DOCTOR OF PHILOSOPHY

State preparation and some applications in quantum optics within the context of quantum information theory

Kok, Pieter

Award date:
2001

Awarding institution:
Bangor University

State Preparation

and

Some Applications

in

Quantum Optics

within

the context

of

Quantum Information Theory

by Pieter Kok

PhD thesis, University of Wales, Bangor.

State Preparation in Quantum Optics

Pieter Kok

How I need a drink, alcoholic of course, after
the heavy chapters involving quantum mechanics.

—Mnemonic for the first fifteen digits of π.

List of Figures

Contents

Summary

Entanglement is perhaps the single-most important resource of quantum information theory. The first part of this thesis deals with the creation of optical event-ready entanglement with a specific class of optical circuits. These circuits include passive components such as beam-splitters and phase-shifters, and active components such as optical parametric down-converters and optical squeezers. Furthermore, the entangled-state preparation may be conditioned on one or more detector outcomes. In this context, I discuss the statistics of down-converters and give a quantitative comparison between realistic detectors and detector cascades, using the *confidence* of the detection. The outgoing states of the optical circuits can be expressed in terms of multi-dimensional Hermite polynomials. Event-ready entanglement cannot be created when the outgoing state is conditioned on two detected photons. For six detected photons using ideal photo-detectors a scheme is known to exist.

Part two of this thesis includes two applications of optical entanglement. First, I discuss quantum teleportation and entanglement swapping using down-conversion. It is shown that higher-order photon-pair production degrades the fidelity of the teleported (or swapped) states. The interpretation of these states proved controversial, and I have attempted to settle this controversy. As a second application, quantum lithography uses optical ('which-way') entanglement of multiple photons to beat the classical diffraction limit. Given a suitable photo-resist, this technique results in sub-wavelength optical resolution and can be used to write features much smaller than is possible with classical lithography. I present classes of states which can be used to create patterns in one and two dimensions with sub-wavelength resolution.

Acknowledgements

I could not have written this thesis without the help of many people. I would like to express my gratitude towards Samuel L. Braunstein for his excellent supervision, and my colleagues Peter van Loock and Arun Pati for the many discussions we've had over the years. Special thanks to prof. Rajiah Simon, who helped me understand the maths of operators and maps, Jonathan Dowling for inviting us to work on quantum lithography and for offering me a job, and Apy Vourdas for the fruitful discussions on multi-dimensional Hermite polynomials. I also thank all the others whom I have had the pleasure to meet and who have generally made me wiser.

During the past three years I have often found consolation and playful adversity from my friends in the Netherlands, especially from Jasper for his long-distance moral and immoral support, and Alex, Gijs and Maarten. The RISK-club rules! Many thanks also to Angèle (mooi, die middeleeuwse kastelen van voor de Renaissance), Jeroen (Slaatje kameraadje), Joep (d'r zit 'n haar in m'n glas) and Klaas-Jan ("mobile construction yard deployed").

I have had an unforgettable time in Bangor, the responsibility of which can be traced almost entirely to my friends here. Many thanks to Barbara (my own drama-queen), Mireia (fly, my pretty), Consuelo (è chica!), Ana (when will we dance again?), Claudia (Mel G. meets tequila), Carlotta (close that cupboard), Martin (*you* are the best man), Marc (I'll see you in Cambridge), Ross (our man in Havana) and Johnny (watch out for that bottle!): pray I don't publish my memoirs ...

Finally, I thank my parents and my brother Joost (architectuur = kunst).

Pieter Kok,
December 2000.

I

INTRODUCTION

The closing decade of the twentieth century has witnessed the coming-of-age of a new field, called *quantum information theory*. This field includes the development of quantum computation and quantum communication. At this point a fully scalable quantum computer has not been built, but there are numerous experimental and theoretical proposals to achieve this [36]. At the same time, the quest for quantum algorithms continues. So far, we have Shor's algorithm to factor large numbers into primes [156], the Deutsch-Jozsa algorithm [47, 48] and Grovers search algorithm [74]. The possibility of quantum error correction was discovered [159], which is very important to any practical application of quantum computation.

Considerable progress has also been made in quantum communication. It is a generic term for communication protocols based on quantum mechanical principles and includes cryptography [12, 57], teleportation [14], entanglement swapping [183], dense coding [13], quantum clock synchronisation [92], entanglement purification [15] and quantum networks [75, 58]. Another recent application of quantum mechanics is quantum lithography [22]. The common divisor of nearly all elements of quantum information is quantum *entanglement* [34]. In this thesis I study the creation of entanglement in quantum optics, and some of its applications.

This introduction will provide the motivation and physical background for the thesis. I discuss entanglement, teleportation and lithography. It will be largely non-mathematical and aimed at an audience of non-specialists. The subsequent chapters will then develop these issues in a rigorous mathematical way.

1 QUANTUM ENTANGLEMENT

In order to explain what quantum entanglement is about, I will first discuss the double slit experiment as presented by Richard Feynman [61]. Suppose we have a gun firing bullets at a screen with two holes which are close to each other. Most of the bullets will hit the screen and fall on the floor, but some of them will pass through the holes and hit a wall of clay. In effect, this wall records the position of impact of the bullets which passed through the holes.

When we inspect the wall, we will see that the bullets are spread around the centre of the clay wall in a straight line behind the gun and the holes in the

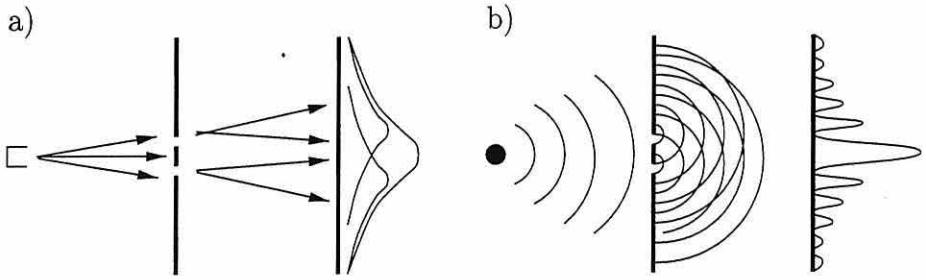

Figure I.1: The double slit experiment a) with particles and b) with waves.

screen. Each bullet must have passed through either hole to make it to the wall. When we record the process with a high-speed camera we can see the bullets going through the holes. In fact, we can mount paint sprayers next to the holes, colouring the bullets which pass through the left hole red, and the bullets which pass through the right hole blue. The clay wall will be peppered with red and blue bullets, with the red bullets shifted slightly to the left and the blue bullets slightly to the right (see figure I.1a).

Let us now repeat this experiment with waves instead of bullets. Suppose we have a shallow tray of water with a screen containing two narrow openings close to each other at the waterline. On one side of the screen a pin is moving up and down in the water, creating a wave which spreads out in all directions. When the wave reaches the screen, the two slits start to act as if they were vertically moving pins *themselves*! The slits thus create *two* waves which spread out in all directions behind the screen.

These two waves will soon start to interfere: when a wave-crest meets another crest, the result will be a crest twice as high; when a trough meets another trough, the result will be a trough twice as deep. And finally, when a crest meets a trough they cancel each other. When we record the vertical displacement of the water at the far end of the tray of water, we will find an *interference* pattern[1] of peaks and troughs (see figure I.1b).

The difference with bullets is obvious: the bullets arrive in a spread area with its bullet density falling off uniformly with the distance from the centre, whereas waves will show an intensity pattern which rises and falls in alternation with increasing distance from the centre. The simple (classical) picture is: waves give interference and particles (bullets) don't.

Now let's take a look at light. Suppose we have again a screen with two slits, and a laser which is aimed at the slits. The light which passes through the slits is recorded on a photographic plate. After development, we will see an interference pattern on the photographic plate: light seems to be a wave.

[1]More precisely, the interference pattern is given by the square of the displacement: the *intensity*.

When we weaken the intensity of the laser enough, we will see (using very sensitive equipment) that the light is no longer a continuous stream, but that instead, it is a succession of small 'bursts'. We call these bursts *photons*, and they are described by quantum theory. It thus seems that light consists of particles which interact to give an interference pattern just like waves.

We now attenuate the laser so that we fire individual photons at the double slit. This way, the photons cannot interact to give an interference pattern since at any time there is only one photon travelling between the laser and the photographic plate. The photons which make it past the double slit will give a dot on the photographic plate, analogous to the bullets in the clay wall. If the photons are truly classical particles, they should pass through either slit, just like the bullets, and they should *not* make an interference pattern.

However, when we develop the photographic plate after a long exposure time we *do* find an interference pattern! We have set up this experiment in such a way that the photons, which seem to behave like particles (indivisible, giving dots on a screen), are passing the slits one at a time so they don't interact with each other. The only way to get an interference pattern is thus when the photon somehow interferes with *itself*. Has the photon gone through both slits simultaneously? Let's test this.

Again, we fire individual photons at a double slit and record the pattern on a photographic plate. But this time we place a detector behind both slits. These detectors tell us through which slit the photon passes. While running this experiment, the detectors are clicking when a photon passes through its corresponding slit, giving us information about the paths of the successive photons. They really go through one slit at a time.

But when we now develop the photographic plate, the interference pattern has gone! Instead, we have a concentration of dots, its density decreasing with increasing distance from the centre. This is the bullet pattern. Apparently, when we *know* through which slit the photons pass, we do not get an interference pattern. When we do not look, it is *meaningless* to ask through which slit the photons pass. In describing the path of the photons without detection, we need to include both possibilities: the path is a *superposition* of going through the left and the right slit.

Feynman elevated this to a general principle: when an event can occur in several different ways, we need to describe the event in terms of a superposition of these ways [61]. The superposition principle is responsible for many of the counterintuitive aspects of quantum mechanics. This simple thought-experiment thus takes us straight into the heart of the theory.

Let's now consider entanglement. Photons have an extra internal property called *polarisation*. A photon which reflects off this paper towards your eye (which can be represented graphically as $\odot$) vibrates in the plane of the paper perpendicular to the direction of travelling ($\updownarrow$ or $\leftrightarrow$, or a combination of these

two. Technically, we also have circular polarisation). The polarisation of the photon is determined by the angle of this vibration direction.

When we want to *measure* the polarisation of a photon, we place a polarising beam-splitter, or *polariser* in the path of the photon. This is essentially a piece of glass which reflects horizontally polarised photons and transmits vertically polarised photons. When we place photo-detectors in the paths of reflected and transmitted photons, a detector click will tell us the polarisation of that photon. When a horizontally (vertically) polarised photon encounters the polariser, it will always be reflected (transmitted). But what if the photon has a diagonal polarisation?

When a diagonally polarised photon encounters the polariser it will be either reflected or transmitted. We can only make a probabilistic prediction as to which path the photon will take. When we *rotate* the polariser so that its horizontal orientation is turned parallel to the (diagonal) polarisation of the photon, the photon will be reflected *with certainty*. We now consider two polarised photons.

Suppose we have two photons originating from a common source and heading off in opposite directions. One photon is received by Alice, and the other by Bob. Furthermore, Alice and Bob are far away from each other, possibly in different galaxies.

First, we consider the case where both photons are horizontally polarised ($\leftrightarrow$). When Alice and Bob measure the polarisation of the photon in the horizontal and vertical direction using polarisers, both will find with certainty that the photons have horizontal polarisation. When Alice rotates her polariser by 45 degrees, the probability that either detector signals the detection of a photon is one half. This situation is similar to the measurement of a single photon since the photons received by Alice and Bob behave completely independent from each other.

Now suppose that the two photons are prepared in the following way: either Alice's photon is horizontally polarised and Bob's photon is vertically polarised, *or* Alice's photon is vertically polarised and Bob's photon is horizontally polarised. Furthermore, the photons are prepared in a *superposition* of these two possibilities. When Alice and Bob measure the polarisation of these photons they will find that their photons always have opposite polarisations: when Alice detects a horizontally polarised photon, Bob will find a vertically polarised photon and vice versa. This means that given a measurement outcome, Bob knows what Alice's measurement outcome will be, even though she might be light years away. The measurement results are said to be correlated.

So far, nothing strange has happened. We know these correlations from classical physics. Suppose Alice and Bob meet in Amsterdam. They blindly draw a marble from a vase containing only one black and one white marble. Alice travels to New York and Bob travels to Tokyo. When Alice looks at her marble and finds that it is white, she immediately knows that Bob's marble is black. These outcomes are also correlated.

There is, however, a difference in the case of polarised photons. Suppose

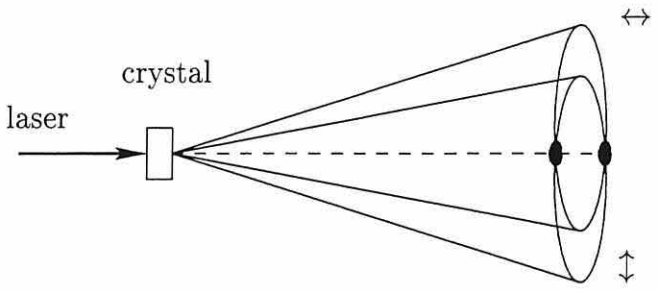

Figure I.2: A down-converter.

Alice and Bob both rotate their polariser over 45 degrees. According to the classical picture, both photons have a 50:50 chance to end up in either detector. That means that with 50% probability the photons have equal polarisation. But this is not what they find: Alice and Bob *always* find that they have opposite polarisations! Clearly, this is not just a classical correlation. The two photons are said to be *entangled.*

The question is now: how can we make entangled photons? One way of doing it is to use a so-called down-converter. In a down-converter, a high-powered laser is sent into a special crystal. A photon of the laser interacts with the crystal and breaks up into two photons with half the energy. The photons will travel away from the central axis (defined by the path of the laser light) under a fixed angle. The photons thus travel on the surface of a *cone* originating from the crystal (see figure I.2).

Furthermore, we can set up the down-converter in such a way that the photons have opposite polarisations. This is where the crystal performs its special trick: the refraction index of the crystal is different for horizontally and vertically polarised photons. This means that the cone corresponding to the possible paths of horizontally polarised photons is tilted upwards from the central axis. Similarly, the cone for vertically polarised photons is tilted slightly downwards.

Due to momentum conservation, the two photons are always travel on different cones along lines opposite of each other with respect to the central axis. The cones intersect each other at two opposite lines, and as a consequence, we find a photon in one of those lines *if and only if* there is a photon in the other line. Furthermore, we cannot tell to which cone the photons on the intersecting lines belong. Either the left photon belongs to the upper cone and the right photon to the lower, or the other way around. By virtue of Feynman's principle we have to take the superposition of these two possibilities.

The down-converter only produces two entangled photons probabilistically; not every laser pulse results in two down-converted photons. Furthermore, since we select only the intersection of the two cones, we lose all the instances where

photons were not produced along the intersecting lines. This means that most of the time we fire the laser into the crystal we do not produce entanglement. In this thesis I study whether and how we can minimise the number of cases where no photons are produced.

2 TELEPORTATION

Another subject of this thesis is quantum teleportation, in particular the teleportation of a photon. In this procedure, the (unknown) polarisation of one photon is transferred to another photon far away. It is not true that the photon itself is magically transported from Alice to Bob, only the polarisation direction (or, more generally, the *state* of the photon) is transferred.

The same is true for other types of matter: we can teleport atoms, but that does not mean we can make an atom appear somewhere in the distance. I will now present the general protocol, using polarised photons.

Suppose Alice received a photon with a polarisation direction which is unknown to her. We assume that she has some way of storing it without disturbance. In other words, she has a device called a 'quantum memory'. Now she wants to transfer the polarisation direction of the photon to Bob. When she has only measurements and a telephone at her disposal to tell Bob the results, she has a problem. Since she doesn't know what the polarisation direction is, she cannot choose her polariser to be parallel to this direction. Therefore, when she measures the polarisation of the photon in a chosen direction and tells Bob the result, his reconstruction of the polarisation direction will generally be off by a certain angle. Faithful teleportation cannot be performed this way.

However, the story changes when Alice and Bob share entanglement (produced, for example, by the down-converter of the previous section). Alice and Bob both hold one part of an entangled photon pair. Remember that these photons are correlated: whatever the polarisation direction measured by Alice, Bob will always find the opposite polarisation. Alice proceeds by making a *joint* measurement of her part of the entanglement and the the incoming photon with unknown polarisation. Such a measurement does not give any information about the individual photons, but determines the relation of the photons relative to each other. It is a carefully chosen measurement which will correlate the incoming photon with unknown polarisation to Alice's part of the entangled photon-pair. In the case of polarisation we consider here, the outcome of Alice's measurement has four possible outcomes. These outcomes correspond to four different ways the two photons can be correlated (technically, we have four *orthogonal* ways).

Let's pause for a second to contemplate the current state of affairs. Alice has just correlated the unknown incoming photon with her half of the entangled photon-pair. She cannot choose or predict how she correlates them, every one of the four possibilities is equally likely. But her half of the entangled photon-pair is

already correlated with the other half. It then follows that the unknown incoming photon is now correlated with Bob's half of the entangled pair.

The only thing Bob does not know is the exact nature of the correlation. Every one of the four possible correlations will give a different polarisation direction in Bob's photon. That is why Alice has to tell him. She picks up the phone and gives one of four possibilities (i.e., she sends two classical bits) corresponding to her measurement outcome. The beautiful thing about teleportation is now that Bob has to perform a polarisation rotation corresponding to the measurement outcome, which is *independent of the unknown polarisation direction of the incoming photon!* His photon now has the same polarisation direction as the incoming photon and teleportation is complete.

We have to note three things. First of all, neither Alice, nor Bob gains any information about the direction of the polarisation of the incoming photon. Secondly, no photon magically appears at Bob's site; he already held a photon during the whole procedure. In this respect, quantum teleportation is quite unlike the Star Trek version. And finally, teleportation cannot be used for superluminal signalling. If Alice does not tell Bob her measurement outcome (which is a classical message and thus restricted by the speed of light), teleportation will fail since Bob does not know what polarisation rotation he has to perform.

The entanglement shared by Alice and Bob is typically produced with a down-converter. In this thesis I study the effects of the down-conversion characteristics on the quality of teleportation.

3 LITHOGRAPHY

The second application of optical quantum entanglement I study in this thesis is quantum lithography. This technique may be used to write components on micro-chips which are smaller than possible with classical optical lithography. It works as follows.

Consider again the double slit experiment with photons, given in section 1. When photons are fired at the slits one at a time without looking through which slit they pass, we obtain an interference pattern on the photographic plate. In this case the photon can travel along two possible paths: either through the left slit or through the right.

Suppose we now fire two photons per shot at the slits. If we assume that all photons pass the slits we now have three possible paths: both photons may pass through the left slit; one may pass through the left and the other though the right; or both may pass through the right slit. In this case the interference pattern will be twice as bright, because we use twice as much light.

But now we can ask what happens if we suppress one of these three possible paths. What will the interference pattern look like when the photons *do not separate*, that is, what happens when the photons either both pass through the left slit, or both through the right slit? The answer is that the interference

pattern, which is an array of bright and dark lines, will become twice as narrow: the distance between two bright lines is halved. The reason why this happens is because the photons 'stick together', thus effectively acting as a single particle with twice the momentum. The De Broglie wavelength (which determines the line spacing of the interference pattern) is inversely proportional to the momentum of the particle. The higher the momentum, the shorter this wavelength and the narrower the interference pattern. Since photons also have momentum, this means that the more photons we can make acting as a single particle, the smaller the interference pattern. Note that in general, we need a special surface which is sensitive to two photons (a 'two-photon resist') to record these patterns.

Classically, light cannot resolve features which are much smaller than its wavelength. As a consequence, classical optical lithography, in which light is used to etch a surface, cannot write features much smaller that its wavelength. This is Rayleigh's diffraction limit. It is derived from the interference between two waves. We have seen that we can narrow this interference pattern using the quantum properties of light, which means that the Rayleigh limit is a *classical* limit. Quantum lithography can therefore be used to create sub-wavelength patterns, to be used in, for example, the micro-chip industry.

So far, quantum lithography is still a theoretical method. Only the two-photon case described above has been experimentally tested. It is not easy to see how more exotic patterns may be produced, and what the requirements for the surface are. Nevertheless, it gives us a new insight in the nature of light. In this thesis I study how we can create arbitrary sub-wavelength patterns in one and two dimensions.

4 Thesis outline

This thesis is organised in two parts. The first part, called 'Quantum State Preparation' is divided in four chapters. Chapter II gives the general quantum mechanical background. It includes the postulates of quantum mechanics, the quantisation of the electro-magnetic field and some topics from quantum information theory such as the Von Neumann entropy and the fidelity.

In chapter III, I study a limited set of optical circuits for creating near maximal polarisation entanglement *without* the usual large vacuum contribution. The optical circuits I consider involve passive interferometers, feed-forward detection, down-converters and squeezers. For input vacuum fields the creation of maximal entanglement using such circuits is impossible when conditioned on two detected auxiliary photons. Furthermore, I derive the statistical properties of down-converters and show that coincidences between photon-pairs from parametric down-conversion automatically probe the non-Poissonian structure of these sources.

So far, the photo-detectors I considered are ideal. In chapter IV I study the use of detection devices in entanglement-based state preparation. In particu-

lar I consider realistic optical detection devices such as single-photon sensitivity detectors, single-photon resolution detectors and detector cascades (with a limited efficiency). I develop an extensive theory for the use of these devices. In entanglement-based state preparation we perform measurements on subsystems, and we therefore need precise bounds on the distinguishability of these measurements. To this end, I introduce the *confidence* of preparation, which may also be used to quantify the performance of detection devices in entanglement-based preparation. I give a general expression for detector cascades of arbitrary size for the detection up to two photons. I show that, contrary to the general belief, cascading does not give a practical advantage over detectors with single-photon resolution in entanglement-based state preparation.

Finally, in chapter V, I study a special class of optical circuits and show that the *outgoing state* leaving the optical circuit can be expressed in terms of so-called multi-dimensional Hermite polynomials and give their recursion and orthogonality relations. I show how quantum teleportation of photon polarisation can be modelled using this description.

The second part is called 'Some Applications' and covers two chapters. In chapter VI, I study the experimental realisation of quantum teleportation as performed by Bouwmeester *et al.* [23] and the adjustments to it suggested by Braunstein and Kimble [32]. These suggestions include the employment of a detector cascade and a relative slow-down of one of the two down-converters. Furthermore, I discuss entanglement swapping and the creation of GHZ states within this context.

Chapter VII gives the theory of quantum lithography. I generalise the lithography procedure in order to create patterns in one and two dimensions. This renders quantum lithography a potentially useful tool in nano-technology.

Quantum State Preparation

II

Quantum Theory

This chapter presents the mathematical background theory for the understanding of this thesis. It does not contain new results. First, I present quantum mechanics in the Hilbert space formalism. The second section discusses the quantisation of the electro-magnetic field and quantum optics, and in the last section I treat some aspects of quantum information theory, such as entropy, fidelity and non-locality.

1 Quantum mechanics in a nutshell

What is quantum mechanics all about? Initially, the theory was developed to describe the physical world of atoms, that is, to explain the observed spectral lines in spectrometers. In 1913, Niels Bohr, then at the Cavendish laboratory in Cambridge, developed what is now called the 'old quantum theory', in which he presented a model explaining the spectral lines of hydrogen [18, 82]. As the theory was developed further (culminating in the work of Schrödinger and subsequently Heisenberg) it became clear that quantum mechanics is a mathematical theory which describes *measurement outcomes*, rather than the underlying physical processes [131].

Von Neumann proved the equivalence of Schrödingers wave mechanics and Heisenberg's matrix mechanics in *Mathematical foundations of quantum mechanics* [122] and introduced the Hilbert space formalism still in use today. Dirac [51] developed his own version of the theory (of which, incidentally, Von Neumann did not approve[1]), and his bracket notation has become the standard. In accordance with the convention, I will follow Von Neumann's framework and use Dirac's bracket notation.

This section is organised as follows: first I will give the postulates for quantum mechanics. Then I discuss mixed states and composite systems. Finally, this section ends with measurement theory according to Von Neumann and its generalisation to projection operator valued measures (POVM's).

a *The postulates of quantum mechanics*

Using some properties of complex vector spaces (see appendix A), we can formulate the postulates of quantum mechanics [28, 45, 85]. More properties of operators on Hilbert spaces can be found in appendix B.

[1] The preface of his book makes very enjoyable reading.

Postulate 1 For every physical system there is a corresponding Hilbert space $\mathscr{H}$. The accessible (pure) states of the system are completely determined by rays with unit length in $\mathscr{H}$.

A ray in Hilbert space is a set of unit vectors which differ only by an arbitrary (complex) phase. A physical state corresponds to a ray. Thus, a state $|\psi\rangle$ is physically equivalent to $e^{i\varphi}|\psi\rangle$ with $0 \leq \varphi \leq 2\pi$. A complete set of orthonormal states (rays) form a basis of $\mathscr{H}$. I will use the terms 'ray', 'vector' and 'state' interchangeably, while remembering that an overall phase does not change the physical state. Later, in section c, this class of states is extended to *mixed* states.

Since the Hilbert space $\mathscr{H}$ is a (complex) vector space, if any two normalised rays $|\psi\rangle$ and $|\phi\rangle$ in $\mathscr{H}$ are accessible states to the system, then their superposition $\alpha|\psi\rangle + \beta|\phi\rangle$ is also an accessible state to the system. Normalisation then requires $|\alpha|^2 + |\beta|^2 = 1$. A superposition of this type is sometimes called a *coherent* superposition. Note that in this case the phase of a ray *does* have a physical meaning. Consider two orthonormal states $|\psi\rangle$ and $|\phi\rangle$ (i.e., $\langle\psi|\psi\rangle = \langle\phi|\phi\rangle = 1$ and $\langle\psi|\phi\rangle = 0$), and consider the two superpositions

$$\begin{aligned} |v_1\rangle &= \frac{1}{\sqrt{2}}\left(|\psi\rangle + |\phi\rangle\right) , \\ |v_2\rangle &= \frac{1}{\sqrt{2}}\left(|\psi\rangle - |\phi\rangle\right) , \end{aligned} \tag{II.1}$$

then it is easy to verify that $\langle v_1|v_2\rangle = 0$, i.e., they are orthogonal (in fact, they are orthonormal). The two superpositions differ only in a *relative* phase, but they yield two physically distinct (orthonormal) states.

Postulate 2 For every physical observable of the system there is a unique corresponding self-adjoint (Hermitian) operator A in $\mathscr{H}$.

An operator A is called self-adjoint if and only if $A^\dagger = A$. An operator is Hermitian if and only if its eigenvalues are real. I will now prove that (for finite-dimensional Hilbert spaces) any operator is self-adjoint if and only if it is Hermitian.

To show that self-adjointness implies Hermiticity, observe that according to the eigenvalue equation $A|\psi\rangle = \alpha|\psi\rangle$ in Eq. (A.1) we have

$$\langle\psi|A|\psi\rangle = \alpha \quad \text{and} \quad \langle\psi|A^\dagger|\psi\rangle = \alpha^* . \tag{II.2}$$

Substituting $A^\dagger = A$ immediately yields $\alpha^* = \alpha$, i.e., a real eigenvalue. The second implication is proved by running the argument backwards. In this thesis I will use the terms Hermitian and self-adjoint interchangeably. Also, I will use the convention that Greek letters $(\alpha, \beta \ldots)$ denote complex numbers and Roman letters $(a, b \ldots)$ denote real numbers. The fact that self-adjoint operators have real eigenvalues lead to the next postulate.

Postulate 3 The only possible measurement outcomes obtainable from the measurement of an observable are the eigenvalues of its corresponding self-adjoint operator A. If the state of the system is $|\psi\rangle$, then the probability $p(a_i)$ of finding the d_{a_i}-fold degenerate eigenvalue a_i of the observable A is equal to the probability of finding the system in the corresponding eigenspace:

$$p(a_i) = \sum_{j=1}^{d_{a_i}} |\langle \alpha_{ij}|\psi\rangle|^2 \,, \tag{II.3}$$

where $|\alpha_{ij}\rangle$ are the eigenvectors corresponding to the d_{a_i}-fold degenerate eigenvalue a_i.

The outcome of a measurement in the laboratory can only yield a real number, and since the measurement outcomes are the eigenvalues of operators, these operators must be Hermitian. Postulate 2 ensures that there is a one-to-one correspondence between self-adjoint operators and physical observables, and postulate 3 determines the possible measurement outcomes for these observables. Eq. (II.3) is the so-called *Born rule* [85].

Postulate 4 The evolution of a system is governed by a unitary transformation U:

$$|\psi(\theta')\rangle = U(\theta,\theta')|\psi(\theta)\rangle \,, \tag{II.4}$$

where θ and θ' are (vectors of) real parameters.

Any operator U on a Hilbert space $\mathscr{H}$ for which $U^\dagger = U^{-1}$ is called a *unitary* operator on $\mathscr{H}$. In general, every unitary operator can be written as

$$U(\theta) = \exp\left(iA\theta\right) \,, \tag{II.5}$$

with A a self-adjoint operator on $\mathscr{H}$. To prove this statement, note that $U^\dagger = \exp(-iA^\dagger\theta) = \exp(-iA\theta)$. Thus $U^\dagger U = UU^\dagger = \mathbb{1}$, and $U^\dagger = U^{-1}$. Unitary operators in matrix representation always have determinant 1, and they can be viewed as rotations in a complex vector space.

A special choice for $U(\theta,\theta')$ in Eq. (II.4) is the infinitesimal time evolution where $\theta = t$ and $\theta' = t + dt$:

$$U(t,t+dt) = \exp\left[iHt/\hbar - iH(t+dt)/\hbar\right] = \exp(-iHdt/\hbar) \,, \tag{II.6}$$

and H the *Hamiltonian* of the system. It is the observable associated with the total energy of the system. Substituting this evolution into Eq. (II.4) and neglecting higher-order powers of dt, we obtain in the Taylor expansion

$$|\psi(t+dt)\rangle \;=\; e^{-iHdt/\hbar}|\psi(t)\rangle \quad \Leftrightarrow$$

$$\begin{aligned} |\psi(t)\rangle + d|\psi(t)\rangle &= \left(1 - \frac{i}{\hbar}Hdt\right)|\psi(t)\rangle \quad \Leftrightarrow \\ i\hbar\frac{d}{dt}|\psi(t)\rangle &= H|\psi(t)\rangle \,. \end{aligned} \tag{II.7}$$

This is the famous Schrödinger equation.

In this thesis I will not use the Schrödinger equation. Instead, I will use the fact that the self-adjoint operator A in Eq. (II.5) acts as a *generator* of the group of unitary evolutions parametrised by θ. This approach will have great benefits in chapter III. The theory of (Lie) groups and their generators is treated in appendix C.

Unitary transformations not only govern the evolution of quantum states, they also constitute basis transformations. If A and U are a linear operator and a unitary transformation on $\mathscr{H}$ respectively, then there exist another linear operator A' on $\mathscr{H}$ such that

$$A' = U^\dagger A U \,. \tag{II.8}$$

In particular, if A is self-adjoint there always exist a unitary transformation such that A' is diagonal.

Postulate 5 When a measurement of an observable A yields the (non-degenerate) eigenvalue a_i, the state of the system immediately after the measurement will be the eigenstate $|\alpha_i\rangle$ corresponding to a_i.

This is the so-called *projection* postulate. It is often referred to as state collapse, since a measurement can induce a discontinuous jump from a superposition to an eigenstate of the measured observable. This postulate has caused severe problems for interpretations of quantum mechanics which assign some form of 'reality' to the state. Such interpretations suffer from what has become generically known as the 'measurement problem' [141, 85]. In this thesis I will ignore this problem, since it does not seem to have any effect on the experimental success of quantum mechanics[2]. The general theory of measurements is discussed in section *d*.

So far, I have presented quantum mechanics in the so-called *Schrödinger picture*. In this picture the time dependence is captured in the *state*: $|\psi(t)\rangle$. Alternatively, we can choose the states to be time independent, and have all the time dependence in the *operators*. This is called the *Heisenberg picture*:

$$A_{\mathrm{H}} = U^\dagger(t) A_{\mathrm{S}} U(t) \,, \tag{II.9}$$

[2]The reader should note that, although I will not discuss the measurement problem, this does not imply that there *is* no measurement problem. This is still very much open to debate [66].

where A_{H} denotes the operator in the Heisenberg picture and A_{S} the operator in the Schrödinger picture. When part of the time dependence is in the states and part is in the operators, we speak of the *interaction picture.*

Another alternative formulation of quantum mechanics is Feynman's *path integral formalism.* This is particularly useful in the formulation of quantum field theories, but I will not discuss it here.

b *The linear harmonic oscillator*

One application of quantum mechanics which deserves attention in the context of this thesis is the description of the linear harmonic oscillator. I will treat this in a telegraphic manner, since this is a well known example. For a full derivation see, for example, Merzbacher [119].

We start by defining a quadratic potential $V(x)$ for a classical particle with mass m, position x and momentum p:

$$V(x) = \frac{1}{2} m\omega^2 x^2 \,, \tag{II.10}$$

where ω is, loosely speaking, the classical frequency of the oscillator. The classical Hamiltonian is then given by the sum of the kinetic and potential energy:

$$H_{\text{classical}} = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} \,. \tag{II.11}$$

In quantum mechanics the observables x and p have to be replaced by self-adjoint operators. This procedure is called 'quantisation'[3]. The quantum mechanical Hamiltonian thus becomes

$$H_{\text{quantum}} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2 x^2}{2} \,, \tag{II.12}$$

where $p \to \hat{p} = -i\hbar d/dx$ and $x \to \hat{x} = x$ with $[\hat{x}, \hat{p}] = i\hbar$. When the quantum mechanical state of the harmonic oscillator is denoted by $\psi(x) = \langle x|\psi\rangle$, with $|x\rangle$ the position eigenvector corresponding to the position x, then we obtain the differential (Schrödinger) equation (see postulate 4):

$$\frac{d^2\psi(x)}{dx^2} - \left(\frac{m\omega}{\hbar}\right)^2 x^2\, \psi(x) = E\psi(x) \,. \tag{II.13}$$

This equation is satisfied by the following class of wave-functions:

$$\psi_n(x) = 2^{-n/2}(n!)^{-1/2} \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right) H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) \,, \tag{II.14}$$

[3] Or *first* quantisation. Indeed, there is something called 'second' quantisation, in which the fields are written in the operators formalism. We will encounter this in section 2, where I introduce quantum optics and the quantisation of the electro-magnetic field.

corresponding to energies

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) . \tag{II.15}$$

The $H_n(x)$ are the so-called *Hermite* polynomials (see appendix G).

The Hamiltonian of the harmonic oscillator can also be expressed in terms of so-called *raising* and *lowering* operators $\hat{a}^\dagger$ and $\hat{a}$ respectively:

$$\hat{a}^\dagger \equiv \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - i\frac{\hat{p}}{m\omega}\right) \qquad \text{and} \qquad \hat{a} \equiv \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + i\frac{\hat{p}}{m\omega}\right) \tag{II.16}$$

(remember that $\hat{x}^\dagger = \hat{x}$ and $\hat{p}^\dagger = \hat{p}$ since position and momentum are physical observables). It is easily found that $[\hat{a}, \hat{a}^\dagger] = 1$. The eigenstate corresponding to the energy E_n of the linear harmonic oscillator is now symbolically denoted by $|n\rangle$, and we have

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle\,, \quad \hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad \text{and} \quad \hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle\,. \tag{II.17}$$

The operator $\hat{a}^\dagger\hat{a}$ in the last equation is also called the *number* operator $\hat{n}$. The Hamiltonian of the linear harmonic oscillator in terms of the raising and lowering operators is then given by

$$H_{\text{LHO}} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) . \tag{II.18}$$

The raising and lowering operators will return in section 2 as creation and annihilation operators.

c *Composite and mixed states*

After this brief, but necessary digression I now return to the definition of states of *composite* systems. Postulate 1 tells us that with every physical system corresponds a Hilbert space. Two systems, 1 and 2, therefore have two Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$. However, the composite system $1+2$ is also a physical system. The question is thus which Hilbert space corresponds to system $1+2$.

Let $\{|\psi_i\rangle_1\}$ be an orthonormal basis for $\mathscr{H}_1$ and let $\{|\phi_j\rangle_2\}$ be an orthonormal basis for $\mathscr{H}_2$. When the two systems are independent of each other, every basis vector in $\mathscr{H}_1$ can be paired with every basis vector in $\mathscr{H}_2$ and still give a mathematically legitimate description of the composite system. Therefore, one possible orthonormal basis for the Hilbert space of the composite system is given by the set of ordered pairs $\{|\psi_i\rangle_1, |\phi_j\rangle_2\}$. This is a basis of the *tensor product*, or *direct* product of the two Hilbert spaces of the subsystems:

$$\mathscr{H}_{1+2} = \mathscr{H}_1 \otimes \mathscr{H}_2\,. \tag{II.19}$$

An orthonormal basis is given by $\{|\psi_i\rangle_1 \otimes |\phi_j\rangle_2\}$.

From postulate 1 and the fact that a Hilbert space is a complex vector space we immediately see that any tensor product of two superpositions is again a superposition of tensor product states (see appendix A):

$$\sum_i \alpha_i |\psi_i\rangle_1 \otimes \sum_j \beta_j |\phi_j\rangle_2 = \sum_{ij} \alpha_i \beta_j |\psi_i\rangle_1 \otimes |\phi_j\rangle_2 \ , \tag{II.20}$$

i.e., the tensor product is linear. Note that the right-hand side can in general *not* be written as a state $|\zeta_k\rangle_1 \otimes |\xi_l\rangle_2$. This exemplifies the fact that two systems need not be independent of each other. This property, called *entanglement* is crucial to quantum information theory. It will be discussed in detail later on in chapter III.

