosons

that have two degrees of freedom. This loss of symmetric properties leads to observable consequences such as the difference of the scattering probabilities seen in Eq. (3.110) and (3.107). To understand how modelling bosons with two degrees of freedom relates to distinguishability, we notice that so far we have used particle labelling without there being physical grounds for it, referring to "first particle", "second particle", and so on, based on an imposed ordering due to the way states are written. For example, for the state

$$|abc\rangle \tag{5.33}$$

we would just naturally read it such that particle "one" is in state $|a\rangle$, particle "two" is in state $|b\rangle$, and particle "three" is in state $|c\rangle$. We could write this explicitly

$$|a\rangle_1 |b\rangle_2 |c\rangle_3 \,, \tag{5.34}$$

but it is also often implied from the above notation if it is not mentioned explicitly. Inevitably, we write particle states in a certain order, imposing the first state read out to be the state of particle one, etc. However, we can see that in an (unnormalized) state such as

$$|a\rangle_1 |b\rangle_2 |c\rangle_3 + |a\rangle_2 |b\rangle_1 |c\rangle_3 + |a\rangle_1 |b\rangle_3 |c\rangle_2 + |a\rangle_2 |b\rangle_3 |c\rangle_1 + |a\rangle_3 |b\rangle_2 |c\rangle_1 + |a\rangle_3 |b\rangle_1 |c\rangle_2 \,, \tag{5.35}$$

there is no specific particle linked to a specific state, moreover the physics of this state stays the same under the action of the symmetric group on particles. The approach taken in this thesis is assigning each particle an extra degree of freedom, referred to as "Label", which models this particle indexing as physical information that can be manipulated like any other. The notation used will be the one introduced in the previous sections, thus the first degree of freedom is referred to as "System", and here we group any information about the state that we have control over. Modeling distinguishability using two degrees of freedom has been researched before [27, 46], although not in combination with representation theory and the simplifications unitary-unitary duality gives us.

To talk further about distinguishable states and how the Schur-Weyl basis is natural, let us take a look at a few examples of states of two bosons with two degrees of freedom and discuss whether they should be considered distinguishable or indistinguishable. Like in the above description of the HOM experiment, we will assume the first degree of freedom can be in state "1" or "2", and this can for example correspond to the path degree of freedom, and the second degree of freedom can be in state "a" or "b", and this can for example correspond to the polarization degree of freedom that we cannot access for some reason. We will put these particles on the balanced beamsplitter (the action of which can be found in Eq. (3.104)) and check the statistics particle counting detectors would give us for the first degree of freedom. If these are different than what we would expect for indistinguishable particles (HOM statistics), then the state we started with must be distinguishable. However, if these match statistics that we would expect from indistinguishable particles, then the

state is a strong candidate for what we should consider to be "indistinguishable" state, and should be included in the definition of distinguishability.

Example 5.3.1.

$$\left| \begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix} \right\rangle = \frac{1}{\sqrt{2}} \left| \boxed{1\,2} \right\rangle_{\mathrm{S}} \left| \boxed{a\,a} \right\rangle_{\mathrm{L}} \tag{5.36}$$

$$\xrightarrow{\mathrm{BS}_{50:50}} \frac{1}{2} \left(\left| \boxed{1\,1} \right\rangle_{\mathrm{S}} - \left| \boxed{2\,2} \right\rangle_{\mathrm{S}} \right) \left| \boxed{a\,a} \right\rangle_{\mathrm{L}}. \tag{5.37}$$

As we saw before, $P_{(2,0)} = 0.5$, $P_{(0,2)} = 0.5$, and $P_{(1,1)} = 0$. Nothing very surprising here, as this is the exact state we consider above in our HOM example.

Example 5.3.2.

$$\left| \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right\rangle = \frac{1}{\sqrt{2}} \left| \boxed{1\,2} \right\rangle_{\mathrm{S}} \left| \boxed{a\,b} \right\rangle_{\mathrm{L}} + \frac{1}{\sqrt{2}} \left| \boxed{\begin{matrix}1\\2\end{matrix}} \right\rangle_{\mathrm{S}} \left| \boxed{\begin{matrix}a\\b\end{matrix}} \right\rangle_{\mathrm{L}} \tag{5.38}$$

$$\xrightarrow{\mathrm{BS}_{50:50}} \frac{1}{2} \left(\left| \boxed{1\,1} \right\rangle_{\mathrm{S}} - \left| \boxed{2\,2} \right\rangle_{\mathrm{S}} \right) \left| \boxed{a\,b} \right\rangle_{\mathrm{L}} - \frac{1}{\sqrt{2}} \left| \boxed{\begin{matrix}1\\2\end{matrix}} \right\rangle_{\mathrm{S}} \left| \boxed{\begin{matrix}a\\b\end{matrix}} \right\rangle_{\mathrm{L}}. \tag{5.39}$$

Again, as before, $P_{(2,0)} = 0.25$, $P_{(0,2)} = 0.25$, and $P_{(1,1)} = 0.5$. As mentioned, this state is entangled, and contains more than just the state from the symmetric subspace. The presence of the antisymmetric state in Eq. (5.39) might be a marker of the distinguishability of the bosons. Similarly the entanglement across the irreps might be another marker for distinguishability.

Example 5.3.3.

$$\frac{1}{\sqrt{2}} \left(\left| \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right\rangle + \left| \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\rangle \right) = \left| \boxed{1\,2} \right\rangle_{\mathrm{S}} \left| \boxed{a\,b} \right\rangle_{\mathrm{L}} \tag{5.40}$$

$$\xrightarrow{\mathrm{BS}_{50:50}} \frac{1}{2} \left(\left| \boxed{1\,1} \right\rangle_{\mathrm{S}} - \left| \boxed{2\,2} \right\rangle_{\mathrm{S}} \right) \left| \boxed{a\,b} \right\rangle_{\mathrm{L}} \tag{5.41}$$

Here, $P_{(2,0)} = 0.5$, $P_{(0,2)} = 0.5$, and $P_{(1,1)} = 0$. This is exactly the statistics found in our HOM example for the indistinguishable state (see Eq. (3.107)), except there, the beginning state was $\left| \boxed{1\,2} \right\rangle_{\mathrm{S}} \left| \boxed{a\,a} \right\rangle_{\mathrm{L}}$. This shows that even though there are two particle labels present, due to the symmetry of the particle labelling we cannot actually differentiate those two particles (for the purposes of a HOM-like experiment, these two particles are indistinguishable). Notice that the state itself is entangled in the Fock basis, however in the Schur-Weyl basis, there is no entanglement. Moreover, the state has support in the symmetric subspace only.