The total state of two systems can always be written in a special form, called the *Schmidt decomposition*. The most general composite state is given by Eq. (II.20), which involves a double sum over the indices i and j. In the Schmidt decomposition the state is written as a single sum [122]:

$$|\Psi\rangle_{12} = \sum_i c_i |\psi_i', \phi_i'\rangle_{12} \ , \tag{II.21}$$

where c_i can be chosen real and $\{|\psi_i'\rangle_1\}$ and $\{|\phi_i'\rangle_2\}$ are two orthonormal bases for the two subsystems. The bases of the subsystems in eqs. (II.20) and (II.21) are transformed into each other by a unitary transformation:

$$|\psi_i'\rangle_1 = \sum_k U_{ik} |\psi_k\rangle_1 \quad \text{and} \quad |\phi_j'\rangle_2 = \sum_l U_{jl} |\psi_l\rangle_2 \ . \tag{II.22}$$

The Schmidt decomposition is unique (up to phase factors) if and only if the c_i are non-degenerate. If the dimensions of the Hilbert spaces of the two subsystems are d_1 and d_2 respectively, the sum in Eq. (II.21) the index i runs up to the dimension of the smallest Hilbert space [122, 131]. A Schmidt decomposition of the state of three or more subsystems exists only in special circumstances [132].

When a state is in a superposition $|\Psi\rangle = \sum_i \alpha_i |\psi_i\rangle$ with $\sum_i |\alpha_i|^2 = 1$, the operator $|\Psi\rangle\langle\Psi|$ is a so-called projection operator (see appendix A):

$$\begin{aligned}
(|\Psi\rangle\langle\Psi|)^2 &= \left(\sum_{ij} \alpha_i \alpha_j^* |\psi_i\rangle\langle\psi_j| \right)^2 = \sum_{ijkl} \alpha_i \alpha_j^* \alpha_k \alpha_l^* |\psi_i\rangle\langle\psi_j|\psi_k\rangle\langle\psi_l| \\
&= \sum_{ijkl} \alpha_i \alpha_j^* \alpha_k \alpha_l^* |\psi_i\rangle\langle\psi_l| \delta_{jk} = \sum_j |\alpha_j|^2 \sum_{il} \alpha_i \alpha_l^* |\psi_i\rangle\langle\psi_l| \\
&= \sum_{il} \alpha_i \alpha_l^* |\psi_i\rangle\langle\psi_l| = |\Psi\rangle\langle\Psi| \ .
\end{aligned} \tag{II.23}$$

We can now extend our notion of states for a system. In particular, suppose that we have a classical probability distribution over a set of states. We write this as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \ , \tag{II.24}$$

where p_i is the probability to find the system in state $|\psi_i\rangle$. This is sometimes called an *incoherent* superposition. Since the p_i are probabilities, we have $\sum_i p_i = 1$. The operator ρ is called the *density operator* of the system. It is also referred to as a *mixed* state. It has the following properties:

1. $\rho^\dagger = \rho$;
2. $\langle\psi|\rho|\psi\rangle \geq 0$ for all $|\psi\rangle$;
3. $\mathrm{Tr}\rho = 1$.

In a complex Hilbert space, properties 1 and 2 are equivalent.

d Measurements

I will now consider the effect of a measurement on the state of a system. According to the projection postulate, immediately after the measurement of an observable A, the state of the system is in the eigenstate $|\psi_i\rangle$ corresponding to the eigenvalue a_i found in the measurement outcome. With the knowledge of projection operators given in appendix A we can now model this as follows.

Suppose that the system under consideration is in a mixed state ρ. Let the eigenvalues of A be given by $\{a_i\}$. Then the probability $p(a_i)$ that we obtain outcome a_i in a measurement of A is given by

$$p(a_i) = \langle\psi_i|\rho|\psi_i\rangle = \mathrm{Tr}\left(\rho|\psi_i\rangle\langle\psi_i|\right) . \tag{II.25}$$

The right-hand side can be shown to equal the centre term by using the cyclic property of the trace. This type of measurement is called a *Von Neumann* measurement or *ideal* measurement [85, 119, 122]. The underlying assumption in this model is that the measurement outcome faithfully identifies the state of the system immediately after the measuring process.

In practice, this is of course not always the case. Instead, due to the imperfections of the measurement apparatus there might be a whole family of projectors $\{|\psi_k\rangle\langle\psi_k|\}$ which, with some probability $\eta_k > 0$, give rise to the measurement outcome $p(a_i)$. Rather than a projection operator $|\psi_i\rangle\langle\psi_i|$ in Eq. (II.25) we include a *projection operator valued measure*, or POVM:

$$P_{|\psi_i\rangle} = |\psi_i\rangle\langle\psi_i| \longrightarrow E_\mu = \sum_k \eta_k^\mu |\psi_k\rangle\langle\psi_k| , \tag{II.26}$$

with $\sum_\mu E_\mu = \mathbb{1}$. When $\{|\psi_k\rangle\}$ is an orthonormal basis, this implies $\sum_\mu \eta_k^\mu = 1$. In chapters IV and VI, I will use these POVM's to model non-ideal measurements in the context of quantum optics. A more formal presentation of POVM's is given in appendix B.

2 QUANTUM OPTICS

In this section I present quantum optics; the quantum theory of light. First, the electro-magnetic field is quantised and given a particle interpretation, yielding the concept of photons. Then I describe various optical components in terms of unitary evolutions and their generators.

a Quantisation of the electro-magnetic field

Quantum mechanics, as presented in the previous section, can describe a particle in an electro-magnetic field given by a vector potential $\mathbf{A}(\mathbf{r}, t)$ by making the following substitution:

$$\hat{\mathbf{p}} \longrightarrow \hat{\mathbf{p}} - e\hat{\mathbf{A}}(\hat{\mathbf{r}}, t) \,, \tag{II.27}$$

where e is the charge of the particle and $\hat{\mathbf{A}}$ is the vector potential operator obtained by replacing the coordinates $\mathbf{r}$ by their corresponding operator $\hat{\mathbf{r}}$.

Alternatively, the quantisation of the electro-magnetic field can be derived from the Maxwell equations for the electric and magnetic field $\mathbf{E}$ and $\mathbf{B}$ respectively[4]:

$$\begin{aligned} \nabla \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = 0 \quad &, \quad \nabla \cdot \mathbf{B} = 0 \,, \\ \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad &, \quad \nabla \cdot \mathbf{E} = 0 \,, \end{aligned} \tag{II.28}$$

with c the velocity of light in free space. For the fully quantum mechanical description of the electro-magnetic field in free space, I will follow the derivation of Scully and Zubairy [148]. Other books on quantum optics include Loudon [112] and Walls and Milburn [171].

Suppose we want to quantise the electro-magnetic field in a cavity with length L and volume V. Classically, we can describe the electric field in terms of the transverse modes in the x-direction:

$$E_x(z, t) = \sum_j A_j q_j(t) \sin k_j z \,, \tag{II.29}$$

where z is the propagation direction, $q_j(t)$ the mode amplitude, $k_j = j\pi/L$ the wave number and A_j a proportionality constant:

$$A_j = \sqrt{\frac{2\nu_j^2 m_j}{V\epsilon_0}} \,. \tag{II.30}$$

[4] I avoid the notation $\mathbf{H}$, because its components may be confused with the Hamiltonian H later on.

The ν_j are the eigenfrequencies of the cavity. The constant m_j is a dummy mass, included to make the subsequent argument more suggestive [148].

From the Maxwell equations (Eq. (II.28)), we can derive the magnetic field (which is only non-zero in the y-direction):

$$B_y(z,t) = \sum_j A_j \dot{q}_j(t) \frac{\epsilon_0}{\mu_0 k_j} \cos k_j z \,, \tag{II.31}$$

where $\dot{q}_j(t)$ denotes the time derivative of the mode amplitude q_j. The classical Hamiltonian then reads

$$H_{\text{classical}} = \frac{1}{2} \int_V dv \left(\epsilon_0 E_x^2 + \mu_0^{-1} B_y^2 \right) \,. \tag{II.32}$$

After substitution of Eqs. (II.29) and (II.31) in the classical Hamiltonian and integrating over the cavity volume we obtain

$$H_{\text{classical}} = \frac{1}{2} \sum_j \left(m_j \nu_j^2 q_j^2(t) + m_j \dot{q}_j^2(t) \right) \,. \tag{II.33}$$

When we write $p_j = m_j \dot{q}_j$, this has exactly the same form as the classical Hamiltonian of the harmonic oscillator. Therefore, when we want to quantise the electro-magnetic field we proceed in a similar fashion as in section 1.c. We replace the variables q_j and p_j by their respective operators $\hat{q}_j$ and $\hat{p}_j = i\hbar\partial/\partial q_j$.

There are, however, several subtleties. The variables q_j are *amplitudes* of the field modes, and *not* coordinates, as is the case in the linear harmonic oscillator. The variables $\dot{q}_j$ are their corresponding conjugate variables, which facilitate the position-momentum interpretation since $[\hat{q}, \dot{\hat{q}}] = 1$. But this is really *field quantisation*, or second quantisation. Secondly, the masses m_j do *not* have any physical meaning. They are removed by changing our description from q_j and $\dot{q}_j$ to creation and annihilation operators $\hat{a}_j^\dagger$ and $\hat{a}_j$, as I shall now show.

b *Creation and annihilation operators*

Starting with the classical Hamiltonian of the free field in Eq. (II.33) and replacing the variables q_j and p_j with the quantum mechanical operators $\hat{q}_j$ and $\hat{p}_j$, we obtain the quantum mechanical Hamiltonian

$$H = \frac{1}{2} \sum_j \left(m_j \nu_j^2 \hat{q}_j^2(t) + \frac{\hat{p}_j^2}{m_j} \right) \,, \tag{II.34}$$

where

$$[\hat{q}_j, \hat{p}_{j'}] = i\hbar\delta_{jj'} \qquad \text{and} \qquad [\hat{q}_j, \hat{q}_{j'}] = [\hat{p}_j, \hat{p}_{j'}] = 0 \,. \tag{II.35}$$

We can now make the canonical transformation to the operators $\hat{a}_j^\dagger$ and $\hat{a}_j$:

$$\begin{aligned} \hat{a}_j e^{-i\nu_j t} &= \frac{1}{\sqrt{2m_j\hbar\nu_j}}\left(m_j\nu_j\hat{q}_j + i\hat{p}_j\right) , \\ \hat{a}_j^\dagger e^{i\nu_j t} &= \frac{1}{\sqrt{2m_j\hbar\nu_j}}\left(m_j\nu_j\hat{q}_j - i\hat{p}_j\right) . \end{aligned} \tag{II.36}$$

The Hamiltonian in terms of the creation and annihilation operators thus becomes

$$H = \hbar\sum_j \nu_j\left(\hat{a}_j^\dagger\hat{a}_j + \frac{1}{2}\right) . \tag{II.37}$$

With every operator $\hat{a}_j^\dagger$ corresponds a mode a_j. This is a Hamiltonian for a *massless* quantum field. At this point I should briefly clarify my notation. Since the creation and annihilation operators are closely related to the modes they act upon, I make the distinction between modes and operators by writing the operators with a hat. Although observables like the Hamiltonian and unitary transformations are also operators, they do not yield such a potential ambiguity, and I will not write them with hats.

Using the canonical commutation relations given by Eq. (II.35) we immediately see that

$$[\hat{a}_j, \hat{a}_{j'}^\dagger] = \delta_{jj'} , \qquad \text{and} \qquad [\hat{a}_j, \hat{a}_{j'}] = [\hat{a}_j^\dagger, \hat{a}_{j'}^\dagger] = 0 . \tag{II.38}$$

The electric and magnetic fields after second quantisation thus read

$$\begin{aligned} \hat{E}_x(z,t) &= \sum_j \mathcal{E}_j\left(\hat{a}_j e^{-i\nu_j t} + \hat{a}_j^\dagger e^{i\nu_j t}\right)\sin k_j z \\ \hat{B}_x(z,t) &= -i\frac{\epsilon_0 c}{\mu_0}\sum_j \mathcal{E}_j\left(\hat{a}_j e^{-i\nu_j t} - \hat{a}_j^\dagger e^{i\nu_j t}\right)\cos k_j z , \end{aligned} \tag{II.39}$$

with the field strength

$$\mathcal{E}_j = \left(\frac{\hbar\nu_j}{\epsilon_0 V}\right)^{1/2} . \tag{II.40}$$

Just as in the case of the linear harmonic oscillator, we can write the energy eigenstates of one mode of the electro-magnetic field as $|n\rangle$:

$$H|n\rangle = \hbar\nu\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) = E_n|n\rangle . \tag{II.41}$$

By applying the annihilation operator on the last two sides of this equation and using the commutation relations we easily find [148]:

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle , \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle , \quad \hat{a}^\dagger\hat{a}|n\rangle = \hat{n}|n\rangle = n|n\rangle , \tag{II.42}$$

with $E_n = \hbar\nu(n + 1/2)$. The eigenstates are orthonormal: $\langle m|n\rangle = \delta_{mn}$.

Rather than interpreting the eigenstates $|n\rangle$ as the energy levels for a fixed system, in quantum optics the state $|n\rangle$ denotes a state of n *quanta*. The quanta corresponding to the electro-magnetic field are the light-quanta or *photons*. The operators $\hat{a}^\dagger$ and $\hat{a}$ thus *create* and *destroy* photons. Generally, in quantum field theory a field or a wave function is quantised, and the excited modes is given a particle interpretation [37, 144].

At this point I would like to stress that a single photon *does not have a wave function* [123]. It is the excitation of the electro-magnetic field (see also Ref. [148]). In the rest of this thesis I let $|n\rangle_{a_j}$ denote the state of the field, giving the number of photons n in mode a_j.

In quantum mechanics, with every physical system corresponds a Hilbert space. An orthonormal basis for a single-mode system a is given by $\{|n\rangle\}$, which spans an infinite dimensional Hilbert space. When we have several modes $\{a_j\}$ in our system, the total Hilbert space of the system is a tensor product of the Hilbert spaces of the separate modes with orthonormal basis $\{|\vec{n}\rangle \equiv |n_1, \ldots, n_N\rangle\}$. This total Hilbert space can be uniquely decomposed into subspaces with fixed photon number. A Hilbert space with this property is called a Fock space $\mathscr{F}$:

$$\mathscr{F} = \mathscr{H}_0 \oplus \mathscr{H}_1 \oplus \mathscr{H}_2 \oplus \cdots , \tag{II.43}$$

where $\mathscr{H}_k$ denotes the subspace spanned by the vectors $|n_1, \ldots, n_N\rangle$ with $\sum_i n_i = k$ for an N-mode system. The states $|n\rangle$ are also called *Fock states*. The subspace $\mathscr{H}_0$ is a one-dimensional subspace, better known as the *vacuum*.

As a last remark, a tensor product state having a total of k photons may involve one or more modes in vacuum. For example $|n, 0, m\rangle$, where $|0\rangle$ denotes the vacuum. The state $|n, 0, m\rangle$ is part of the basis spanning the subspace $\mathscr{H}_{k=n+m}$. The second mode is also said to be in the vacuum state. This ambiguity is further explored in section *d*.

c Coherent and squeezed states

The creation and annihilation operators are very important in quantum optics. We can therefore ask what the eigenstates of, e.g., the annihilation operator are. Consider the eigenvalue equation for the annihilation operator:

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle . \tag{II.44}$$

The eigenstate can be expanded in terms of number states [148]:

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle , \tag{II.45}$$

which can be written as

$$|\alpha\rangle = e^{-|\alpha|^2/2} e^{\alpha\hat{a}^\dagger} |0\rangle . \tag{II.46}$$

The corresponding *displacement* operator can then be written as

$$|\alpha\rangle = D(\alpha)|0\rangle \equiv e^{-|\alpha|^2/2} e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}|0\rangle \,, \tag{II.47}$$

since the second term in the exponential does not change its behaviour when applied to the vacuum. The operator $D(\alpha)$ is unitary, with $D^\dagger(\alpha) = D(-\alpha) = D^{-1}(\alpha)$. When acting on a creation or annihilation operator, we have

$$\begin{aligned} D^{-1}(\alpha)\hat{a}D(\alpha) &= \hat{a} + \alpha \,, \\ D^{-1}(\alpha)\hat{a}^\dagger D(\alpha) &= \hat{a}^\dagger + \alpha^* \,. \end{aligned} \tag{II.48}$$

The states $|\alpha\rangle$ are called *coherent states*. On a single mode, there are no two coherent states which are orthogonal:

$$\langle\alpha|\alpha'\rangle = \exp\left(-\frac{1}{2}|\alpha|^2 + \alpha'\alpha^* - \frac{1}{2}|\alpha'|^2\right) \,. \tag{II.49}$$

This is only zero when $\alpha = \alpha' = 0$.

The creation and annihilation operators do not commute, and as a consequence, (exponential) functions of these operators generally cannot be rewritten according to the rules of normal arithmetic. In particular, when $[A, B] \neq 0$ we have $e^{A+B} \neq e^A e^B$. From a computational point of view, it is often convenient to deal with the annihilation operators first, and then the creation operators. Especially when the state acted upon is the vacuum, the annihilation operators will yield zero, thus simplifying the task. A function of these operators which is written as

$$f(\hat{a}_1^\dagger, \hat{a}_1, \ldots, \hat{a}_N^\dagger, \hat{a}_N) = \sum_j g_j(\hat{a}_1^\dagger, \ldots, \hat{a}_N^\dagger)\, h_j(, \hat{a}_1, \ldots, \hat{a}_N) \tag{II.50}$$

is said to be in *normal ordered form*. For every term in the sum, the annihilation operators are placed on the right and the creation operators on the left. When the positions of these operators are reversed (i.e., creation operators on the right), we speak of *anti-normal ordering*.

Let me consider a simple example. The second order term in the displacement operator (with α real for simplicity) is proportional to $(\hat{a}^\dagger - \hat{a})^2$. In normal ordered form, this is equal to $(\hat{a}^\dagger)^2 + \hat{a}^2 - 2\hat{a}^\dagger\hat{a} - 1$. Note the -1 in this expression. For higher order terms, the 'non-arithmetic' addition becomes more complicated, until finally, we arrive at [148]:

$$D(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} = e^{-|\alpha|^2/2} e^{\alpha\hat{a}^\dagger} e^{-\alpha^*\hat{a}} \,. \tag{II.51}$$

Later in this section I will return to the normal ordering in more general terms.

Just as the displacement operator creates coherent states, we can construct a squeezing operator which creates so-called *squeezed* states [170, 113, 148]:

$$S(\xi)|0\rangle = \exp\left(\frac{1}{2}\xi^*\hat{a}^2 - \frac{1}{2}\xi(\hat{a}^\dagger)^2\right)|0\rangle = |\xi\rangle \,. \tag{II.52}$$

This operator is also unitary: $S^\dagger(\xi) = S(-\xi) = S^{-1}(\xi)$. It transforms the creation and annihilation operators according to ($\xi = re^{i\theta}$)

$$\begin{aligned} S^{-1}(\xi)\,\hat{a}\,S(\xi) &= \hat{a}\cosh r - \hat{a}^\dagger e^{i\theta}\sinh r\,, \\ S^{-1}(\xi)\,\hat{a}^\dagger\,S(\xi) &= \hat{a}^\dagger\cosh r - \hat{a}\,e^{-i\theta}\sinh r\,. \end{aligned} \tag{II.53}$$

The squeezing operator can also be written in normal ordered form [60]:

$$\begin{aligned} S(\xi) &= \exp\left[\frac{e^{i\theta}}{2}\tanh r(\hat{a}^\dagger)^2\right]\exp\left[-\ln(\cosh r)\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)\right] \\ &\quad\times\exp\left(-\frac{e^{-i\theta}}{2}\tanh r\hat{a}^2\right)\,. \end{aligned} \tag{II.54}$$

Rather than deriving this formula, I will now concentrate on the so-called Baker-Campbell-Hausdorff formula.

We have seen that for non-commuting operators A and B we have $e^Ae^B \neq e^{A+B}$. The natural question to ask is then: what *is* e^Ae^B? The relationship between the two is given by a Baker-Campbell-Hausdorff formula. There are several ways in which we can write this formula, and here I will give two (without proof; the interested reader is referred to, e.g., Gilmore [69]):

$$\begin{aligned} e^Ae^B &= e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}[A,[A,B]]+\frac{1}{12}[[A,B],B]+\cdots}\,, \\ e^{-B}Ae^B &= A+[A,B]+\frac{1}{2!}[[A,B],B]+\frac{1}{3!}[[[A,B],B],B]+\cdots\,. \end{aligned} \tag{II.55}$$

When $A = \alpha\hat{a}^\dagger$ and $B = \alpha^*\hat{a}$, the first BCH formula immediately gives the normal ordered form for the displacement operator: $[A,B] = |\alpha|^2$, which is a constant. Therefore, repeated commutators are zero and the BCH formula terminates.

The sum over repeated commutators does *not* terminate in general, in particular when A and B form a Lie algebra (possibly with a set of other operators C, D, ...). When A and B generate an $su(1,1)$ or an $su(2)$ algebra (with $C = \pm[A,B]/2$ the third generator of the respective algebras) the BCH formula consists of an infinite number of terms, which converge to exponential functions of the generators A, B and C [164]. But even this convergence is not guaranteed. In some cases it is just not possible to write a function of operators in a concise normal ordered form. In appendix D I further discuss the squeezing operator in connection with the $su(1,1)$ algebra. Because of the particular non-compactness for the group $SU(1,1)$, we recognise a one-to-one correspondence between squeezing and $SU(1,1)$.

More properties of the squeezing and displacement operators can be found in Ref. [148]. I will now turn my attention to so-called *multi-mode* squeezing.

Eq. (II.52) is defined for a single mode a. However, we can apply this single-mode squeezing operator to several distinct modes

$$|\vec{\xi}\rangle \equiv S(\vec{\xi})|0\rangle = S(\xi_1)\times\ldots\times S(\xi_N)|0\rangle$$

$$= \exp\left[\frac{1}{2}\sum_{j=1}^{N}\left(\xi_j^*\hat{a}_j^2 - \xi_j(\hat{a}_j^\dagger)^2\right)\right]|0\rangle\ . \tag{II.56}$$

After an N-mode basis transformation U we obtain

$$|\vec{\xi}'\rangle = U^\dagger S(\vec{\xi})U|0\rangle = S'(\vec{\xi})|0\rangle\ . \tag{II.57}$$

The last operator S' can in general be written as

$$S'(\vec{\xi}) = \exp\left[\frac{1}{2}\sum_{j,k=1}^{N}\left(A_{jk}^*\hat{a}_j\hat{a}_k + B_{jk}\hat{a}_j^\dagger\hat{a}_k - A_{jk}\hat{a}_j^\dagger\hat{a}_k^\dagger\right)\right]\ , \tag{II.58}$$

with A and B complex symmetric matrices. This is multi-mode squeezing. It has been studied among others by Caves [40], Barnett and Knight [10], and Caves and Schumaker [41, 147]. The normal-ordering of these operators has been studied by Yuen [180], Fisher *et al.* [60] and Truax [164].

Multi-mode squeezing is of fundamental importance to this thesis. In the next chapter I study whether operators of the form of Eq. (II.58) can yield so-called *event-ready entanglement.* In chapter V I determine the general state of Eq. (II.58) when, in addition, conditional measurements are included. Also parametric down-conversion, a technique which will appear frequently in this thesis, can be described by this evolution. It is now also clear why I present unitary evolutions in terms of generators rather than the Schrödinger equation. Eq. (II.58) is not necessarily an interaction Hamiltonian, but we can still consider the evolution it yields.

d Optical components

The multi-mode displacement and squeezing operators are not just mathematical inventions, they correspond to physical devices. A coherent displacement of the vacuum yields a state which can be generated by a laser. Squeezed states can be generated by, for instance, optical parametric oscillators [171] or parametric down-conversion [109].

Another type of optical components is given by unitary evolutions, the generator of which leaves the photon number invariant. The most important one is the *beam-splitter.* Physically, the beam-splitter consists of a semi-reflective mirror: when light falls on this mirror part will be reflected and part will be transmitted.

Let the two incoming modes be denoted by $\hat{a}_{\text{in}}$ and $\hat{b}_{\text{in}}$ respectively. The outgoing modes are denoted by $\hat{a}_{\text{out}}$ and $\hat{b}_{\text{out}}$. There are four *global* modes, depicted in figure II.1.

When a photon is incident on a beam-splitter, it has a certain probability of being reflected and a certain probability that it is transmitted. When we

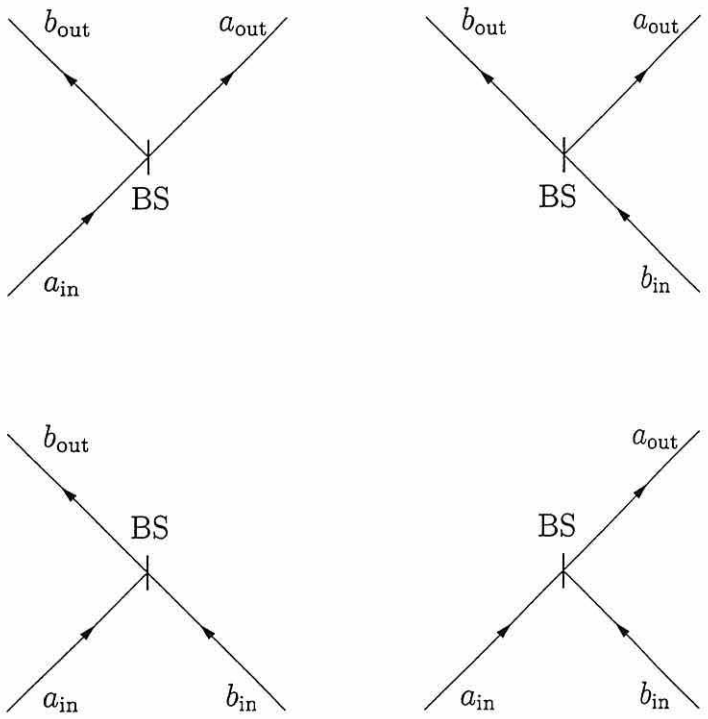

Figure II.1: The four global modes of the beam-splitter.

parametrise the probability amplitudes of these possibilities as $\cos\theta$ and $\sin\theta$, then in operator form the beam-splitter yields an evolution

$$\begin{aligned} \hat{a}^\dagger_{\rm out} &= \cos\theta\,\hat{a}^\dagger_{\rm in} + \sin\theta\,\hat{b}^\dagger_{\rm in}\,, \\ \hat{b}^\dagger_{\rm out} &= \sin\theta\,\hat{a}^\dagger_{\rm in} - \cos\theta\,\hat{b}^\dagger_{\rm in}\,, \end{aligned} \tag{II.59}$$

and similar relations for the annihilation operators. The reflection and transmission coefficients R and T of the beam-splitter are $R = \cos^2\theta$ and $T = 1 - R = \sin^2\theta$. The relative phase shift in the second relation ensures that the transformation is unitary. This means that the beam-splitter is an asymmetric device.

Alternatively, we can write the beam-splitter evolution in terms of a unitary operator generated by an Hermitian operator. Eq. (II.59) can be interpreted as the beam-splitter version of Eqs. (II.48) and (II.53). The question is therefore what the corresponding unitary transformation is (analogous to $D(\alpha)$ and $S(\xi)$). Using the second line in Eq. (II.55) we can easily verify that

$$\begin{aligned} \hat{a}^\dagger_{\rm out} &= e^{\theta(\hat{a}^\dagger_{\rm in}\hat{b}_{\rm in} - \hat{a}_{\rm in}\hat{b}^\dagger_{\rm in})}\,\hat{a}^\dagger_{\rm in}\,e^{-\theta(\hat{a}^\dagger_{\rm in}\hat{b}_{\rm in} - \hat{a}_{\rm in}\hat{b}^\dagger_{\rm in})} = \cos\theta\,\hat{a}^\dagger_{\rm in} + \sin\theta\,\hat{b}^\dagger_{\rm in}\,, \\ \hat{b}^\dagger_{\rm out} &= e^{\theta(\hat{a}^\dagger_{\rm in}\hat{b}_{\rm in} - \hat{a}_{\rm in}\hat{b}^\dagger_{\rm in})}\,\hat{b}^\dagger_{\rm in}\,e^{-\theta(\hat{a}^\dagger_{\rm in}\hat{b}_{\rm in} - \hat{a}_{\rm in}\hat{b}^\dagger_{\rm in})} = \sin\theta\,\hat{a}^\dagger_{\rm in} - \cos\theta\,\hat{b}^\dagger_{\rm in}\,. \end{aligned} \tag{II.60}$$

In general, the generator $H_{\rm BS}$ of the beam-splitter evolution $U = \exp(iH_{\rm BS})$ is

given by

$$H_{\mathrm{BS}} = -i\lambda \hat{a}_{\mathrm{in}}^{\dagger} \hat{b}_{\mathrm{in}} + i\lambda^{*} \hat{a}_{\mathrm{in}} \hat{b}_{\mathrm{in}}^{\dagger} \,. \tag{II.61}$$

Since the photon-number is conserved in the beam-splitter, the operator H_{BS} commutes with the number operator: $[H_{\mathrm{BS}}, \hat{n}] = 0$. Furthermore, H_{BS} is a generator of an $su(2)$ algebra.

The same mathematical description applies to the evolution due to a *polarisation rotation.* Instead of having two different spatial modes a_{in} and b_{in}, the two incoming modes have different polarisations. We write $\hat{a}_{\mathrm{in}} \to \hat{a}_x$ and $\hat{b}_{\mathrm{in}} \to \hat{a}_y$, for some rectilinear set of coordinates x and y. The parameter θ is now the angle of rotation:

$$\begin{aligned} \hat{a}_{x'}^{\dagger} &= \cos\theta\, \hat{a}_x^{\dagger} + \sin\theta\, \hat{a}_y^{\dagger} \,, \\ \hat{a}_{y'}^{\dagger} &= \sin\theta\, \hat{a}_x^{\dagger} - \cos\theta\, \hat{a}_y^{\dagger} \,. \end{aligned} \tag{II.62}$$

This evolution has the same generator as the beam-splitter, except that the angle θ is real for the polarisation rotation, whereas for the beam-splitter we admit complex λ's in Eq. (II.61).

Another important optical component is the single-mode *phase shift*:

$$\hat{a}_{\mathrm{out}}^{\dagger} = e^{i\varphi} \hat{a}_{\mathrm{in}}^{\dagger} \,. \tag{II.63}$$

It is easily verified that

$$e^{i\varphi \hat{a}_{\mathrm{in}}^{\dagger} \hat{a}_{\mathrm{in}}}\, \hat{a}_{\mathrm{in}}^{\dagger}\, e^{-i\varphi \hat{a}_{\mathrm{in}}^{\dagger} \hat{a}_{\mathrm{in}}} = e^{i\varphi} \hat{a}_{\mathrm{in}}^{\dagger} \,. \tag{II.64}$$

The corresponding generator is given by $H_{\varphi} = \varphi\, \hat{a}_{\mathrm{in}}^{\dagger} \hat{a}_{\mathrm{in}}$. It also commutes with the number operator.

In general, when some unitary evolution leaves the photon number invariant, that evolution corresponds to some *passive* optical circuit. Alternatively, when a unitary evolution does not conserves the photon number, we speak of *active* optical devices or *photon sources.*

3 Quantum information

Quantum information theory delivers the foundations for quantum computation, and in order to perform a quantum computation we need to be able to, among other things, prepare certain quantum states. In this thesis, I will not consider any quantum computation algorithms, but I *do* consider state preparation. In later chapters, I will need some concepts from quantum information theory, such as the *fidelity*, to assess various aspects of a state preparation process. Furthermore, since this thesis revolves around states, I have to develop an understanding of what subtleties are involved when talking about quantum states.

a *The computational basis and alphabets*

Suppose we have a single system with a corresponding Hilbert space of dimension N. Rather than giving an orthonormal basis of a system as wave functions in configuration space (like, for instance the eigenstates of the harmonic oscillator in Eq. (II.14)), we can ignore the particular spatial behaviour of these states and *enumerate* them from 0 to $N-1$. The corresponding basis $\{|j\rangle\}$, with $j \in \{0, \ldots, N-1\}$, is then called the *computational* basis. When we have a two-level system ($N = 2$), the computational basis states are $|0\rangle$ and $|1\rangle$, and we speak of a *qubit*.

The advantage of the computational basis is that it is independent of the physical representation. *Any* quantum mechanical two-level system is a qubit, for example an electron in a magnetic field, a polarised photon or a SQUID with clockwise or counter-clockwise current.

When we have two or more systems, the computational basis can be extended accordingly. If we have M systems, we can choose a computational basis $\{|j_1, \ldots, j_M\}$, where $j_i \in \{0, \ldots, N_i - 1\}$, with N_i the dimensionality of the i^{th} system. For instance, the computational basis for two qubits is given by $\{|0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle\}$. In most of the rest of this thesis, I will concentrate on qubits.

As a final point in this section, I present the concept of an *alphabet* of states. It is a finite set of possibly non-orthogonal states. It may be over-complete or it may not span the total Hilbert space. Furthermore, an alphabet of states can generate a POVM, corresponding to a generalised measurement. I am now ready to discuss some information-theoretic aspects of quantum states.

b *Shannon entropy and quantum information*

Prior to the measurement of a system (in, for example, the computational basis), we have a probability distribution $\{p_j\}$ with $\sum_j p_j = 1$ over all possible outcomes. We do *not* know the measurement outcome beforehand. Can we quantify our ignorance of this measurement outcome?

Obviously, when one probability p_k is 1 and the others are all 0, there is no ignorance about the measurement outcome: we will find the system in the state $|k\rangle$, corresponding to the probability $p_k = 1$. On the other hand, when all p_j's are equal, our ignorance about the measurement outcome is maximal. We are looking for a function of the set of probabilities $\{p_j\}$ which is zero when one of the p_j's is zero, and maximal when all p_j's are equal. Such a function is given by

$$S_{\text{Shannon}} = -\sum_{j=0}^{N-1} p_j \log p_j \,. \tag{II.65}$$

This is called the *Shannon entropy* of a probability distribution [151, 131]. It is immediately verified that $S_{\text{Shannon}} = 0$ if all p_j's are 0, except $p_k = 1$, and

by differentiation $\partial S_{\text{Shannon}}/\partial p_k$ we find that the only extremum (a maximum) occurs when all p_j's are equal.

The Shannon entropy is a classical entropy. Quantum mechanically, we can also define an entropy which gives a measure for our ignorance of a state. Suppose the state can be written as a mixture ρ:

$$\rho = \sum_{j=0}^{N-1} p_j |j\rangle\langle j| \,, \tag{II.66}$$

with $\{|j\rangle\}$ some suitable basis. The Shannon entropy can be calculated according to Eq. (II.65) using the probability distribution $\{p_i\}$. Quantum mechanically, our ignorance of the state is given by the *Von Neumann entropy*:

$$S_{\text{VonNeumann}} = -\text{Tr}\,(\rho \log \rho) \,, \tag{II.67}$$

When ρ is pure, it is easy to verify that $S_{\text{VonNeumann}} = 0$ by using the fact that basis transformations inside the trace leave $S_{\text{VonNeumann}}$ invariant. Therefore $S_{\text{Shannon}} \geq S_{\text{VonNeumann}}$, the Von Neumann entropy is a *lower bound* on our ignorance of the state.

As an example, consider a source which produces pure right-handedly polarised photons. In the linear polarisation basis this state is given by

$$|\circlearrowright\rangle = \frac{1}{\sqrt{2}} \left(|\leftrightarrow\rangle + i|\updownarrow\rangle\right) \,. \tag{II.68}$$

A measurement in the basis $\{|\leftrightarrow\rangle, |\updownarrow\rangle\}$ would yield a probability distribution over the measurement outcomes $p_{\leftrightarrow} = p_{\updownarrow} = 1/2$, so the Shannon entropy is maximal. However, since we *could have* measured in the circular basis, the probability distribution would have been $\{p_{\circlearrowright} = 1, p_{\circlearrowleft} = 0\}$, with a corresponding $S_{\text{Shannon}} = 0$. In both cases the Von Neumann entropy is 0, corresponding to the lower bound of the Shannon entropy.

Massar and Popescu [118] proved that the minimal ignorance about a state after a measurement is given by the Von Neumann entropy. For more details on the Shannon entropy in quantum theory I refer the reader to Peres [131].

c *Fidelity and the partition ensemble fallacy*

In the previous section, we used the knowledge of the probability distribution to quantify our ignorance of the measurement outcomes prior to the measurement. However, in general this probability distribution is not known. When we measure the polarisation of a photon we will find definite outcomes, not probabilities. In this section, I ask the question how much information can be gained in a single-shot measurement when the state of the system is not known beforehand. One measure of the information we can extract from a state is given by the *fidelity* [83, 86, 63, 64, 118].

Suppose we want to measure an unknown state and use the knowledge gained by the measurement outcome to reconstruct that state. We then need a measure quantifying the accuracy of the reconstruction. Such a measure is given by the *fidelity*. Suppose further that the initial unknown state is given by $|\psi\rangle$. We now measure this state along $|\phi\rangle\langle\phi|$[5], which is our *estimate*. We can define the measure of success of our estimate by [118]

$$F_{\phi\psi} = |\langle\phi|\psi\rangle|^2 \,. \tag{II.69}$$

This is not the only possible measure, but for our present purposes it is the simplest. When we have a probability distribution over a set of initial states (i.e., we have a mixed state $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$), and a POVM $E_k = \sum_j \mu_{jk}|\phi_j\rangle\langle\phi_j|$, we can average $F_{\phi\psi}$ over the two alphabets of states:

$$F = \sum_j \langle\phi_j| \left(\sum_i p_i|\psi_i\rangle\langle\psi_i|p_{j|i} \right) |\phi_j\rangle \,, \tag{II.70}$$

where $p_{j|i}$ is the probability of estimating the state $|\phi_j\rangle$ when the prepared state is $|\psi_i\rangle$. F is called the average fidelity of state reconstruction.

Consider a source which creates either randomly polarised photons, or linearly polarised photons $|\leftrightarrow\rangle$ and $|\updownarrow\rangle$ with equal probabilities in a given coordinate frame. When we measure the polarisation in $\{|\leftrightarrow\rangle, |\updownarrow\rangle\}$, the outcomes will always be a horizontally or vertically polarised photon for both randomly and linearly polarised photons. We now reconstruct the photon state according to this measurement outcome (creating a photon in the direction corresponding to the measurement outcome). If the photons are linearly polarised, the fidelity of the reconstruction is equal to 1, whereas in the case of randomly polarised photons the fidelity is 2/3 [117]. In the last case we recognise the state very well ($F = 1$), and in the former, we recognise the state quite badly. Therefore, the *same* measurement with the *same* outcomes (with the *same* relative frequencies) yield a *different* fidelity. As a consequence, this fidelity is a measure of the information we extract from the state [118].

Let's now ask a slightly different question: what is the probability that a state ρ is mistaken for another 'estimated' state $|\phi\rangle$? This probability is given by the overlap between the two states:

$$F = \mathrm{Tr}[\rho|\phi\rangle\langle\phi|] \,. \tag{II.71}$$

This equation gives the definition of fidelity in a different context. It corresponds to the lower bound for the probability of mistaking ρ for $|\phi\rangle$ in any possible (single) measurement [63]. When ρ is an exact replica of $|\phi\rangle$ then $F = 1$, and when ρ is an imprecise copy of $|\phi\rangle$ then $F < 1$. Finally, when ρ is completely orthogonal to $|\phi\rangle$ the fidelity is zero.

[5] In general, $|\phi\rangle\langle\phi|$ is part of a POVM, yielding a generalised measurement.

In the above discussion I constructed a mixed state ρ for randomly polarised photons. Consider a polarisation state

$$|\psi(\theta)\rangle = \cos\theta|\leftrightarrow\rangle + \sin\theta|\updownarrow\rangle \,. \tag{II.72}$$

The mixed state of randomly polarised photons is obtained by integrating over all θ which yield a different state $|\psi(\theta)\rangle$:

$$\rho = \frac{1}{\pi}\int_0^\pi d\theta\, |\psi(\theta)\rangle\langle\psi(\theta)| = \frac{1}{2}|\leftrightarrow\rangle\langle\leftrightarrow| + \frac{1}{2}|\updownarrow\rangle\langle\updownarrow| \,. \tag{II.73}$$

But this mixture is equal to that which we would have obtained by randomly choosing only horizontally and vertically polarised photons. In other words, a mixed state does not contain information about the preparation process! In general, we can construct infinitely many physically different sources which generate the same mixed outgoing state. Or equivalently, there exist infinitely many decompositions, or *partitions*, of any given mixed state.

Another example. Consider the state ρ of the form of

$$\rho = \alpha|\phi_1\rangle\langle\phi_1| + \beta|\phi_2\rangle\langle\phi_2| \,. \tag{II.74}$$

It is the sum of two pure states. Again, this is not a unique partition. Whereas in a chemical mixture of, say, nitrogen and oxygen there is a unique partition (into N_2 and O_2), a quantum mixture can be decomposed in many ways. For instance, ρ can equally be written in terms of

$$|\psi_1\rangle = \alpha|\phi_1\rangle + \beta|\phi_2\rangle \quad \text{and} \quad |\psi_2\rangle = \alpha|\phi_1\rangle - \beta|\phi_2\rangle \tag{II.75}$$

as

$$\rho_{\text{out}} = \frac{1}{2}|\psi_1\rangle\langle\psi_1| + \frac{1}{2}|\psi_2\rangle\langle\psi_2| \,. \tag{II.76}$$

This is just one of an infinite number of possible decompositions. Quantum mechanics dictates that all partitions are equivalent to each other [131]. They are indistinguishable. To elevate one partition over another is to commit the 'Partition Ensemble Fallacy.'

Why is this so important? Suppose we have a mixture of the vacuum $|0\rangle\langle 0|$ and a single-photon state $|1\rangle\langle 1|$, yielding $\rho = \alpha|0\rangle\langle 0| + \beta|1\rangle\langle 1|$. It is very tempting to interpret such a mixture as: 'with probability $|\beta|^2$ there *is* a photon, and with probability $|\alpha|^2$ there *is no* photon'. However, quantum theory does not say anything about what *is* without referring to measurement outcomes. If we were to measure an observable whose eigenstates are *not* number states (like, for instance, coherent states), the outcome would not involve any reference to photon numbers. Therefore, in the context of quantum mechanics, the above statement is meaningless. These considerations will become important in chapter VI.

d Non-locality issues

Quantum theory is a local theory, in the sense that space-like and time-like separated operators $A(x_\mu)$ and $B(x'_\mu)$ always commute: $[A(x_\mu), B(x'_\mu)] \propto \delta^4(x_\mu - x'_\mu)$ [76]. However, when one seeks a *classical deterministic* underlying explanation for the correlations observed in quantum mechanics, one has to allow non-local influences. This was first noted by Bell [11], who formulated his now famous inequalities [85, 141].

Let me set up a simple version of Bell's argument. Alice and Bob, who are sufficiently far away from each other, both receive a photon with some unknown polarisation. Alice randomly chooses a polarisation measurement out of two possible directions $\mathbf{a}$ and $\mathbf{a}'$. Similarly, Bob randomly chooses a polarisation measurement out of $\mathbf{b}$ and $\mathbf{b}'$. Let's denote the two possible measurement outcomes of a polarisation measurement by ± 1. Then the eigenvalues a, a', b and b' are all either $+1$ or -1. We repeat this procedure a large number of times.

We now define the expression [141]

$$\gamma_n \equiv a_n b_n + a'_n b_n + a_n b'_n - a'_n b'_n = a_n(b_n + b'_n) + a'_n(b_n - b'_n) \,, \tag{II.77}$$

where the subscript n indicates the n^{th} trial. The value of γ_n is an integer between -2 and $+2$. The absolute value of the average of γ_n over all the N trials is given by

$$\left| \frac{1}{N} \sum_{n=1}^{N} \gamma_n \right| = \left| \frac{1}{N} \sum_{n=1}^{N} (a_n b_n + a'_n b_n + a_n b'_n - a'_n b'_n) \right| \leq 2 \,. \tag{II.78}$$

When we define the correlation coefficients

$$c(\mathbf{x}, \mathbf{y}) \equiv \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} x_n y_n \,, \tag{II.79}$$

the above inequality becomes

$$|c(\mathbf{a}, \mathbf{b}) + c(\mathbf{a}', \mathbf{b}) + c(\mathbf{a}, \mathbf{b}') - c(\mathbf{a}', \mathbf{b}')| \leq 2 \,. \tag{II.80}$$

This is one form of the Bell inequality. We can calculate these correlation coefficients for the case where the two photons are part of the singlet state $(|\leftrightarrow, \updownarrow\rangle - |\updownarrow, \leftrightarrow\rangle)/\sqrt{2}$, which yields $c(\mathbf{a}, \mathbf{b}) = -\cos(\theta_{ab}/2)$. For suitably chosen angles θ_{ab}, the Bell inequality is violated by quantum mechanics.

What does this mean? All I assumed in the above derivation of the inequality was statistical independence of the measurement outcomes obtained by Alice and Bob. The violation therefore implies that the measurements performed by Alice and Bob, though possibly in different galaxies, and thus well and truly separated, can not be considered statistically independent! This has led to wild

speculations about superluminal signalling, but all quantum mechanics predicts are correlations which cannot be given a local realistic interpretation. If we want a classical picture, we therefore have to give up either realism or locality. The choice is yours.

Suppose we have a bi-partite state which violates a Bell inequality. Then that state is said to be *entangled*. The contrary is not necessarily true: a state which is entangled does not have to violate any Bell inequalities [136] (see also appendix B). Several different Bell inequalities have been experimentally verified by many groups, the first of which was led by Aspect [7, 8]. Nowadays, experimental tests of the violation of a Bell inequality is used mostly to indicate whether a state is entangled. Alternatively, tests of non-locality without Bell inequalities have been proposed by DiGiuseppe and Boschi [50, 19].

In the context of quantum optics, there are nonlocal effects in Fock space which are of some interest in this thesis. Hardy and Peres showed that a single photon can exhibit non-local properties, following the work by Tan, Walls and Collett [160, 161, 145]. Here, I will follow Peres' argument [78, 79, 165, 72, 133].

Consider a pure one-particle state

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle_a |1\rangle_b - |1\rangle_a |0\rangle_b \right) \ , \tag{II.81}$$

where $|0\rangle$ is the vacuum and $|1\rangle$ a single-photon state. The subscripts a and b denote the different modes, possibly spatially separated over a large distance. Note that this state has the same mathematical structure as a singlet state. Alice and Bob can demonstrate a violation of a Bell inequality[6] using the strategy described above: Alice and Bob both randomly choose an observable from a set of two non-commuting observables. The first observable for both parties is obviously the one spanned by $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$, i.e., whether there is a photon in mode a or b respectively. Let P_a denote the projector $|1\rangle_a\langle 1|$ and P_b the projector $|1\rangle_b\langle 1|$.

Another observable for Alice might include the projector $P_{a'}$ along the eigenvector $(|1\rangle_a + \sqrt{3}|0\rangle_a)/2$ and Bob can choose to measure along the projector $P_{b'}$ along the eigenvector $(|1\rangle_b - \sqrt{3}|0\rangle_b)/2$. Given Eq. (II.81) quantum theory yields [133, 73]:

$$\begin{aligned} \langle P_{a'} \rangle = \langle P_{b'} \rangle = 0.5 \ , \\ \langle P_a P_b \rangle = 0 \ , \\ \langle P_a P_{b'} \rangle = \langle P_{a'} P_b \rangle = 0.375 \ , \\ \langle P_{a'} P_{b'} \rangle = 0.375 \ , \end{aligned} \tag{II.82}$$

[6] More precisely, a Clauser-Horne inequality.

which violates the inequality

$$0 \leq \langle P_{a'} + P_{b'} - P_{a'}P_{b'} - P_{a'}P_b - P_aP_{b'} + P_aP_b \rangle \leq 1 \,. \qquad \text{(II.83)}$$

This proves non-locality. However, the fact that $P_{a'}$ and $P_{b'}$ do not conserve photon number means that active detection devices, i.e., detectors which can create photons, have to be used. This has provoked many comments [145, 166, 72, 73], but treating them all would lead me too far from the main subject of this thesis. Let us therefore move on to the creation of maximal entanglement in quantum optics, the subject of the next chapter.

III

Creation of Maximal Entanglement

That's the wacky thing about these entangled photon pairs—They're sort of the Bill Clinton and Monica Lewinsky of the quantum world: they're heavily entangled until somebody 'looks' at them.

—Jonathan P. Dowling

Entanglement is one of the key ingredients in quantum communication and information. For instance, quantum protocols such as dense coding [13], quantum error correction [159, 155] and quantum teleportation [14] rely on the non-classical correlations provided by entanglement. Currently, substantial efforts are being made to use *optical* implementations for quantum communication.

The advantages of this are obvious: light travels at high speed and it weakly interacts with the environment. However, exactly this weak interaction poses serious drawbacks. The fact that photons do not interact with each other makes it hard to manipulate them. For example, it has recently been shown that it is impossible to perform so-called complete Bell measurements on two-mode polarisation states in linear quantum optics [114, 167] (although theoretical schemes involving Kerr media [149] and atomic coherence [128] have been reported). Furthermore, maximally polarisation-entangled two-photon states have not been unconditionally produced. In this chapter I investigate the possibility of creating such states with linear optics and a specific class of non-linear elements.

Before that, however, I will have to introduce the terminology I will use in this chapter (and throughout this thesis). In the next section I will discuss various issues connected to entanglement, such as separability, maximal entanglement, multi-partite entanglement and purification. In section 2 I will study parametric down-conversion, currently the most common entanglement source in quantum optics. Finally, I give limitations for the creation of maximal entanglement with a special class of optical circuits. This chapter is based on Kok and Braunstein [99, 100].

1 Separability and entanglement

In this section I discuss the concept of entanglement. First, I define separable and entangled states, and then I introduce *event-ready* entanglement. Three-particle

entanglement is briefly considered, and finally, I discuss the entanglement measure for pure states and entanglement *purification*.

a *What is maximal entanglement?*

Two quantum systems in a pure state, labelled by x_1 and x_2 respectively, are called entangled when the state $\Psi(x_1, x_2)$ describing the total system cannot be factorised into two separate states $\psi_1(x_1)$ and $\psi_2(x_2)$:

$$\Psi(x_1, x_2) \neq \psi_1(x_1)\psi_2(x_2) \; . \tag{III.1}$$

All possible states $\Psi(x_1, x_2)$ accessible to the combined state of the two quantum systems form a set $\mathcal{S}$. These states are generally entangled. Only in extreme cases is $\Psi(x_1, x_2)$ *separable*, i.e., it can be written as a product of states describing the separate systems. The set of separable states form a subset of $\mathcal{S}$ with measure zero.

We arrive at another extremum when the states $\Psi(x_1, x_2)$ are *maximally* entangled. The set of maximally entangled states also forms a subset of $\mathcal{S}$ with measure zero. I will now give a definition of maximal entanglement for two finite-dimensional systems.

Definition: Two N-level systems are called *maximally entangled* when their total state $|\Psi\rangle$ in the Schmidt decomposition can be written as

$$|\Psi\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{i\theta_k} |n_k, m_k\rangle \; , \tag{III.2}$$

with $\{|n_k\rangle\}$ and $\{|m_k\rangle\}$ two orthonormal bases and $\{\theta_k\}$ a set of arbitrary phases.

Suppose we have two (not necessarily identical) two-level systems, 1 and 2, whose states can be written in the orthonormal basis $\{|0\rangle_k, |1\rangle_k\}$ (where $k = 1, 2$). This is defined as the computational basis for these two systems. Physically, those systems could be for example polarised photons or electrons in a magnetic field. Every possible state of the two systems together can be written on the basis of four orthonormal states $|0,0\rangle_{12}$, $|0,1\rangle_{12}$, $|1,0\rangle_{12}$ and $|1,1\rangle_{12}$. These basis states generate a four-dimensional Hilbert space. Another possible basis for this space is given by the so-called Bell states:

$$\begin{aligned} |\Psi^{\pm}\rangle_{12} &= (|0,1\rangle_{12} \pm |1,0\rangle_{12})/\sqrt{2} \; , \\ |\Phi^{\pm}\rangle_{12} &= (|0,0\rangle_{12} \pm |1,1\rangle_{12})/\sqrt{2} \; . \end{aligned} \tag{III.3}$$

These states are also orthonormal. They are examples of *maximally* entangled states. The Bell states are not the only maximally entangled states (as an alternative, we can include a relative phase $e^{i\theta}$ in one of the branches; see also

the definition on this page), but they are a convenient and common choice. All maximally entangled states can be transformed into each other by a local unitary transformation (see Appendix E).

If we want to conduct an experiment which makes use of maximal entanglement, in particular $|\Psi^-\rangle_{12}$, we would most straightforwardly like to have a source which produces these states at the push of a button. In practice, this might be a bit much to ask. A second option might be to have a source which only produces $|\Psi^-\rangle_{12}$ randomly, but flashes a red light when it happens. Such a source would create so-called *event-ready* entanglement[1]: it produces $|\Psi^-\rangle_{12}$ only part of the time, but when it does, it tells you.

More formally, the outgoing state $|\psi_{\text{out}|\text{red light flashes}}\rangle$ conditioned on the red light flashing is said to exhibit event-ready entanglement if it can be written as

$$|\psi_{\text{out}|\text{red light flashes}}\rangle_{12} \simeq |\Psi^-\rangle_{12} + O(\xi) \,, \tag{III.4}$$

where $\xi \ll 1$. In what follows I shall omit the subscript '|red light flashes' since it is clear that we can only speak of event-ready entanglement conditioned on the red light flashing.

Non-maximal entanglement has been created in the context of quantum optics by means of parametric down-conversion [153]. Rather than a (near) maximally entangled state, as in Eq. (III.4), this process produces states with a large vacuum contribution. Only a minor part consists of an entangled photon state. Every time parametric down-conversion is employed, there is only a small probability[2] of creating an entangled photon-pair (notice my use of PEF; see chapter II). We will call this *randomly produced* entanglement.

b Tri-partite entanglement

So far, I have only considered the entanglement of two systems. But quantum mechanics does not give a limit to the number of systems which can be entangled. For instance, we can define a maximally entangled state for three systems 1, 2 and 3:

$$|\text{GHZ}\rangle_{123} = \frac{1}{\sqrt{2}} \left(|0,0,0\rangle_{123} + |1,1,1\rangle_{123}\right) \,. \tag{III.5}$$

Such multi-partite entangled states are called Greenberger-Horne-Zeilinger- or GHZ-states [71]. They also represent a particular state of three maximally entangled systems. Post-selected three-particle GHZ entanglement was observed experimentally by Bouwmeester *et al.* in 1999 [26]. In chapter VI I will extensively discuss the post-selected nature of this and related experiments.

[1] The term first seems to appear in the context of detector efficiencies [183] in 1993, and subsequently with the meaning used here by Pavičić [130] in 1996.

[2] This probability is kept small so that the occurrence of higher order double photon-pairs is negligible.

The state $|\mathrm{GHZ}\rangle$ can also be interpreted as a Schmidt decomposition for three systems (i.e., it can be written as a single sum over orthonormal basis states). Contrary to the bi-partite case, it is *not* true that a Schmidt decomposition exists for any three-partite state. The Schmidt decomposition for two systems follows from the existence of the unitary transformations U_1 and U_2 which diagonalise a matrix T according to $\Lambda = U_1 T U_2$, giving [1]

$$|\Psi\rangle_{12} = \sum_{i,j} T_{ij}|i\rangle_1|j\rangle_2 \rightarrow \sum_{\mu} \lambda_\mu |\mu\rangle_1|\mu\rangle_2 \ , \tag{III.6}$$

with $\{|i\rangle\}$, $\{|j\rangle\}$ and $\{|\mu\rangle\}$ orthonormal bases and λ_μ the eigenvalues of the diagonal matrix Λ. For the general three-system case this is no longer true:

$$|\Psi\rangle_{123} = \sum_{i,j,k} T_{ijk}|i\rangle_1|j\rangle_2|k\rangle_3 \rightarrow \sum_{\mu,\nu} \lambda_{\mu\nu}|\mu\rangle_1|\nu\rangle_{23} \ , \tag{III.7}$$

by virtue of the Schmidt decomposition (again with $\{|\mu\rangle\}$ and $\{|\nu\rangle\}$ orthonormal bases). This is converted to a single sum *if and only if* [132]

$$|\nu\rangle_{23} = \sum_{j,k} (\Omega_\mu)_{jk}|j\rangle_2|k\rangle_3 = |\zeta_\mu\rangle_2|\xi_\mu\rangle_3 \ , \tag{III.8}$$

or $\Omega_\mu^\dagger \Omega_\nu = \Omega_\mu \Omega_\nu^\dagger = 0$ with $(\mu \neq \nu)$.

For multi-partite entanglement, we can no longer completely define maximal entanglement in terms of the Schmidt decomposition. For example, it can easily be verified that the state

$$|W\rangle = \frac{1}{\sqrt{3}}\left(|0,0,1\rangle + |0,1,0\rangle + |1,0,0\rangle\right) \ , \tag{III.9}$$

although representing three maximally entangled systems, cannot be written in terms of the Schmidt decomposition in Eq. (III.5) (the number of linear independent terms exceeds the number of orthonormal basis states of the separate systems). More specifically, Dür *et al.* prove that ensembles of three-partite entangled states of qubits can be transformed either to the state $|\mathrm{GHZ}\rangle$ or to the state $|W\rangle$ by stochastic[3] local operations and classical communication (SLOCC) alone [55]. That is, under SLOCC, $|\mathrm{GHZ}\rangle$ and $|W\rangle$ generate two distinct invariant subspaces of the total Hilbert space $(\mathscr{H}_2^{\otimes 3})^{\otimes \infty}$ spanned by the three qubits. In conclusion, the definition of maximally entangled states in terms of the Schmidt decomposition given on page 38 only works for bi-partite systems. However, in the rest of this chapter (and indeed, this thesis) I will concentrate on entanglement between two systems, and this definition is sufficient.

[3] 'Stochastic' meaning that the transformation is successful with non-zero probability.

c *Purification*

In this section, I will define a measure of entanglement E for (pure) states which are non-maximally entangled [15, 16][4]. The natural measure of entanglement is defined by the Von Neumann entropy of the reduced density matrix of the subsystems. Suppose we have two systems 1 and 2 held by Alice and Bob respectively in a (non-maximally) entangled state $|\Psi\rangle_{12}$. The reduced density matrices of the two subsystems are

$$\rho_1 \equiv \mathrm{Tr}_2\left[|\Psi\rangle_{12}\langle\Psi|\right] \qquad \text{and} \qquad \rho_2 \equiv \mathrm{Tr}_1\left[|\Psi\rangle_{12}\langle\Psi|\right] \ . \tag{III.10}$$

In chapter II we defined the Von Neumann entropy of a density matrix ρ as

$$S(\rho) = -\mathrm{Tr}\left[\rho \log \rho\right] \ . \tag{III.11}$$

The measure of entanglement $E(\Psi)$ is now defined as

$$E(\Psi) \equiv S(\rho_1) = S(\rho_2) \ . \tag{III.12}$$

This measure has a number of pleasant properties: it is zero for separable states and maximal for maximally entangled states. Furthermore, E remains the same whether we trace out system 1 or system 2.

Can we in some way increase the entanglement in non-maximally entangled (pure) states? The answer to this question is the domain of *entanglement purification.* I will now briefly discuss the general idea behind purification.

Suppose we distribute a set of singlet states among Alice and Bob for quantum communication purposes. After the distribution, however, the actual states held by Alice and Bob will in general no longer be maximally entangled. This is because noise and decoherence in the distribution process degrade the entanglement. In order to obtain singlet states again we have to *purify* the ensemble of states shared between Alice and Bob with only local operations [15]. Because operations of this kind cannot increase the amount of entanglement, the total entanglement shared by Alice and Bob remains the same or decreases. However, we still have the possibility of converting several copies of poorly entangled states into a few highly entangled states. In general, when we have N non-maximally entangled initial states $|\Psi\rangle$ with entanglement content $E(\Psi)$, we wish to obtain $M < N$ entangled states $|\Phi\rangle$ with $E(\Phi) > E(\Psi)$ using a restricted class of operations. The local operations can be divided into three groups [27]:

1. *Local transformations and measurements.* These can be modelled in general by local POVM's, i.e., POVM's acting on only one subsystem,

2. *classical communication.* This creates the opportunity to classically correlate the actions of the two (distant) parties holding the entanglement,

[4]There are subtleties in defining measures of entanglement for mixed states; see Ref. [27] and references therein.

3. *post-selection.* Depending on the outcome of local measurements and classical communication, we can select a subset from our ensemble of states.

A procedure which manages to increase the entanglement in a subset of the N initial states using only operations from these three classes is called an *entanglement purification* protocol.

As an example, I consider the original purification protocol[5] [15]. We distribute (at least) two singlets, written in the computational basis $\{|0\rangle_k, |1\rangle_k\}$ with $k = 1, 2$, between Alice and Bob:

$$|\Psi^-\rangle_{12} = \frac{1}{\sqrt{2}} \left(|0,1\rangle_{12} - |1,0\rangle_{12}\right) . \tag{III.13}$$

Similarly, the other Bell states are given by $|\Psi^+\rangle_{12}$ and $|\Phi^\pm\rangle_{12}$. A simple model of the noise due to the distribution implies that upon arrival the singlets have become Werner states[6] [174] ρ:

$$\rho_{12} = \varepsilon |\Psi^-\rangle_{12}\langle\Psi^-| + (1-\varepsilon)\frac{\mathbb{1}}{4} . \tag{III.14}$$

Note that there is a singlet contribution in the maximally mixed part $\mathbb{1}$ as well. Two states with equal density matrices cannot be distinguished in *any* physical way, which means that the method of preparation of ρ is irrelevant. In other words, we do not care *how* decoherence took place in the distribution process, the occurrence of $\mathbb{1}$ defines a class of noise types which give rise to Eq. (III.14). The noise gain is parametrised by ε.

In order to purify ρ_{12} we need at least two such systems in a total state $\rho_{12} \otimes \rho_{12}$. As a first step, Bob operates with the Pauli matrix σ_{2y} on his part of the density matrix, yielding a transformation

$$\rho'_{12} = (\sigma_{2y} \otimes \sigma_{2y})\, \rho \otimes \rho\, (\sigma_{2y} \otimes \sigma_{2y})^\dagger , \tag{III.15}$$

This transformation is equivalent to the symbol swapping $\Psi^\pm \leftrightarrow \Phi^\mp$ in the Bell states given in Eq. (III.3).

Next, Bob applies the 'controlled NOT' to his two subsystems, after which both Alice and Bob measure the target in the computational basis. They compare their measurement outcome by means of classical communication, and if the two outcomes are the same (00 or 11), Bob again applies the σ_y operator to his remaining state. If the measurement outcomes are not the same, the purification failed. This is the post-selection stage. Note the *probabilistic* character of the purification protocol. After a successful purification run, Alice and Bob share a mixed state ρ''_{12}.

[5] In chapter VI we will encounter another purification protocol.

[6] Decoherence tends to evolve pure states towards (maximally) mixed states.

Before purification the fidelity of the distributed system was

$$F = \langle \Psi^- | \rho_{12} | \Psi^- \rangle = \frac{1+3\varepsilon}{4} . \tag{III.16}$$

After a successful purification round the new fidelity F' is [15]

$$F' = \langle \Psi^- | \rho''_{12} | \Psi^- \rangle = \frac{1+2\varepsilon+5\varepsilon^2}{4(1+\varepsilon)} . \tag{III.17}$$

For purification to be meaningful we need $F' > F$, or $\varepsilon > \frac{1}{3}$. In other words, if the decoherence is too strong, this type of purification cannot increase the entanglement in a subset of the distributed ensemble.

Numerous other purification protocols have been proposed (see Ref. [27] and references therein). In general, it provides a procedure to obtain (near) maximal entanglement. Also called distillation, entanglement purification reduces an ensemble of poorly entangled states to a few highly entangled states.

By contrast, in this thesis I study the creation of event-ready entanglement, in particular with quantum optics. The difference with purification is that it is *dynamic*: an event-ready entangler does not need to store entanglement (something which is very difficult for photons), it tries a (large) number of times and 'flashes a red light' upon success. In the next section I will study a more modest device, called a parametric down-converter. It produces entanglement randomly.

2 Entanglement sources in quantum optics

In this section I will look at a particular class of devices capable of creating entanglement in quantum optics. These devices are commonly known as *down-converters*, since they convert a high-energy photon into two lower-energy photons. Physically, in parametric down-conversion a crystal is pumped by a high-intensity laser, which we will treat classically (the parametric approximation). The crystal is special in the sense that it has different refractive indices for horizontally and vertically polarised light. In the case of degenerate (type II) parametric down-conversion a photon from the pump is split into two photons with half the energy of the pump photon. Furthermore, the process can be set up such that the two photons have orthogonal polarisations. The outgoing modes of the crystal constitute two intersecting cones with orthogonal polarisations $|\updownarrow\rangle$ and $|\leftrightarrow\rangle$ as depicted in Fig. III.1.

Due to the conservation of momentum, the two produced photons are always in opposite modes with respect to the central axis (determined by the direction of the pump). In the two spatial modes where the different polarisation cones intersect we can no longer infer the polarisation of the photons, and as a consequence the two photons become entangled in their polarisation. Parametric

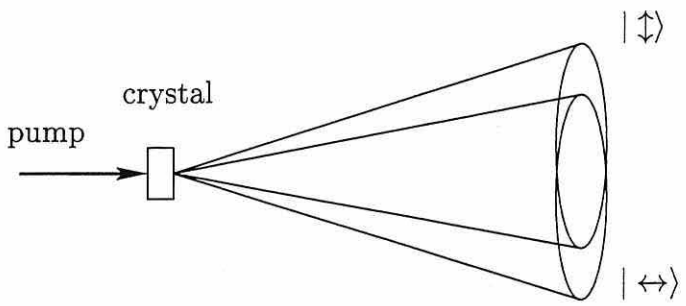

Figure III.1: A schematic representation of type II parametric down-conversion. A high-intensity laser pumps a non-linear crystal. With some probability a photon in the pump beam will be split into two photons with orthogonal polarisation $|\updownarrow\rangle$ and $|\leftrightarrow\rangle$ along the surface of the two respective cones. Depending on the optical axis of the crystal, the two cones are slightly tilted from each other. Selecting the spatial modes at the intersection of the two cones yields the outgoing state $(1-\xi^2)|0\rangle + \xi|\Psi^-\rangle + O(\xi^2)$.

down-conversion as a device to create entangled photons was introduced by Shih and Alley in 1988 [153] and is being continuously improved [154, 109, 111, 94, 125].

However, parametric down-converters do *not* produce pure Bell-states [183, 32, 99]. Because of the spontaneous nature of the down-conversion process, only a in a small number of cases (i.e., a fraction of the trials) will a photon from the pump be split into two photons. We thus have randomly produced entanglement. Furthermore, there is an even smaller probability of creating more photon pairs, originating from several pump photons. The probability of this happening decreases with the number of created pairs. I will derive the value of these probabilities in due course.

The outgoing state of the parametric down-converter I am interested in here is

$$|\psi_{\text{out}}\rangle = (1-\xi^2)|0\rangle + \xi|\Psi^-\rangle + O(\xi^2)\,, \tag{III.18}$$

where $\xi \ll 1$ is a parameter indicating the strength of the down-conversion. Note that $|0\rangle$ denotes the *vacuum* here, rather than a computational basis state. In the next section I will give a mathematical description of down-converters and subsequently I will determine the statistical properties of these devices.

a The physics of down-converters

In this section I will describe the physical properties of parametric down-conversion. Consider a down-converter with outgoing field modes a_i and b_j. The indices denote the particular polarisation along the x- and y-axis of a given coordinate system. We are working in the interaction picture of the Hamiltonian

which governs the dynamics of creating two entangled field modes a and b using weak parametric down-conversion. In the rotating wave approximation this Hamiltonian reads ($\hbar = 1$):

$$H = i\kappa(\hat{a}_x^\dagger \hat{b}_y^\dagger - \hat{a}_y^\dagger \hat{b}_x^\dagger) + \text{H.c.} \tag{III.19}$$

In this equation H.c. means Hermitian conjugate, and κ is the product of the pump amplitude and the coupling constant between the electro-magnetic field and the crystal. The operators $\hat{a}_i^\dagger$, $\hat{b}_i^\dagger$ and $\hat{a}_i$, $\hat{b}_i$ are creation and annihilation operators for polarisations $i \in \{x, y\}$ respectively. They satisfy the following commutation relations:

$$\begin{aligned} [\hat{a}_i, \hat{a}_j^\dagger] &= \delta_{ij} \ , & [\hat{a}_i, \hat{a}_j] &= [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0, \\ [\hat{b}_i, \hat{b}_j^\dagger] &= \delta_{ij} \ , & [\hat{b}_i, \hat{b}_j] &= [\hat{b}_i^\dagger, \hat{b}_j^\dagger] = 0 \ , \end{aligned} \tag{III.20}$$

where $i, j \in \{x, y\}$. The time evolution due to this Hamiltonian is given by

$$U(t) \equiv \exp(-iHt) \ , \tag{III.21}$$

where t is the time it takes for the pulse to travel through the crystal. By applying this unitary transformation to the vacuum $|0\rangle$ the state $|\Psi_{\text{src}}\rangle$ is obtained:

$$|\Psi_{\text{src}}\rangle = U(t)|0\rangle = \exp(-iHt)|0\rangle \ . \tag{III.22}$$

We are interested in the properties of $|\Psi_{\text{src}}\rangle$. Define the L_+ and the L_- operator to be

$$L_+ = \hat{a}_x^\dagger \hat{b}_y^\dagger - \hat{a}_y^\dagger \hat{b}_x^\dagger = L_-^\dagger \ . \tag{III.23}$$

This will render Eqs. (III.19) and (III.21) into:

$$H = i\kappa L_+ - i\kappa^* L_- \qquad \text{and} \qquad U(t) = \exp[\kappa t L_+ - \kappa^* t L_-] \ . \tag{III.24}$$

Applying L_+ to the vacuum will yield a singlet state (up to a normalisation factor) in modes a and b:

$$\begin{aligned} L_+|0\rangle &= |\leftrightarrow, \updownarrow\rangle_{ab} - |\updownarrow, \leftrightarrow\rangle_{ab} \\ &= |1,0;0,1\rangle_{a_x a_y b_x b_y} - |0,1;1,0\rangle_{a_x a_y b_x b_y} \ , \end{aligned} \tag{III.25}$$

we henceforth use the latter notation where $|i,j;k,l\rangle_{a_x a_y b_x b_y}$ is shorthand for $|i\rangle_{a_x} \otimes |j\rangle_{a_y} \otimes |k\rangle_{b_x} \otimes |l\rangle_{b_y}$, a tensor product of photon number states. Applying this operator n times gives a state $|\Phi^n\rangle$ (where we have included a normalisation factor N_n, so that $\langle\Phi^n|\Phi^n\rangle = 1$):

$$|\Phi^n\rangle \equiv N_n L_+^n|0\rangle = \sum_{m=0}^{n} (-1)^m \sqrt{\frac{n!}{(n+1)!}} |m_x, (n-m)_y; (n-m)_x, m_y\rangle_{ab} \ , \tag{III.26}$$

where the normalisation constant N_n is given by

$$N_n^2 = \frac{1}{n!(n+1)!} \,. \tag{III.27}$$

We interpret $|\Phi^n\rangle$ as the state of n entangled photon-pairs on two spatial modes a and b.

We want the unitary operator $U(t)$ in Eq. (III.24) to be in a normal ordered form, because then the annihilation operators will 'act' on the vacuum first, in which case Eq. (III.22) simplifies. In order to obtain the normal ordered form of $U(t)$ we examine the properties of L_+ and L_-. Given the commutation relations (III.20), it is straightforward to show that:

$$\begin{aligned} [L_-, L_+] &= \hat{a}_x^\dagger \hat{a}_x + \hat{a}_y^\dagger \hat{a}_y + \hat{b}_x^\dagger \hat{b}_x + \hat{b}_y^\dagger \hat{b}_y + 2 \equiv 2L_0 \quad \text{and} \\ [L_0, L_\pm] &= \pm L_\pm \,. \end{aligned} \tag{III.28}$$

An algebra which satisfies these commutation relations (together with the properties $L_- = L_+^\dagger$ and $L_0 = L_0^\dagger$) is an $su(1,1)$ algebra[7]. The normal ordering for this algebra is known [164] (with $\hat{\tau} = \tau/|\tau|$) (see also appendix D for more details):

$$e^{\tau L_+ - \tau^* L_-} = e^{\hat{\tau} \tanh|\tau| L_+} e^{-2\ln(\cosh|\tau|) L_0} e^{-\hat{\tau}^* \tanh|\tau| L_-} \,. \tag{III.29}$$

The scaled time τ is defined as $\tau \equiv \kappa t$. Without loss of generality we can take τ to be real. Since the 'lowering' operator L_- is placed on the right, it will yield zero when applied to the vacuum and the exponential reduces to the identity. Similarly, the exponential containing L_0 will yield a c-number, contributing only an overall phase.

Parametric down-conversion is an example of so-called *multi-mode squeezed vacuum*. The photon-statistics of two-mode squeezed states have been studied in Refs. [42, 6, 146]. Here, I study the particular case of the down-conversion process used to create randomly produced maximal entanglement.

Are the pairs formed in parametric down-conversion independent of each other? If they are, the number of pairs should give a Poisson distribution. I will now calculate whether this is the case.

Suppose $P_{\text{PDC}}(n)$ is the probability of creating n photon-pairs with parametric down-conversion and let

$$r \equiv \tanh\tau \qquad \text{and} \qquad q \equiv 2\ln(\cosh\tau) \,, \tag{III.30}$$

then the probability of finding n entangled photon-pairs is:

$$P_{\text{PDC}}(n) \equiv |\langle \Phi^n | \Psi_{\text{src}} \rangle|^2$$

[7] The interaction Hamiltonian from Eq. (III.19) thus generates unitary evolutions which are closely related to the group elements of $SU(1,1)$.

$$
\begin{aligned}
&= |\langle 0| \left(L_-^n N_n\right) \left(e^{rL_+} e^{-qL_0} e^{-rL_-}\right) |0\rangle|^2 \\
&= e^{-2q} |\langle 0| L_-^n N_n \left[\sum_{l=0}^{\infty} \frac{r^l}{l!} L_+^l\right] |0\rangle|^2 \\
&= (n+1) r^{2n} e^{-2q} \,. \qquad \text{(III.31)}
\end{aligned}
$$

It should be noted that this is a normalised probability distribution in the limit of $r, q \to 0$.

Given Eqs. (III.30) $P_{\text{PDC}}(n)$ deviates from the Poisson distribution, and the pairs are therefore not independent. For weak sources, however, one might expect that $P_{\text{PDC}}(n)$ approaches the Poisson distribution sufficiently closely. This hypothesis can be tested by studying the distinguishability of the two distributions.

b Statistical properties of down-converters

Here, I study the distinguishability between the pair distribution calculated in the previous section and the Poisson distribution. The Poisson distribution for independently created objects is given by

$$P_{\text{poisson}}(n) = \frac{p^n e^{-p}}{n!} \,. \qquad \text{(III.32)}$$

Furthermore, rewrite the pair distribution in Eq. (III.31) as

$$P_{\text{PDC}}(n) = (n+1) \left(\frac{p}{2}\right)^n e^{-p} \,, \qquad \text{for } p \ll 1 \,, \qquad \text{(III.33)}$$

using $q \approx r^2$ and $p \equiv 2r^2 = 2\tanh^2 \tau$ for small scaled times. Here $p/2$ is the probability of creating one entangled photon-pair. Are these probability distributions distinguishable? Naively one would say that for sufficiently weak down-conversion (i.e., when $p \ll 1$) these distributions largely coincide, so that instead of the complicated pair-distribution (III.33) we can use the Poisson distribution, which is much easier from a mathematical point of view. The distributions are distinguishable when the 'difference' between them is larger than the size of an average statistical fluctuation of the difference. This fluctuation depends on the number of samplings.