Example 5.3.4.

$$\frac{1}{\sqrt{2}} \left(\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \Big\rangle - \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \Big\rangle \right) = \Big\Vert \begin{smallmatrix} 1 \\ 2 \end{smallmatrix} \Big\rangle_{\mathrm{S}} \Big\Vert \begin{smallmatrix} a \\ b \end{smallmatrix} \Big\rangle_{\mathrm{L}} \tag{5.42}$$

$$\xrightarrow{\mathrm{BS}_{50:50}} -\frac{1}{\sqrt{2}} \Big\Vert \begin{smallmatrix} 1 \\ 2 \end{smallmatrix} \Big\rangle_{\mathrm{S}} \Big\Vert \begin{smallmatrix} a \\ b \end{smallmatrix} \Big\rangle_{\mathrm{L}}. \tag{5.43}$$

Now we have $P_{(2,0)} = 0$, $P_{(0,2)} = 0$, and $P_{(1,1)} = 1$. This state exhibits exactly the statistics expected for two indistinguishable fermions (as seen in papers trying to mimic fermion statistics using carefully designed states of bosons [220]), but not indistinguishable bosons (they are contradictory to the HOM experiment). As the previous example, here the state is entangled in the Fock basis, but not in the Schur-Weyl basis. Notice there is also nothing we can do to this starting state to make it give the right statistics for the HOM experiment, as the only amplitude is in the anti-symmetric subspace which does not have support in the $\{|11\rangle, |22\rangle\}$ subspace.

Example 5.3.5.

$$\frac{1}{\sqrt{2}} \left(\begin{vmatrix} 2 & 0 \\ 0 & 0 \end{vmatrix} \Big\rangle - \begin{vmatrix} 0 & 0 \\ 0 & 2 \end{vmatrix} \Big\rangle \right) = \frac{1}{2} \Big\Vert \boxed{1\,1} \Big\rangle_{\mathrm{S}} \Big\Vert \boxed{a\,a} \Big\rangle_{\mathrm{L}} - \frac{1}{2} \Big\Vert \boxed{2\,2} \Big\rangle_{\mathrm{S}} \Big\Vert \boxed{b\,b} \Big\rangle_{\mathrm{L}} \tag{5.44}$$

$$\xrightarrow{\mathrm{BS}_{50:50}} \frac{1}{2} \left(\frac{1}{\sqrt{2}} \Big\Vert \boxed{1\,1} \Big\rangle_{\mathrm{S}} + \Big\Vert \boxed{1\,2} \Big\rangle_{\mathrm{S}} + \frac{1}{\sqrt{2}} \Big\Vert \boxed{2\,2} \Big\rangle_{\mathrm{S}} \right) \Big\Vert \boxed{a\,a} \Big\rangle_{\mathrm{L}} \tag{5.45}$$

$$- \frac{1}{2} \left(\frac{1}{\sqrt{2}} \Big\Vert \boxed{1\,1} \Big\rangle_{\mathrm{S}} - \Big\Vert \boxed{1\,2} \Big\rangle_{\mathrm{S}} + \frac{1}{\sqrt{2}} \Big\Vert \boxed{2\,2} \Big\rangle_{\mathrm{S}} \right) \Big\Vert \boxed{b\,b} \Big\rangle_{\mathrm{L}} \tag{5.46}$$

$$\tag{5.47}$$

Here we have, $P_{(2,0)} = 0.25$, $P_{(0,2)} = 0.25$, and $P_{(1,1)} = 0.5$. This is the same statistics we found with our "standard" distinguishable state in Eq. (3.110). Compare that to the state

$$\frac{1}{\sqrt{2}} \left(\Big\Vert \boxed{1\,1} \Big\rangle_{\mathrm{S}} - \Big\Vert \boxed{2\,2} \Big\rangle_{\mathrm{S}} \right) \Big\Vert \boxed{a\,a} \Big\rangle_{\mathrm{L}} \xrightarrow{\mathrm{BS}_{50:50}} \Big\Vert \boxed{1\,2} \Big\rangle_{\mathrm{S}} \Big\Vert \boxed{a\,a} \Big\rangle_{\mathrm{L}}, \tag{5.48}$$

which would have statistics $P_{(2,0)} = 0$, $P_{(0,2)} = 0$, and $P_{(1,1)} = 1$. After applying the balanced beamsplitter, we see discrepancies in what we would get from the state with just one label present, and with two labels present. The state does not display the statistics we would expect from a state with indistinguishable bosons. While this state is entangled, there is actually no way to identify which particle is which from measuring either degree of freedom. That is, if we find a particle in state "1" in the first degree of freedom, the second degree of freedom is "a", and similarly if a particle is in state "2" in the first degree of freedom, the second degree of freedom is "b". However, neither measurement actually tells us which particle is which. Labels "a"/"b" do not actually "label" a particle. Unlike, for example, the state $|12\rangle |ab\rangle + |21\rangle |ba\rangle$, where detecting a particle in the state 1 tells us the label is a and vice versa. So while this state does not display the statistics we would expect

from a state of indistinguishable bosons, there does not seem to be a way to discern the two particles from each other. However, we are comparing it to the state in Equation (5.48), which is possibly not a fair comparison. After tracing out the Label modes, the state looks like a mixed state of indistinguishable bosons (Fock states on System modes). This does not seem to answer the problem of defining the distinguishability of this state, because we now introduce the complication of how exactly is this mixed state created. If it is possible to create it out of indistinguishable bosons, then there should be a pure state in this full System-Label picture to represent it properly.

Further, we cannot change the occupation of the Label Fock marginal, without applying some unitary action on it, but we already defined that the only action we are allowing are the ones that can be carried out on the System. Therefore, if we are given a state with starting occupation $\underline{n}_{\mathrm{L}}$, this will stay the same throughout. This does not mean we cannot have all the possible irreps and the inner multiplicities that match this occupation – indeed, see Example 5.3.2 for a state with occupation $\underline{n}_{\mathrm{L}} = (1,1)$ in both symmetric and antisymmetric subspace of the Label degree of freedom. However, in our last example we have two different occupations for our label states, $\underline{n}_{\mathrm{L},1} = (2,0)$ and $\underline{n}_{\mathrm{L},2} = (0,2)$. While it is not impossible to imagine how to construct such a state, its meaning for the definition of distinguishability, especially in the manner we are focusing on, is less clear.

- In Examples 5.3.1, 5.3.2, 5.3.3, and 5.3.4, the statistics before application of a balanced beamsplitter would all show the presence of two particles in two spatial modes, "1" and "2", that is $P_{(1,1)} = 1$. After applying the balanced beamsplitter, we see that Examples 5.3.1 and 5.3.3 are the only ones that are still consistent with probabilities that we would expect to see from fully symmetric particles. It is interesting that in the Example 5.3.3, the particles behave as we would expect for identical bosons, even though there is more than one label in the Label degree of freedom. We conclude if we have a state with no entanglement (the two degrees of freedom are decoupled), with support only on the symmetric subspace, this state will behave as a state of indistinguishable System bosons.

- From Example 5.3.4 we find another state with two decoupled degrees of freedom with fermion statistics, so it is possible that more generally these decoupled states act as indistinguishable particles (that is, lack of entanglement might imply indistinguishability, just not necessarily boson indistinguishability).

- From Example 5.3.5 we see that not all states that only have symmetric support necessairily pass the HOM experiment.

- Due to unitary-unitary duality, the presence of any other mixed symmetry state in addition to the symmetric state will imply entanglement between two degrees of freedom. In some situations when there is entanglement between the visible and hidden degree of freedom, we can then extract some information about the labelling of the particles from the measurements on the visible degree of freedom.