Consider two nearby discrete probability distributions $\{p_j\}$ and $\{p_j + dp_j\}$. A natural difference between these distributions is given by the so-called (infinitesimal) *statistical distance* ds [177, 29, 84] (see also appendix F):

$$ds^2 = \sum_j \frac{dp_j^2}{p_j} \,. \qquad \text{(III.34)}$$

When the typical statistical fluctuation after N samplings is $1/\sqrt{N}$, the two probability distributions are distinguishable if:

$$ds \gtrsim \frac{1}{\sqrt{N}} \quad \Leftrightarrow \quad N ds^2 \gtrsim 1 \,. \qquad \text{(III.35)}$$

The statistical distance between (III.32) and (III.33), and therefore the distinguishability criterion is:

$$ds^2 \propto \frac{p^2}{8} \quad \rightarrow \quad N \gtrsim \frac{8}{p^2} \,. \tag{III.36}$$

On the other hand, the average number of trials in the teleportation experiment required to get one photon-pair from both down-converters is:

$$N = \frac{1}{p^2} \,. \tag{III.37}$$

The minimum number of trials in the experiment thus almost immediately renders the two probability distributions distinguishable, and we therefore cannot approximate the actual probability distribution with the Poisson distribution.

Since the Poisson distribution in Eq. (III.32) is derived by requiring statistical independence of n pairs and the pair distribution is distinguishable from the Poisson distribution, the photon-pairs cannot be considered to be independently produced, *even in the weak limit.*

This concludes my study of parametric down-conversion here. I will now return to the problem of maximal entanglement creation in quantum optics.

3 The creation of maximal entanglement

Now that I have investigated the properties of one of the most common entanglement sources, i.e., down-conversion, I am ready to consider the creation of maximal, or event-ready, entanglement. The optical circuits discussed here consist of so-called *passive* and *active* components. The passive components leave the photon-number invariant, i.e., they correspond to unitary operators which commute with the number operator. Examples of such components are beam-splitters, phase-shifters and polarisation rotators.

Active components correspond to unitary evolutions which do *not* commute with the photon-number operator, like the parametric down-converter. Other examples of active components are single-mode squeezers and pumped $\chi^{(3)}$ media [5]. They can generally be characterised by an interaction Hamiltonian which is a polynomial function of creation and annihilation operators. Later in this chapter I will restrict the discussion to interaction Hamiltonians which are quadratic in these operators.

In this section, I will first show that the creation of maximal entanglement with only passive components from a pure separable state is impossible. Then, a general condition for an optical setup is derived, which should be satisfied in order to yield event-ready entanglement. I subsequently examine this condition for a specific class of optical circuits.

a Passive optical components

So far, I have hardly paid attention to passive optical components. In this section I will show that they cannot transform a completely separable state of two photons into a maximally entangled state.

Suppose we have a linear interferometer which consists only of passive components. Such an interferometer is described by a unitary matrix U [140], which transforms the creation operators of the electro-magnetic field according to

$$\hat{a}_i \rightarrow \sum_j u_{ij}\hat{b}_j \qquad \text{and} \qquad \hat{a}_i^\dagger \rightarrow \sum_j u_{ij}^*\hat{b}_j^\dagger \,, \tag{III.38}$$

where the u_{ij} are the components of U and i, j enumerate both the modes and polarisations. There is no mixing between the creation and annihilation operators, because photons do not interact with each other. I will now show that we cannot create maximal (event-ready) entanglement with such linear interferometers when the input state is a separable state.

Without loss of generality I consider the separable state $|x, y\rangle = \hat{a}_x^\dagger \hat{b}_y^\dagger |0\rangle$. In order to create maximal entanglement, the creation operators should be transformed according to

$$\hat{a}_x^\dagger \hat{b}_y^\dagger \;\rightarrow\; (\hat{c}_x^\dagger \hat{d}_y^\dagger - \hat{c}_y^\dagger \hat{d}_x^\dagger)/\sqrt{2} \,. \tag{III.39}$$

Relabel the modes a_x, a_y, b_x and b_y as a_1 to a_4 respectively. Without loss of generality (and leaving the normalisation aside for the moment) we can then write

$$\hat{a}_1^\dagger \hat{a}_2^\dagger \;\rightarrow\; \hat{b}_1^\dagger \hat{b}_2^\dagger - \hat{b}_3^\dagger \hat{b}_4^\dagger \,. \tag{III.40}$$

Substituting Eq. (III.38) into Eq. (III.40) generates ten equations for eight variables u_{ij}^*:

$$\begin{aligned}
u_{11}^* u_{21}^* = u_{12}^* u_{22}^* = u_{13}^* u_{23}^* = u_{14}^* u_{24}^* &= 0 \\
(u_{11}^* u_{23}^* + u_{13}^* u_{21}^*) = (u_{11}^* u_{24}^* + u_{14}^* u_{21}^*) &= 0 \\
(u_{12}^* u_{23}^* + u_{13}^* u_{22}^*) = (u_{12}^* u_{24}^* + u_{14}^* u_{22}^*) &= 0 \\
(u_{11}^* u_{22}^* + u_{12}^* u_{21}^*) &= 1 \\
(u_{13}^* u_{24}^* + u_{14}^* u_{23}^*) &= -1 \,.
\end{aligned} \tag{III.41}$$

It can be easily verified that there are no solutions for the u_{ij}^* which satisfy these ten equations simultaneously: since $u_{11}^* u_{21}^* = 0$, choose $u_{11}^* = 0$ and $u_{21}^* \neq 0$. From the fourth line above follows that $u_{12}^* \neq 0$. Hence $u_{22}^* = 0$. The second line (second equation) then determines $u_{14}^* = 0$, which (third line, second equation) implies that $u_{12}^* u_{24}^* = 0$. We already set $u_{12}^* \neq 0$, so we obtain $u_{24}^* = 0$. We now derive a contradiction between $u_{14}^* = 0$, $u_{24}^* = 0$ and the last line of the set of equations above. This means that there is no passive interferometer which transforms pure separable states into maximal (event-ready) entangled states.

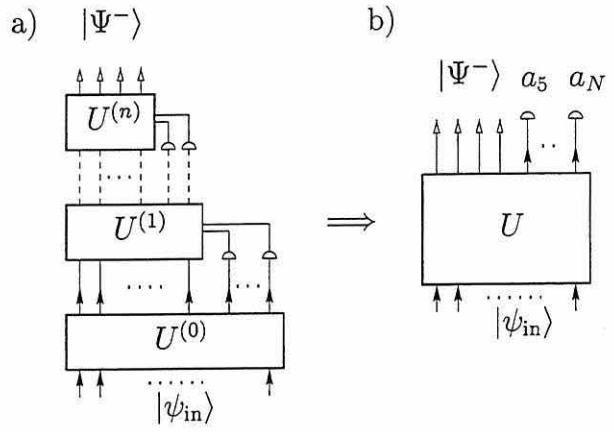

Figure III.2: If an optical circuit with feed-forward detection (a) produces a specific state, the same output can be obtained by an optical circuit where detection of the *auxiliary* modes takes place at the end (b). The efficiency of the latter, however, will generally be smaller.

b General optical circuits

In order to make $|\Psi^-\rangle$, I will assume that we have several resources at our disposal. The class of elements will consist of beam-splitters, phase-shifters, photo-detectors and non-linear components such as down-converters, squeezers, etc. These elements are then arranged to give a specific optical circuit (see Fig. III.2). Part of this setup might be so-called *feed-forward* detection. In this scheme the outcome of the detection of a number of modes dynamically chooses the internal configuration of the subsequent optical circuit based on the interim detection results (see also Ref. [114]). Conditioned on these detections we want to obtain a freely propagating $|\Psi^-\rangle$ Bell state in the remaining undetected modes.

I now introduce two simplifications for such an optical circuit. First, I will show that we can discard feed-forward detection. Secondly, we only have to consider the detection of modes with at most one photon.

Theorem 1: In order to show that it is *possible* to produce a specific outgoing state, any optical circuit with feed-forward detection can be replaced by a *fixed* optical circuit where detection only takes place at the end.

Proof: Suppose a feed-forward optical circuit (like the one depicted in Fig. III.2a) giving $|\Psi^-\rangle$ exists. That means that the circuit creates $|\Psi^-\rangle$ conditioned on one of potentially many patterns of detector responses. It is sufficient to consider a single successful pattern. We can then take every interferometer to be fixed and postpone all detections of the auxiliary modes to the very end (Fig. III.2b). Note that this procedure selects generally only *one* setup in which entanglement is produced, whereas a feed-forward optical circuit

potentially allows more setups. It therefore might reduce the efficiency of the process. However, since we are only interested in the *possibility* of creating $|\Psi^-\rangle$, the efficiency is irrelevant. □

Theorem 2: Suppose an optical circuit produces a specific outgoing state conditioned on n_1 detected photons in mode 1, n_2 detected photons in mode 2, etc. (with $n_i = 0, 1, 2, \ldots$). The same output can be obtained by a circuit where in every detected mode *at most* one photon is found.

Proof: If there are more photons in a mode, we can replace the corresponding detector by a so-called detector *cascade* [103]. This device splits the mode into many modes which are all detected (see also chapter IV). For a sufficiently large cascade there is always a non-vanishing probability to have at most one photon in each outgoing mode. In that case, the same state is created while at most one photon enters each detector. Note that this again yields a lower efficiency. □

Applying these results to the creation of $|\Psi^-\rangle$, it is sufficient to consider a single *fixed* interferometer acting on an incoming state, followed at the end by detection of the so-called *auxiliary* modes. $|\Psi^-\rangle$ is signalled by at least one fixed detection pattern with at most one photon in each detector.

How do I proceed in trying to make the $|\Psi^-\rangle$ Bell state? Let the time independent interaction Hamiltonian H_I incorporate both the interferometer U and the creation of $|\psi_{\text{in}}\rangle$ (see Fig. III.2b). The outgoing state prior to the detection can be formally written as

$$|\psi_{\text{out}}\rangle = U|\psi_{\text{in}}\rangle \equiv \exp\left(-itH_I\right)|0\rangle\,, \tag{III.42}$$

with $|0\rangle$ the vacuum. This defines an effective Hamiltonian H_I which is generally not unique.

c *The Bargmann representation*

At this point it is useful to change the description. Since the creation and annihilation operators satisfy the same commutation relations as c-numbers and their derivatives, we can make the substitution $\hat{a}_i^\dagger \to \alpha_i$ and $\hat{a}_i \to \partial_i$, where $\partial_i \equiv \partial/\partial\alpha_i$. Furthermore, we define $\vec{\alpha} = (\alpha_1, \ldots, \alpha_N)$. Quantum states are then represented by functions of c-numbers and their derivatives. This is called the Bargmann representation [9].

Furthermore, suppose we can normal order the operator $\exp(-itH_I)$ in Eq. (III.42). This would yield a function of only the creation operators, acting on the vacuum. In the Bargmann representation we then obtain a function of complex numbers without their derivatives. In particular, an optical circuit consisting of N distinct modes (for notational convenience I treat distinct polarisations like,

for instance, x and y as separate modes), can be written as a function $\psi_{\text{out}}(\vec{\alpha})$ after the unitary evolution U and normal ordering. The normal ordering of the evolution operator in conjunction with the vacuum input state is crucial, since it allows a significant simplification of the problem.

I now treat the (ideal) detection of the auxiliary modes in the Bargmann representation. Suppose the outgoing state after the detection of M photons emerges in modes a_1, a_2, a_3 and a_4. After a suitable reordering of the detected modes the state which is responsible for the detector coincidence indicating success can be written as $|1_5, \ldots, 1_{M+4}, 0_{M+5}, \ldots\rangle$ (possibly on a countably infinite number of modes). We then obtain the post-selected state $|\psi_{\text{post}}\rangle$

$$\begin{aligned} |\psi_{\text{post}}\rangle_{1..4} &\propto \langle 1_5, \ldots, 1_{M+4}, 0_{M+5}, \ldots | \psi_{\text{out}}\rangle \\ &= \langle 0 | \hat{a}_5 \cdots \hat{a}_{M+4} | \psi_{\text{out}}\rangle \ . \end{aligned} \tag{III.43}$$

In the Bargmann representation the right-hand side of Eq. (III.43) is

$$\partial_5 \cdots \partial_{M+4}\ \psi_{\text{out}}(\vec{\alpha})|_{\vec{\alpha}'=0} \ , \tag{III.44}$$

where I have written $\vec{\alpha}' = (\alpha_5, \ldots, \alpha_{M+4}, \ldots)$.

Writing out the entanglement explicitly in the four modes (treating the polarisation implicitly), I arrived at the following condition for the creation of two photons in the antisymmetric Bell state:

$$\partial_5 \cdots \partial_{M+4}\ \psi_{\text{out}}(\vec{\alpha})|_{\vec{\alpha}'=0} \propto \alpha_1\alpha_2 - \alpha_3\alpha_4 + O(\xi) \ . \tag{III.45}$$

The term $O(\xi)$ will allow for a small pollution ($\xi \ll 1$) in the outgoing state. I will show that for certain special classes of interaction Hamiltonians this condition is very hard (if not impossible) to satisfy. This renders the experimental realisation of two maximally polarisation entangled photons at least highly impractical.

d *Physical limitations on event-ready entanglement*

I am now ready to shape ψ_{out} in more detail. Consider optical circuits including mode-mixing, squeezers and down-converters. The corresponding interaction Hamiltonians H_I are *quadratic* in the creation operators. There are no linear terms, so there are no coherent displacements. More formally

$$H_I = \sum_{i,j=1}^{N} \hat{a}_i^\dagger A_{ij}^{(1)} \hat{a}_j^\dagger + \sum_{i,j=1}^{N} \hat{a}_i^\dagger A_{ij}^{(2)} \hat{a}_j + \text{H.c.} \ . \tag{III.46}$$

With $A^{(1)}$ and $A^{(2)}$ complex matrices (see also Appendix D for more details about the dependence of H on A). According to Braunstein [35], such an active interferometer is equivalent to a passive interferometer V, followed by a set of single-mode squeezers and another passive interferometer U'. The photon source

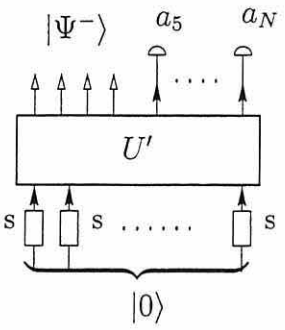

Figure III.3: The unitary interferometer U' with conditional photodetection and single-mode squeezers which should transform $|0\rangle$ into $|\Psi^-\rangle$.

described by Eq. (III.46) can be viewed as an active bilinear component of an interferometer. For vacuum input and after normal ordering [164], the optical setup then gives rise to

$$\psi_{\text{out}}(\vec{\alpha}) = \exp\left[(\vec{\alpha}, B\vec{\alpha})\right] , \tag{III.47}$$

with $(\vec{\alpha}, B\vec{\alpha}) = \sum_{ij}^{N} \alpha_i B_{ij} \alpha_j$. Such an optical setup would correspond to a collection of single-mode squeezers acting on the vacuum, followed by a passive optical interferometer U'. Here, B is a complex symmetric matrix determined by the interaction Hamiltonian H_I and the interferometer U'. We take B to be proportional to a common coupling constant ξ. The outgoing auxiliary modes a_5 to a_N are detected (see Fig. III.3). I will now investigate whether the production of $|\Psi^-\rangle$ conditioned on a given number of detected photons is possible.

In the case of a bilinear interaction Hamiltonian (see Eq. (III.46)), photons are always created in pairs. In addition, we seek to create *two* maximally entangled photons. An odd number of detected photons can never give $|\Psi^-\rangle$ and the number of detected photons should therefore be even. The lowest even number is zero. In this case no photons are detected and ψ_{out} in Eq. (III.47) is proportional to $1 + O(\xi)$, which corresponds to the vacuum state.

The next case involves two detected photons. To have entanglement in modes α_1 to α_4 after detecting *two* photons requires

$$\partial_5 \partial_6 \, e^{(\vec{\alpha}, B\vec{\alpha})}\big|_{\vec{\alpha}'=0} \propto \alpha_1\alpha_2 - \alpha_3\alpha_4 + O(\xi) . \tag{III.48}$$

The left-hand side of Eq. (III.48) is equal to

$$\left(B_{56} + \sum_{i,j=1}^{4} \alpha_i B_{i5} B_{j6} \alpha_j\right) e^{(\vec{\alpha}, B\vec{\alpha})}\big|_{\vec{\alpha}'=0} . \tag{III.49}$$

To satisfy Eq. (III.48), the vacuum contribution B_{56} would have to be negligible. I now investigate whether the second term can give us entanglement. The

right hand side of Eq. (III.48) can be rewritten according to $\alpha_1\alpha_2 - \alpha_3\alpha_4 = \sum_{i,j=1}^{4} \alpha_i E_{ij} \alpha_j$, where E_{ij} are the elements of a symmetric matrix E:

$$E = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} , \tag{III.50}$$

from which it is immediate that seen that $\det E = 1$.

Let $M_{ij} = B_{i5}B_{j6}$. Since only the symmetric part of M contributes, consider $\widetilde{M}_{ij} = (M_{ij} + M_{ji})/2$. The condition for two detected photons now yields

$$\sum_{i,j=1}^{4} \alpha_i \widetilde{M}_{ij} \alpha_j = \sum_{i,j=1}^{4} \alpha_i E_{ij} \alpha_j + O(\xi) , \tag{III.51}$$

If this equality is to hold, we need $\det E = \det \widetilde{M} + O(\xi) = 1$. However, it can be shown that $\det \widetilde{M} = 0$. $\widetilde{M}$ can therefore never have the same form as E for small ξ, so it is *not* possible to create maximal polarisation entanglement conditioned upon two detected photons.

Finally, consider the outgoing state conditioned on four detected photons. Define $X_i \equiv \sum_j B_{ij}\alpha_j$. The left-hand side of Eq. (III.45) for four detected photons then gives

$$\begin{aligned}(&B_{56}B_{78} + B_{57}B_{68} + B_{58}B_{67} + B_{56}X_7X_8 + B_{57}X_6X_8 + B_{58}X_6X_7 \\ &+B_{67}X_5X_8 + B_{68}X_5X_7 + B_{78}X_5X_6 + X_5X_6X_7X_8)\, e^{(\vec{\alpha},B\vec{\alpha})}\big|_{\vec{\alpha}'=0} .\end{aligned}$$

I have not been able either to prove or disprove that $|\Psi^-\rangle$ can be made this way. The number of terms which contribute to the bilinear part in α rapidly increases for more detected photons.

Suppose we *could* create maximal entanglement conditioned upon four detected photons, how efficient would this process be? For four detected photons yielding $|\Psi^-\rangle$ we need at least three photon-pairs. These are created with a probability of the order of $|\xi|^6$. Currently, $|\xi|^2$, the probability per mode, has a value of 10^{-4} [173]. For experiments operating at a repetition rate of 100 MHz using ideal detectors, the procedure conditioned on four detected photons will amount to approximately one maximally entangled pair every few hours. For realistic detectors this is much less.

So far, there have been no experiments which exceeded the detection of more than two *auxiliary* photons (not including the actual detection of the maximally entangled state). This, and the estimation of the above efficiency appears to place strong practical limitations on the creation of maximal entanglement.

e Six detected photons

Recently, Knill, Laflamme and Milburn have discovered a method which allows us to create event-ready entanglement conditioned on *six* detected photons [98]. This method involves the construction of the C-SIGN operator

$$U_{\text{C-SIGN}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \tag{III.52}$$

on the basis $\{|0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle\}$. In quantum optics, these two qubits are defined on four distinct modes a_1, a_2, a_3 and a_4.

The C-NOT $U_{\text{C-NOT}}$ is then defined using the Hadamard transform H on modes a_3 and a_4 as

$$\begin{aligned} U_{\text{C-NOT}} &= H^{\dagger}_{a_3,a_4} U_{\text{C-SIGN}} H_{a_3,a_4} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} . \end{aligned} \tag{III.53}$$

Applying the C-NOT and the Hadamard transformation on a separable state yields a maximally entangled state.

How do we construct the C-SIGN operator? Following Knill *et al.* this amounts to the construction of the operator

$$A : \alpha_0|0\rangle + \alpha_1|1\rangle + \alpha_2|2\rangle \longrightarrow \alpha_0|0\rangle + \alpha_1|1\rangle - \alpha_2|2\rangle . \tag{III.54}$$

First, we apply a beam-splitter B_θ:

$$B_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} , \tag{III.55}$$

with $\theta = \pi/4$ to modes a_1 and a_3, which transforms $|1,1\rangle$ into $(|2,0\rangle - |0,2\rangle)/\sqrt{2}$. Subsequently, we apply the operator A to modes a_1 and a_3, and finally, we apply a beam-splitter $B_{-\pi/4}$ again to these modes.

The operator A is defined as a unitary transformation U_A on three modes, one main mode and two auxiliary modes which are detected:

$$U_A = \begin{pmatrix} 1-\sqrt{2} & \frac{1}{\sqrt[4]{2}} & \sqrt{\frac{3}{\sqrt{2}} - 2} \\ \frac{1}{\sqrt[4]{2}} & \frac{1}{2} & \frac{1}{2} - \frac{1}{\sqrt{2}} \\ \sqrt{\frac{3}{\sqrt{2}} - 2} & \frac{1}{2} - \frac{1}{\sqrt{2}} & \sqrt{2} - \frac{1}{2} \end{pmatrix} . \tag{III.56}$$

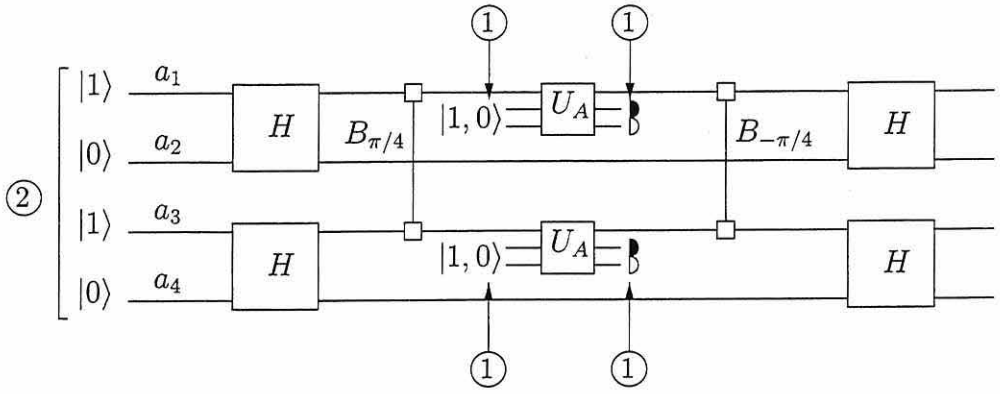

Figure III.4: Circuit for event-ready entanglement conditioned on six detected photons. Here, U_A is given by Eq. (III.56), H is the Hadamard transform and B_θ is a beam-splitter. The encircled numbers denote the number of detected photons needed to create the corresponding states.

This transformation can be explicitly constructed using the techniques developed by Reck *et al.* [140]. The input state on the two auxiliary modes b_1 and b_2 is $|1,0\rangle$, and the operator is conditioned on a state $|1,0\rangle$ in the outgoing auxiliary modes. This post-selection means that A is a probabilistic operator with a probability of success of $1/4$.

Event-ready entanglement can now be created using the setup shown in figure III.4. The incoming state is given by $|1,0,1,0\rangle$, which can be made conditioned on two detected photons. The input states on the auxiliary modes can be created conditioned on one detected photon and the outgoing auxiliary modes also involve one detected photon. Therefore the total number of detected photons is six.

4 Summary

I have demonstrated strong limitations on the possibility of creating maximal entanglement with quantum optics. To this end, I introduced two simplifications to the hypothetical optical circuit: I replaced feed-forward detection by a fixed set of detectors at the end, and secondly, every detector needs to detect at most one photon. Conditioned on two detected photons, multi-mode squeezed vacuum fails to create maximal entanglement.

What happens when we have a combination of squeezing and coherent displacements? In that case the approach taken here fails due to the more complex normal ordering of the interaction Hamiltonian. Also, I have only considered ideal detections, but how do realistic detectors affect the outgoing state? This is the subject of the next chapter.

IV

Auxiliary Resources: Detection Devices

Wouldn't it be nice to have a machine which creates the quantum states of your choice at the push of a button? Unfortunately, these machines do not yet exist[1]. There are currently machines which create certain specific states, like for instance lasers (creating coherent states) and down-converters (creating squeezed states), but notwithstanding their importance for scientific and technological applications, these devices create only a limited class of quantum states.

When we want to create more exotic quantum states, we need to extend our resources: in addition to the devices mentioned above we may use passive transformations (like, for instance, beam-splitters and phase-shifters in quantum optics) and *measurements*. With this new set of tools we can build more sophisticated state preparation devices, or 'circuits'. As I have shown in the previous chapter, depending on the particular physical implementation of these circuits we can create more exotic quantum states. State preparation has been studied among others by Vogel *et al.* [168], Harel *et al.* [81], Dakna *et al.* [46] and Rubin [143].

Having extended our resources to state preparation circuitry, the next issue is the *quality* of the state preparation. Suppose we want to create a particular state. In practice, we can never obtain this state perfectly, due to uncontrollable effects like decoherence and measurement errors. Nevertheless, we want our maximise the quality of the state preparation process.

Formulating this more precisely, we want to prepare a *single* (pure) state $|\phi\rangle$ by means of some process, and we want the resulting state ρ to be as 'close' to $|\phi\rangle$ as possible. In chapter II we have seen that a measure of resemblance between states is given by the fidelity F:

$$F = \mathrm{Tr}[\rho|\phi\rangle\langle\phi|] \,. \tag{IV.1}$$

The quality of a state preparation process can therefore be measured by the fidelity. When $F = 1$, the process gives exactly $|\phi\rangle$ and when $F = 0$, the prepared state is orthogonal to $|\phi\rangle$. In practice, the fidelity will not reach these extreme measures, but will lie between 0 and 1.

In short, we have a state preparation circuit which creates states with some fidelity. Generally, the preparation process is *conditioned* on measurements [97].

[1] Quantum computers will be able to make such states for qubits.

For example, if we want to prepare a single-photon state $|1\rangle$ in quantum optics we can use the following process: a parametric down-converter creates a state $|\psi\rangle_{ab}$ on two spatial modes a and b (see chapter III):

$$|\psi\rangle_{ab} \propto |0\rangle_a|0\rangle_b + \xi|1\rangle_a|1\rangle_b + O(\xi^2) , \tag{IV.2}$$

where $|0\rangle$ denotes the vacuum state and we assume $\xi \ll 1$. The higher order terms (included in $O(\xi^2)$) consist of states with more than two photons. We now place a photo-detector in mode a, which 'clicks' when it sees one or more photons (typically, standard detectors can see single photons, but fail to distinguish between one and two photons). Conditioned on such a click, mode b will be in a state

$$\rho \propto |1\rangle_b\langle 1| + O(|\xi|^2) . \tag{IV.3}$$

The fidelity of this process is high: $F = \langle 1|\rho|1\rangle \simeq 1$, and this is therefore typically a very good single-photon state preparation process (although the situation changes drastically when multiple down-converters are considered [32, 99]). Due to the large vacuum contribution, however, the probability of the detector giving a 'click' will be small (of order $O(|\xi|^2)$). When the detector does not click, that particular trial is dismissed, hence the *conditional* character of the detection.

In this example the outcome of the detection is used to either accept or reject a particular run of the state preparation device. However, in general the outcome of the detector can be used to determine a more complicated operation on the remainder of the state preparation process. This is detection plus *feed-forward*, since the outcome is used further on in the process. An example of this is quantum teleportation, where the outcome of the Bell measurement determines the unitary transformation needed to retrieve the original input state.

When the measurements in the state preparation process are prone to errors, the state we want to create may not be the state we actually create. This means that errors in the detection devices can lead to reduced fidelities. In this chapter I study the effect of detection errors on state preparation. To this end I introduce the concept of the *confidence* of preparation. Using this measure I evaluate different types of detection devices. This chapter is based on Kok and Braunstein [103].

1 Confidence

Consider a preparation device which prepares a state conditioned on a single measurement. For simplicity, I employ two subsystems. One subsystem will be measured, leaving a quantum state in the other. It is clear that prior to the measurement the two systems have to be entangled. Otherwise conditioning on the measurement does not have any effect on the state of the second system.

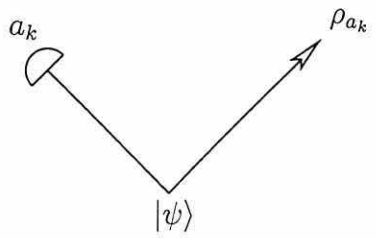

Figure IV.1: A schematic representation of state preparation conditioned on a measurement. One branch of the entanglement $|\psi\rangle$ is detected, yielding an eigenvalue a_k. The other branch is now in a state ρ_{a_k}.

We can write the total state $|\psi\rangle_{12}$ prior to the measurement in the Schmidt decomposition:

$$|\psi\rangle_{12} = \sum_k c_k |a_k\rangle_1 |b_k\rangle_2 \,, \tag{IV.4}$$

with $\{|a_k\rangle\}$ and $\{|b_k\rangle\}$ orthonormal sets of states for system 1 and 2 respectively. These states correspond to eigenstates of observables A and B with sets of eigenvalues $\{a_k\}$ and $\{b_k\}$ respectively. We now measure the observable A in system 1, yielding an outcome a_k (see Fig. IV.1).

We can model this measurement using so-called projection operator valued measures, or POVM's for short. For *ideal* measurements, we can describe the measurement of mode 1 as a projection $P_k = |a_k\rangle\langle a_k|$ operating on the state $|\psi\rangle_{12}$. When we trace out the first system the (normalised) state of the second system will be

$$\rho_{a_k} = \frac{\mathrm{Tr}_1[(P_k \otimes \mathbb{1})|\psi\rangle_{12}\langle\psi|]}{\mathrm{Tr}_1[(P_k \otimes \mathbb{1})|\psi\rangle_{12}\langle\psi|]} = |b_k\rangle\langle b_k| \,. \tag{IV.5}$$

For non-ideal measurements we do not use a projection operator, but rather a *projection operator valued measure*. In general, a POVM E_ν can be written as

$$E_\nu = \sum_\mu d_{\mu\nu} \mathcal{P}_\mu \geq 0 \,, \tag{IV.6}$$

where the $\mathcal{P}_\mu$'s form a set[2] of projection operators $\{|\mu\rangle\langle\mu|\}_\mu$. We also require a completeness relation

$$\sum_\nu E_\nu = \mathbb{1} \,. \tag{IV.7}$$

a more general definition of POVM's is given by (see appendix B):

$$E_\nu = \sum_\mu d_{\mu\nu} \mathcal{P}_\mu = \sum_\mu d_{\mu\nu} |\mu\rangle\langle\mu| = \sum_\mu d_{\mu\nu} |\mu\rangle\langle\nu|\nu\rangle\langle\mu|$$

[2]This set is possibly over-complete, hence the difference in notation from P_k.

$$= \sum_{\mu} \left(u^*_{\mu\nu}|\mu\rangle\langle\nu|\right)\left(u_{\mu\nu}|\nu\rangle\langle\mu|\right) = \sum_{\mu} \mathcal{A}^{\dagger}_{\mu\nu}\mathcal{A}_{\mu\nu} \,. \tag{IV.8}$$

The operator $\mathcal{A}_{\mu\nu}$ is generally not unique. These POVM's are used to model non-ideal measurements.

As mentioned before, a measurement outcome a_k in mode 1 gives rise to an outgoing state ρ_{a_k} in mode 2. We cannot describe a non-ideal measurement with the projection $P_k = |a_k\rangle\langle a_k|$. Instead, we have a POVM E_k (corresponding to the outcome a_k), which reduces to P_k in the case of an ideal measurement. Let $\rho_{12} = |\psi\rangle_{12}\langle\psi|$, the entangled state prior to the measurement. The outgoing state in mode b will then be

$$\rho_{a_k} = \frac{\mathrm{Tr}_1[(E_k \otimes \mathbb{1})\rho_{12}]}{\mathrm{Tr}[(E_k \otimes \mathbb{1})\rho_{12}]} \,, \tag{IV.9}$$

where the total trace over both systems in the denominator gives the proper normalisation.

If we had an ideal detector (corresponding to $E_k = |a_k\rangle\langle a_k|$), the outgoing state would be $\rho_{a_k} = |b_k\rangle\langle b_k|$. However, with the general POVM E_k, this will not be the case. The resulting state will be different. In order to quantify the reliability of a state preparation process I introduce the *confidence* of a process.

Definition: The *confidence* in the preparation of a particular state is given by the fidelity of the preparation process.

That means that using Eqs. (IV.4) and (IV.9) the confidence C is given by

$$C = \frac{\mathrm{Tr}[(E_k \otimes |b_k\rangle\langle b_k|)\rho_{12}]}{\mathrm{Tr}[(E_k \otimes \mathbb{1})\rho_{12}]} = \frac{|c_k|^2\langle a_k|E_k|a_k\rangle}{\sum_l |c_l|^2\langle a_l|E_k|a_l\rangle} \,, \tag{IV.10}$$

where the $|c_l|^2$ are the diagonal elements of the density matrix. The confidence C can be interpreted as the probability of obtaining outcome a_k from the 'branch' containing $|a_k\rangle$ in Eq. (IV.4) divided by the unconditional probability of obtaining outcome a_k. We will also call this the 'confidence of state preparation'.

This interpretation suggests that there does not need to be a second system to give the idea of confidence meaning. Suppose, for instance, that we have an 'electron factory' which produces electrons with random spin. A Stern-Gerlach apparatus in the path of such an electron will make a spin measurement along a certain direction $\mathbf{r}$. Suppose we find that the electron has spin 'up' along $\mathbf{r}$. Before this measurement the electron was in a state of random spin ($\rho_{\mathrm{in}} = \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow|$), and after the measurement the electron is in the 'spin up' state ($\rho_{\mathrm{out}} = |\uparrow\rangle\langle\uparrow|$). The state of the electron has *collapsed* into the 'spin up' state. I will now investigate how we can define the confidence of the detection of a single system.

Formally, we can model state collapse by means of the super-operator $\hat{\mathcal{F}}_{a_k}$, where a_k is again the outcome of the measurement of observable A ('spin up'

in the above example). In general, a super-operator yields a (non-normalised) mapping $\rho \rightarrow \hat{\mathcal{F}}_\mu(\rho)$. In the POVM representation used above (see Eq. (IV.8)) we can write this as

$$\hat{\mathcal{F}}_\mu : \rho \longrightarrow \sum_\nu \mathcal{A}_{\mu\nu}\, \rho\, \mathcal{A}_{\mu\nu}^\dagger \,. \tag{IV.11}$$

When the eigenstate corresponding to a_k is given by $|a_k\rangle$, we can define the confidence of this measurement as

$$C_{\mathrm{m}} = \frac{\langle a_k|\hat{\mathcal{F}}_{a_k}(\rho)|a_k\rangle}{\mathrm{Tr}[\hat{\mathcal{F}}_{a_k}(\rho)]} = \frac{\mathrm{Tr}[\hat{\mathcal{F}}_{a_k}(\rho)|a_k\rangle\langle a_k|]}{\mathrm{Tr}[\hat{\mathcal{F}}_{a_k}(\rho)]} \,, \tag{IV.12}$$

with $\mathrm{Tr}[\hat{\mathcal{F}}_{a_k}(\rho)]$ the proper normalisation. However, this expression depends strongly on the details of the family of operators $\mathcal{A}_{\mu\nu}$. This is a more complicated generalisation than the POVM's E_k. The confidence of state preparation, on the other hand, is a function of the POVM E_k. Furthermore, C_{m} will in general *not* be equal to the confidence of state preparation derived in Eq. (IV.10).

In conclusion, there are two distinct versions of the confidence: the confidence of *measurement* and the confidence of *state preparation*. Later in this chapter I will use the concept of the confidence to make a quantitative comparison between different detection devices. This suggests that we need to calculate the confidence of measurement with all its difficulties. One way to circumvent this problem is to calculate the the confidence of state preparation using a fixed state. Instead of concentrating on the state preparation process we now choose a standard input state and calculate the confidence for different types of measurement devices. One such choice might be the maximally entangled state

$$|\Psi\rangle_{12} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |a_k, a_k\rangle \,. \tag{IV.13}$$

When $N \rightarrow \infty$, this is perhaps not the ideal choice and another state may be preferred. For any choice, the confidence offers a quantitative measure of performance for different types of measurement devices.

2 Optical detection devices

Having set the stage for state preparation conditioned on measurement outcomes, I will now restrict the remainder of this chapter to optical implementations. Let's consider the measurement of optical Fock states using photo-detectors. In order to classify different types of detectors I use the following terminology: a detector is said to have a *single-photon sensitivity* when it is sensitive enough to detect a single-photon wave-packet. When a detector can distinguish between n- and $(n+1)$-photon wave-packets, it is said to have a *single-photon resolution.*

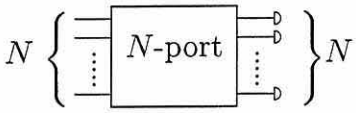

Figure IV.2: An N-port with unit-efficiency, non-resolving detectors. The N incoming modes are unitarily transformed into N output modes. The N-ports considered here consist of mirrors and beam-splitters and do not mix creation operators with annihilation operators.

Real detectors have a variety of characteristics. Most common detectors do not have single-photon resolution, although they can distinguish between a few and many photons. When small photon numbers are detected, however, these are single-photon sensitivity detectors to a good approximation. There are also single-photon resolution detectors [93, 162]. Currently, these detectors require demanding operating conditions.