In this thesis we mostly guide ourselves with the idea that the states with no entanglement present and with support only in the fully symmetric subspace are those of indistinguishable bosons. Understanding existing metrics of distinguishability and defining a good metric of distinguishability in the model presented here is left as future work.

[1] S. Stanisic, "Universal quantum computation by linear optics." unpublished, 2015.

[2] S. Stanisic, N. Linden, A. Montanaro, and P. S. Turner, "Generating entanglement with linear optics," *Physical Review A*, vol. 96, no. 4, p. 043861, 2017.

[3] S. Stanisic, N. Linden, A. Montanaro, and P. S. Turner, "Data from entanglement in linear optics." https://doi.org/10.5523/bris.1o09f3qf75gnc2018ko8kxklnf, 2017.

[4] S. Stanisic and P. S. Turner, "Discriminating distinguishability," *Physical Review A*, vol. 98, no. 4, p. 043839, 2018.

[5] S. Stanisic and P. S. Turner, "Data from distinguishability." https://doi.org/10.5523/bris.3bj7o4rqo2kxd2f8nnfdx5gw6p, 2018.

[6] J. Adcock, E. Allen, M. Day, S. Frick, J. Hinchliff, M. Johnson, S. Morley-Short, S. Pallister, A. Price, and S. Stanisic, "Advances in quantum machine learning," *arXiv:1512.02900*, 2015.

[7] P. Sibson, J. E. Kennard, S. Stanisic, C. Erven, J. L. O'Brien, and M. G. Thompson, "Integrated silicon photonics for high-speed quantum key distribution," *Optica*, vol. 4, no. 2, p. 172, 2017.

[8] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. OBrien, "Quantum computers," *Nature*, vol. 464, no. 7285, p. 45, 2010.

[9] A. Montanaro, "Quantum algorithms: an overview," *npj Quantum Information*, vol. 2, p. 15023, 2016.

[10] J. L. O'Brien, A. Furusawa, and J. Vučković, "Photonic quantum technologies," *Nature Photonics*, vol. 3, no. 12, p. 687, 2010.

[11] C.-K. Hong, Z.-Y. Ou, and L. Mandel, "Measurement of subpicosecond time intervals between two photons by interference," *Physical Review Letters*, vol. 59, no. 18, p. 2044, 1987.

[12] J. L. O'Brien, "Optical quantum computing," *Science*, vol. 318, no. 5856, p. 1567, 2007.

[13] T. Rudolph, "Why I am optimistic about the silicon-photonic route to quantum computing," *APL Photonics*, vol. 2, no. 3, p. 030901, 2017.

[14] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature*, vol. 409, no. 6816, p. 46, 2001.

[15] R. Raussendorf and H. J. Briegel, "A One-Way Quantum Computer," *Physical Review Letters*, vol. 86, no. 22, p. 5188, 2001.

[16] D. E. Browne and T. Rudolph, "Resource-efficient linear optical quantum computation," *Physical Review Letters*, vol. 95, no. 1, p. 010501, 2004.

[17] K. Kieling, T. Rudolph, and J. Eisert, "Percolation, renormalization, and quantum computing with nondeterministic gates," *Physical Review Letters*, vol. 99, no. 13, p. 130501, 2007.

[18] M. Gimeno-Segovia, *Towards practical linear optical quantum computing*. PhD thesis, Imperial College London, 2015.

[19] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, 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*, vol. 349, no. 6249, p. 711, 2015.

[20] M. Streshinsky, R. Ding, Y. Liu, A. Novack, C. Galland, A.-J. Lim, P. G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "The road to affordable, large-scale silicon photonics," *Optics and Photonics News*, vol. 24, no. 9, pp. 32–39, 2013.

[21] J. W. Silverstone, D. Bonneau, J. L. O'Brien, and M. G. Thompson, "Silicon quantum photonics," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 22, no. 6, pp. 390–402, 2016.

[22] T. Baehr-Jones, T. Pinguet, P. L. Guo-Qiang, S. Danziger, D. Prather, and M. Hochberg, "Myths and rumours of silicon photonics," *Nature Photonics*, vol. 6, no. 4, p. 206, 2012.

[23] Y. Li, P. C. Humphreys, G. J. Mendoza, and S. C. Benjamin, "Resource costs for fault-tolerant linear optical quantum computing," *Physical Review X*, vol. 5, no. 4, p. 041007, 2015.

[24] R. H. Hadfield, "Single-photon detectors for optical quantum information applications," *Nature Photonics*, vol. 3, no. 12, p. 696, 2009.

[25] M. Gimeno-Segovia, P. Shadbolt, D. E. Browne, and T. Rudolph, "From three-photon Greenberger–Horne–Zeilinger states to ballistic universal quantum computation," *Physical Review Letters*, vol. 115, no. 2, p. 020502, 2015.

[26] M. C. Tichy, "Interference of identical particles from entanglement to boson-sampling," *Journal of Physics B: Atomic, Molecular and Optical Physics*, vol. 47, no. 10, p. 103001, 2014.

[27] M. C. Tichy, "Sampling of partially distinguishable bosons and the relation to the multidimensional permanent," *Physical Review A*, vol. 91, no. 2, p. 022316, 2015.

[28] A. J. Menssen, A. E. Jones, B. J. Metcalf, M. C. Tichy, S. Barz, W. S. Kolthammer, and I. A. Walmsley, "Distinguishability and many-particle interference," *Physical Review Letters*, vol. 118, no. 15, p. 153603, 2017.

[29] V. S. Shchesnovich, "Partial indistinguishability theory for multiphoton experiments in multiport devices," *Physical Review A*, vol. 91, no. 1, p. 013844, 2015.

[30] V. S. Shchesnovich, "Partial distinguishability and photon counting probabilities in multiport devices," *arXiv:1712.03191*, 2017.

[31] M. Tillmann, S. H. Tan, S. E. Stoeckl, B. C. Sanders, H. De Guise, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, "Generalized multiphoton quantum interference," *Physical Review X*, vol. 5, no. 4, p. 041015, 2015.

[32] V. Tamma and S. Laibacher, "Multiboson correlation interferometry with arbitrary single-photon pure states," *Physical Review Letters*, vol. 114, no. 24, p. 243601, 2015.

[33] P. P. Rohde, "Boson sampling with photons of arbitrary spectral structure," *Physical Review A*, vol. 91, no. 1, p. 012307, 2015.

[34] J. J. Renema, A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley, "Efficient classical algorithm for boson sampling with partially distinguishable photons," *Physical Review Letters*, vol. 120, no. 22, p. 220502, 2018.

[35] T. Brünner, G. Dufour, A. Rodrıguez, and A. Buchleitner, "Signatures of indistinguishability in bosonic many-body dynamics," *Physical Review Letters*, vol. 120, no. 21, p. 210401, 2018.

[36] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, "Linear optical quantum computing with photonic qubits," *Reviews of Modern Physics*, vol. 79, no. 1, p. 135, 2007.

[37] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, "Experimental realization of any discrete unitary operator," *Physical Review Letters*, vol. 73, no. 1, p. 58, 1994.