When we need single-photon resolution but do not have the resources to employ single-photon resolution detectors, we can use a so-called *detector cascade* [158]. In a detector cascade an incoming mode (populated by a number of photons) is split into N output modes with equal amplitude which are all detected with single-photon sensitivity detectors. The idea is to choose the number of output modes large enough, so that the probability that two photons enter the same detector becomes small. In general, an optical setup which transforms N incoming modes into N outgoing modes is called an N-port (see Fig. IV.2) [140]. A detector cascade is a symmetric N-port with detectors at the outgoing modes and vacuum states in all input modes except the first mode. In the next section I will study the statistics of symmetric N-ports, but first I need to elaborate on the types of errors which occur in detectors.

There are two sources of errors for a detector: it might fail to detect a photon, or it might give a signal although there wasn't actually a photon present. The former may be characterised as a 'detector loss' and the latter as a 'dark count'. Here, the emphasis will be on detector losses. Later I will give a model for incorporating dark counts in a realistic detector model. In some experiments (like the Innsbruck teleportation experiment [23]) the detectors operate within short gated time intervals. This greatly reduces the effect of dark counts.

Detector losses are not so easily dismissed. Every photon entering a detector has a certain probability of triggering it. This probability is called the *efficiency* of the detector. For the purposes of brevity, when a detector is perfectly efficient, we will call it a *unit-efficiency* detector. When it has some lower efficiency, we speak of a *finite-efficiency* detector. Here, I study detector cascading with unit-efficiency detectors, as well as cascading with finite-efficiency detectors [181]. I am interested in the case where cascading distinguishes between photon-number states $|k\rangle$ and $|k'\rangle$ with $k \simeq k'$.

3 N-PORTS

In this section I treat the properties of detector cascades, or symmetric N-ports with single-photon sensitivity detectors in the outgoing modes. Symmetric N-ports yield a (unitary) transformation U of the spatial field modes a_k, with $j, k = 1, \ldots, N$:

$$\hat{b}_k \to \sum_{j=1}^{N} U_{jk}\hat{a}_j \qquad \text{and} \qquad \hat{b}_k^\dagger \to \sum_{j=1}^{N} U_{jk}^*\hat{a}_j^\dagger \,, \tag{IV.14}$$

where the incoming modes of the N-port are denoted by a_j and the outgoing modes by b_j. Here, $\hat{a}_j^\dagger$ and $\hat{a}_j$ are the respective creation and annihilation operators of mode a_j. Similarly for mode b_k. The unitary matrix U can be chosen to be

$$U_{jk} = \frac{1}{\sqrt{N}} \exp[2\pi i(j-1)(k-1)/N] \tag{IV.15}$$

without loss of generality up to an overall phase-factor. Paul *et al.* have studied such devices in the context of tomography and homodyne detection [157, 4, 129].

Here, I study N-ports in the context of optical state preparation, where only one copy of a state is given, instead of an ensemble. I will use the concept of the *confidence*, introduced in section 1.

a *Statistics of N-ports*

Suppose we have a detector cascade, consisting of a symmetric N-port with single-photon sensitivity detectors in the outgoing modes. According to Eqs. (IV.14) and (IV.15) incoming photons will be redistributed over the outgoing modes. In this section I study the photon statistics of this device. In particular, I study the case where k photons enter a single input mode of the N-port, with vacuum in all other input modes. This device (i.e., the detector cascade) will act as a sub-ideal single-photon resolution detector since there is a probability that some of the photons end up in the same outgoing mode, thus triggering the same detector.

To quantify the single-photon resolution of the cascade we use the confidence given by Eq. (IV.10). Suppose we have two spatially separated entangled modes of the electro-magnetic field a and b with number states $|m\rangle$ in a and some other orthogonal states $|\phi_m\rangle$ in b:

$$|\Psi\rangle = \sum_m \gamma_m |m\rangle_a |\phi_m\rangle_b \,, \tag{IV.16}$$

where the second mode is used only to give the confidence an operational meaning. The POVM governing the detection can be written as $E_k = \sum_m p_N(k|m)|m\rangle\langle m|$,

since we assume that the photons are not lost in the N-port. In this expression $p_N(k|m)$ is the probability that m incoming photons cause a k-fold detector coincidence in the N-port cascade. The confidence can then be written as

$$C = \frac{|\gamma_k|^2\langle k|E_k|k\rangle}{\sum_l |\gamma_m|^2\langle m|E_k|m\rangle} = \frac{|\gamma_k|^2 p_N(k|k)}{\sum_m |\gamma_m|^2 p_N(k|m)} \,. \tag{IV.17}$$

In order to find the confidence, I therefore first have to calculate the probability distribution p_N. This will allow us to compare single-photon resolution detectors with various arrangements (N-ports) of single-photon sensitivity detectors.

Suppose k photons enter the first input mode and all other input modes are in the vacuum state. The density matrix of the pure input state $\rho_0 = |k\rangle\langle k|$ will be transformed according to $\rho = U_N \rho_0 U_N^\dagger$ with U_N the unitary transformation associated with the symmetric N-port. Let $\vec{n}$ be the N-tuple of the photon number in every outgoing mode: $\vec{n} = (n_1, n_2, \ldots, n_N)$. The probability of finding n_1 photons in mode 1 and n_2 photons in mode 2, et cetera, is given by $p_{\vec{n}} = \langle \vec{n}|\rho|\vec{n}\rangle$. Using the N-port transformation this probability yields

$$p_{\vec{n}} = \langle \vec{n}|U_N \rho_0 U_N^\dagger|\vec{n}\rangle = |\langle \vec{n}|U_N|\vec{k}\rangle|^2 \,, \tag{IV.18}$$

where $\vec{k} = (k, 0, \ldots, 0)$, since only the first input mode inhabits photons and the rest are vacuum. From Refs. [53] and [54] we find that this can be rewritten as

$$p_{\vec{n}} = \frac{\left[H^R_{\vec{k}\vec{n}}(0)\right]^2}{n_1! \cdots n_N! k!} \,. \tag{IV.19}$$

Here, $H^R_{\vec{k}\vec{n}}(\vec{x})$ is a so-called multi-dimensional Hermite polynomial (MDHP) [52] (this is a non-trivial result; see appendix G for a comprehensive treatment of multi-dimensional Hermite polynomials) and the matrix R is defined as

$$R \equiv \begin{pmatrix} 0 & -U^\dagger \\ -U^\dagger & 0 \end{pmatrix} \,. \tag{IV.20}$$

For our present purposes it is convenient to characterise the N-port by its transformation of the field modes given by Eqs. (IV.14) and (IV.15). I therefore concentrate on U rather than U_N.

Since there is a one-to-one correspondence between the N-port (U) and the matrix R, knowledge of U is sufficient to calculate the confidence of a given event using the N-port. The MDHP for N input modes with k photons in the first mode and zero in the others (giving an N-tuple $\vec{k}$) and N output modes $\vec{n}$ is given by

$$H^R_{\vec{k}\vec{n}}(\vec{x}) = (-1)^{2k}\, e^{\frac{1}{2}\vec{x}R\vec{x}^T}\, \nabla^{2k}_{\vec{k}\vec{n}}\, e^{-\frac{1}{2}\vec{x}R\vec{x}^T} \,, \tag{IV.21}$$

where $\vec{x}\,R\,\vec{x}^T = \sum_{ij} x_i R_{ij} x_j$, $\vec{x} = (x_1, \ldots, x_{2N})$ and

$$\nabla^{2k}_{\vec{k}\vec{n}} \equiv \frac{\partial^{2k}}{\partial x_1^k \partial x_{N+1}^{n_1} \cdots \partial x_{2N}^{n_N}} .$$

The number of photons in the input mode is equal to the total number of photons in the output modes. The dimension of $\vec{x}$ obeys $\dim \vec{x} = \dim \vec{k} + \dim \vec{n} = 2N$. For example, for a two-photon input state we have

$$e^{\frac{1}{2}\vec{x}\,R\,\vec{x}^T} \frac{\partial^4}{\partial x_1^2 \partial x_l \partial x_k} e^{-\frac{1}{2}\vec{x}\,R\,\vec{x}^T}\Big|_{\vec{x}=0} = 2R_{1l}R_{1k} . \qquad \text{(IV.22)}$$

There are many different ways in which k incoming photons can trigger a k-fold detector coincidence. These different ways correspond to different photon distributions in the outgoing (detected) modes, and are labelled by $\vec{n}_r$. The probability that all k photons enter a different detector is found by determining the $p_{\vec{n}_r}$s where every n_i in $\vec{n}_r$ is at most one. The sum over all these $p_{\vec{n}_r}$'s is equal to the probability $p_N(k|k)$ of a k-fold coincidence in an N-port conditioned on k incoming photons:

$$p_N(k|k) = \sum_{\vec{n}_r} p_{\vec{n}_r} = \frac{k!}{N^k} \binom{N}{k} . \qquad \text{(IV.23)}$$

Finally, in order to find the probability of a k-fold detector coincidence conditioned on m photons in the input state (with $m \geq k$) we need to sum all probabilities in Eq. (IV.19) with k non-zero entries in the N-tuple $\vec{n}$:

$$p_N(k|m) = \sum_{\vec{n}\in\mathcal{S}_k} \frac{\left[H^R_{\vec{m}\vec{n}}(0)\right]^2}{n_1! \cdots n_N! m!} , \qquad \text{(IV.24)}$$

where $\mathcal{S}_k$ is the set of all $\vec{n}$ with exactly k non-zero entries.

b *Realistic N-ports*

I now consider a symmetric N-port cascade with finite-efficiency single-photon sensitivity detectors. Every one of the N detectors has a certain loss, which means that some photons do not trigger the detector they enter. We can model this situation by putting a beam-splitter with intensity transmission coefficient η^2 in front of the ideal detectors [181]. The reflected photons are sent into the environment and can be associated with the loss. The transmitted photons are detected (see Fig. IV.3).

Before I continue with the description of realistic detector cascades, let me return to the question of the dark counts. In the model for detectors with finite efficiency I assumed a beam-splitter with intensity transmission coefficient η^2 and

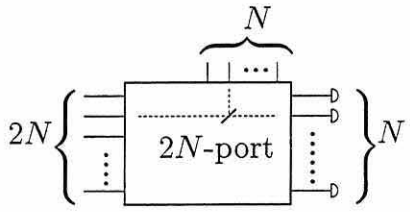

Figure IV.3: A $2N$-port with N modes which are detected with ideal detectors and N undetected modes. These modes are associated with the detector losses.

vacuum in the second input mode. We can now model dark counts by replacing this vacuum state with a thermal input state ρ_{th} [148]:

$$\rho_{\text{th}} = \sum_{n=0}^{\infty} \int_0^{\infty} d\nu f(\nu) \left[1 - \exp\left(\frac{-\hbar\nu}{k_B T_{\text{eff}}}\right)\right] \exp\left(\frac{-n\hbar\nu}{k_B T_{\text{eff}}}\right) |n\rangle\langle n| \,, \qquad \text{(IV.25)}$$

with ν the frequency, $f(\nu)$ the appropriate frequency distribution ($\int_0^{\infty} d\nu f(\nu) = 1$), k_B Boltzmann's constant and T_{eff} the effective temperature 'seen' by the detector. However, for the remainder of this chapter I will assume that there are no dark counts ($T_{\text{eff}} = 0$).

Let us now consider cascades with finite-efficiency detectors. The implementation of the beam-splitters responsible for the detector losses transform our N-port into a $2N$-port and the unitary transformation U of the field modes in this N-port now becomes a $2N \times 2N$ unitary matrix $U \rightarrow U \otimes \mathbb{1}_2$ (where $\mathbb{1}_2$ is the two dimensional unit matrix). Applying a transformation V_η to implement the beam-splitters with transmission coefficient η^2 will give a new unitary transformation governing the behaviour of the $2N$-port. Although nothing holds us from considering detectors with different efficiencies, for simplicity I will assume that all detectors have the same efficiency η^2. In terms of the original unitary matrix U from Eq. (IV.15) the new unitary matrix $\widetilde{U}$ becomes

$$U \rightarrow \widetilde{U} = \begin{pmatrix} \eta\, U & \sqrt{1-\eta^2}\, U \\ -\sqrt{1-\eta^2}\, U & \eta\, U \end{pmatrix} .$$

This changes the matrix R of the MDHP accordingly:

$$R \rightarrow \widetilde{R} = \begin{pmatrix} 0 & -\widetilde{U}^{\dagger} \\ -\widetilde{U}^{\dagger} & 0 \end{pmatrix} \qquad \text{(IV.26)}$$

and $\widetilde{R}$ is now a $4N \times 4N$ matrix dependent on η. The probability of finding a k-fold detector coincidence in an N-port cascade with finite-efficient detectors

then becomes

$$p_N(k|m) = \sum_{\vec{n}\in\mathcal{S}_k} \frac{\left[H^{\tilde{R}}_{\vec{m}\vec{n}}(0)\right]^2}{n_1!\cdots n_{2N}!m!}\,, \tag{IV.27}$$

where $\mathcal{S}_k$ is the set of all $\vec{n}$ with exactly k non-zero entries in the detected modes (note that I still call it an N-port although technically it is a $2N$-port). The confidence of having a total of k photons in a k-fold detector coincidence is again given by Eq. (IV.17). The variables of the MDHP will be a $2N$-tuple $\vec{k} = (k, 0, \ldots 0)$. The output photon number $2N$-tuple can now be written as $\vec{n} = (n_1^d, n_2^d, \ldots n_N^d, n_1^u, \ldots n_N^u)$, where the superscripts d and u again denote the detected and undetected modes respectively. Furthermore we have $\sum_{i=1}^N n_i^d \equiv N_d$ and $\sum_{i=1}^N n_i^u \equiv N_u$.

Using Eq. (IV.23) and observing that every detected photon carries a factor η^2 it is quite straightforward to obtain the probability that k photons give a k-fold coincidence in an efficient N-port cascade:

$$p_N(k|k) = \frac{\eta^{2k}N!}{N^k(N-k)!}\,. \tag{IV.28}$$

c The single-photon resolution of N-ports

Having determined the probability distribution p_N, I can now calculate the confidence of detector cascading. First of all, in order to obtain a high confidence in the outcome of a detector cascade, the possible number of photons should be much smaller than the number of modes in the cascade: $N \gg k$. In practice there is a limit to the number of detectors we can build a cascade with, so I only look at the lowest order: distinguishing between one and two photons.

I will calculate the confidence of having outgoing state $|\phi_1\rangle$ conditioned a single detector giving a 'click' in the detector cascade when the input state is given by

$$|\Psi\rangle_{12} = \alpha|0\rangle_1|\phi_0\rangle_2 + \beta|1\rangle_1|\phi_1\rangle_2 + \gamma|2\rangle_1|\phi_2\rangle_2\,. \tag{IV.29}$$

This state corresponds, for example, to the output of a down-converter when we ignore higher-order terms. The confidence is then

$$C_N(1, |\Psi\rangle_{12}) = \frac{|\beta|^2 p_N(1|1)}{|\alpha|^2 p_N(1|0) + |\beta|^2 p_N(1|1) + |\gamma|^2 p_N(1|2)}\,. \tag{IV.30}$$

Eqs. (IV.27) and (IV.22) allow us to calculate the probabilities of a zero-, one- and two-fold detector coincidence conditioned on one or two incoming photons:

$$p_N(0|0) = 1 \tag{IV.31}$$

$$p_N(1|0) = 0 \tag{IV.32}$$

$$p_N(0|1) = 1-\eta^2 \tag{IV.33}$$
$$p_N(1|1) = \eta^2 \tag{IV.34}$$

$$p_N(0|2) = (1-\eta^2)^2 \tag{IV.35}$$
$$p_N(1|2) = \frac{\eta^4}{N} + 2\eta^2(1-\eta^2) \tag{IV.36}$$
$$p_N(2|2) = \frac{N-1}{N}\eta^4\,, \tag{IV.37}$$

For example, using these probabilities, together with Eq. (IV.29), gives us an expression for the confidence that a single detector hit was triggered by one photon ($\delta = |\gamma|^2/|\beta|^2$):

$$C = \frac{N}{N+\delta[\eta^2+2N(1-\eta^2)]}\,, \tag{IV.38}$$

where, for simplicity, we omitted the functional dependence of C on the incoming state, the size of the cascade and the order of the detector coincidence.

A close look at Eq. (IV.36) shows us that $p_N(1|2)$ includes a term which is independent of the number of modes in the N-port cascade. This term takes on a maximum value of $1/2$ for $\eta^2 = \frac{1}{2}$. However, the confidence is a monotonously increasing function of η^2. As expected, for small δ's the confidence $C_N(1,|\Psi\rangle)$ approaches 1. Detector cascading thus turns a collection of single-photon sensitivity detectors into a device with *some* single-photon resolution. In the next section I will give a quantitative estimation of this resolution.

4 Comparing detection devices

Let's return again to the schematic state preparation process depicted in figure IV.1. There we had two modes, one of which was detected, giving the prepared outgoing state in the other. I argued that different detection devices yield different output states, and the comparison of these states with the ideal case (where we used an ideal detector) led to the introduction of the confidence of a state preparation process. Here, I will use the confidence to make a comparison of different *detection devices*, rather than output states. This can be done by choosing a fixed entangled input state. The confidence then quantifies the performance of these detection devices.

Consider the state preparation process in the setting of quantum optics. We have two spatial modes of the electro-magnetic field, one of which is detected. In this thesis I am mostly interested in states containing a few photons, and the detection devices I consider therefore include single-photon sensitivity detectors, single-photon resolution detectors and detector cascades. As an example, I set the task of distinguishing between one and two photons. Since single-photon

sensitivity detectors are not capable of doing this, I will compare the performance of detector cascading with that of a single-photon resolution detector. Let the state prior to the detection be given by

$$|\Psi\rangle = \frac{1}{\sqrt{3}}\left(|0\rangle|\phi_0\rangle + |1\rangle|\phi_1\rangle + |2\rangle|\phi_2\rangle\right) . \tag{IV.39}$$

This state is maximally entangled and will serve as our 'benchmark' state. It corresponds to the choice $\delta = 1$ in the previous section. Suppose the outgoing state conditioned on a 'one-photon' indication in the detection device is ρ. The confidence is then again given by $C = \langle\phi_1|\rho|\phi_1\rangle$.

First, consider the single-photon resolution detector described in Refs. [93, 162]. This detector can distinguish between one and two photons very well, but it does suffer from detector losses (the efficiency was determined at 88%). That means that a two-photon state can be identified as a single-photon state when one photon is lost. The confidence of this detector is therefore not perfect.

In order to model the finite efficiency of the single-photon resolution detector we employ the beam-splitter model from section *b*. We write the input state as

$$|\Psi\rangle = \frac{1}{\sqrt{3}}\left(|0\rangle|\phi_0\rangle + \hat{a}^\dagger|0\rangle|\phi_1\rangle + \frac{(\hat{a}^\dagger)^2}{\sqrt{2}}|0\rangle|\phi_2\rangle\right) . \tag{IV.40}$$

When we make the substitution $\hat{a}^\dagger \to \eta\hat{b}^\dagger + \sqrt{1-\eta^2}\hat{c}^\dagger$ we obtain a state ρ. The outgoing density matrix conditioned on a single photon in mode b is then

$$\rho_{\text{out}} = \frac{\text{Tr}_{bc}[(|1\rangle_b\langle 1| \otimes \mathbb{1}_c)\rho]}{\text{Tr}[(|1\rangle_b\langle 1| \otimes \mathbb{1}_c)\rho]} = \frac{\eta^2}{4-3\eta^2}|\phi_1\rangle\langle\phi_1| + \frac{4(1-\eta^2)}{4-3\eta^2}|\phi_2\rangle\langle\phi_2| . \tag{IV.41}$$

With $\eta^2 = 0.88$ the confidence of the single-photon resolution detector is easily calculated to be $C = 0.65$.

Now we consider a detector cascade with single-photon sensitivity detectors. In Fig. IV.4 the confidence of a single-photon detection with N-port cascades is depicted. When the cascade consists of four detectors ($N = 4$) it can be easily calculated from Eq. (IV.38) that the detectors need an efficiency of 0.84 to achieve a 0.65 confidence. In the case of infinite cascading ($N = \infty$) the single-photon confidence of 0.65 is met only if the efficiency is roughly 0.73. This puts a severe practical limit on the efficiency of the single-photon sensitivity detectors in the cascade.

Detector cascading would be practically useful if a reasonably small number of finite-efficiency detectors yields a high confidence. In particular when cascading is viewed as an economical alternative to a detector with single-photon resolution the number of detectors in the cascade should be small. Additionally, cascading should yield a confidence similar to single-photon resolution detectors. Unfortunately, as a practical application, detector cascading only appears to yield a

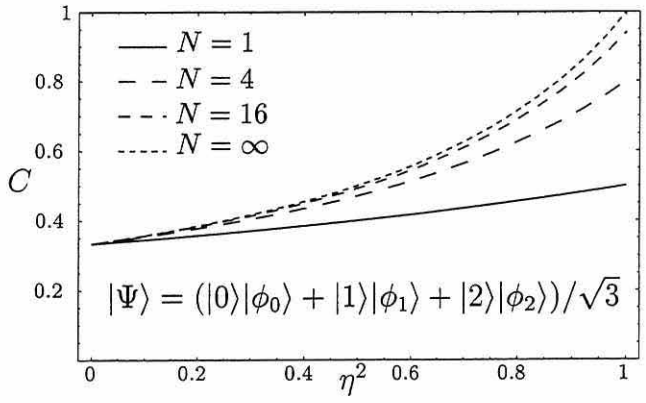

Figure IV.4: The single-photon confidence C [Eq. (IV.38)] as a function of the detector efficiency η^2. The solid line corresponds to a single-detector cascade (no cascading: $N = 1$), the dashed lines correspond to $N = 4$, $N = 16$ and $N = \infty$ in ascending order. We consider a maximally entangled input state $|\Psi\rangle = (|0\rangle|\phi_0\rangle + |1\rangle|\phi_1\rangle + |2\rangle|\phi_2\rangle)/\sqrt{3}$ to serve as a benchmark.

modest boost in resolution, unless the detectors with single-photon sensitivity have a very high efficiency. Real single-photon resolution detectors are therefore superior to detector cascading with currently available detectors, notwithstanding the demanding operating conditions.

5 Summary

In this chapter I have studied the use of detection devices in entanglement based state preparation. In particular I considered optical devices such as single-photon sensitivity detectors, single-photon resolution detectors and detector cascades.

Detector cascading has generally been regarded as a good way to enhance single-photon resolution and consequently the fidelity of a state preparation process [158]. However, an extensive theory for the use of these detection devices has not been available so far. The statistics of N-ports have been considered in the context of tomography [129], which relies on the availability of a large number of copies of a quantum state. In state preparation, however, we perform measurements on single systems, and we therefore need precise bounds on the distinguishability of these measurements.

To this end, I introduced the confidence of preparation, which can also be used to quantify the performance of a detection device. Thus, I compared a single-photon resolution detector with a cascade of single-photon sensitivity detectors and found that cascading *does not give a practical advantage* over detectors with single-photon resolution.

V

Mathematical Description of Optical Circuits

In the previous chapters, I have discussed state preparation in quantum optics with realistic detectors. I now ask what the general outgoing state of an optical circuit is.

Suppose we have an optical circuit, that is, a collection of connected optical components. It is usually important to know what the outgoing state of this circuit is. In this chapter, I give a description of the outgoing state for a special class of optical circuits. First, in section 1, I define this class of optical circuits and show that they can be described by so-called multi-dimensional Hermite polynomials. In section 2, I give an example of this description. Section 3 discusses the Hermite polynomials, and finally, in section 4, I briefly consider the effect of imperfect detectors on the outgoing state. This chapter is based on Kok and Braunstein [106].

1 The Optical Circuit

What do we mean by an optical circuit? We can think of a black box with incoming and outgoing modes of the electro-magnetic field. The black box transforms a state of the incoming modes into a (different) state of the outgoing modes. The black box is what we call an *optical circuit.* We can now take a more detailed look inside the black box. We will consider three types of components.

First, the modes might be mixed by beam-splitters, or they may pick up a relative phase shift or polarisation rotation. These operations all belong to a class of optical components which preserve the photon number. We call them *passive* optical components.

Secondly, we may find optical components such as lasers, down-converters or (optical) parametric amplifiers in the black box. These components can be viewed as photon sources, since they do not leave the photon number invariant. We will call these components *active* optical components.

And finally, the box will generally include measurement devices, the outcomes of which may modify optical components on the remaining modes depending on the detection outcomes. This is called *feed-forward* detection. We can immediately simplify optical circuits using feed-forward detection, by considering the family of fixed circuits corresponding to the set of measurement outcomes (see

also Ref. [100]). In addition, we can postpone the measurement to the end, where all the optical components have 'acted' on the modes.

These three component types have their own characteristic mathematical description. A passive component yields a unitary evolution U_i, which can be written as

$$U_i = \exp\left(-i\kappa \sum_{jk} c_{jk}\hat{a}_j\hat{a}_k^\dagger - \text{H.c.}\right) , \tag{V.1}$$

where H.c. denotes the Hermitian conjugate. This unitary evolution commutes with the total number operator $\hat{n} = \sum_j \hat{a}_j^\dagger \hat{a}_j$.

Active components also correspond to unitary transformations, which can be written as $\exp(-itH_I^{(j)})$. Here $H_I^{(j)}$ is the interaction Hamiltonian associated with the j^{th} active component in a sequence. This Hamiltonian does not necessarily commute with the total number operator. To make a typographical distinction between passive and active components, we denote the i^{th} passive component by U_i, and the j^{th} active component by its evolution in terms of the interaction Hamiltonian. The mathematical description of the (ideal) measurement will correspond to taking the inner product of the outgoing state prior to the measurement with the eigenstate corresponding to the measurement.

a *The state prior to detection*

Now that we have the components of an optical circuit of N modes, we have to combine them into an actual circuit. Mathematically, this corresponds to applying the unitary evolutions of the successive components to the input state. Let $|\psi_{\text{in}}\rangle$ be the input state and $|\psi_{\text{prior}}\rangle$ the output state *prior* to the measurement. We then have (with $K > 0$ some integer)

$$|\psi_{\text{prior}}\rangle = U_K e^{-itH_I^{(K)}} \ldots U_1 e^{-itH_I^{(1)}} U_0 |\psi_{\text{in}}\rangle , \tag{V.2}$$

where it should be noted that U_i might be the identity operator $\mathbb{1}$ or a product of unitary transformations corresponding to passive components:

$$U_i = \prod_k U_{i,k} . \tag{V.3}$$

When the (multi-mode) eigenstate corresponding to the measurement outcome for a limited set of modes labelled $1, \ldots, M$ with $M < N$ is given by $|\gamma\rangle = |n_1, n_2, \ldots, n_M\rangle$ with M the number of detected modes out of a total of N modes, and n_i the number of photons found in mode i, the state leaving the optical circuit in the undetected modes is given by

$$|\psi_{\text{out}}\rangle_{M+1,\ldots,N} = {}_{1,\ldots,M}\langle\gamma|\psi_{\text{prior}}\rangle_{1,\ldots,N} . \tag{V.4}$$

In this chapter, I study the outgoing states $|\psi_{\text{out}}\rangle$ for a special class of optical circuits. First, I assume that the input state is the vacuum on all modes. Thus, I effectively study optical circuits as state preparation devices. Secondly, our class of optical circuits include all possible passive components, but *only* active components with quadratic interaction Hamiltonians:

$$H_I^{(j)} = \sum_{kl} \hat{a}_k^\dagger R_{kl}^{(j)} \hat{a}_l^\dagger + \sum_{kl} \hat{a}_k R_{kl}^{(j)*} \hat{a}_l \,, \tag{V.5}$$

where $R^{(j)}$ is some complex symmetric matrix. This matrix determines the behaviour of the j^{th} active component, which can be any combination of down-converters and squeezers. Finally, we consider ideal photo-detection, where the eigenstate corresponding to the measurement outcome can be written as $|\gamma\rangle = |n_1, \ldots, n_M\rangle$.

The class of optical circuits I consider here is not the most general class, but it still includes important experiments like quantum teleportation [23], entanglement swapping [126] and the demonstration of GHZ correlations [26]. In section 2, I show how teleportation can be modelled using the methods presented here.

The state $|\psi\rangle$ prior to the photo-detection can be written in terms of the components of the optical circuit as

$$|\psi\rangle = U_K e^{-it\mathcal{H}_I^{(K)}} \ldots U_1 e^{-itH_I^{(1)}} |0\rangle \,. \tag{V.6}$$

The creation and annihilation operators $\hat{a}_i^\dagger$ and $\hat{a}_i$ for mode i satisfy the standard canonical commutation relations

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij} \quad \text{and} \quad [\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0 \,, \tag{V.7}$$

with $i, j = 1 \ldots N$.

For any unitary evolution U, we have the relation

$$U e^R U^\dagger = \sum_{l=0}^{\infty} \frac{U R^l U^\dagger}{l!} = \sum_{l=0}^{\infty} \frac{(URU^\dagger)^l}{l!} = e^{URU^\dagger} \,. \tag{V.8}$$

Furthermore, if U is due to a collection of only passive components, such an evolution leaves the vacuum invariant: $U|0\rangle = |0\rangle$. Using these two properties it can be shown that Eq. (V.6) can be written as

$$|\psi\rangle = \exp\left[-\frac{1}{2} \sum_{i,j=1}^{N} \hat{a}_i^\dagger A_{ij}^\dagger \hat{a}_j^\dagger + \frac{1}{2} \sum_{i,j=1}^{N} \hat{a}_i A_{ij} \hat{a}_j + \frac{1}{2} \sum_{i,j=1}^{N} \hat{a}_i^\dagger B_{ij} \hat{a}_j \right] |0\rangle \,, \tag{V.9}$$

where A is some complex symmetric matrix and $B = B^\dagger$. I will now simplify this expression by normal-ordering this evolution.

Define $(\vec{a}, A\vec{a}) \equiv \sum_{ij} \hat{a}_i A_{ij} \hat{a}_j$. As shown by Braunstein [35], we can rewrite Eq. (V.9) using two passive unitary transformations U and V as:

$$|\psi\rangle = U e^{-\frac{1}{2}(\vec{a}^\dagger, \Lambda^\dagger \vec{a}^\dagger) + \frac{1}{2}(\vec{a}, \Lambda \vec{a})} V^T |0\rangle \,, \tag{V.10}$$

where Λ is a diagonal matrix with real non-negative eigenvalues λ_i. This means that, starting from vacuum, the class of optical circuits I consider here is equivalent to a set of single-mode squeezers, followed by a passive unitary transformation U and photo-detection. Since Λ is diagonal, we can write Eq. (V.10) as

$$|\psi\rangle = U \left(\prod_{i=1}^{N} \exp\left[-\frac{\lambda_i^*}{2} (\hat{a}_i^\dagger)^2 + \frac{\lambda_i}{2} \hat{a}_i^2 \right] \right) |0\rangle \,. \tag{V.11}$$

We can now determine the normal ordering of every factor $\exp[-\frac{\lambda_i^*}{2}(\hat{a}_i^\dagger)^2 + \frac{\lambda_i}{2}\hat{a}_i^2]$ separately. Note that the operators $\hat{a}_i^2$, $(\hat{a}_i^\dagger)^2$ and $2\hat{a}_i^\dagger \hat{a}_i + 1$ generate an su(1,1) algebra. According to Refs. [180, 60, 164], this may be normal-ordered as

$$e^{-\frac{\lambda_i^*}{2}(\hat{a}_i^\dagger)^2 + \frac{\lambda_i}{2}\hat{a}_i^2} = e^{-\hat{\lambda_i}^* \tanh|\frac{\lambda_i}{2}|(\hat{a}_i^\dagger)^2} e^{-2\ln(\cosh|\frac{\lambda_i}{2}|)\hat{a}_i^\dagger \hat{a}_i} e^{\hat{\lambda_i} \tanh|\frac{\lambda_i}{2}|\hat{a}_i^2} \,, \tag{V.12}$$

where $\hat{\lambda_i} = \lambda_i/|\lambda_i|$. In general, when $L_\pm$ and L_0 are generators of an $su(1,1)$ algebra (i.e., when A is unitary) we find [164]

$$e^{-\frac{1}{2}(\tau L_+ - \tau^* L_-)} = e^{-\hat{\tau} \tanh|\tau| L_+} e^{-2\ln(\cosh|\tau|) L_0} e^{\hat{\tau} \tanh|\tau| L_-} \,, \tag{V.13}$$

with τ a complex coupling constant and $\hat{\tau}$ its orientation in the complex plane. When we now apply this operator to the vacuum, the annihilation operators will vanish, leaving only the exponential function of the creation operators. We thus have

$$|\psi\rangle = U e^{-\frac{1}{2}(\vec{a}^\dagger, \Lambda^* \vec{a}^\dagger)} V^T |0\rangle = e^{-\frac{1}{2}(\vec{a}^\dagger, B\vec{a}^\dagger)} |0\rangle \,, \tag{V.14}$$

with $B \equiv U\Lambda^* U^\dagger$, again by virtue of the invariance property of the vacuum. This is the state of the interferometer prior to photo-detection. It corresponds to *multi-mode squeezed vacuum.*

b *Photo-detection and Bargmann representation*

The photo-detection itself can be modelled by successive application of annihilation operators. Every annihilation operator $\hat{a}_i$ removes a photon in mode i from the state $|\psi\rangle$. Suppose the optical circuit employs N distinct modes. We will now detect M modes, finding $n_1 + \ldots + n_M = N_{\text{tot}}$ photons (with $M < N$). These modes can be relabelled 1 to M. The vector $\vec{n}$ denotes the particular detector 'signature': $\vec{n} = (n_1, \ldots, n_M)$ means that n_1 photons are detected in mode 1, n_2

in mode 2, and so on. The freely propagating outgoing state $|\psi_{\vec{n}}\rangle$ can then be described as

$$|\psi_{\vec{n}}\rangle = {}_{1..M}\langle n_1, \ldots, n_M|\psi\rangle_{1..N} = c_{\vec{n}}\langle 0|\hat{a}_1^{n_1}\cdots\hat{a}_M^{n_M}|\psi\rangle \ . \tag{V.15}$$

Here, $c_{\vec{n}} = (n_1!\cdots n_M!)^{-\frac{1}{2}}$.

At this point it is convenient to introduce the N-mode Bargmann representation we encountered in chapter III [9]. The creation and annihilation operators obey the commutation relations given in Eq. (V.7). We can replace these operators with c-numbers and their derivatives according to

$$\hat{a}_i^\dagger \to \alpha_i \quad \text{and} \quad \hat{a}_i \to \partial_i \equiv \frac{\partial}{\partial\alpha_i} \ . \tag{V.16}$$

The commutation relations then read

$$[\partial_i, \alpha_j] = \delta_{ij} \quad \text{and} \quad [\partial_i, \partial_j] = [\alpha_i, \alpha_j] = 0 \ . \tag{V.17}$$

Note that the actual values of α_i are irrelevant (the creation and annihilation operators do not have numerical values either); what matters here is the functional relationship between α_i and ∂_{α_i}.

The state created by the optical circuit in this representation (prior to the detections, analogous to Eq. (V.14)) in the Bargmann representation is

$$\psi(\vec{\alpha}) = \exp\left[-\frac{1}{2}(\vec{\alpha}, B\vec{\alpha})\right] = \exp\left[-\frac{1}{2}\sum_{ij}\alpha_i B_{ij}\alpha_j\right] \ . \tag{V.18}$$

Returning to Eq. (V.15), we can write the freely propagating state after detection of the auxiliary modes in the Bargmann representation as

$$\psi_{\vec{n}}(\vec{\alpha}) \propto c_{\vec{n}}\ \partial_1^{n_1}\cdots\partial_M^{n_M}\ e^{-\frac{1}{2}(\vec{\alpha}, B\vec{\alpha})}\Big|_{\vec{\alpha}'=0} \ , \tag{V.19}$$

up to some normalisation factor, where $\vec{\alpha}' = (\alpha_1, \ldots, \alpha_M)$. By setting $\vec{\alpha}' = 0$ we ensure that *no more* than n_i photons are present in mode i. It plays the role of the vacuum bra in Eq. (V.15).

c *The outgoing state in terms of Hermite polynomials*

Now that we have an expression for the freely propagating state emerging from our optical setup after detection, we seek to simplify it. We can multiply $\psi_{\vec{n}}(\vec{\alpha})$ by the identity operator $\mathbb{1}$, written as

$$\mathbb{1} = (-1)^{2N_{\text{tot}}}\exp\left[-\frac{1}{2}(\vec{\alpha}, B\vec{\alpha})\right]\ \exp\left[\frac{1}{2}(\vec{\alpha}, B\vec{\alpha})\right] \ , \tag{V.20}$$

where N_{tot} is the total number of detected photons. We then find the following expression for the unnormalised freely propagating state created by our optical circuit:

$$\psi_{\vec{n}}(\vec{\alpha}) \propto c_{\vec{n}}(-1)^{N_{\text{tot}}} H^B_{\vec{n}}(\vec{\alpha})\, e^{-\frac{1}{2}(\vec{\alpha},B\vec{\alpha})}\Big|_{\vec{\alpha}'=0} . \tag{V.21}$$

Now I introduce the so-called *multi-dimensional* Hermite polynomial, or MDHP for short:

$$H^B_{\vec{n}}(\vec{\alpha}) = (-1)^{N_{\text{tot}}} e^{\frac{1}{2}(\vec{\alpha},B\vec{\alpha})} \frac{\partial^{n_1}}{\partial\alpha_1^{n_1}} \cdots \frac{\partial^{n_M}}{\partial\alpha_M^{n_M}}\, e^{-\frac{1}{2}(\vec{\alpha},B\vec{\alpha})} . \tag{V.22}$$

The use of multi-dimensional Hermite polynomials and Hermite polynomials of two variables have previously been used to describe N-dimensional first-order systems [52, 96] and photon statistics [169, 53, 103] (see also chapter IV). Here, I have shown that the lowest order of the *outgoing* state of optical circuits with quadratic components (as described by Eq. (V.9)) and conditional photo-detection can be expressed directly in terms of an MDHP.