[38] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, "Optimal design for universal multiport interferometers," *Optica*, vol. 3, no. 12, p. 1460, 2016.

[39] H. de Guise, O. Di Matteo, and L. L. Sánchez-Soto, "Simple factorization of unitary transformations," *Physical Review A*, vol. 97, no. 2, p. 022328, 2018.

[40] A. Bouland and S. Aaronson, "Generation of universal linear optics by any beam splitter," *Physical Review A*, vol. 89, no. 6, p. 062316, 2014.

[41] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nature Photonics*, vol. 5, no. 4, p. 222, 2011.

[42] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Reviews of Modern Physics*, vol. 74, no. 1, p. 145, 2002.

[43] M. Cattaneo, M. G. A. Paris, and S. Olivares, "Hybrid quantum key distribution using coherent states and photon-number-resolving detectors," *Physical Review A*, vol. 98, no. 1, p. 012333, 2018.

[44] S. Aaronson and A. Arkhipov, "The computational complexity of linear optics," in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, pp. 333–342, ACM, 2011.

[45] S. Scheel, "Permanents in linear optical networks," *arXiv:0406127*, 2004.

[46] M. C. Tichy, *Entanglement and interference of identical particles*. PhD thesis, The University of Freiburg, 2013.

[47] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, pp. 212–219, 1996.

[48] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM review*, vol. 41, no. 2, p. 303, 1999.

[49] J. Preskill, "Quantum Computing: Pro and Con," in *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 469–486, The Royal Society, 1998.

[50] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Physical review A*, vol. 52, no. 4, p. R2493, 1995.

[51] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum Error Correction via Codes over GF(4)," *IEEE Transactions on Information Theory*, vol. 44, no. 4, p. 1369, 1998.

[52] D. P. DiVincenzo, "The physical implementation of quantum computation," *Fortschritte der Physik: Progress of Physics*, vol. 48, no. 9-11, p. 771, 2000.

[53] P. Kok, *Lecture notes on Optical Quantum Computing.* Springer, 2010.

[54] Y. Shih and C. O. Alley, "New type of einstein-podolsky-rosen-bohm experiment using pairs of light quanta produced by optical parametric down conversion," *Physical Review Letters*, vol. 61, no. 26, p. 2921, 1988.

[55] Z. Ou and L. Mandel, "Violation of bell's inequality and classical probability in a two-photon correlation experiment," *Physical review letters*, vol. 61, no. 1, p. 50, 1988.

[56] S. Popescu, L. Hardy, and M. Żukowski, "Revisiting bell's theorem for a class of down-conversion experiments," *Physical Review A*, vol. 56, no. 6, p. R4353, 1997.

[57] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, "Surface codes: Towards practical large-scale quantum computation," *Physical Review A*, vol. 86, no. 3, p. 032324, 2012.

[58] R. Raussendorf and J. Harrington, "Fault-tolerant quantum computation with high threshold in two dimensions," *Physical Review Letters*, vol. 98, no. 19, p. 190504, 2007.

[59] N. Cerf, C. Adami, and P. Kwiat, "Optical simulation of quantum logic," *Physical Review A*, vol. 57, no. 3, p. R1477(R), 1998.

[60] J. F. Clauser and J. P. Dowling, "Factoring integers with Young's N-slit interferometer," *Physical Review A*, vol. 53, no. 6, p. 4587, 1996.

[61] P. G. Kwiat, J. R. Mitchell, P. D. D. Schwindt, and A. G. G. White, "Grover's search algorithm: An optical approach," *Journal of Modern Optics*, vol. 47, no. 2-3, p. 257, 2000.

[62] D. Gottesman and I. L. Chuang, "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations," *Nature*, vol. 402, no. 6760, p. 390, 1999.

[63] E. Knill, "Quantum gates using linear optics and postselection," *Physical Review A*, vol. 66, no. 5, p. 052306, 2002.

[64] F. M. Spedalieri, H. Lee, and J. P. Dowling, "High-fidelity linear optical quantum computing with polarization encoding," *Physical Review A*, vol. 73, no. 1, p. 012334, 2006.

[65] R. Raussendorf, "Measurement-based quantum computation with cluster states," *International Journal of Quantum Information*, vol. 7, no. 06, p. 1053, 2009.

[66] N. Yoran and B. Reznik, "Deterministic linear optics quantum computation with single photon qubits," *Physical Review Letters*, vol. 91, no. 3, p. 037903, 2003.

[67] M. A. Nielsen, "Optical quantum computation using cluster states," *Physical Review Letters*, vol. 93, no. 4, p. 040503, 2004.

[68] D. G. Glynn, "The permanent of a square matrix," *European Journal of Combinatorics*, vol. 31, no. 7, p. 1887, 2010.

[69] A. Nijenhuis and H. S. Wilf, *Combinatorial algorithms: for computers and calculators*. Elsevier, 2014.

[70] A. Neville, C. Sparrow, R. Clifford, E. Johnston, P. M. Birchall, A. Montanaro, and A. Laing, "Classical boson sampling algorithms with superior performance to near-term experiments," *Nature Physics*, vol. 13, no. 12, p. 1153, 2017.

[71] P. Clifford and R. Clifford, "The classical complexity of boson sampling," in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 146–155, Society for Industrial and Applied Mathematics, 2018.

[72] L. Gurvits, "On the complexity of mixed discriminants and related problems," in *International Symposium on Mathematical Foundations of Computer Science*, pp. 447–458, Springer, 2005.

[73] A. Arkhipov and G. Kuperberg, "The bosonic birthday paradox," *Geometry & Topology Monographs*, vol. 18, p. 1, 2012.

[74] L. Latmiral, N. Spagnolo, and F. Sciarrino, "Towards quantum supremacy with lossy scattershot boson sampling," *New Journal of Physics*, vol. 18, no. 11, p. 113008, 2016.

[75] H. Wang, Y. He, Y.-H. Li, Z.-E. Su, B. Li, H.-L. Huang, X. Ding, M.-C. Chen, C. Liu, J. Qin, J.-P. Li, Y.-M. He, C. Schneider, M. Kamp, C.-Z. Peng, S. Höfling, C.-Y. Lu, and P. Jian-Wei, "High-efficiency multiphoton boson sampling," *Nature Photonics*, vol. 11, no. 6, p. 361, 2017.

[76] H. Wang, W. Li, X. Jiang, Y.-M. He, Y.-H. Li, X. Ding, M.-C. Chen, J. Qin, C.-Z. Peng, C. Schneider, M. Kamp, W.-J. Zhang, H. Li, L.-X. You, Z. Wang, J. P. Dowling, S. Höfling, C.-Y. Lu, and J.-W. Pan, "Toward scalable boson sampling with photon loss," *Physical Review Letters*, vol. 120, no. 23, p. 230502, 2018.

[77] M. Bentivegna, N. Spagnolo, C. Vitelli, F. Flamini, N. Viggianiello, L. Latmiral, P. Mataloni, D. J. Brod, E. F. Galvão, A. Crespi, R. Ramponi, R. Osellame, and F. Sciarrino, "Experimental scattershot boson sampling," *Science Advances*, vol. 1, no. 3, 2015.