In physical systems, the coupling constants (the λ_i's) are usually very small (i.e., $\lambda_i \ll 1$ or possibly $\lambda_i \lesssim 1$). This means that for all practical purposes only the first order term in Eq. (V.21) is important (i.e., for small λ_i's we can approximate the exponential by 1). Consequently, studying the multi-dimensional Hermite polynomials yields knowledge about the typical states we can produce using Gaussian sources without coherent displacements. In section 3 I take a closer look at these polynomials, but first I consider the description of quantum teleportation in this representation.

2 Example: Quantum Teleportation

As an example of how to determine the outgoing state of an optical circuit, consider the teleportation experiment by Bouwmeester *et al.* [23]. The optical circuit corresponding to this experiment consists of eight incoming modes, all in the vacuum state. Physically, there are four spatial modes a, b, c and d, all with two polarisation components x and y. Two down-converters create entangled polarisation states; they belong to the class of active Gaussian components without coherent displacements. Mode a undergoes a polarisation rotation over an angle θ and modes b and c are mixed in a 50:50 beam-splitter. Finally, modes b and c emerging from the beam-splitter are detected with polarisation insensitive detectors and mode a is detected using a polarisation sensitive detector. The state which is to be teleported is therefore given by

$$|\Psi\rangle = \cos\theta|x\rangle - \sin\theta|y\rangle . \tag{V.23}$$

The state prior to the detection and normal ordering (corresponding to Eq. (V.2)) is given by (τ is a coupling constant)

$$|\psi_{\text{prior}}\rangle = U_{\text{BS}}U_\theta e^{\tau(\vec{u}^\dagger, L\vec{u}^\dagger)/2+\tau^*(\vec{u},L\vec{u})/2+\tau(\vec{v}^\dagger,L\vec{v}^\dagger)/2+\tau^*(\vec{v},L\vec{v})/2}|0\rangle\ , \tag{V.24}$$

with

$$L = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{V.25}$$

and $\vec{u}^\dagger = (\hat{a}_x^\dagger, \hat{a}_y^\dagger, \hat{b}_x^\dagger, \hat{b}_y^\dagger)$, $\vec{v}^\dagger = (\hat{c}_x^\dagger, \hat{c}_y^\dagger, \hat{d}_x^\dagger, \hat{d}_y^\dagger)$. This can be written as

$$|\psi_{\text{prior}}\rangle = \exp\left[\frac{\tau}{2}(\vec{a}^\dagger, A\vec{a}^\dagger) + \frac{\tau^*}{2}(\vec{a}, A\vec{a})\right]|0\rangle\ , \tag{V.26}$$

with $\vec{a} \equiv (\hat{a}_x, \ldots, \hat{d}_y)$ and A the (symmetric) matrix

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & -\sin\theta & \cos\theta & -\sin\theta & \cos\theta & 0 & 0 \\ & 0 & \cos\theta & \sin\theta & \cos\theta & \sin\theta & 0 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & -1 \\ & & & 0 & 0 & 0 & -1 & 0 \\ & & & & 0 & 0 & 0 & 1 \\ & & & & & 0 & 1 & 0 \\ & & & & & & 0 & 0 \\ & & & & & & & 0 \end{pmatrix}\ . \tag{V.27}$$

We now have to find the normal ordering of Eq. (V.26). Since A is unitary, the polynomial $(\vec{a}^\dagger, A\vec{a}^\dagger)$ is a generator of an $su(1,1)$ algebra. According to Truax [164], the normal ordering of the exponential thus yields a state

$$|\psi_{\text{prior}}\rangle = \exp\left[\frac{\xi}{2}(\vec{a}^\dagger, A\vec{a}^\dagger)\right]|0\rangle\ , \tag{V.28}$$

with $\xi = (\tau \tanh|\tau|)/|\tau|$. The lowest order contribution after three detected photons is due to the term $\xi^2(\vec{a}^\dagger, A\vec{a}^\dagger)^2/8$. However, first I write Eq. (V.28) in the Bargmann representation:

$$\psi_{\text{prior}}(\vec{\alpha}) = \exp\left[\frac{\xi}{2}(\vec{\alpha}, A\vec{\alpha})\right]\ , \tag{V.29}$$

where $\vec{\alpha} = (\alpha_{a_x}, \ldots, \alpha_{d_y})$ and $\vec{\alpha}' = (\alpha_{a_x}, \ldots, \alpha_{c_y})$. The polarisation independent photo-detection (the Bell measurement) is then modelled by the differentiation

$(\partial_{b_x}\partial_{c_y} - \partial_{b_y}\partial_{c_x})$. Given a detector hit in mode a_x, the polarisation sensitive detection of mode a is modelled by ∂_{a_x}:

$$\psi_{\text{out}}(\vec{\alpha}) = \partial_{a_x}\left(\partial_{b_x}\partial_{c_y} - \partial_{b_y}\partial_{c_x}\right)\exp\left[\frac{\xi}{2}(\vec{\alpha}, A\vec{\alpha})\right]\Bigg|_{\vec{\alpha}'=0} . \tag{V.30}$$

The outgoing state in the Bargmann representation is thus given by

$$\psi_{\text{out}}(\vec{\alpha}) = \left(\cos\theta\,\alpha_{d_x} + \sin\theta\,\alpha_{d_y}\right) e^{\frac{\xi}{2}(\vec{\alpha}, A\vec{\alpha})} , \tag{V.31}$$

which is the state teleported from mode a to mode d in the Bargmann representation. This procedure essentially amounts to evaluating the multi-dimensional Hermite polynomial $H^A_{\vec{n}}(\vec{\alpha})$. Note that the polarisation independent Bell-detection of modes b and c yield a *superposition* of the MDHP's.

3 The Hermite Polynomials

The one-dimensional Hermite polynomials are of course well known from the description of the linear harmonic oscillator in quantum mechanics. These polynomials may be obtained from a generating function G (see appendix G). Furthermore, there exist two recursion relations and an orthogonality relation between them. The theory of multi-dimensional Hermite polynomials with real variables has been developed by Appell and Kempé de Fériet [3] and in the Bateman project [59]. Mizrahi derived an expression for real MDHP's from an n-dimensional generalisation of the Rodriguez formula [121]. I will now give the generating function for the complex MDHP's given by Eq. (V.22) and consecutively derive the recursion relations and the orthogonality relation (see also Ref. [96]).

a *Generating functions and recursion relations*

Define the generating function $G_B(\vec{\alpha}, \vec{\beta})$ to be

$$G_B(\vec{\alpha}, \vec{\beta}) = e^{(\vec{\alpha}, B\vec{\beta}) - \frac{1}{2}(\vec{\beta}, B\vec{\beta})} = \sum_{\vec{n}} \frac{\beta_1^{n_1}}{n_1!} \cdots \frac{\beta_M^{n_M}}{n_M!} H^B_{\vec{n}}(\vec{\alpha}) . \tag{V.32}$$

$G_B(\vec{\alpha}, \vec{\beta})$ gives rise to the MDHP in Eq. (V.22), which determines this particular choice. Note that the inner product $(\vec{\alpha}, B\vec{\beta})$ does not involve any complex conjugation. If complex conjugation was involved, we would have obtained different polynomials (which we could also have called multi-dimensional Hermite polynomials, but they would not bear the same relationship to optical circuits).

In the rest of the chapter I use the following notation: by $\vec{n} - e_j$ I mean that the j^{th} entry of the vector $\vec{n} = (n_1, \ldots, n_M)$ is lowered by one, thus becoming

$n_j - 1$. By differentiation of both sides of the generating function in Eq. (V.32) we can thus show that the first recursion relation becomes

$$\frac{\partial}{\partial \alpha_i} H^B_{\vec{n}}(\vec{\alpha}) = \sum_{j=1}^{M} B_{ij} n_j H^B_{\vec{n}-e_j}(\vec{\alpha}) \; . \tag{V.33}$$

The second recursion relation is given by

$$H^B_{\vec{n}+e_i}(\vec{\alpha}) - \sum_{j=1}^{M} B_{ij}\alpha_j H^B_{\vec{n}}(\vec{\alpha}) + \sum_{j=1}^{M} B_{ij} n_j H^B_{\vec{n}-e_j}(\vec{\alpha}) = 0 \; , \tag{V.34}$$

which can be proved by mathematical induction using

$$\sum_{k=1}^{M} B_{ik} n_k H^B_{\vec{n}-e_k+e_i}(\vec{\alpha}) - B_{ii} H^B_{\vec{n}}(\vec{\alpha}) = \sum_{k=1}^{M} B_{ik} m_k H^B_{\vec{m}+e_i}(\vec{\alpha}) \; . \tag{V.35}$$

Here, I have set $\vec{m} = \vec{n} - e_k$.

b Orthogonality relation

The orthogonality relation is somewhat more involved. Ultimately, we want to use this relation to determine the normalisation constant of the states given by Eq. (V.21). To find this normalisation we have to evaluate the integral

$$\int_{\mathbb{C}^N} d\vec{\alpha}\, \psi^*_{\vec{n}}(\vec{\alpha}) \psi_{\vec{m}}(\vec{\alpha}) \; .$$

The state $\psi_{\vec{n}}$ includes $|_{\vec{\alpha}'=0}$, which translates into a delta-function $\delta(\vec{\alpha}')$ in the integrand. The relevant integral thus becomes

$$\int_{\mathbb{C}^N} d\vec{\alpha}\, e^{-\mathrm{Re}(\vec{\alpha}, B\vec{\alpha})} \left[H^B_{\vec{n}}(\vec{\alpha})\right]^* H^B_{\vec{m}}(\vec{\alpha})\, \delta(\vec{\alpha}') \; .$$

From the orthonormality of different quantum states we know that this integral must be proportional to $\delta_{\vec{n},\vec{m}}$.

Since in the Bargmann representation we are only concerned with the *functional* relationship between α_i and ∂_{α_i} and not the actual values, we can choose α_i to be real. To stress this, we write $\alpha_i \to x_i$. The orthogonality relation is thus derived from

$$\int_{\mathbb{R}^N} d\vec{x}\, \psi^*_{\vec{n}}(\vec{x}) \psi_{\vec{m}}(\vec{x}) = \int_{\mathbb{R}^N} d\vec{x}\, e^{-(\vec{x}, \mathrm{Re}(B)\vec{x})} H^{B^*}_{\vec{n}}(\vec{x}) H^B_{\vec{m}}(\vec{x})\, \delta(\vec{x}') \; . \tag{V.36}$$

where $\delta(\vec{x}')$ is the real version of $\delta(\vec{\alpha}')$. Following Klauderer [96] we find that

$$\int d\vec{x}\, e^{-(\vec{x}, \mathrm{Re}(B)\vec{x})} H^{B^*}_{\vec{n}}(\vec{x}) H^B_{\vec{m}}(\vec{x}) = (-1)^{N_{\mathrm{tot}}} \int d\vec{x}\, e^{-\frac{1}{2}(\vec{x}, B\vec{x})} \partial^{\vec{n}}_{\vec{x}} \left[e^{-\frac{1}{2}(\vec{x}, B^*\vec{x})}\right] H^B_{\vec{m}}(\vec{x}) \; , \tag{V.37}$$

where $\partial_{\vec{x}}^{\vec{n}}$ is the differential operator $\partial_{x_1}^{n_1} \cdots \partial_{x_M}^{n_M}$ acting solely on the exponential function. We now integrate the right-hand side by parts, yielding

$$(-1)^{N_{\text{tot}}} \int d\vec{x}\, e^{-\frac{1}{2}(\vec{x},B\vec{x})} \partial_{\vec{x}}^{\vec{n}}\, e^{-\frac{1}{2}(\vec{x},B^*\vec{x})} H_{\vec{m}}^{B}(\vec{x}) =$$
$$(-1)^{N_{\text{tot}}} \int d'\vec{x}\, e^{-\frac{1}{2}(\vec{x},B\vec{x})} \partial_{\vec{x}}^{\vec{n}-e_i}\, e^{-\frac{1}{2}(\vec{x},B^*\vec{x})} H_{\vec{m}}^{B}(\vec{x}) \Big|_{x_i=-\infty}^{+\infty}$$
$$- (-1)^{N_{\text{tot}}} \int d\vec{x}\, e^{-\frac{1}{2}(\vec{x},B\vec{x})} \partial_{\vec{x}}^{\vec{n}-e_i}\, e^{-\frac{1}{2}(\vec{x},B^*\vec{x})} \partial_{x_i} H_{\vec{m}}^{B}(\vec{x}) \,, \quad \text{(V.38)}$$

with $d'\vec{x} = dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_N$. The left-hand term is equal to zero when $\mathrm{Re}(B)$ is positive definite, i.e., when $(\vec{x}, \mathrm{Re}(B)\vec{x}) > 0$ for all non-zero $\vec{x}$. Repeating this procedure n_i times yields

$$\int d\vec{x}\, e^{-(\vec{x},\mathrm{Re}(B)\vec{x})} H_{\vec{n}}^{B^*}(\vec{x}) H_{\vec{m}}^{B}(\vec{x}) =$$
$$(-1)^{N_{\text{tot}}+n_i} \int d\vec{x}\, e^{-\frac{1}{2}(\vec{x},B\vec{x})} \partial_{\vec{x}}^{\vec{n}-n_i e_i}\, e^{-\frac{1}{2}(\vec{x},B^*\vec{x})} \partial_{x_i}^{n_i} H_{\vec{m}}^{B}(\vec{x}) \,. \quad \text{(V.39)}$$

When there is at least one $n_i > m_i$, differentiating the MDHP n_i times to x_i will yield zero. Thus we have

$$\int d\vec{x}\, e^{-(\vec{x},\mathrm{Re}(B)\vec{x})} H_{\vec{n}}^{B^*}(\vec{x}) H_{\vec{m}}^{B}(\vec{x}) = 0 \quad \text{for } \vec{n} \neq \vec{m} \quad \text{(V.40)}$$

when $\mathrm{Re}(B)$ is positive definite and $n_i \neq m_i$ for any i. The case where $\vec{n}$ equals $\vec{m}$ is given by

$$\int d\vec{x}\, e^{-\frac{1}{2}(\vec{x},\mathrm{Re}(B)\vec{x})} H_{\vec{n}}^{B^*}(\vec{x}) H_{\vec{m}}^{B}(\vec{x}) = \delta_{\vec{n}\vec{m}} \mathcal{N} \,, \quad \text{(V.41)}$$

where $\delta_{\vec{n}\vec{m}}$ denotes the product of $\delta_{n_i m_i}$ with $1 \leq i \leq N$. Here, $\mathcal{N}$ is equal to

$$\mathcal{N} \equiv 2^{N_{\text{tot}}} B_{11}^{n_1} \cdots B_{NN}^{n_N} n_1! \cdots n_N! \left|\pi^{-1} B\right|^{-\frac{1}{2}} \,. \quad \text{(V.42)}$$

For the proof of this identity I refer to Ref. [96].

4 Imperfect Detectors

So far, I only considered the use of ideal photo-detection. That is, I assumed that the detectors tell us exactly and with unit efficiency how many photons were present in the detected mode. However, in reality such detectors do not exist. In particular we have to incorporate losses (non-perfect efficiency) and dark counts (see chapter IV). Furthermore, we have to take into account the

fact that most detectors do not have a single-photon resolution (i.e., they cannot distinguish a single photon from two photons) [103].

This model is not suitable when we want to include dark counts. These unwanted light sources provide thermal light, which is not of the form of Eq. (V.9) but given by Eq. (IV.25). In single-shot experiments, however, dark counts can be neglected when the detectors operate only within a narrow time interval.

We can model the efficiency of a detector by placing a beam-splitter with transmission amplitude η in front of a perfect detector [103]. The part of the signal which is reflected by the beam-splitter (and which will therefore never reach the detector) is the loss due to the imperfect detector. Since beam-splitters are part of the set of optical devices we allow, we can make this generalisation without any problem. We now trace out all the reflected modes (they are truly 'lost'), and end up with a mixture in the remaining undetected modes.

Next, we can model the lack of single-photon resolution by using the relative probabilities $p(n|k)$ and $p(m|k)$ of the actual number n or m of detected photons conditioned on the indication of k photons in the detector (as described in Ref. [103] and chapter IV). We can determine the pure states according to n and m detected photons, and add them with relative weights $p(n|k)$ and $p(m|k)$. This method is trivially generalised for more than two possible detected photon numbers.

Finally, we should note that my description of this class of optical circuits (in terms of multi-dimensional Hermite polynomials) is essentially a one-way function. Given a certain setup, it is relatively straightforward to determine the outgoing state of the circuit. The other way around, however, is very difficult. As exemplified by our efforts in Ref. [100] and chapter III, it is almost impossible to obtain the matrix B associated with an optical circuit which produces a particular predetermined state from a Gaussian source.

5 Summary

In this chapter, I have derived the general form of squeezed multi-mode vacuum states conditioned on photo-detection of some of the modes. To lowest order, the outgoing states in the Bargmann representation are proportional to multi-dimensional Hermite polynomials. As an example, I showed how teleportation can be described this way.

Some Applications

VI

Teleportation and Entanglement Swapping

In this chapter I study the experimental realisations of quantum teleportation by Bouwmeester *et al.* [23], entanglement swapping by Pan *et al.* [126] and the observation of three-photon GHZ-entanglement by Bouwmeester *et al.* [26]. I will show that these experiments heavily relied on *post-selection.*

In section 1 I briefly discuss the issues concerned with post-selection. Then, in section 2 the quantum teleportation experiment performed in Innsbruck will be studied. This section is based on Ref. [99]. Section 3 is based on Refs. [101, 102], and discusses entanglement swapping and entanglement purification. Finally, I briefly consider the experimental observation of three-photon GHZ-entanglement in section 4. This chapter is based on Kok and Braunstein [99, 101, 102].

1 Post-selection in quantum optics

In this section, I will discuss the concept of *post-selection.* Suppose we measure an observable A with respect to an ensemble of systems in a state $|\phi\rangle$. In general we have a set of different measurement outcomes $\{a_k\}$, where the a_k's denote the eigenvalues of A. We speak of post-selection when a *subset* of the set of outcomes $\{a_k\}$ is discarded. The remaining *post-selected* set of outcomes may be used for subsequent data-analysis.

For example, if we had a tri-partite optical system in the state

$$|\phi\rangle \propto |0,1,1\rangle + |1,0,1\rangle + |1,1,0\rangle + |1,1,1\rangle \,, \tag{VI.1}$$

we could place three photo-detectors[1] in the three outgoing modes. For ideal detectors, there are four possible measurement outcomes: one photon in any two of the three detectors, or one photon in all three detectors. When we repeat this 'experiment' a large number of times, we might discard all the measurements which do not yield a three-fold detector coincidence. This would correspond to post-selecting our data set on a three-fold coincidence. (Admittedly, this is not a very interesting experiment. However, later we will see that using post-selection we can even partially perform a Bell measurement.)

Note that there is a fundamental difference between a *conditional measurement* and post-selection. In chapter IV I discussed entanglement-based state

[1] Here we use three measurements of the same observable, i.e., photon number. There is, however, no reason why we can't measure three different observables in the respective modes.

preparation, in which one subsystem was measured, the outcome of which was used to accept or reject the state of the remaining system. The crucial property of such a conditional measurement is that at the end of the procedure, there is a physically propagating state remaining. Post-selection offers a completely different type of control to the experimenter. Since all the subsystems are measured there is no physically propagating state left over, but a subset of the data can be selected for further analysis.

The question is now whether in our 'experiment' above we have demonstrated the *existence* of the state $|1,1,1\rangle$. The answer has to be 'no': immediately before the measurement the state had the form of Eq. (VI.1), whereas afterwards, there was no state left at all.

Post-selection can be very powerful, though. In section 4 we will see that non-local correlations can be inferred from post-selected data which was obtained in an experiment designed to create a three-photon GHZ-state. In addition, it does *not necessarily* mean that a post-selected state cannot be used further in, say, a quantum computer. As long as the post-selection can be made in the end, the relevant branch (or branches) in the superposition undergo(es) the quantum computation. With this in mind, we can now consider the experimental demonstration of quantum teleportation and entanglement swapping.

2 Quantum teleportation

We speak of quantum teleportation when a (possibly unknown) quantum state $|\phi\rangle$ held by Alice is sent to Bob without actually traversing the intermediate space. The protocol uses an entangled state of two systems which is shared between Alice and Bob. In the next section I will present the teleportation protocol for discrete variables. In the subsequent sections I confine the discussion to the teleportation experiment performed by Bouwmeester *et al.*, and study the difficulties which arise using its particular experimental setup.

a The discrete teleportation protocol

Quantum teleportation was first introduced by Bennett et al. in 1993 [14]. In this protocol, a quantum state held by Alice is sent to Bob by means of what the authors called 'dual classical and Einstein-Podolsky-Rosen channels'. How does this work?

Suppose we have a set of three two-level systems, or qubits, the states of which can be written in the computational basis $\{|0\rangle_k, |1\rangle_k\}$, where $k = 1,2,3$ denotes the system. Let Alice and Bob share a maximally entangled state (the Einstein-Podolsky-Rosen channel [56]), for instance one of the Bell states $|\Psi^-\rangle$ in systems 2 and 3:

$$|\Psi^-\rangle_{23} = \frac{1}{\sqrt{2}} \left(|0,1\rangle_{23} - |1,0\rangle_{23}\right) . \tag{VI.2}$$

System 1 is in an unknown state $|\phi\rangle_1$, which can be written as

$$|\phi\rangle_1 = \alpha|0\rangle + \beta|1\rangle \; . \tag{VI.3}$$

The other three Bell states are given by

$$\begin{aligned} |\Psi^+\rangle &= \frac{1}{\sqrt{2}}\left(|0,1\rangle + |1,0\rangle\right) \\ |\Phi^-\rangle &= \frac{1}{\sqrt{2}}\left(|0,0\rangle - |1,1\rangle\right) \\ |\Phi^+\rangle &= \frac{1}{\sqrt{2}}\left(|0,0\rangle + |1,1\rangle\right) \; . \end{aligned} \tag{VI.4}$$

We can write the total state of the three systems as

$$\begin{aligned} |\phi\rangle_1|\Psi^-\rangle_{23} = \frac{1}{\sqrt{2}}\left(\alpha|0,0,1\rangle_{123} - \alpha|0,1,0\rangle_{123}\right. \\ \left. + \beta|1,0,1\rangle_{123} - \beta|1,1,0\rangle_{123}\right) \; . \end{aligned} \tag{VI.5}$$

The computational basis states of two qubits can also be written in the Bell basis:

$$\begin{aligned} |0,0\rangle &= \frac{1}{\sqrt{2}}\left(|\Phi^+\rangle + |\Phi^-\rangle\right) \; , \\ |0,1\rangle &= \frac{1}{\sqrt{2}}\left(|\Psi^+\rangle + |\Psi^-\rangle\right) \; , \\ |1,0\rangle &= \frac{1}{\sqrt{2}}\left(|\Psi^+\rangle - |\Psi^-\rangle\right) \; , \\ |1,1\rangle &= \frac{1}{\sqrt{2}}\left(|\Phi^+\rangle - |\Phi^-\rangle\right) \; . \end{aligned} \tag{VI.6}$$

When we make this substitution for qubits 1 and 2, Eq. (VI.5) becomes

$$\begin{aligned} |\phi\rangle_1|\Psi^-\rangle_{23} = \frac{1}{2}\left[|\Phi^+\rangle_{12}\left(\alpha|1\rangle_3 - \beta|0\rangle_3\right) + |\Phi^-\rangle_{12}\left(\alpha|1\rangle_3 + \beta|0\rangle_3\right)\right. \\ \left. -|\Psi^+\rangle_{12}\left(\alpha|0\rangle_3 - \beta|1\rangle_3\right) - |\Psi^-\rangle_{12}\left(\alpha + \beta|1\rangle_3\right)\right] \; . \end{aligned} \tag{VI.7}$$

Alice is in possession of qubits 1 and 2, while Bob holds qubit 3. When Alice now performs a Bell measurement, Bob's qubit is transformed into the unknown state *up to one of four unitary transformations.* Alice's measurement outcome determines which one of these transformations should be inverted on Bob's qubit to return it to the original state $|\phi\rangle$. This completes the teleportation protocol.

Quantum teleportation is not restricted to qubits. For example, suppose we have an N-level system[2] in the state

$$|\phi\rangle_1 = \sum_j \alpha_j |j\rangle_1 \; , \tag{VI.8}$$

[2]Sometimes called a 'qu$N$it' or 'qudit'.

and a maximally entangled state shared between Alice and Bob:

$$|\Psi\rangle_{23} = \frac{1}{\sqrt{N}} \sum_j |j,j\rangle_{23} \,. \tag{VI.9}$$

We measure system 1 and 2, held by Alice, in the basis $\{|\psi_{nm}\rangle_{12}\}$, with

$$|\psi_{nm}\rangle_{12} = \frac{1}{\sqrt{N}} \sum_k e^{2\pi ijn/N} |j, j \oplus m\rangle_{12} \,. \tag{VI.10}$$

In this notation we have $j \oplus m = j + m \mod N$. Conditioned on the measurement outcome (n, m) corresponding to $|\psi_{nm}\rangle_{12}$, Bob's system 3 is transformed into $|\phi\rangle_3$ after a transformation [14]:

$$U_{nm} = \sum_k e^{2\pi ikn/N} |k\rangle_3 \langle k \oplus m| \,. \tag{VI.11}$$

Thus the state of system 1 is transferred to system 3. System 1 can itself be mixed or part of an entangled state. Note that, since Bob needs the measurement outcome, Alice has to send a classical message of $2 \log_2 N$ classical bits. Sending this classical message is, like all classical communication, bounded by the speed of light. Therefore, quantum teleportation does not yield an information transfer faster than light.

After the invention of discrete quantum teleportation, Vaidman and Braunstein and Kimble introduced teleportation for states of dynamical variables with continuous spectra [165, 33]. In 1997, teleportation was experimentally realised by Bouwmeester *et al.* in Innsbruck [23] and Boschi *et al.* in Rome [20], followed by Furusawa *et al.* in Pasadena [67] in 1998. This last experiment involved the teleportation of continuous variables. Quantum teleportation was also reported using nuclear magnetic resonance by Nielsen *et al.* in 1998 [124]. In 2000, Kim *et al.* performed quantum teleportation of polarised single-photon states using complete Bell detection [95].

In this chapter, however, I will focus mainly on the teleportation experiment of Bouwmeester *et al.*

b *The 'Innsbruck Experiment'*

In this section, I study the experimental realisation of quantum teleportation of a single polarised photon as performed in Innsbruck, henceforth called the 'Innsbruck experiment' (Bouwmeester *et al.* [23]). In the Innsbruck experiment, parametric down-conversion is used to create two entangled photon-pairs. One pair constitutes the entangled state shared between Alice and Bob, while the other is used by Victor to create an 'unknown' single-photon polarisation state $|\phi\rangle$: Victor detects mode a, shown in figure VI.1 to prepare the single-photon

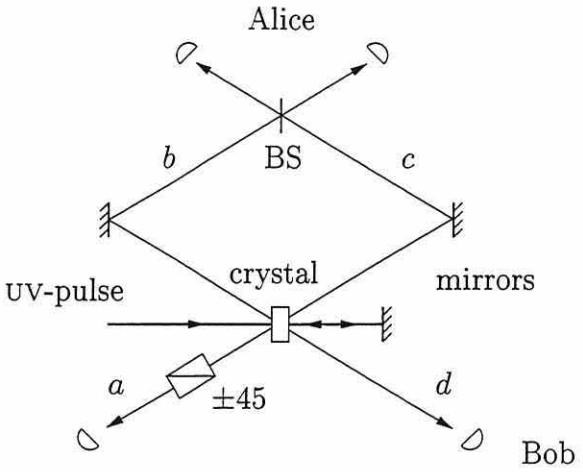

Figure VI.1: Schematic representation of the teleportation experiment conducted in Innsbruck. A UV-pulse is sent into a non-linear crystal, thus creating an entangled photon-pair. The UV-pulse is reflected by a mirror and returned into the crystal again. This reflected pulse creates the second photon-pair. Photons b and c are sent into a beam-splitter and are detected. This is the Bell measurement. Photon a is detected to prepare the input state and photon d is the teleported output state Bob receives. In order to rule out the possibility that there are no photons in mode d, Bob detects this mode as well.

input state in mode b. This mode is sent to Alice. A coincidence in the detection of the two outgoing modes of the beam-splitter (Alice's — incomplete — Bell measurement) tells us that Alice's two photons are in a $|\Psi^-\rangle$ Bell state [172, 30, 31]. The remaining photon (held by Bob) is now in the same unknown state as the photon prepared by Victor because in this case the unitary transformation Bob has to apply coincides with the identity, i.e., doing nothing. Bob verifies this by detecting his state along the same polarisation axis which was used by Victor. A four-fold coincidence in the detectors of Victor's state preparation, Alice's Bell measurement and Bob's outgoing state indicate that quantum teleportation of a single-photon state is complete.

There is however a complication which gave rise to a different interpretation of the experiment [32, 24, 25, 99]. Analysis shows that the state before detection by Bob (but conditioned on the other three detector 'hits') is a mixture of the vacuum and the original state [23, 32] (to lowest order). This vacuum contribution occurs when the down-converter responsible for creating the input state $|\phi\rangle$ yields two photon-pairs, while the other gives nothing. The detectors used in the experiment cannot distinguish between one or several photons coming in, so Victor's detection of mode a in figure VI.1 will not reveal the presence of more than one photon. A

three-fold coincidence in the detectors of Victor and Alice alone is still possible, but Bob has not received a photon and quantum teleportation has not been achieved. Bob therefore needs to detect his state in order to identify successful quantum teleportation. Were Victor to use a detector which can distinguish between one or several photons this problem would disappear. However, currently such detectors require an operating environment of roughly 6K [108, 110, 93, 162].

I evaluate the suggestions to 'improve' the experiment in order to yield non-post-selected operation, as made by Braunstein and Kimble [32] (I will discuss the reply by Bouwmeester *et al.* [24, 25] in section *g*). These suggestions include the employment of a detector cascade (as proposed in chapter IV) in the state preparation mode, and enhancement of the down-converter responsible for the entanglement channel (see chapter III) relative to the one responsible for the initial state preparation. Subsequently, I hope to clarify some of the differences in the interpretation of the Innsbruck experiment [99].

As pointed out by Braunstein and Kimble [32], to lowest order the teleported state in the Innsbruck experiment is a mixture of the vacuum and a single-photon state. However, we cannot interpret this *state* as a low-efficiency teleported state, where sometimes a photon emerges from the apparatus and sometimes not. This reasoning is based on the so-called 'Partition Ensemble Fallacy', or PEF[3] for short. It will be studied more extensively in section *g*. PEF relies on a particular partition of the outgoing density matrix, and this is not consistent with quantum mechanics [131]. Circumventing PEF leads to the notion of *post-selected* teleportation, in which the teleported state is detected. The post-selected teleportation indeed has a high fidelity and a low efficiency. Although generally PEF is harmless (it might even be considered a useful tool in understanding aspects of quantum theory), to my knowledge, this is the first instance where it leads to a *quantitatively* different evaluation of an experiment.

It will turn out that the suggested improvements require near perfect efficiency photo-detectors or a considerable increase in the time needed to run the experiment. The remaining practical alternative in order to obtain *non-post-selected* quantum teleportation (i.e., teleportation *without* the need for detecting the teleported photon) is to employ a single-photon resolution detector in the state-preparation mode (a technology currently requiring approximately 6K operating conditions) [93, 162] (see also chapter IV).

c The generalised experiment

In the rest of this section I consider a generalised scheme for the Innsbruck experiment which enables us to establish the requirements to obtain non-post-selected quantum teleportation (based on a three-fold coincidence of Victor and Alice's detectors). The generalisation consists of a detector cascade (Chapter IV and Ref.

[3]This term was coined by Samuel L. Braunstein and first appeared in Kok and Braunstein [99].

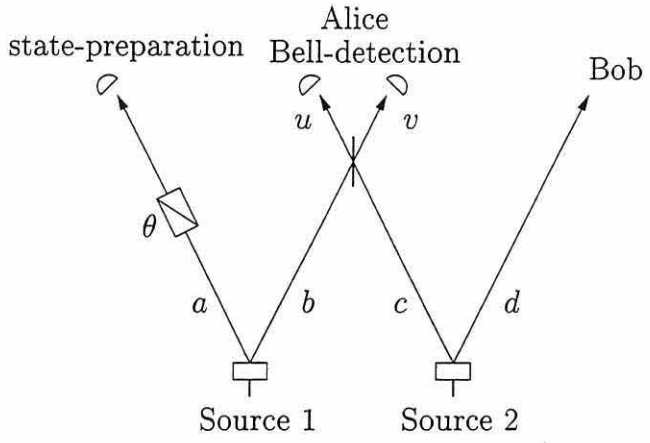

Figure VI.2: Schematic 'unfolded' representation of the teleportation experiment with two independent down-converters (Source 1 and Source 2) and a polarisation rotation θ in mode a. The state-preparation detector is actually a detector cascade and Bob does not detect the mode he receives.

[158] for Victor's state preparation detection and parametric down-converters with different specifications, rather than two identical down-converters. Furthermore, an arbitrary polarisation rotation in the state-preparation mode allows us to consider any superposition of x- and y-polarisation. I calculate the output state and give an expression for the teleportation fidelity in terms of the detector efficiencies and down-conversion rates. To this end, I consider a simplified 'unfolded' schematic representation of the experiment, shown in figure VI.2.

d Detectors

As explained in chapter IV, there are two sources of errors for a detector: losses and dark counts. Dark counts are negligible in the teleportation experiment because the UV-pump is fired during very short time intervals and the probability of finding a dark count in such a small interval is negligible. Consequently, the model for real, finite-efficiency detectors I presented in chapter IV only takes into account detector losses. Furthermore, the detectors cannot distinguish between one or several photons. In my terminology: *finite-efficiency single-photon sensitivity* detectors (see page 61).

To simulate a realistic detector I make use of projection operator valued measures, or POVM's for short [107] (see also appendix B). Consider a beam-splitter in the mode which is to be detected so that part of the signal is reflected (see figure VI.3). The second incoming mode of the beam-splitter is the vacuum (I neglect higher photon number states because they hardly contribute at room temperature). The transmitted signal c is sent into an ideal detector. We identify mode d with the detector loss.

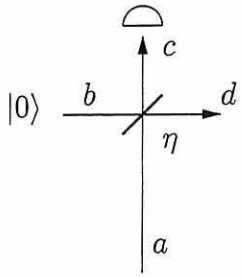

Figure VI.3: A model of an inefficient detector. The beam-splitter with transmission amplitude η will reflect part of the incoming mode a to mode d, which is thrown away. The transmitted part c will be sent into a ideal detector. Mode b is vacuum.

Suppose in mode a there are n x-polarised and m y-polarised photons. Furthermore, let these photons all be reflected by the beam-splitter. The projector for finding these photons in the d-mode is given by:

$$E_d = |n,m\rangle_{d_x d_y}\langle n,m| = \frac{1}{n!m!}(\hat{d}_x^\dagger)^n(\hat{d}_y^\dagger)^m|0,0\rangle_{d_x d_y}\langle 0,0|\hat{d}_x^n\hat{d}_y^m \,. \tag{VI.12}$$

The beam-splitter equations are taken to be ($\widetilde{\eta} \equiv \sqrt{1-\eta^2}$):

$$\hat{c} = \eta\hat{a} + \widetilde{\eta}\hat{b} \qquad \text{and} \qquad \hat{d} = \widetilde{\eta}\hat{a} - \eta\hat{b} \,. \tag{VI.13}$$

Substituting these equations in (VI.12), summing over all n and m and using the binomial expansion yields

$$\begin{aligned} E_{ab} \;=\; & \sum_{n,m}\binom{n}{k}^2\binom{m}{l}^2\frac{(-1)^{2(k+l)}}{n!m!}(\widetilde{\eta}\hat{a}_x^\dagger)^{n-k}(\eta\hat{b}_x^\dagger)^k(\widetilde{\eta}\hat{a}_y^\dagger)^{m-l}(\eta\hat{b}_y^\dagger)^l\;|0\rangle_{ab} \\ & \times\langle 0|(\widetilde{\eta}\hat{a}_x)^{n-k}(\eta\hat{b}_x)^k(\widetilde{\eta}\hat{a}_y)^{m-l}(\eta\hat{b}_y)^l \,. \end{aligned} \tag{VI.14}$$

Since the b-mode is the vacuum, the only contributing term is $k = l = 0$. So the POVM $E_a^{(0)}$ of finding *no* detector counts in mode a is

$$E_a^{(0)} = \sum_{n,m}\frac{\widetilde{\eta}^n(\hat{a}_x^\dagger)^n\widetilde{\eta}^m(\hat{a}_y^\dagger)^m}{n!m!}\;|0\rangle_{a_x a_y}\langle 0|\;\widetilde{\eta}^n\hat{a}_x^n\widetilde{\eta}^m\hat{a}_y^m = \sum_{n,m}\widetilde{\eta}^{2(n+m)}|n,m\rangle_{a_x a_y}\langle n,m| \,. \tag{VI.15}$$

The required POVM for finding a detector count is

$$E_a^{(1)} = \mathbb{1} - E_a^{(0)} = \sum_{n,m}[1-\widetilde{\eta}^{2(n+m)}]|n,m\rangle_{a_x a_y}\langle n,m| \,, \tag{VI.16}$$

where $\mathbb{1}$ is the identity operator, η^2 is the detector efficiency and $\widetilde{\eta}^2 \equiv 1-\eta^2$ the detector loss. When we let $E_a^{(1)}$ act on the total state and trace out mode a, we

have modelled the inefficient detection of this mode. In the case of continuous detection we need a more elaborate model (see for example Ref. [175]).

In order for Victor to distinguish between one or more photons in the state preparation mode a, I consider a detector cascade (Victor doesn't have a detector which can distinguish between one or several photons coming in). When there is a detector coincidence in the cascade, more than one photon was present in mode a, and the event should be dismissed. In the case of ideal detectors, this will improve the fidelity of the teleportation up to an arbitrary level (we assume there are no beam-splitter losses). Since we employ the cascade in the a-mode (which was used by Victor to project mode b onto a superposition in the polarisation basis) we need to perform a *polarisation sensitive* detection.