[78] S. Aaronson and D. J. Brod, "Bosonsampling with lost photons," *Physical Review A*, vol. 93, no. 1, p. 012335, 2016.

[79] M. Walschaers, "Signatures of many-particle interference," *arXiv:1908.08370*, 2019.

[80] D. J. Brod, E. F. Galvão, A. Crespi, R. Osellame, N. Spagnolo, and F. Sciarrino, "Photonic implementation of boson sampling: a review," *Advanced Photonics*, vol. 1, no. 3, p. 034001, 2019.

[81] A. E. Moylett, R. García-Patrón, J. J. Renema, and P. S. Turner, "Classically simulating near-term partially-distinguishable and lossy boson sampling," *Quantum Science and Technology*, 2019.

[82] D. Bonneau, G. J. Mendoza, J. L. O'Brien, and M. G. Thompson, "Effect of loss on multiplexed single-photon sources," *New Journal of Physics*, vol. 17, no. 4, p. 043057, 2015.

[83] S. Morley-Short, *Towards realistic architectures for linear optical quantum computing.* PhD thesis, University of Bristol, 2018.

[84] M. Oszmaniec and D. J. Brod, "Classical simulation of photonic linear optics with lost particles," *New Journal of Physics*, vol. 20, no. 9, p. 092002, 2018.

[85] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Single-photon sources and detectors," *Review of Scientific Instruments*, vol. 82, no. 7, p. 071101, 2011.

[86] J. G. Rarity and P. R. Tapster, "Fourth-order interference in parametric downconversion," *JOSA B*, vol. 6, no. 6, p. 1221, 1989.

[87] J. G. Rarity, P. R. Tapster, E. Jakeman, T. Larchuk, R. A. Campos, M. C. Teich, and B. E. A. Saleh, "Two-photon interference in a Mach-Zehnder interferometer," *Physical Review Letters*, vol. 65, no. 11, p. 1348, 1990.

[88] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental quantum teleportation," *Nature*, vol. 390, no. 6660, p. 575, 1997.

[89] T. E. Keller, M. H. Rubin, and Y. Shih, "Two-photon interference from separate pulses," *Physics Letters A*, vol. 244, no. 6, p. 507, 1998.

[90] C. Santori, D. Fattal, J. Vučković, G. S. Solomon, and Y. Yamamoto, "Indistinguishable photons from a single-photon device," *Nature*, vol. 419, no. 6907, p. 594, 2002.

[91] H. de Riedmatten, I. Marcikic, W. Tittel, H. Zbinden, and N. Gisin, "Quantum interference with photon pairs created in spatially separated sources," *Physical Review A*, vol. 67, no. 2, p. 022301, 2003.

[92] R. Kaltenbaek, B. Blauensteiner, M. Żukowski, M. Aspelmeyer, and A. Zeilinger, "Experimental interference of independent photons," *Physical Review Letters*, vol. 96, no. 24, p. 240502, 2006.

[93] S. P. Walborn, A. N. De Oliveira, S. Pádua, and C. H. Monken, "Multimode Hong–Ou–Mandel interference," *Physical Review Letters*, vol. 90, no. 14, p. 143601, 2003.

[94] R. Lopes, A. Imanaliev, A. Aspect, M. Cheneau, D. Boiron, and C. I. Westbrook, "Atomic Hong–Ou–Mandel experiment," *Nature*, vol. 520, no. 7545, p. 66, 2015.

[95] A. M. Kaufman, B. J. Lester, C. M. Reynolds, M. L. Wall, M. Foss-Feig, K. R. A. Hazzard, A. M. Rey, and C. A. Regal, "Two-particle quantum interference in tunnel-coupled optical tweezers," *Science*, vol. 345, no. 6194, pp. 306–309, 2014.

[96] R. B. Patel, A. J. Bennett, K. Cooper, P. Atkinson, C. A. Nicoll, D. A. Ritchie, and A. J. Shields, "Postselective two-photon interference from a continuous nonclassical stream of photons emitted by a quantum dot," *Physical Review Letters*, vol. 100, no. 20, p. 207405, 2008.

[97] S. Ates, S. M. Ulrich, S. Reitzenstein, A. Löffler, A. Forchel, and P. Michler, "Post-selected indistinguishable photons from the resonance fluorescence of a single quantum dot in a microcavity," *Physical Review Letters*, vol. 103, no. 16, p. 167402, 2009.

[98] E. B. Flagg, A. Muller, S. V. Polyakov, A. Ling, A. Migdall, and G. S. Solomon, "Interference of single photons from two separate semiconductor quantum dots," *Physical Review Letters*, vol. 104, no. 13, p. 137401, 2010.

[99] R. B. Patel, A. J. Bennett, I. Farrer, C. A. Nicoll, D. A. Ritchie, and A. J. Shields, "Two-photon interference of the emission from electrically tunable remote quantum dots," *Nature Photonics*, vol. 4, no. 9, p. 632, 2010.

[100] J. Beugnon, M. P. Jones, J. Dingjan, B. Darquié, G. Messin, A. Browaeys, and P. Grangier, "Quantum interference between two single photons emitted by independently trapped atoms," *Nature*, vol. 440, no. 7085, p. 779, 2006.

[101] 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," *Nature Physics*, vol. 3, no. 8, p. 538, 2007.

[102] A. Kiraz, M. Ehrl, T. Hellerer, O. E. Müstecaploğlu, C. Bräuchle, and A. Zumbusch, "Indistinguishable photons from a single molecule," *Physical Review Letters*, vol. 94, no. 22, p. 223602, 2005.

[103] H. Bernien, L. Childress, L. Robledo, M. Markham, D. Twitchen, and R. Hanson, "Two-photon quantum interference from separate nitrogen vacancy centers in diamond," *Physical Review Letters*, vol. 108, no. 4, p. 043604, 2012.

[104] Y. Y. Gao, B. J. Lester, Y. Zhang, C. Wang, S. Rosenblum, L. Frunzio, L. Jiang, S. M. Girvin, and R. J. Schoelkopf, "Programmable interference between two microwave quantum memories," *Physical Review X*, vol. 8, no. 2, p. 021073, 2018.

[105] G. Di Martino, Y. Sonnefraud, M. S. Tame, S. Kéna-Cohen, F. Dieleman, c. K. Özdemir, M. S. Kim, and S. A. Maier, "Observation of quantum interference in the plasmonic Hong–Ou–Mandel effect," *Physical Review Applied*, vol. 1, no. 3, p. 034004, 2014.

[106] C. Lang, C. Eichler, L. Steffen, J. M. Fink, M. J. Woolley, A. Blais, and A. Wallraff, "Correlations, indistinguishability and entanglement in Hong–Ou–Mandel experiments at microwave frequencies," *Nature Physics*, vol. 9, no. 6, p. 345, 2013.

[107] E. Bocquillon, V. Freulon, J. M. Berroir, P. Degiovanni, B. Plaçais, A. Cavanna, Y. Jin, and G. Fève, "Coherence and indistinguishability of single electrons emitted by independent sources," *Science*, vol. 339, no. 6123, p. 1054, 2013.

[108] T. Rom, T. Best, D. Van Oosten, U. Schneider, S. Fölling, B. Paredes, and I. Bloch, "Free fermion antibunching in a degenerate atomic Fermi gas released from an optical lattice," *Nature*, vol. 444, no. 7120, p. 733, 2006.

[109] M. Iannuzzi, A. Orecchini, F. Sacchetti, P. Facchi, and S. Pascazio, "Direct experimental evidence of free-fermion antibunching," *Physical Review Letters*, vol. 96, no. 8, p. 080402, 2006.

[110] R. C. Liu, B. Odom, Y. Yamamoto, and S. Tarucha, "Quantum interference in electron collision," *Nature*, vol. 391, no. 6664, p. 263, 1998.

[111] J.-W. Pan, Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski, "Multiphoton entanglement and interferometry," *Reviews of Modern Physics*, vol. 84, no. 2, p. 777, 2012.

[112] G. Weihs, M. Reck, H. Weinfurter, and A. Zeilinger, "Two-photon interference in optical fiber multiports," *Physical Review A*, vol. 54, no. 1, p. 893, 1996.

[113] R. A. Campos, "Three-photon Hong–Ou–Mandel interference at a multiport mixer," *Physical Review A*, vol. 62, no. 1, p. 013809, 2000.

[114] M. Koniorczyk and J. Janszky, "Photon number conservation and photon interference," in *First International Workshop on Classical and Quantum Interference*, vol. 4888, pp. 1–9, International Society for Optics and Photonics, 2002.

[115] S.-H. Tan, Y. Y. Gao, H. de Guise, and B. C. Sanders, "SU (3) quantum interferometry with single-photon input pulses," *Physical Review Letters*, vol. 110, no. 11, p. 113603, 2013.

[116] H. de Guise, S.-H. Tan, I. P. Poulin, and B. C. Sanders, "Coincidence landscapes for three-channel linear optical networks," *Physical Review A*, vol. 89, no. 6, p. 063819, 2014.

[117] N. Spagnolo, C. Vitelli, L. Aparo, P. Mataloni, F. Sciarrino, A. Crespi, R. Ramponi, and R. Osellame, "Three-photon bosonic coalescence in an integrated tritter," *Nature Communications*, vol. 4, p. 1606, 2013.

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

[119] Z. Chaboyer, T. Meany, L. G. Helt, M. J. Withford, and M. J. Steel, "Tunable quantum interference in a 3D integrated circuit," *Scientific reports*, vol. 5, p. 9601, 2015.

[120] S. Mährlein, J. Von Zanthier, and G. S. Agarwal, "Complete three photon Hong–Ou–Mandel interference at a three port device," *Optics Express*, vol. 23, no. 12, p. 15833, 2015.

[121] J. Wu, H. de Guise, and B. C. Sanders, "Coincidence landscapes for polarized bosons," *Physical Review A*, vol. 98, no. 1, p. 013817, 2018.

[122] S. Agne, T. Kauten, J. Jin, E. Meyer-Scott, J. Z. Salvail, D. R. Hamel, K. J. Resch,

G. Weihs, and T. Jennewein, "Observation of genuine three-photon interference," *Physical Review Letters*, vol. 118, no. 15, p. 153602, 2017.

[123] P. Kolenderski, A. Raczyński, J. Zaremba, and S. Zielińska-Raczyńska, "Three-photon interference with stored light," *Physical Review A*, vol. 96, no. 6, p. 063809, 2017.

[124] Y. L. Lim and A. Beige, "Generalized Hong–Ou–Mandel experiments with bosons and fermions," *New Journal of Physics*, vol. 7, no. 1, p. 155, 2005.

[125] R. A. Campos, B. E. A. Saleh, and M. C. Teich, "Quantum-mechanical lossless beam splitter: SU (2) symmetry and photon statistics," *Physical Review A*, vol. 40, no. 3, p. 1371, 1989.

[126] F. Laloë and W. J. Mullin, "Quantum properties of a single beam splitter," *Foundations of Physics*, vol. 42, no. 1, p. 53, 2012.

[127] Z. Y. Ou, "Temporal distinguishability of an N-photon state and its characterization by quantum interference," *Physical Review A*, vol. 74, no. 6, p. 063808, 2006.

[128] Z. Y. Ou, "Characterizing temporal distinguishability of an N-photon state by a generalized photon bunching effect with multiphoton interference," *Physical Review A*, vol. 77, no. 4, p. 043829, 2008.

[129] Y.-S. Ra, M. C. Tichy, H.-T. Lim, O. Kwon, F. Mintert, A. Buchleitner, and Y.-H. Kim, "Nonmonotonic quantum-to-classical transition in multiparticle interference," in *Proceedings of the National Academy of Sciences*, vol. 110, pp. 1227–1231, National Acad Sciences, 2013.

[130] M. C. Tichy, H.-T. Lim, Y.-S. Ra, F. Mintert, Y.-H. Kim, and A. Buchleitner, "Four-photon indistinguishability transition," *Physical Review A*, vol. 83, no. 6, p. 062111, 2011.

[131] P. P. Rohde, W. Mauerer, and C. Silberhorn, "Spectral structure and decompositions of optical states, and their applications," *New Journal of Physics*, vol. 9, no. 4, p. 91, 2007.

[132] V. S. Shchesnovich, "Partial indistinguishability theory for multiphoton experiments in multiport devices," *Physical Review A*, vol. 91, no. 1, p. 013844, 2015.

[133] M. C. Tichy, "Sampling of partially distinguishable bosons and the relation to the multidimensional permanent," *Physical Review A*, vol. 91, no. 2, p. 022316, 2015.

[134] V. S. Shchesnovich, "Sufficient condition for the mode mismatch of single photons for scalability of the boson-sampling computer," *Physical Review A*, vol. 89, no. 2, p. 022333, 2014.

[135] V. Tamma, "Sampling of bosonic qubits," *International Journal of Quantum Information*, vol. 12, no. 07n08, p. 1560017, 2014.

[136] V. S. Shchesnovich, "Tight bound on the trace distance between a realistic device with partially indistinguishable bosons and the ideal bosonsampling," *Physical Review A*, vol. 91, no. 6, p. 063842, 2015.

[137] H. de Guise, D. Spivak, J. Kulp, and I. Dhand, "D-functions and immanants of unitary matrices and submatrices," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 9, p. 09LT01, 2016.

[138] I. Dhand, B. C. Sanders, and H. de Guise, "Algorithms for SU (n) boson realizations and D-functions," *Journal of Mathematical Physics*, vol. 56, no. 11, p. 111705, 2015.

[139] A. Khalid, D. Spivak, B. C. Sanders, and H. de Guise, "Permutational symmetries for coincidence rates in multimode multiphotonic interferometry," *Physical Review A*, vol. 97, no. 6, p. 063802, 2018.

[140] V. S. Shchesnovich and M. E. O. Bezerra, "Collective phases of identical particles interfering on linear multiports," *Physical Review A*, vol. 98, no. 3, p. 033805, 2018.

[141] D. J. Brod, E. F. Galvão, N. Viggianiello, F. Flamini, N. Spagnolo, and F. Sciarrino, "Witnessing genuine multi-photon indistinguishability," *Physical Review Letters*, vol. 122, no. 6, p. 063602, 2019.