In order to model this I separate the incoming state $|n,m\rangle_{a_x a_y}$ of mode a into two spatially separated modes $|n\rangle_{a_x}$ and $|m\rangle_{a_y}$ by means of a polarisation beam-splitter. The modes a_x and a_y will now be detected. The POVM's corresponding to inefficient detectors are derived along the same lines as in the previous section and read:

$$\begin{aligned} E^{(0)}_{a_j} &= \sum_n \widetilde{\eta}^{2n} |n\rangle_{a_j}\langle n| \qquad \text{and} \\ E^{(1)}_{a_j} &= \sum_n [1-\widetilde{\eta}^{2n}] |n\rangle_{a_j}\langle n| \,. \end{aligned} \tag{VI.17}$$

with $j \in \{x,y\}$. we choose to detect the x-polarised mode. This means that we only have to make sure that there are no photons in the y-mode. The output state will include a product of the two POVM's: one for finding a photon in mode a_x, and one for finding *no* photons in mode a_y: $E^{(1)}_{a_x} E^{(0)}_{a_y}$.

To make a cascade with two detectors in a_x and one in a_y employ another 50:50 beam-splitter in mode a_x and repeat the above procedure of detecting the outgoing modes c and d (VI.17). Since we can detect a photon in either one of the modes, we have to include the sum of the corresponding POVM's, yielding a transformation $E^{(1)}_{c_x} E^{(0)}_{d_x} + E^{(0)}_{c_x} E^{(1)}_{d_x}$. This is easily expandable to larger cascades by using more beam-splitters and summing over all possible detector hits.

e *Output state*

In this section I incorporate the finite-efficiency detectors and the detector cascade in the calculation of the undetected teleported output state. This calculation includes the creation of two photon-pairs (lowest order) and three photon pairs (higher order corrections due to four or more photon-pairs in the experiment are highly negligible). A formula for the vacuum contribution to the teleportation fidelity is given for double-pair production (lowest order).

Let the two down-converters in the generalised experimental setup yield evolutions U_{src1} and U_{src2} on modes a, b and c, d respectively (see figures VI.1 and VI.2) according to Eq. (III.22). The beam-splitter which transforms modes b and

c into u and v (see figure VI.2) is incorporated by a suitable unitary transformation U_{BS}, as is the polarisation rotation U_θ over an angle θ in mode a. The N-cascade will be modelled by $N-1$ beam-splitters in the x-polarisation branch of the cascade, and can therefore be expressed in terms of a unitary transformation $U_{a_1 \ldots a_N}$ on the Hilbert space corresponding to modes a_1 to a_N (i.e., replace mode a with modes a_1 to a_N):

$$|\Psi_\theta\rangle\langle\Psi_\theta| = U_{a_1 \ldots a_N} U_\theta U_{\mathrm{BS}} U_{\mathrm{scr1}} U_{\mathrm{scr2}} |0\rangle\langle 0| U_{\mathrm{scr1}}^\dagger U_{\mathrm{scr2}}^\dagger U_{\mathrm{BS}}^\dagger U_\theta^\dagger U_{a_1 \ldots a_N}^\dagger \ . \qquad \text{(VI.18)}$$

Detecting modes $a_1 \ldots a_N$, u and v with real (inefficient) detectors means taking the partial trace over the detected modes, including the POVM's derived in section d:

$$\rho_{\mathrm{out}} = \mathrm{Tr}_{a_1 \ldots a_N uv} \left[E_{N\text{-cas}} E_u^{(1)} E_v^{(1)} \ |\widetilde{\Psi}_\theta\rangle_{a_1 \ldots a_N uvd} \langle \widetilde{\Psi}_\theta| \right] \ , \qquad \text{(VI.19)}$$

with $E_{N\text{-cas}}$ the superposition of POVM's for a polarisation sensitive detector cascade having n detectors with finite efficiency. In the case $N = 2$ this expression reduces to the 2-cascade POVM-superposition derived in the previous section. Eq. (VI.19) is an analytic expression of the undetected outgoing state in the generalisation of the Innsbruck experiment.

The evolutions U_{src1} and U_{src2} are exponentials of creation operators. In the computer simulation (using MATHEMATICA, see appendix H) I truncated these exponentials at first and second order. The terms that remain correspond to double and triple pair production in the experimental setup. To preserve the order of the creation operators we put them as arguments in a function f. I defined the following algebraic rules for f (see appendix H):

```
f[x__, y__ + w__, z__]     :=   f[x, y, z] + f[x, w, z]
f[x__, n_ a__, y__]        := n f[x, a, y]
f[x__, n_ adagger__, y__] := n f[x, adagger, y]
```

where `x,y,z` and `w` are arbitrary expressions including creation and annihilation operators (`adagger` and `a`) and `n` some expression *not* depending on creation or annihilation operators. The last entry of f is always a photon number state (including the initial vacuum state).

Since we now have functions of creation and annihilation operators, it is quite straightforward to define (lists of) substitution rules for a beam-splitter (see also Eq. (VI.13)), polarisation rotation, POVM's and the trace operation. I then use these substitution rules to 'build' a model of the generalised experimental setup.

f Results

The probability of creating one entangled photon-pair using the weak parametric down-conversion source 1 or 2 is p_1 or p_2 respectively (see figure VI.2). I calculated

the output state both for an N-cascade up to order p^2 (i.e. p_1^2 or p_1p_2) and for a 1-cascade up to the order p^3 (p_1^3, $p_1^2p_2$ or $p_1p_2^2$). The results are given below. For brevity, we take:

$$\begin{aligned} |\Psi_\theta\rangle &= \cos\theta|0,1\rangle + e^{i\varphi}\sin\theta|1,0\rangle \qquad \text{and} \\ |\Psi_\theta^\perp\rangle &= e^{i\varphi}\sin\theta|0,1\rangle - \cos\theta|1,0\rangle \end{aligned} \tag{VI.20}$$

as the ideally prepared state and the state orthogonal to it. Suppose η_u^2 and η_v^2 are the efficiencies of the detectors in mode u and v respectively, and η_c^2 the efficiency of the detectors in the cascade (for simplicity I assume that the detectors in the cascade have the same efficiency). Define $g_{uvc} = \eta_u^2\eta_v^2\eta_c^2$. The detectors in modes u and v are polarisation insensitive, whereas the cascade consists of polarisation sensitive detectors. Bearing this in mind, we have up to order p^2 for an N-cascade in mode a_x and finding no detector click in the a_y-mode:

$$\rho_{\text{out}} \propto \frac{p_1}{8} g_{uvc}\left\{\frac{p_1}{N}[1+(5N-3)(1-\eta_c^2)]|0\rangle\langle 0| + p_2|\Psi_\theta\rangle\langle\Psi_\theta|\right\} + O(p^3)\,, \tag{VI.21}$$

where the vacuum contribution formula was calculated and found to be correct for $N \leq 4$ (and $N \neq 0$).

In order to have non-post-selected quantum teleportation, the fidelity F must be larger than 3/4 [117, 118, 65]. Since I only estimated the two lowest order contributions (to p^2 and p^3), the fidelity is also correct up to p^2 and p^3, and I write $F^{(2)}$ and $F^{(3)}$ respectively. Using Eqs. (VI.35) and (VI.21) we have:

$$F^{(2)} = \frac{Np_2}{p_1[1+(5N-3)(1-\eta_c^2)] + Np_2} \geq \frac{3}{4}\,, \tag{VI.22}$$

$$\iff \qquad \eta_c^2 \geq \frac{(15N-6)p_1 - Np_2}{(15N-9)p_1}\,. \tag{VI.23}$$

This means that in the limit of infinite detector cascading ($N \to \infty$) and $p_1 = p_2$ the efficiency of the detectors must be better than $\frac{14}{15}$ or 93.3% to achieve non-post-selected quantum teleportation. When we have detectors with efficiencies of 98%, we need at least four detectors in the cascade to get unequivocal quantum teleportation. The necessity of a lower bound on the efficiency of the detectors used in the cascade might seem surprising, but this can be explained as follows. Suppose the detector efficiencies become smaller than a certain value x. Then upon a two-photon state entering the detector, finding only one click becomes more likely than finding a coincidence, and 'wrong' events end up contributing to the output state. Eq. (VI.23) places a severe limitation on the practical use of detector cascades in this situation.

In the experiment in Innsbruck, no detector cascade was employed and also the a_y-mode was left undetected. The state entering Bob's detector therefore was

(up to order p^2):

$$\rho_{\text{out}} \propto \frac{p^2}{8} g_{uvc} \left[(3 - \eta_c^2) |0\rangle\langle 0| + |\Psi_\theta\rangle\langle\Psi_\theta| \right] + O(p^3) \,. \tag{VI.24}$$

Remember that $p_1 = p_2$ since the experiment involves one source which is pumped twice. The detector efficiency η_c^2 in the Innsbruck experiment was 10% [173], and the fidelity without detecting the outgoing mode therefore would have been $F^{(2)} \simeq 26\%$ (conditioned only on successful Bell detection and state preparation). This clearly exemplifies the need for Bob's detection. Braunstein and Kimble [32] predicted a theoretical maximum of 50% for the teleportation fidelity, which was conditioned upon (perfect) detection of both the a_x- and the a_y-mode.

Rather than improving the detector efficiencies and using a detector cascade, Eq. (VI.22) can be satisfied by adjusting the probabilities p_1 and p_2 of creating entangled photon-pairs [32]. From Eq. (VI.22) we have

$$p_1 \leq \frac{N}{3[1 + (5N - 3)(1 - \eta_c^2)]} p_2 \,. \tag{VI.25}$$

Experimentally, p_1 can be diminished by employing a beam-splitter with a suitable reflection coefficient rather than a mirror to reverse the pump beam (see figure VI.1). Bearing in mind that κ is proportional to the pump amplitude, the equation $p_i = 2\tanh^2(\kappa_i t)$ [see the discussion following Eq. (III.33) with $i = 1, 2$] gives a relation between the pump amplitude and the probability of creating a photon-pair. In particular when $p_2 = x p_1$:

$$\frac{\tanh(\kappa_2 t)}{\tanh(\kappa_1 t)} = \sqrt{x} \,. \tag{VI.26}$$

Decreasing the production rate of one photon-pair source will increase the time needed to run the experiment. In particular, we have from Eq. (VI.24) that

$$p_2 \geq 3(3 - \eta_c^2) p_1 \,. \tag{VI.27}$$

With $\eta_c^2 = 10\%$, we obtain $p_2 \geq 8.7 p_1$. Using Eq. (III.37) I estimated that diminishing the probability p_1 by a factor 8.7 will increase the running time by that same factor (i.e., running the experiment about nine days, rather than twenty four hours).

The third-order contribution to the outgoing density matrix without cascading and without detecting the a_y-mode is

$$\begin{aligned} \rho_{\text{out}} \propto \frac{p_1}{8} g_{uvc} (4 - \eta_u^2 - \eta_v^2) \frac{1}{16} \big[& 6p_1^2 (6 - 4\eta_c^2 + \eta_c^4) \, |0\rangle\langle 0| \\ & + 2p_1 p_2 (2 - \eta_c^2) \left(|\Psi_\theta\rangle\langle\Psi_\theta| + |\Psi_\theta^\perp\rangle\langle\Psi_\theta^\perp| \right) + \\ & 8p_1 p_2 (3 - \eta_c^2) \rho_1 + 12 p_2^2 \rho_2 \big] \end{aligned} \tag{VI.28}$$

with (we assume from now on that the phase factor $e^{i\varphi}$ in $|\Psi_\theta\rangle$ is real)

$$\rho_1 = \frac{1}{2}\left(|1,0\rangle\langle 1,0| + |0,1\rangle\langle 0,1|\right) , \tag{VI.29}$$

$$\begin{aligned}\rho_2 \;=\; & \frac{1}{6}\left[(2+\cos 2\theta)|0,2\rangle\langle 0,2| + (2-\cos 2\theta)|2,0\rangle\langle 2,0|\right. \\ & + 2|1,1\rangle\langle 1,1| + \frac{1}{2}\sqrt{2}\sin 2\theta\left(|2,0\rangle\langle 1,1|\right. \\ & \left.\left. + |1,1\rangle\langle 2,0| + |0,2\rangle\langle 1,1| + |1,1\rangle\langle 0,2|\right)\right] .\end{aligned} \tag{VI.30}$$

I have explicitly extracted the state which is to be teleported ($|\Psi_\theta\rangle\langle\Psi_\theta|$) from the density matrix contribution ρ_1 (this is not *necessarily* the decomposition with the largest $|\Psi_\theta\rangle\langle\Psi_\theta|$ contribution). As expected, this term is less important in the third order than it is in the second[4].

The teleportation fidelity including the third-order contribution (VI.28) can be derived along the same lines as (VI.22). Assuming that all detectors have the same efficiency η^2 and $p_1 = p_2 = p$, the teleportation fidelity up to third order is

$$F^{(3)} = \frac{4 + p(2-\eta^2)^2}{4(4-\eta^2) + p(80 - 76\eta^2 + 34\eta^4 - 3\eta^6)} . \tag{VI.31}$$

With $p = 10^{-4}$ and a detector efficiency of $\eta^2 = 0.1$, this fidelity differs from (VI.22) with only a few parts in ten thousand:

$$\frac{F^{(2)} - F^{(3)}}{F^{(2)}} \;\propto\; p \;\sim\; 10^{-4} . \tag{VI.32}$$

[4]The density matrix consists of several distinct parts: a vacuum contribution, a contribution due to one photon in mode d, two photons, and so on. Suppose there are n photon-pairs created in the whole system, and m photon-pairs out of n are produced by the second source (modes c and d). The outgoing mode must then contain m photons. Reversing this argument, when we find m photons in the outgoing mode the probability of creating this particular contribution must be proportional to $p_1^{n-m}p_2^m$. Expanding the n-th order output state into parts of definite photon number we can write

$$\rho_{\text{out}}^{(n)} = \sum_{m=0}^{n-1} p_1^{n-m} p_2^m \rho_m^{(n)} ,$$

where $\rho_m^{(n)}$ is the (unnormalised) n-th order contribution containing all terms with m photons.

An immediate corollary of this argument is that all the cross-terms between different photon number states in the density matrix must vanish. The cross-terms *are* present in Eq. (VI.18), and I must therefore show that the partial trace in Eq. (VI.19) makes them vanish. Suppose there are n photons in the total system. A cross-term in the density matrix will have the form

$$|j,k,l,m\rangle_{auvd}\langle j',k',l',m'| ,$$

with $m \neq m'$. We also know that $j+k+l+m = j'+k'+l'+m' = n$, so that at least one of the other modes must have the cross-term property as well. Suppose k is not equal to k'. Since we have $\text{Tr}[|k\rangle\langle k'|] = \delta_{k,k'}$, the cross-terms must vanish.

On the other hand, let me compare two *gedanken* experiments in which the cascades have different detector efficiencies (but all the detectors in one cascade still have the same efficiency). The ratio between the teleportation fidelity with detector efficiencies η_-^2 and η_+^2 (with η_-^2 and η_+^2 the lower and higher detector efficiencies respectively) up to lowest order is

$$\frac{F^{(2)}_{95\%} - F^{(2)}_{10\%}}{F^{(2)}_{95\%}} \propto \frac{\Delta\eta^2}{2-\eta_-^2} \sim 0.1\,, \tag{VI.33}$$

where $\Delta\eta^2$ is the difference between these efficiencies. This shows that detector efficiencies have a considerably larger influence on the teleportation fidelity than the higher-order pair production, as expected.

To summarise my results, I have found that detector cascading is only useful for this realisation of quantum teleportation when the detectors in the cascade have near unit efficiency, in accordance with the results of chapter IV. In particular, there is a lower bound to the efficiency below which an increase in the number of detectors in the cascade actually *decreases* the ability to distinguish between one or several photons entering the cascade. Finally, enhancement of the photon-pair source responsible for the entanglement channel relative to the one responsible for the state preparation increases the time needed to run the experiment by roughly an order of magnitude.

g Fidelity versus efficiency

In the context of the Innsbruck experiment, the fidelity is used to distinguish between quantum teleportation and teleportation which could have been achieved 'classically'. Here, classical teleportation is the disembodied transport of some quantum state from Alice to Bob by means of a classical communication channel alone. There is *no* shared entanglement between Alice and Bob. Since classical communication can be duplicated, such a scheme can lead to many copies of the transported output state (so-called *clones*). Classical teleportation with perfect fidelity (i.e., $F = 1$) would then lead to the possibility of *perfect* cloning, thus violating the no-cloning theorem [178, 49]. This means that the maximum fidelity for classical teleportation has an upper bound which is less than one.

Quantum teleportation, on the other hand, can achieve perfect fidelity (and circumvents the no-cloning theorem by disrupting the original). To demonstrate *quantum* teleportation therefore means that the teleported state should have a higher fidelity than possible for a state obtained by any scheme involving classical communication *alone*[5].

For classical teleportation of randomly sampled polarisations, the maximum attainable fidelity is $F = 2/3$ [63, 65]. When only linear polarisations are to be

[5]The fidelity captures this one particular feature of quantum teleportation very well and is already extensively studied.

teleported, the maximum attainable fidelity is $F = 3/4$ [117, 118]. These are the values which the fidelity of true quantum teleportation should exceed.

In the case of the Innsbruck experiment, $|\phi\rangle$ denotes the 'unknown' linear polarisation state of the photon issued by Victor. I can write the undetected outgoing state (to lowest order and conditioned on a successful Bell state measurement) as

$$\rho_{\text{out}} \propto |\alpha|^2 |0\rangle\langle 0| + |\beta|^2 |\phi\rangle\langle\phi| \,, \tag{VI.34}$$

where $|0\rangle$ is the vacuum state. The overlap between $|\phi\rangle$ and ρ_{out} is given by Eq. (II.71). In the Innsbruck experiment the fidelity F is then given by

$$F \equiv \text{Tr}[\rho_{\text{out}} |\phi\rangle\langle\phi|] = \frac{|\beta|^2}{|\alpha|^2 + |\beta|^2} \,. \tag{VI.35}$$

This should be larger than 3/4 in order to demonstrate quantum teleportation. The vacuum contribution in Eq. (VI.34) arises from the fact that Victor cannot distinguish between one or several photons entering his detector, i.e., Victor's inability to properly prepare a single-photon state.

As pointed out by Braunstein and Kimble [32], the fidelity of the Innsbruck experiment remains well below the lower bound of 3/4 due to the vacuum contribution. Replying to this, Bouwmeester *et al.* [24, 25] argued that 'when a photon appears, it has all the properties required by the teleportation protocol'. The vacuum contribution in Eq. (VI.34) should therefore only affect the efficiency of the experiment, with a consequently high fidelity. However, this is a potentially ambiguous statement. If by 'appear' we mean 'appearing in a photo-detector', I agree that a high fidelity (and low efficiency) can be inferred. However, this yields a so-called *post-selected* fidelity, where the detection destroys the teleported state. The fidelity prior to (or without) Bob's detection is called the *non*-post-selected fidelity. The question is now whether we can say that a photon appears when *no* detection is made, thus yielding a high *non*-post-selected fidelity.

This turns out not to be the case. Making an *ontological* distinction between a photon and *no* photon in a mixed state (without a detection) is based on what we call the 'Partition Ensemble Fallacy,' introduced in chapter II. In the absence of Bob's detection, the density matrix of the teleported state (i.e., the *non*-post-selected state) may be decomposed into an infinite number of partitions. These partitions do not necessarily include the vacuum state at all. It would therefore be incorrect to say that teleportation did or did not occur except through some operational means (e.g., a detection performed by Bob).

Bob's detection thus leads to a high post-selected fidelity. However, the vacuum term in Eq. (VI.34) contributes to the *non*-post-selected fidelity, decreasing it well below the lower bound of 3/4. Due to this vacuum contribution, the Innsbruck experiment did *not* demonstrate *non*-post-selected quantum teleportation. Nonetheless, teleportation was demonstrated using post-selected data obtained

by detecting the teleported state. By selecting events where a photon was observed in the teleported state, a post-selected fidelity higher than 3/4 could be inferred (estimated at roughly 80% [25])[6].

3 Entanglement swapping and purification

a *Teleportation of entanglement: swapping*

In the previous section, I discussed quantum teleportation [14], in which a quantum state is sent from Alice to Bob using (maximal) entanglement. If this quantum state is itself part of an entangled state, i.e., if Alice's system is entangled with Charlie's system, this entanglement is 'transferred' from Alice to Bob. In other words, Bob's system becomes entangled with Charlie's system, even though these two systems might never have physically interacted. This is called *entanglement swapping* [183].

For example, suppose we have a system of two independent polarisation entangled photon-pairs in modes a, b and c, d respectively. If we restrict ourselves to the Bell states, we have

$$|\Psi\rangle_{abcd} = |\Psi^-\rangle_{ab} \otimes |\Psi^-\rangle_{cd} \,. \tag{VI.36}$$

However, on a different basis this state can be written as:

$$\begin{aligned} |\Psi\rangle_{abcd} \;=\; & \frac{1}{2}|\Psi^-\rangle_{ad} \otimes |\Psi^-\rangle_{bc} + \frac{1}{2}|\Psi^+\rangle_{ad} \otimes |\Psi^+\rangle_{bc} \\ & + \frac{1}{2}|\Phi^-\rangle_{ad} \otimes |\Phi^-\rangle_{bc} + \frac{1}{2}|\Phi^+\rangle_{ad} \otimes |\Phi^+\rangle_{bc} \,. \end{aligned} \tag{VI.37}$$

This can be easily checked by writing out the Bell states. The non-cancelling terms can be rewritten as Eq. (VI.36).

If we make a Bell measurement on modes b and c, we can see from Eq. (VI.37) that the undetected remaining modes a and d become entangled. For instance, when we find modes b and c in a $|\Phi^+\rangle$ Bell state, the remaining modes a and d must be in the $|\Phi^+\rangle$ state as well. In appendix E I show that a suitably chosen unitary transformation of Bob's branch can return the state to $|\Psi^-\rangle_{ad}$, just as in the teleportation of a single state.

Entanglement swapping was performed in Innsbruck by Pan *et al.* in 1998 [126]. In this experiment two parametric down-converters were employed to create polarisation entanglement[7]. The schematics of the experimental setup are depicted in Fig. VI.4. One branch of each down-converter is sent into a 50:50

[6] We recall that this entire discussion is restricted to the subset of events where successful Bell-state and state-preparation have occurred.

[7] In the experiment, the two down-converters were implemented by a single BBO crystal pumped twice in opposite directions. The experiment thus closely resembled the quantum teleportation experiment six months earlier [23] (see also chapter II).

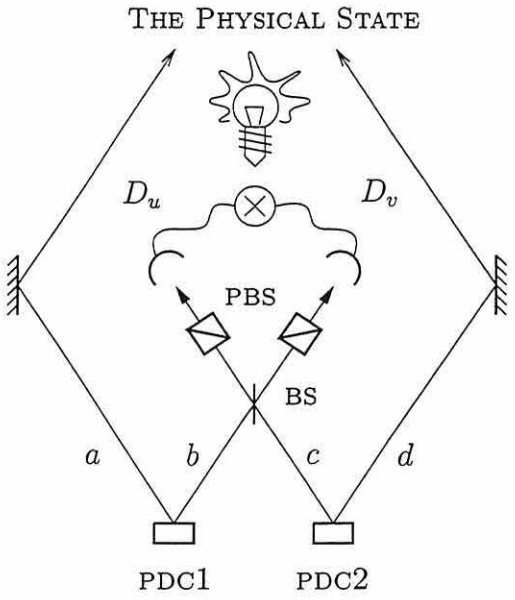

Figure VI.4: A schematic representation of the entanglement swapping setup. Two parametric down-converters (PDC) create states which exhibit polarisation entanglement. One branch of each source is sent into a beam splitter (BS), after which the polarisation beam splitters (PBS) select particular polarisation settings. A coincidence in detectors D_u and D_v ideally identify the $|\Psi^-\rangle$ Bell state. However, since there is a possibility that one down-converter produces two photon-pairs while the other produces nothing, the detectors D_u and D_v no longer constitute a Bell-detection, and the freely propagating PHYSICAL STATE is no longer a pure Bell state.

beam-splitter. The outgoing modes of the beam-splitter are detected. A detector coincidence indicates that the state $|\Psi^-\rangle$ was present, and thus acts as an (incomplete) Bell measurement. I have included two polarisation beam-splitters in the outgoing modes of the beam-splitter. These were not present in the actual experiment, but they play an important rôle in the subsequent discussion [184].

Since entanglement swapping is formally the teleportation of one branch of an entangled state, it should not come as a surprise that the entanglement swapping experiment performed by Pan *et al.* suffers from the same complication as the quantum teleportation experiment performed by Bouwmeester *et al.* [23]: apart from both down-converters creating a single pair, there is also the possibility that one of the down-converters creates two pairs, while nothing happens in the other.

Now I use the extra information about the detected photons due to the polarisation beam-splitters in figure VI.4. Conditioned on the polarisation (j, k), with

$j, k \in \{x, y\}$, of the detected photons we obtain the outgoing states

$$\begin{aligned} |\Upsilon_{(x,x)}\rangle_{14} &= \frac{1}{\sqrt{2}}\left(|0, y^2\rangle - |y^2, 0\rangle\right) , \\ |\Upsilon_{(x,y)}\rangle_{14} &= \frac{1}{2}\left(|xy, 0\rangle - |x, y\rangle + |y, x\rangle - |0, xy\rangle\right) , \\ |\Upsilon_{(y,x)}\rangle_{14} &= \frac{1}{2}\left(|xy, 0\rangle + |x, y\rangle - |y, x\rangle - |0, xy\rangle\right) , \\ |\Upsilon_{(y,y)}\rangle_{14} &= \frac{1}{\sqrt{2}}\left(|0, x^2\rangle - |x^2, 0\rangle\right) . \end{aligned} \tag{VI.38}$$

The outgoing state of the entanglement swapping experiment (without the polarisation beam-splitter) is a random mixture of these four states.

Let me define the following states:

$$\begin{aligned} |\Phi_{xy}\rangle &\equiv \frac{1}{\sqrt{2}}\left(|xy, 0\rangle - |0, xy\rangle\right) , \\ |\Phi_{x^2}\rangle &\equiv \frac{1}{\sqrt{2}}\left(|x^2, 0\rangle - |0, x^2\rangle\right) , \\ |\Phi_{y^2}\rangle &\equiv \frac{1}{\sqrt{2}}\left(|y^2, 0\rangle - |0, y^2\rangle\right) , \\ |\Psi^-\rangle &\equiv \frac{1}{\sqrt{2}}\left(|x, y\rangle - |y, x\rangle\right) . \end{aligned} \tag{VI.39}$$

After some involved, but essentially straightforward algebra it can be shown that the outgoing state of the entanglement swapping experiment performed by Pan *et al.* can also be written as the mixed state ρ:

$$\rho = \frac{1}{4}\left(|\Phi_{xy}\rangle_{ad}\langle\Phi_{xy}| + |\Phi_{x^2}\rangle_{ad}\langle\Phi_{x^2}| + |\Phi_{y^2}\rangle_{ad}\langle\Phi_{y^2}| + |\Psi^-\rangle_{ad}\langle\Psi^-|\right) \tag{VI.40}$$

to lowest order. *Conditioned* on detected photons in the outgoing modes a high entanglement swapping fidelity can be inferred ($F \sim 1$). However, the fidelity for non-post-selected entanglement swapping is $F = \frac{1}{4}$. This argument is completely analogous to the argument presented in section 2*g*.

b Entanglement swapping as purification

The non-post-selected fidelity $F = \frac{1}{4}$ of having a maximally entangled state $|\Psi^-\rangle$ as the output of the entanglement swapping experiment is much higher than the that of the down-converter output state, where $F \sim 10^{-4}$. This suggests that entanglement swapping can be viewed as a purification protocol (see also chapter III). Indeed, this has been suggested by Bose *et al.* [21]. This protocol was subsequently extended by Shi *et al.* [152].

The Bose protocol works as follows: let x and y denote photons with polarisations in the x- and y-direction of a Cartesian coordinate system. We consider

(an ensemble of) non-maximally entangled states for two systems 1 and 2

$$|\Phi(\theta)\rangle_{12} = \cos\theta|x,x\rangle_{12} + \sin\theta|y,y\rangle_{12}\ , \tag{VI.41}$$

and similarly for systems 3 and 4:

$$|\Phi(\theta)\rangle_{34} = \cos\theta|x,x\rangle_{34} + \sin\theta|y,y\rangle_{34}\ . \tag{VI.42}$$

The purification protocol now employs entanglement swapping from systems 1, 2 and 3, 4 to the two systems 1 and 4. These two systems were previously unentangled. Making a Bell measurement of system 2 and 3 entangles the remaining systems 1 and 4.

There are four different outcomes of the Bell measurement, which give rise to four different entangled states in systems 1 and 4. These are [21]

$$\begin{aligned}
|\Phi^+\rangle_{23}: &\quad |\Psi_{\text{out}}\rangle_{14} = \frac{1}{N}\left(\cos^2\theta|x,x\rangle_{14} + \sin^2\theta|y,y\rangle_{14}\right)\ ,\\
|\Phi^-\rangle_{23}: &\quad |\Psi_{\text{out}}\rangle_{14} = \frac{1}{N}\left(\cos^2\theta|x,x\rangle_{14} - \sin^2\theta|y,y\rangle_{14}\right)\ ,\\
|\Psi^+\rangle_{23}: &\quad |\Psi_{\text{out}}\rangle_{14} = \frac{1}{\sqrt{2}}\left(|x,y\rangle_{14} + |y,x\rangle_{14}\right)\ ,\\
|\Psi^-\rangle_{23}: &\quad |\Psi_{\text{out}}\rangle_{14} = \frac{1}{\sqrt{2}}\left(|x,y\rangle_{14} - |y,x\rangle_{14}\right)\ .
\end{aligned} \tag{VI.43}$$

The normalisation factor N is given by $N = \sqrt{\cos^4\theta + \sin^4\theta}$. It is easily seen that the measurement outcomes $|\Phi^+\rangle_{23}$ and $|\Phi^-\rangle_{23}$ actually *degrade* the entanglement compared to the entanglement of the systems 1 and 2 or 3 and 4. In the case of measurement outcomes $|\Psi^+\rangle_{23}$ and $|\Psi^-\rangle_{23}$, however, the resulting (pure) states are *maximally* entangled. With probability $\sqrt{2}\cos\theta\sin\theta$ we will obtain a maximally entangled state, and with probability $\sqrt{\cos^4\theta + \sin^4\theta}$ we degrade the entanglement.

When we compare this protocol with the swapping experiment by Pan *et al.*, we note that there is a crucial difference: The outgoing state of the experiment is not confined to the Hilbert space spanned by the basis $\{|x,x\rangle, |x,y\rangle, |y,x\rangle, |y,y\rangle\}$, contrary to the protocol by Bose *et al.* In addition, we have to include states like the vacuum ($|0\rangle$) and two-photon states ($|x^2\rangle$, $|y^2\rangle$ and $|xy\rangle$).

However, if the swapping protocol used by Pan *et al.* can increase the entanglement content upon repetition of the procedure in this larger Hilbert space (or, more precisely, this truncated Fock space), we can still call it a purification protocol. I will now investigate this.

First, I have to determine precisely what I mean by a repetition of the swapping procedure. It means that the *outgoing states* of two distinct entanglement swapping setups are again used as an entanglement source for a swapping experiment. Repeating this N times, we can depict this as a string of down-converters connected by beam-splitters (see figure VI.5). Such a string can consist of an even

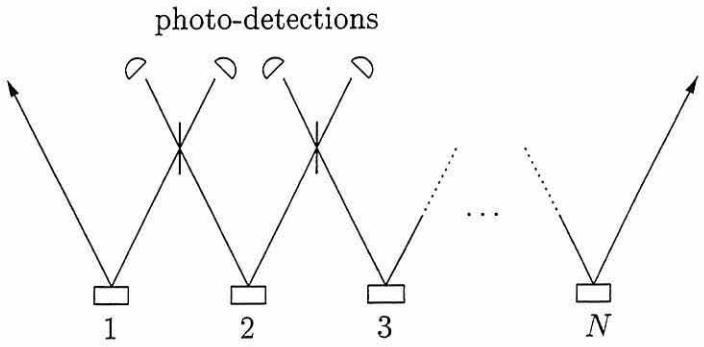

Figure VI.5: A series of parametric down-converters 1 to N, of which the outgoing modes are connected by beam-splitters to form a string. The photo-detections are essentially *polarisation sensitive* photo-detectors (an incomplete Bell measurement would require the loss of the polarisation information). This can be interpreted as repeated entanglement swapping. However, is it also a repeated purification protocol?

or an odd number of down-converters. When N is odd, we have a string of 'entanglement swappers' the outgoing states of which are again used in entanglement swapping.

Instead of Bell detections, we consider polarisation sensitive photo-detection. This allows us to condition the outgoing state on x- and y-polarised photons in the detectors. In the case $N = 1$ this led to the outgoing state

$$|\Upsilon_{(x,y)}\rangle = \frac{1}{\sqrt{2}} \left(|xy, 0\rangle - |0, xy\rangle \right) . \tag{VI.44}$$

In the case of N down-converters, we can keep track of the single- and double-pair production in the following table:

PDC	1	2	3		N
# pairs	1	1	1	...	1
	0	2	0	...	0/2 (odd/even)
	2	0	2	...	2/0 (odd/even)

In the top row the parametric down-converters are enumerated (in accordance with figure VI.5). The entries in the lower rows identify the number of photon-pairs created by the associated down-converter. These rows identify the only three possibilities in which the detectors (from left to right) signal the detection of polarised photons in the direction x, y, x, $y \ldots$

There are several things to be noted. First of all, depending on the parity of N, the bottom three rows correspond to different orders of pair-creation. The lowest order corresponds to no created pairs (vacuum), the next order is one created pair, and so on. Since parametric down-conversion has such a small

probability of creating an entangled photon-pair, the lowest order is always by far the leading order.

Secondly, we can easily verify that the possibility of every down-converter creating exactly *one* photon-pair never occurs alone. We can always construct a different photon-pair configuration of the same order which triggers the detectors in the same way, thus preventing the creation of maximal entanglement (this is, of course, not the proof I was looking for in section 3 of chapter III). At the same time, it is easily verified that (for ideal detections) the possibility that one down-converter creates *three* pairs is dismissed, since it would mean that some detectors see at least two photons. Furthermore, this procedure can be immediately repeated for other polarisation choices in the detectors.

Finally, we should note that the entanglement content alternates between almost nothing ($F \sim 10^{-4}$) and about one quarter ($F = \frac{1}{4}$). This behaviour occurs because for an odd number of down-converters, the detectors in the setup can be triggered by $(N-1)/2$ down-converters creating a double-pair. The total number of pairs is then $N-1$, which is the lowest order.

When we have an even number of down-converters we have three possibilities given in the table above. When every down-converter creates one photon-pair, the outgoing state will be (up to lowest order) the anti-symmetric Bell state. In this case, all down-converter modes are connected by means of the beam-splitter operations and entanglement swapping is successful. On the other hand when the down-converters create two and zero photon pairs in alternation, the left-hand outgoing mode is independent from the right-hand outgoing mode. There is only a classical correlation between them: if on the left there are two photons, then we have the vacuum state on the right and vice versa.

The outgoing state is

$$|\Upsilon\rangle = \frac{1}{2}\left(|0, xy\rangle - |x, y\rangle + |y, x\rangle - |xy, 0\rangle\right) , \tag{VI.45}$$

conditioned on a detector sequence x, y, x, y ... This is independent of the number of down-converters, as long as it is even. As a consequence, we cannot interpret entanglement swapping as performed by Pan *et al.* as a purification protocol.

c *Entanglement content of output states*

I will now return to the states given by Eq. (VI.38). These states can also be obtained by running the states $|x, x\rangle$, $|x, y\rangle$, $|y, x\rangle$ and $|y, y\rangle$ through a 50:50 beam-splitter. This raises the question what the entanglement content of the states of Eq. (VI.38) is. After all, the states $|j, k\rangle$ (with $j, k \in \{x, y\}$) are *separable*.

The non-locality of single photons has been studied by Hardy [79, 80] and Peres [133] (see also chapter II). Let a and b denote different spatial modes, $|0\rangle$ is

the vacuum and $|1\rangle$ is a single-photon state. The essential idea is that the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle_a|0\rangle_b - |0\rangle_a|1\rangle_b\right) \tag{VI.46}$$

can be used to construct the violation of the Clauser-Horne-Shimony-Holt, or CHSH inequality [44, 131] (this is a variant of a Bell inequality [11]). Consequently, a single-photon state can exhibit non-local properties. This point was debated by Vaidman [166] and Greenberger, Horne and Zeilinger [72, 73].

Here, I will discuss the entanglement content of the two-photon states which are obtained by mixing two single-photon states at a 50:50 beam-splitter. To this end I use the so-called Peres-Horodecki partial transpose criterion for density matrices [87, 134] (see appendix B). Using the partial transpose criterion I will examine the entanglement content of the state $(|x^2,0\rangle - |0,x^2\rangle)/\sqrt{2}$ and the density matrix ρ, which is a mixture of the states given in Eq. (VI.38).

We can write the density matrix of the state $(|x^2,0\rangle - |0,x^2\rangle)/\sqrt{2}$ as

$$\rho = \frac{1}{2}\left(|0,x^2\rangle\langle 0,x^2| - |0,x^2\rangle\langle x^2,0| - |x^2,0\rangle\langle 0,x^2| + |x^2,0\rangle\langle x^2,0|\right) . \tag{VI.47}$$

We obtain the partial transpose by exchanging the second entries of the bras and kets, yielding

$$\rho' = \frac{1}{2}\left(|0,x^2\rangle\langle 0,x^2| - |0,0\rangle\langle x^2,x^2| - |x^2,x^2\rangle\langle 0,0| + |x^2,0\rangle\langle x^2,0|\right) . \tag{VI.48}$$

In matrix representation on the basis $\{|0,0\rangle, |0,x^2\rangle, |x^2,0\rangle, |x^2,x^2\rangle,\}$ this becomes[8]

$$\rho' = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{VI.49}$$

The eigenvalues of this matrix are $1/2$ (with multiplicity 3) and $-1/2$. As proved in appendix B, the negative eigenvalues imply that the state $(|0,x^2\rangle - |x^2,0\rangle)/\sqrt{2}$ is entangled.