[142] T. Giordani, D. J. Brod, C. Esposito, N. Viggianiello, M. Romano, F. Flamini, G. Carvacho, N. Spagnolo, E. F. Galvão, and F. Sciarrino, "Experimental quantification of genuine four-photon indistinguishability," *arXiv:1907.01325*, 2019.

[143] E. F. Galvão and D. J. Brod, "Quantum and classical bounds for unknown two-state overlaps," *arXiv:1902.11039*, 2019.

[144] M. C. Tichy and K. Mølmer, "Extending exchange symmetry beyond bosons and fermions," *Physical Review A*, vol. 96, no. 2, p. 022119, 2017.

[145] M. C. Tichy, M. Tiersch, F. de Melo, F. Mintert, and A. Buchleitner, "Zero-transmission law for multiport beam splitters," *Physical Review Letters*, vol. 104, no. 22, p. 220405, 2010.

[146] C. Dittel, G. Dufour, M. Walschaers, G. Weihs, A. Buchleitner, and R. Keil, "Totally destructive many-particle interference," *Physical Review Letters*, vol. 120, no. 24, p. 240404, 2018.

[147] M. C. Tichy, M. Tiersch, F. Mintert, and A. Buchleitner, "Many-particle interference beyond many-boson and many-fermion statistics," *New Journal of Physics*, vol. 14, no. 9, p. 093015, 2012.

[148] A. Crespi, "Suppression laws for multiparticle interference in Sylvester interferometers," *Physical Review A*, vol. 91, no. 1, p. 013811, 2015.

[149] A. Perez-Leija, R. Keil, A. Kay, H. Moya-Cessa, S. Nolte, L.-C. Kwek, B. M. Rodríguez-Lara, A. Szameit, and D. N. Christodoulides, "Coherent quantum transport in photonic lattices," *Physical Review A*, vol. 87, no. 1, p. 012309, 2013.

[150] S. Weimann, A. Perez-Leija, M. Lebugle, R. Keil, M. Tichy, M. Gräfe, R. Heilmann, S. Nolte, H. Moya-Cessa, G. Weihs, D. N. Christodoulides, and A. Szameit, "Implementation of quantum and classical discrete fractional fourier transforms," *Nature Communications*, vol. 7, p. 11027, 2016.

[151] C. Dittel, R. Keil, and G. Weihs, "Many-body quantum interference on hypercubes," *Quantum Science and Technology*, vol. 2, no. 1, p. 015003, 2017.

[152] C. Dittel, G. Dufour, M. Walschaers, G. Weihs, A. Buchleitner, and R. Keil, "Totally destructive interference for permutation-symmetric many-particle states," *Physical Review A*, vol. 97, no. 6, p. 062116, 2018.

[153] M. Saffman, T. G. Walker, and K. Mølmer, "Quantum information with Rydberg atoms," *Reviews of Modern Physics*, vol. 82, no. 3, p. 2313, 2010.

[154] Z.-l. Xiang, S. Ashhab, J. Q. You, and F. Nori, "Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems," *Reviews of Modern Physics*, vol. 85, no. 2, p. 623, 2013.

[155] R. Jozsa and N. Linden, "On the role of entanglement in quantum-computational speed-up," in *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 459, pp. 2011–2032, 2003.

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

[157] J. S. Bell, "On the Einstein–Podolsky–Rosen paradox," *Physics*, vol. 1, no. 3, p. 195, 1964.

[158] A. Aspect, "Proposed experiment to test separable hidden-variable theories," *Physics Letters A*, vol. 54, no. 2, p. 117, 1975.

[159] S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," *Foundations of Physics*, vol. 24, no. 3, p. 379, 1994.

[160] J. Joo, P. L. Knight, J. L. O'Brien, and T. Rudolph, "One-way quantum computation with four-dimensional photonic qudits," *Physical Review A*, vol. 76, no. 5, p. 052326, 2007.

[161] Q. Zhang, X.-H. Bao, C.-Y. Lu, X.-Q. Zhou, T. Yang, T. Rudolph, and J.-W. Pan, "Demonstration of a scheme for the generation of "event-ready" entangled photon pairs from a single-photon source," *Physical Review A*, vol. 77, no. 6, p. 062316, 2008.

[162] S. Aaronson and A. Arkhipov, "The computational complexity of linear optics," *Theory of Computing*, vol. 9, no. 4, p. 143, 2013.

[163] F. Benatti, R. Floreanini, and U. Marzolino, "Bipartite entanglement in systems of identical particles: the partial transposition criterion," *Annals of Physics*, vol. 327, no. 5, p. 1304, 2012.

[164] F. Benatti, R. Floreanini, and U. Marzolino, "Entanglement robustness and geometry in systems of identical particles," *Physical Review A*, vol. 85, no. 4, p. 042329, 2012.

[165] J. C. Adcock, *Generating Optical Graph Statesscomputing*. PhD thesis, University of Bristol, 2019.

[166] E. Knill, "Bounds on the probability of success of postselected nonlinear sign shifts implemented with linear optics," *Physical Review A*, vol. 68, no. 6, p. 064303, 2003.

[167] S. Scheel and N. Lütkenhaus, "Upper bounds on success probabilities in linear optics," *New Journal of Physics*, vol. 6, no. 1, p. 51, 2004.

[168] J. Eisert, "Optimizing linear optics quantum gates," *Physical Review Letters*, vol. 95, no. 4, p. 040502, 2005.

[169] S. Scheel and K. M. R. Audenaert, "Scaling of success probabilities for linear optics gates," *New Journal of Physics*, vol. 7, no. 1, p. 149, 2005.

[170] D. B. Uskov, L. Kaplan, A. M. Smith, S. D. Huver, and J. P. Dowling, "Maximal

success probabilities of linear-optical quantum gates," *Physical Review A*, vol. 79, no. 4, p. 042326, 2009.

[171] N. M. VanMeter, P. Lougovski, D. B. Uskov, K. Kieling, J. Eisert, and J. P. Dowling, "General linear-optical quantum state generation scheme: Applications to maximally path-entangled states," *Physical Review A*, vol. 76, no. 6, p. 063808, 2007.

[172] K. Kieling, *Linear optics quantum computing–construction of small networks and asymptotic scaling.* PhD thesis, Imperial College London, 2008.

[173] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," 2001–2017.

[174] W. Vogel and J. Sperling, "Unified quantification of nonclassicality and entanglement," *Physical Review A*, vol. 89, no. 5, p. 052302, 2014.

[175] J. P. Olson, K. R. Motes, P. M. Birchall, N. M. Studer, M. LaBorde, T. Moulder, P. P. Rohde, and J. P. Dowling, "Linear optical quantum metrology with single photons — Experimental errors, resource counting, and quantum Cramér-Rao bounds," *Physical Review A*, vol. 96, no. 1, p. 013810, 2017.

[176] M. Horodecki, "Entanglement Measures," *Quantum Information & Computation*, vol. 1, no. 1, p. 3, 2001.

[177] A. Hurwitz, "Ueber die erzeugung der invarianten durch integration," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1897, p. 71, 1897.

[178] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Reviews of Modern Physics*, vol. 84, no. 2, p. 621, 2012.

[179] 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*, vol. 339, no. 6121, p. 798, 2013.

[180] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, "Photonic boson sampling in a tunable circuit," *Science*, vol. 339, no. 6121, p. 794, 2013.

[181] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, "Integrated multimode interferometers with arbitrary designs for photonic boson sampling," *Nature Photonics*, vol. 7, no. 7, p. 545, 2013.

[182] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, "Experimental boson sampling," *Nature Photonics*, vol. 7, no. 7, p. 540, 2013.

[183] N. Viggianiello, F. Flamini, M. Bentivegna, N. Spagnolo, A. Crespi, D. J. Brod, R. Osellame, and F. Sciarrino, "Optimal photonic indistinguishability tests in multimode networks," *Science Bulletin*, vol. 63, no. 22, p. 1470, 2018.

[184] S. Popescu, "Knill–Laflamme–Milburn linear optics quantum computation as a

measurement-based computation," *Physical Review Letters*, vol. 99, no. 25, p. 250501, 2007.

[185] R. B. A. Adamson, P. S. Turner, M. W. Mitchell, and A. M. Steinberg, "Detecting hidden differences via permutation symmetries," *Physical Review A*, vol. 78, no. 3, p. 033832, 2008.

[186] A. E. Moylett and P. S. Turner, "Quantum simulation of partially distinguishable boson sampling," *Physical Review A*, vol. 97, no. 6, p. 062329, 2018.

[187] D. J. Rowe, M. J. Carvalho, and J. Repka, "Dual pairing of symmetry and dynamical groups in physics," *Reviews of Modern Physics*, vol. 84, no. 2, p. 711, 2012.

[188] B.-G. Englert, D. Kaszlikowski, L. C. Kwek, and W. H. Chee, "Wave-particle duality in multi-path interferometers: general concepts and three-path interferometers," *International Journal of Quantum Information*, vol. 6, no. 01, p. 129, 2008.

[189] J. C. Garcia-Escartin and P. Chamorro-Posada, "Swap test and Hong-Ou-Mandel effect are equivalent," *Physical Review A*, vol. 87, no. 5, p. 052330, 2013.

[190] L. Rigovacca, C. Di Franco, B. J. Metcalf, I. A. Walmsley, and M. S. Kim, "Non-classicality criteria in multiport interferometry," *Physical Review Letters*, vol. 117, no. 21, p. 213602, 2016.

[191] H. Weyl, *The classical groups: their invariants and representations*, vol. 1. Princeton university press, 1946.

[192] H. Weyl, *The theory of groups and quantum mechanics*. Courier Corporation, 1950.

[193] W. Fulton and J. Harris, *Representation theory: a first course*, vol. 129. Springer, 1991.

[194] S. Sternberg, *Group theory and physics*. Cambridge University Press, 1995.

[195] R. Goodman and N. R. Wallach, *Representations and invariants of the classical groups*, vol. 68. Cambridge University Press, 2000.

[196] A. W. Harrow, *Applications of coherent classical communication and the Schur transform to quantum information theory*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2005.

[197] M. Christandl, *The structure of bipartite quantum states-insights from group theory and cryptography*. PhD thesis, University of Cambridge, Cambridge, UK, 2006.

[198] A. W. Harrow, "The church of the symmetric subspace," *arXiv:1308.6595*, 2013.

[199] M. Keyl and R. F. Werner, "Estimating the spectrum of a density operator," in *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*, pp. 458–467, World Scientific, 2005.

[200] K. Matsumoto and M. Hayashi, "Universal distortion-free entanglement concentration," *Physical Review A*, vol. 75, no. 6, p. 062338, 2007.

[201] M. Hayashi, "Exponents of quantum fixed-length pure-state source coding," *Physical Review A*, vol. 66, no. 3, p. 032321, 2002.

[202] P. P. Rohde, "Optical quantum computing with photons of arbitrarily low fidelity and purity," *Physical Review A*, vol. 86, no. 5, p. 052321, 2012.

[203] D. E. Littlewood and A. R. Richardson, "Group characters and algebra," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 233, p. 99, 1934.

[204] B. Kostant, "Immanant inequalities and 0-weight spaces," *Journal of the American Mathematical Society*, vol. 8, no. 1, p. 181, 1995.

[205] R. Bhatia, *Matrix analysis. Graduate Texts in Mathematics.* Springer-Verlag, 1997.

[206] A. Chefles, "Unambiguous discrimination between linearly independent quantum states," *Physics Letters A*, vol. 239, no. 6, p. 339, 1998.

[207] J. A. Bergou, U. Herzog, and M. Hillery, "11 discrimination of quantum states," in *Quantum state estimation*, pp. 417–465, Springer, 2004.

[208] S. M. Barnett and S. Croke, "Quantum state discrimination," *Advances in Optics and Photonics*, vol. 1, no. 2, p. 238, 2009.

[209] C. W. Helstrom, "Quantum detection and estimation theory," *Journal of Statistical Physics*, vol. 1, no. 2, p. 231, 1969.

[210] J. Bae and L.-C. Kwek, "Quantum state discrimination and its applications," *Journal of Physics A: Mathematical and Theoretical*, vol. 48, no. 8, p. 083001, 2015.

[211] I. D. Ivanovic, "How to differentiate between non-orthogonal states," *Physics Letters A*, vol. 123, no. 6, p. 257, 1987.

[212] D. Dieks, "Overlap and distinguishability of quantum states," *Physics Letters A*, vol. 126, no. 5-6, p. 303, 1988.

[213] A. Peres, "How to differentiate between non-orthogonal states," *Physics Letters A*, vol. 128, no. 1-2, p. 19, 1988.

[214] G. Jaeger and A. Shimony, "Optimal distinction between two non-orthogonal quantum states," *Physics Letters A*, vol. 197, no. 2, p. 83, 1995.

[215] A. Chefles, "Quantum state discrimination," *Contemporary Physics*, vol. 41, no. 6, p. 401, 2000.

[216] J. A. Bergou, "Quantum state discrimination and selected applications," in *Journal of Physics: Conference Series*, vol. 84, p. 012001, IOP Publishing, 2007.

[217] T. Rudolph, R. W. Spekkens, and P. S. Turner, "Unambiguous discrimination of mixed states," *Physical Review A*, vol. 68, no. 1, p. 010301, 2003.

[218] S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, "Maximum confidence quantum measurements," *Physical Review Letters*, vol. 96, no. 7, p. 070401, 2006.

[219] U. Herzog and J. A. Bergou, "Distinguishing mixed quantum states: Minimum-error discrimination versus optimum unambiguous discrimination," *Physical Review A*, vol. 70, no. 2, p. 022302, 2004.

[220] J. C. F. Matthews, K. Poulios, J. D. A. Meinecke, A. Politi, A. Peruzzo, N. Ismail, K. Wörhoff, M. G. Thompson, and J. L. O'Brien, "Observing fermionic statistics with photons in arbitrary processes," *Scientific reports*, vol. 3, p. 1539, 2013.