In the experiment performed by Pan *et al.*, no polarisation beam-splitters were used, and the outgoing state before post-selection was a mixture of the states given in Eq. (VI.39). The density matrix can be written as

$$\begin{aligned}\rho \;=\; \frac{1}{8}\Big[& |0,x^2\rangle\langle 0,x^2| - |0,x^2\rangle\langle x^2,0| - |x^2,0\rangle\langle 0,x^2| \\ & + |x^2,0\rangle\langle x^2,0| + |0,y^2\rangle\langle 0,y^2| - |0,y^2\rangle\langle y^2,0| - |y^2,0\rangle\langle 0,y^2|\end{aligned}$$

[8] This is a matrix on a truncated Fock space corresponding to the given basis.

$$
\begin{aligned}
&+|y^2,0\rangle\langle y^2,0| + |xy,0\rangle\langle xy,0| - |xy,0\rangle\langle 0,xy| - |0,xy\rangle\langle xy,0| \\
&+|0,xy\rangle\langle 0,xy| + |x,y\rangle\langle x,y| - |x,y\rangle\langle y,x| \\
&-|y,x\rangle\langle x,y| + |y,x\rangle\langle y,x|\Big] \,. \qquad \text{(VI.50)}
\end{aligned}
$$

The partial transpose then becomes

$$
\begin{aligned}
\rho' \;=\; \frac{1}{8}\Big[&|0,x^2\rangle\langle 0,x^2| - |0,0\rangle\langle x^2,x^2| - |x^2,x^2\rangle\langle 0,0| \\
&+|x^2,0\rangle\langle x^2,0| + |0,y^2\rangle\langle 0,y^2| - |0,0\rangle\langle y^2,y^2| - |y^2,y^2\rangle\langle 0,0| \\
&+|y^2,0\rangle\langle y^2,0| + |xy,0\rangle\langle xy,0| - |xy,xy\rangle\langle 0,0| - |0,0\rangle\langle xy,xy| \\
&+|0,xy\rangle\langle 0,xy| + |x,y\rangle\langle x,y| - |x,x\rangle\langle y,y| \\
&-|y,y\rangle\langle x,x| + |y,x\rangle\langle y,x|\Big] \,. \qquad \text{(VI.51)}
\end{aligned}
$$

In matrix representation on the basis

$$
\begin{aligned}
\{&|0,0\rangle, |0,x^2\rangle, |x^2,0\rangle, |x^2,x^2\rangle, |0,y^2\rangle, |y^2,0\rangle, |y^2,y^2\rangle, \\
&|xy,0\rangle, |0,xy\rangle, |xy,xy\rangle, |x,y\rangle, |y,x\rangle, |x,x\rangle, |y,y\rangle\} \,,
\end{aligned}
$$

the partial transpose ρ' becomes[9]

$$
\rho' = \frac{1}{8}\begin{pmatrix}
0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0
\end{pmatrix} . \qquad \text{(VI.52)}
$$

The eigenvalues of this matrix are given by 0 (multiplicity 2), $1/8$ (multiplicity 9), $-1/8$ (multiplicity 1), $\sqrt{3}/8$ (multiplicity 1) and $-\sqrt{3}/8$ (multiplicity 1). Since ρ' has negative eigenvalues, ρ is an entangled state [134, 87]. This means that entanglement swapping as it was originally proposed really does work, although maximally entangled (Bell) states can only be seen in a post-selected manner.

This teaches us something interesting about entanglement. It is generally believed that when two systems are entangled, they have somehow interacted

[9] Again on a truncated Fock space.

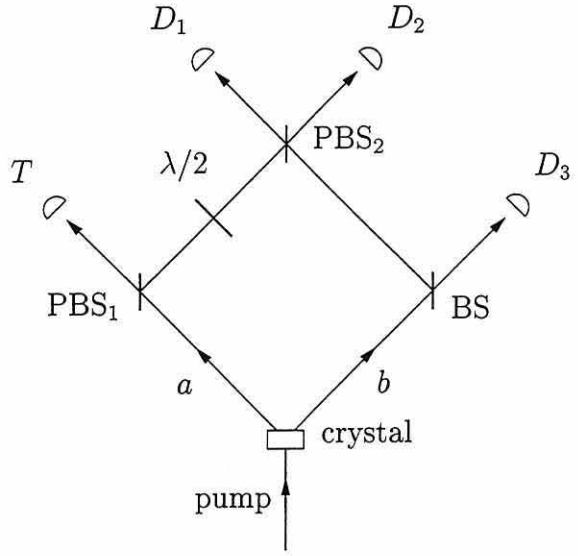

Figure VI.6: Schematic representation of the experimental setup which was used to demonstrate the existence of three-photon GHZ-states in a post-selected manner. A BBO crystal is pumped to create *two* photon pairs. The subsequent interferometer is arranged such that conditioned on a detection event in detector T, the detectors D_1, D_2 and D_3 signal the detection of a GHZ-state. Furthermore, the interferometer includes polarisation beam-splitters (PBS$_1$ and PBS$_2$), a beam-splitter (BS) and a $\lambda/2$ phase plate which transforms $|y\rangle$ into $(|x\rangle + |y\rangle)/\sqrt{2}$.

in the past. However, in the beam-splitter the two photons do *not* interact with each other, and yet the outgoing state (given a separable input state) is entangled. This shows that entanglement does not necessarily originates from an interaction.

4 Three-particle entanglement

Three-particle maximally entangled states have also been produced in a post-selected manner [26]. In figure VI.6 a schematic representation of this experiment is shown [182]. As in the teleportation and entanglement swapping experiments, a non-linear crystal is pumped with a short-pulsed high-intensity laser. However, this time the setup is chosen such that the down-converter directly creates two photon-pairs in modes a and b. The lowest order of the created state which can trigger the four detectors (T, D_1, D_2 and D_3) is given by

$$|\Psi\rangle_{ab} \simeq \frac{1}{\sqrt{3}} \left(|x^2, y^2\rangle - |xy, xy\rangle + |y^2, x^2\rangle \right) + O(\xi) \,, \qquad \text{(VI.53)}$$

with $\xi \ll 1$.

It is clear that only the branch $|xy, xy\rangle_{ab}$ can trigger all detectors due to the polarisation beam-splitter PBS_1 (up to lowest order). This polarisation beam-splitter transmits x-polarised photons (which trigger T) and the retardation plate transforms the reflected photon $|y\rangle$ into $(|x\rangle + |y\rangle)/\sqrt{2}$. The part of the outgoing state (i.e., before detection) due to the input $|xy, xy\rangle_{ab}$ and conditioned on a photon in detector T is thus

$$\begin{aligned} |\Phi\rangle \quad \propto \quad & |xy, x, 0\rangle + |x, x, y\rangle + |xy, 0, x\rangle + |x, 0, xy\rangle \\ & + |y, xy, 0\rangle + |0, xy, y\rangle + |y, y, x\rangle + |0, y, xy\rangle \; . \end{aligned} \tag{VI.54}$$

From this state it is easily seen that three-photon entanglement can only be observed in a post-selected manner by discarding the branches which include the vacuum $|0\rangle$. A three-fold coincidence in the detectors D_1, D_2 and D_3 can thus only come from the branches $|x, x, y\rangle$ and $|y, y, x\rangle$.

As I argued in section 1, this experiment does *not* demonstrate the existence of the GHZ-state $|x, x, y\rangle + |y, y, x\rangle$. However, using post-selection of the data-set on a four-fold detector coincidence, non-local correlations could be inferred. This was demonstrated by Pan *et al.* in 2000 [127]. In the first experiment [26] two tests were made: first it was shown that *only* the branches $|x, x, y\rangle$ and $|y, y, x\rangle$ contributed to a four-fold coincidence. Secondly it had to be shown that the two branches were in a coherent superposition, since the post-selected events could be due to a statistical mixture of the two branches. This was done by rotating the polarisation over $\pm 45°$ before detection. The observed visibility in these experiments was 75% [26].

5 Summary

In this chapter I studied the optical implementations of quantum teleportation, entanglement swapping and the creation of three-photon Greenberger-Horne-Zeilinger entanglement. All these experiments succeeded in a *post-selected* manner and demonstrated their respective non-local features.

The undetected outgoing state of the teleportation experiment is a mixture of the teleported state and the vacuum. As we have shown, this vacuum contribution degrades the non-post-selected fidelity, rather than the efficiency of the experiment. The outgoing state of the entanglement swapping experiment is more complicated, since it is not simply a mixture of the vacuum and the swapped state. Here, the photo-detection post-selects particular branches from a superposition.

The same happens in the creation of the three-photon GHZ-state. Furthermore, since no physically propagating state left the apparatus after detection, we cannot say that the state $|x, x, y\rangle + |y, y, x\rangle$ was created. However, non-local correlations using post-selected data can be observed.

VII

Quantum Lithography

Optical lithography is a widely used printing method. In this process light is used to etch a substrate. The (un)exposed areas on the substrate then define the pattern. In particular, the micro-chip industry uses lithography to produce smaller and smaller processors. However, classical optical lithography can only achieve a resolution comparable to the wavelength of the used light [38, 115, 116]. It therefore limits the scale of the patterns. To create smaller patterns we need to venture beyond this classical boundary [179]. Here, I investigate how we can beat this boundary. This chapter is based on a collaboration with Agedi N. Boto, Daniel S. Abrams, Colin P. Williams and Jonathan P. Dowling at the Jet Propulsion Laboratory, Pasadena [22, 104, 105]. Recently, similar work was done by Björk, Sánchez Soto and Söderholm [17].

In Ref. [22] we introduced a procedure called *quantum* lithography which offers an increase in resolution without an upper bound. This enables us to use quantum lithography to write closely spaced lines in one dimension. However, for practical purposes (like, e.g., optical surface etching) we need the ability to create more complicated patterns in both one and two dimensions. Here, we study how quantum lithography allows us to create these patterns.

This chapter is organised as follows: first I derive the classical resolution limit in section 1. Section 2 reiterates the method introduced in Ref. [22]. Then, in section 3, I consider a generalised version of this procedure and show how we can tailor arbitrary one-dimensional patterns. Section 4 shows how a further generalisation of this procedure leads to arbitrary patterns in two dimensions. Finally, section 5 addresses the issues concerning the physical implementation of quantum lithography.

1 Classical resolution limit

Classically, we can only resolve details of finite size. In this section we give a derivation of this classical resolution limit using the so-called Rayleigh criterion [139]. Suppose two plane waves characterised by $\vec{k}_1$ and $\vec{k}_2$ hit a surface under an angle θ from the normal vector. The wave vectors are given by

$$\vec{k}_1 = k(\cos\theta, \sin\theta) \quad \text{and} \quad \vec{k}_2 = k(\cos\theta, -\sin\theta)\,, \tag{VII.1}$$

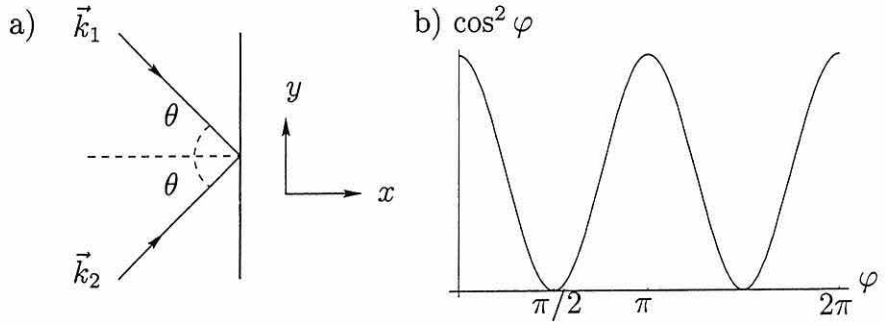

Figure VII.1: a) Schematic representation of two light beams $\vec{k}_1$ and $\vec{k}_2$ incident on a surface, yielding an interference pattern. b) The interference pattern for $\varphi = ky\sin\theta$.

where we used $|\vec{k}_1| = |\vec{k}_2| = k$. The wave number k is related to the wavelength of the light according to $k = 2\pi/\lambda$.

In order to find the interference pattern in the intensity, we sum the two plane waves at position $\vec{r}$ at the amplitude level:

$$I(\vec{r}) \propto \left| e^{i\vec{k}_1\cdot\vec{r}} + e^{i\vec{k}_2\cdot\vec{r}} \right|^2 = 4\cos^2\left[\frac{1}{2}(\vec{k}_1 - \vec{k}_2)\cdot\vec{r}\right] . \tag{VII.2}$$

When we calculate the inner product $(\vec{k}_1 - \vec{k}_2)\cdot\vec{r}/2$ from Eq. (VII.1) we obtain the expression

$$I(x) \propto \cos^2(kx\sin\theta) \tag{VII.3}$$

for the intensity along the substrate in direction x.

The Rayleigh criterion states that the minimal resolvable feature size Δx corresponds to the distance between an intensity maximum and an adjacent minimum (see figure VII.1). From Eq. (VII.3) we obtain

$$k\Delta x\sin\theta = \frac{\pi}{2} . \tag{VII.4}$$

This means that the maximum resolution is given by

$$\Delta x = \frac{\pi}{2k\sin\theta} = \frac{\pi}{2\left(\frac{2\pi}{\lambda}\sin\theta\right)} = \frac{\lambda}{4\sin\theta} , \tag{VII.5}$$

where λ is the wavelength of the light. The maximum resolution is therefore proportional to the wavelength and inversely proportional to the sine of the angle between the incoming plane waves and the normal. The resolution is thus maximal (Δx is minimal) when $\sin\theta = 1$, or $\theta = \pi/2$. This is the grazing limit. The classical diffraction limit is therefore $\Delta x = \lambda/4$.

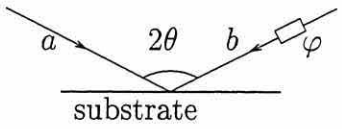

Figure VII.2: Two light beams a and b cross each other at the surface of a photosensitive substrate. The angle between them is 2θ and they have a relative phase difference φ. We consider the limit case of $2\theta \to \pi$.

2 Introduction to Quantum Lithography

In this section we briefly reiterate our method of Ref. [22]. Suppose we have two intersecting light beams a and b. We place some suitable substrate at the position where the two beams meet, such that the interference pattern is recorded. For simplicity, we consider the grazing limit in which the angle θ off axis for the two beams is $\pi/2$ (see figure VII.2). Classically, the interference pattern on the substrate has a resolution of the order of $\lambda/4$, where λ is the wavelength of the light. However, by using entangled photon-number states (i.e., inherently *non-classical* states) we can increase the resolution well into the sub-wavelength regime [90, 138, 163, 62].

How does quantum lithography work? Let the two counter-propagating light beams a and b be in the combined entangled number state

$$|\psi_N\rangle_{ab} = \left(|N,0\rangle_{ab} + e^{iN\varphi}|0,N\rangle_{ab}\right)/\sqrt{2}\,, \tag{VII.6}$$

where $\varphi = kx/2$, with $k = 2\pi/\lambda$. We define the mode operator $\hat{e} = (\hat{a} + \hat{b})/\sqrt{2}$ and its adjoint $\hat{e}^\dagger = (\hat{a}^\dagger + \hat{b}^\dagger)/\sqrt{2}$. The deposition rate Δ on a substrate sensitive to N photons with wavelength λ (a so-called N-photon resist) is then given by

$$\Delta_N = \langle\psi_N|\hat{\delta}_N|\psi_N\rangle \qquad \text{with} \qquad \hat{\delta}_N = \frac{(\hat{e}^\dagger)^N \hat{e}^N}{N!}\,, \tag{VII.7}$$

i.e., we look at the higher moments of the electric field operator [70, 91, 135]. The deposition rate Δ is measured in units of intensity. Leaving the substrate exposed for a time t to the light source will result in an exposure pattern $P(\varphi) = \Delta_N t$. After a straightforward calculation we see that

$$\Delta_N \propto (1 + \cos N\varphi)\,. \tag{VII.8}$$

We interpret this as follows. A path-differential phase-shift φ in light beam b results in a displacement x of the interference pattern on the substrate. Using two classical waves, a phase-shift of 2π will return the pattern to its original position. However, according to Eq. (VII.8), one cycle is completed after a shift of $2\pi/N$. This means that a shift of 2π will displace the pattern N times. In other words, we have N times more maxima in the interference pattern. These need

to be closely spaced, yielding an effective Rayleigh resolution of $\Delta x = \lambda/4N$, a factor of N below the classical interferometric result of $\Delta x = \lambda/4$ [38].

Physically, we can interpret this result as follows: instead of having a state of N single photons, Eq. (VII.6) describes an N-photon state. Since the momentum of this state is N times as large as the momentum for a single photon, the corresponding De Broglie wavelength is N times smaller. The interference of this N-photon state with itself on a substrate thus gives a periodic pattern with a characteristic resolution dimension of $\Delta x = \lambda/4N$.

3 General Patterns in 1D

So far, we have described a method to print a simple pattern of evenly spaced lines of sub-wavelength resolution. However, for any practical application we need the ability to produce more complicated patterns. To this end, we introduce the state

$$|\psi_{Nm}\rangle_{ab} = \left(e^{im\varphi}|N-m,m\rangle_{ab} + e^{i(N-m)\varphi}e^{i\theta_m}|m,N-m\rangle_{ab}\right)/\sqrt{2}\,. \qquad \text{(VII.9)}$$

This is a generalised version of Eq. (VII.6). In particular, Eq. (VII.9) reduces to Eq. (VII.6) when $m = 0$ and $\theta_m = 0$. Note that we included a relative phase $e^{i\theta_m}$, which will turn out to be crucial in the creation of arbitrary one-dimensional patterns.

We can calculate the deposition rate again according to the procedure in section 2. As we shall see later, in general, we can have superpositions of the states given by Eq. (VII.9). We therefore have to take into account the possibility of different values of m, yielding a quantity

$$\Delta_{Nm}^{Nm'} = \langle\psi_{Nm}|\hat{\delta}_N|\psi_{Nm'}\rangle\,. \qquad \text{(VII.10)}$$

Note that this deposition rate depends not only on the parameter φ, but also on the relative phases θ_m and $\theta_{m'}$. The deposition rate then becomes

$$\begin{aligned}\Delta_{Nm}^{Nm'} \propto \sqrt{\binom{N}{m}\binom{N}{m'}}\Big[&e^{i(m'-m)\varphi} + e^{i(N-m-m')\varphi}e^{i\theta_{m'}} + \\ &e^{-i(N-m-m')\varphi}e^{-i\theta_m} + e^{-i(m'-m)\varphi}e^{i(\theta_{m'}-\theta_m)}\Big]\,. \qquad \text{(VII.11)}\end{aligned}$$

Obviously, $\langle\psi_{Nm}|\hat{\delta}_l|\psi_{N'm'}\rangle = 0$ when $l \notin \{N, N'\}$. For $m = m'$, the deposition rate takes on the form

$$\Delta_{Nm} \propto \binom{N}{m}\{1+\cos[(N-2m)\varphi+\theta_m]\}\,, \qquad \text{(VII.12)}$$

which, in the case of $m = 0$ and $\theta_m = 0$, coincides with Eq. (VII.8). When θ_m is suitably chosen, we see that we also have access to deposition rates $(1-\cos N\varphi)$

and $(1 \pm \sin N\varphi)$. Apart from this extra phase freedom, Eq. (VII.12) does not look like an improvement over Eq. (VII.8), since $N - 2m \leq N$, which means that the resolution decreases. However, we will show later how these states *can* be used to produce non-trivial patterns.

First, we look at a few special cases of θ_m and $\theta_{m'}$. When we write $\Delta_{Nm}^{Nm'} = \Delta_{Nm}^{Nm'}(\theta_m, \theta_{m'})$ we have

$$\Delta_{nm}^{Nm'}(0,0) \propto \cos\Big(\frac{N-2m}{2}\varphi\Big)\cos\Big(\frac{N-2m'}{2}\varphi\Big)\,, \tag{VII.13}$$

$$\Delta_{Nm}^{Nm'}(0,\pi) \propto \cos\Big(\frac{N-2m}{2}\varphi\Big)\sin\Big(\frac{N-2m'}{2}\varphi\Big)\,, \tag{VII.14}$$

$$\Delta_{Nm}^{Nm'}(\pi,0) \propto \sin\Big(\frac{N-2m}{2}\varphi\Big)\cos\Big(\frac{N-2m'}{2}\varphi\Big)\,, \tag{VII.15}$$

$$\Delta_{Nm}^{Nm'}(\pi,\pi) \propto \sin\Big(\frac{N-2m}{2}\varphi\Big)\sin\Big(\frac{N-2m'}{2}\varphi\Big)\,. \tag{VII.16}$$

These relations give the dependence of the matrix elements $\Delta_{Nm}^{Nm'}$ on θ_m and $\theta_{m'}$ in a more intuitive way than Eq. (VII.11) does. Finally, when $\theta_m = \theta_{m'} = \theta$ we obtain

$$\Delta_{Nm}^{Nm'} \propto \cos\left[\frac{(N-2m)\varphi+\theta}{2}\right]\cos\left[\frac{(N-2m')\varphi-\theta}{2}\right]\,. \tag{VII.17}$$

So far we have only considered generalised deposition rates given by Eq. (VII.9), with special values of their parameters. We will now turn our attention to the problem of creating more arbitrary patterns.

Note that there are two main, though fundamentally different, ways we can superpose the states given by Eq. (VII.9). We can superpose states with different photon numbers n and a fixed distribution m over the two modes:

$$|\Psi_m\rangle = \sum_{n=0}^{N} \alpha_n |\psi_{nm}\rangle\,. \tag{VII.18}$$

Alternatively, we can superpose states with a fixed photon number N, but with different distributions m:

$$|\Psi_N\rangle = \sum_{m=0}^{\lfloor N/2\rfloor} \alpha_m |\psi_{Nm}\rangle\,, \tag{VII.19}$$

where $\lfloor N/2\rfloor$ denotes the largest integer l with $l \leq N/2$.

These two different superpositions can be used to tailor patterns which are more complicated than just closely spaced lines. We will now study these two different methods.

a The Pseudo-Fourier Method

The first method, corresponding to the superposition given by Eq. (VII.18), we will call the pseudo-Fourier method (this choice of name will become clear shortly). When we calculate the deposition rate Δ_m according to the state $|\Psi_m\rangle$ we immediately see that branches with different photon numbers n and n' do not exhibit interference:

$$\Delta_m = \sum_{n=0}^{N} |\alpha_n|^2 \langle \psi_{nm}|\hat{\delta}_n|\psi_{nm}\rangle = \sum_{n=0}^{N} |\alpha_n|^2 \Delta_{nm} \; . \tag{VII.20}$$

Using Eq. (VII.12) the exposure pattern $P(\varphi) = \Delta_m t$ becomes

$$P(\varphi) = t \sum_{n=0}^{N} c_n \left\{1 + \cos[(n-2m)\varphi + \theta_n]\right\} \; , \tag{VII.21}$$

where t is the exposure time and the c_n are positive. Since $m < n$ and m is fixed, we have $m = 0$. I will now prove that this is a Fourier series up to a constant.

A general Fourier expansion of $p(\varphi)$ can be written as

$$P(\varphi) = \sum_{n=0}^{N} (a_n \cos n\varphi + b_n \sin n\varphi) \; . \tag{VII.22}$$

Writing Eq. (VII.21) as

$$P(\varphi) = t \sum_{n=0}^{N} c_n + t \sum_{n=0}^{N} c_n \cos(n\varphi + \theta_n) \; , \tag{VII.23}$$

where $t \sum_{n=0}^{N} c_n$ is a constant. If we ignore this constant (its contribution to the deposition rate will give a general uniform background exposure of the substrate, since it is independent of φ) we see that we need

$$c_n \cos(n\varphi + \theta_n) = a_n \cos n\varphi + b_n \sin n\varphi \tag{VII.24}$$

with c_n positive, $\theta_n \in [0, 2\pi)$ and a_n, b_n real. Expanding the left-hand side and equating terms in $\cos n\varphi$ and $\sin n\varphi$ we find

$$a_n = c_n \cos\theta_n \qquad \text{and} \qquad b_n = c_n \sin\theta_n \; . \tag{VII.25}$$

This is essentially a co-ordinate change from Cartesian to polar co-ordinates. Thus, Eq. (VII.21) is equivalent to a Fourier series up to an additive constant. Since in the limit of $N \to \infty$ a Fourier series can converge to any well-behaved pattern $P(\varphi)$, this procedure allows us to approximate arbitrary patterns in one dimension (up to a constant). It is now clear why we call this procedure the pseudo-Fourier method.

However, there is a drawback with this procedure. The deposition rate Δ is a positive definite quantity, which means that once the substrate is exposed at a particular Fourier component, there is no way this can be undone. Technically, Eq. (VII.21) can be written as

$$P(\varphi) = Q \cdot t + t \sum_{n=0}^{N} (a_n \cos n\varphi + b_n \sin n\varphi) \,, \tag{VII.26}$$

where Q is the uniform background 'penalty exposure rate' $Q = \sum_{n=0}^{N} c_n$ we mentioned earlier. The second term on the right-hand side is a true Fourier series. Thus in the pseudo-Fourier method there is always a minimum exposure of the substrate. Ultimately, this penalty can be traced to the absence of interference between the terms with different photon number in Eq. (VII.18). Next, we will investigate whether our second method of tailoring patterns can remove this penalty exposure.

b *The Superposition Method*

We will now study our second method of tailoring patterns, which we call the 'superposition method' (lacking a better name). Here we keep the total number of photons N constant, and change how the photons are distributed between the two beams in each branch [see Eq. (VII.19)]. A distinct advantage of this method is that it *does* exhibit interference between the different branches in the superposition, which eliminates the uniform background penalty exposure.

Take for instance a superposition of two distinct terms

$$|\Psi_N\rangle = \alpha_m |\psi_{Nm}\rangle + \alpha_{m'} |\psi_{Nm'}\rangle \,, \tag{VII.27}$$

with $|\alpha_m|^2 + |\alpha_{m'}|^2 = 1$ and $|\psi_{nm}\rangle$ given by Eq. (VII.11). After some algebraic manipulation the deposition rate can be written as

$$\begin{aligned} \Delta_N \;\propto\; & |\alpha_m|^2 \binom{N}{m} \{1 + \cos[(N-2m)\varphi + \theta_m]\} \\ & + |\alpha_{m'}|^2 \binom{N}{m'} \{1 + \cos[(N-2m')\varphi + \theta_{m'}]\} \\ & + 8 r_m^{m'} \sqrt{\binom{N}{m}\binom{N}{m'}} \cos\left(\frac{\theta_{m'}}{2} - \frac{\theta_m}{2} + \xi_m^{m'}\right) \\ & \times \cos\frac{1}{2}[(N-2m)\varphi + \theta_m] \cos\frac{1}{2}[(N-2m')\varphi + \theta_{m'}] \,, \end{aligned} \tag{VII.28}$$

where the deposition rate Δ is now a function of α_m and $\alpha_{m'}$, where we have chosen the real numbers $r_m^{m'}$ and $\xi_m^{m'}$ to satisfy $\alpha_m^* \alpha_{m'} \equiv r_m^{m'} \exp(i\xi_m^{m'})$. For the special values $N = 20$, $m = 9$, $m' = 5$ and $\theta_m = \theta_{m'} = 0$ we obtain the pattern

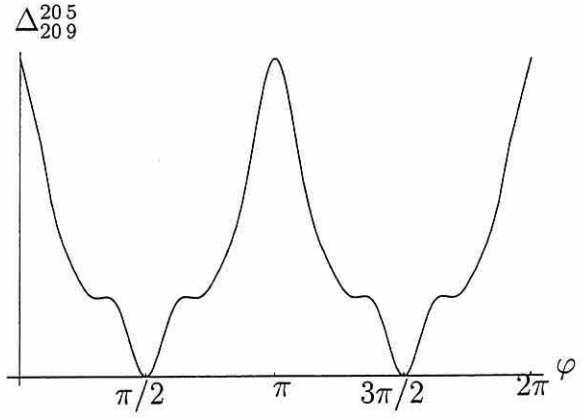

Figure VII.3: A simple superposition of two states containing 20 photons with distributions $m = 9$ and $m' = 5$ ($\theta_m = \theta_{m'} = 0$). The deposition rate at $\varphi = \pi/2$ and $\varphi = 3\pi/2$ is zero, which means that there is no general uniform background exposure using the superposition method.

shown in figure VII.3. Clearly, there is no uniform background penalty exposure here.

For more than two branches in the superposition this becomes a complicated function, which is not nearly as well understood as a Fourier series. The general expression for the deposition rate can be written as

$$\Delta_N \propto \sum_{m=0}^{\lfloor N/2 \rfloor} \sum_{m'=0}^{\lfloor N/2 \rfloor} r_m^{m'} \sqrt{\binom{N}{m}\binom{N}{m'}} \cos\left(\frac{\theta_{m'}}{2} - \frac{\theta_m}{2} + \xi_m^{m'}\right) \times \cos\frac{1}{2}\left[(N-2m)\varphi + \theta_m\right] \cos\frac{1}{2}\left[(N-2m')\varphi + \theta_{m'}\right] \quad \text{(VII.29)}$$

where we have chosen $r_m^{m'}$ and $\xi_m^{m'}$ real to satisfy $\alpha_m^* \alpha_{m'} \equiv r_m^{m'} \exp(i\xi_m^{m'})$. Note that $\xi_m^m = 0$.

If we want to tailor a pattern $F(\varphi)$, it might be the case that this type of superposition will also converge to the required pattern. We will now compare the superposition method with the pseudo-Fourier method.

c *Comparing the two methods*

So far, we discussed two methods of creating non-trivial patterns in one dimension. The pseudo-Fourier method is simple but yields a uniform background penalty exposure. The superposition method is far more complicated, but seems to get around the background exposure. Before we make a comparison between the two methods we will discuss the creation of 'arbitrary' patterns.

It is well known that any sufficiently well-behaved periodic function can be written as an infinite Fourier series (we ignore such subtleties which arise when

two functions differ only at a finite number of points, etc.). However, when we create patterns with the pseudo-Fourier lithography method we do not have access to every component of the Fourier expansion, since this would involve an infinite number of photons ($n \to \infty$). This means that we can only employ truncated Fourier series, and these can merely approximate arbitrary patterns.

The Fourier expansion has the nice property that when a series is truncated at N, the remaining terms still give the best Fourier expansion of the function up to N. In other words, the coefficients of a truncated Fourier series are equal to the first N coefficients of a full Fourier series. If the full Fourier series is denoted by F and the truncated series by F_N, we can define the normed-distance quantity D_N:

$$D_N \equiv \int_0^{2\pi} |F(\varphi) - F_N(\varphi)|^2 d\varphi \,, \tag{VII.30}$$

which can be interpreted as a distance between F and F_N. If quantum lithography yields a pattern $P_N(\varphi) = \Delta_N t$, we can introduce the following definition: quantum lithography can approximate arbitrary patterns if

$$\int_0^{2\pi} |F(\varphi) - P_N(\varphi)|^2 d\varphi \leq \varepsilon D_N \,, \tag{VII.31}$$

with ε some proportionality constant. This definition gives the concept of approximating patterns a solid basis.

We compare the pseudo-Fourier and the superposition method for one special case. We choose the test function

$$F(\varphi) = \begin{cases} h \text{ if } -\frac{\pi}{2} < \varphi < \frac{\pi}{2} \,, \\ 0 \text{ otherwise} \,. \end{cases} \tag{VII.32}$$

With up to ten photons, we ask how well the pseudo-Fourier and the superposition method approximate this pattern.

In the case of the pseudo-Fourier method the solution is immediate. The Fourier expansion of the 'trench' function given by Eq. (VII.32) is well known:

$$F(\varphi) = \sum_{q=0}^{\infty} \frac{(-1)^q}{2q+1} \cos[(2q+1)\varphi] \,. \tag{VII.33}$$

Using up to $n = 10$ photons we include terms up to $q = 4$, since $2q+1 \leq 10$. The Fourier method thus yields a pattern $P(\varphi)$ (the two patterns $P(\varphi)$ and $F(\varphi)$ are generally not the same) which can be written as

$$P(\varphi) = \sum_{q=0}^{4} \frac{c_q t}{2q+1} \left(1 + \cos\left[(2q+1)\varphi + \pi\kappa_q\right]\right) \,, \tag{VII.34}$$

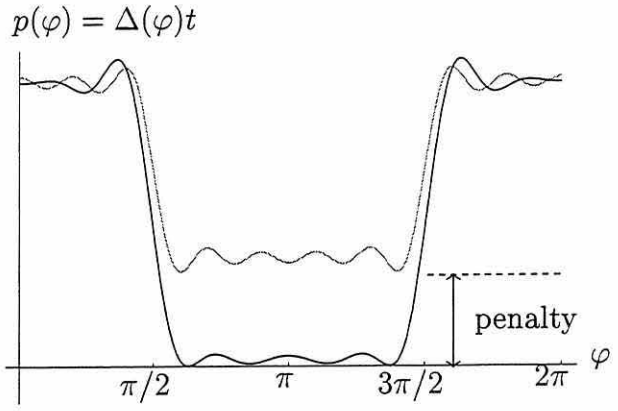

Figure VII.4: The deposition rate on the substrate resulting from a superposition of states with $n = 10$ and different m (black curve) and resulting from a superposition of states with different n and $m = 0$ (grey curve). The coefficients of the superposition yielding the black curve are optimised using a genetic algorithm [137], while the grey curve is a truncated Fourier series. Notice the 'penalty' (displaced from zero) deposition rate of the Fourier series between $\pi/2$ and $3\pi/2$.

where c_q is a constant depending on the proportionality constant of Δ_{2q+1}, the rate of production of $|\psi_{nm}\rangle$ and the coupling between the light field and the substrate. The term κ_q is defined to accommodate for the minus signs in Eq. (VII.33): it is zero when q is even and one when q is odd. Note the uniform background penalty exposure rate $\sum_{q=0}^{4} c_q/(2q+1)$. The result of this method is shown in figure VII.3.

Alternatively, the superposition method employs a state

$$|\Psi_N\rangle = \sum_{m=0}^{\lfloor N/2 \rfloor} \alpha_m |\psi_{Nm}\rangle \; . \tag{VII.35}$$

The procedure of finding the best fit with the test function is more complicated. We have to minimise the absolute difference between the deposition rate $\Delta_N(\vec{\alpha})$ times the exposure time t and the test function $F(\varphi)$. We have chosen $\vec{\alpha} = (\alpha_0, \ldots, \alpha_{n/2})$. Mathematically, we have to evaluate the $\vec{\alpha}$ and t which minimise d_N:

$$d_N = \int_0^{2\pi} |F(\varphi) - \Delta_N(\vec{\alpha})t|^2 d\varphi \; ,$$

with

$$\Delta_N(\vec{\alpha}) = \langle \Psi_N | \hat{\delta}_N | \Psi_N \rangle \; . \tag{VII.36}$$

We have to fit both t and $\vec{\alpha}$. Using a genetic optimalisation algorithm [137] (with $h = 1$, a normalised height of the test function; see appendix I) we found

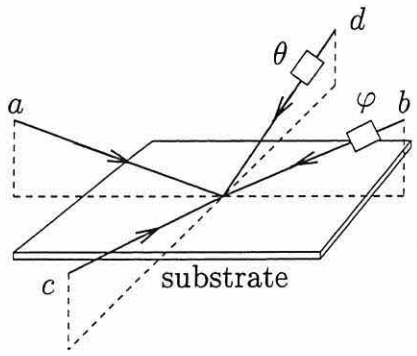

Figure VII.5: Four light beams a, b, c and d cross each other at the surface of a photosensitive substrate. The angles between a and b and c and d are again taken in the grazing limit of π. The relative phase difference between a and b is φ and the relative phase difference between c and d is θ.

that the deposition rate is actually very close to zero in the interval $\pi/2 \leq \varphi \leq 3\pi/2$, unlike the pseudo-Fourier method, where we have to pay a uniform background penalty. This result implies that in this case a superposition of different photon distributions m, given a fixed total number of photons N, works better than a superposition of different photon number states (see figure VII.3). In particular, *the fixed photon number method allows for the substrate to remain virtually unexposed in certain areas.*

We stress that this is merely a comparison for a specific example, namely that of the trench target function $F(\varphi)$. We conjecture that the superposition method can approximate other arbitrary patterns equally well, but we have not yet found a proof. Besides the ability to fit an arbitrary pattern, another criterion of comparison between the pseudo-Fourier method and the superposition method, is the time needed to create the N-photon entangled states.

Until now, we have only considered sub-wavelength resolution in one direction, namely parallel to the direction of the beams. However, for practical applications we would like sub-wavelength resolution in both directions on the substrate. This is the subject of the next section.

4 General Patterns in 2D

In this section we study how to create two-dimensional patterns on a suitable substrate using the quantum lithography techniques developed in the previous sections. As we have seen, the phase shift φ, in the setup given by figure VII.1, acts as a parametrisation for the deposition rate in one dimension. Let's call this the x-direction.

We can now do the same for the y-direction, employing two counter-propagating beams (c and d) in the y-direction (see figure VII.4). The same conditions apply:

we consider the limit where the spatial angle θ off axis approaches $\pi/2$, thus grazing along the substrate's surface.

Consider the region where the four beams a, b, c and d overlap. For real lithography we have to take into account the mode shapes, but when we confine ourselves to an area with side lengths λ (where λ is the wavelength of the used light) this problem does not arise.

The class of states on modes a to d that we consider here are of the form

$$|\psi^k_{Nm}\rangle = \frac{1}{2}\Big[e^{im\varphi}|N-m,m;0,0\rangle + e^{i(N-m)\varphi}e^{i\zeta_m}|m,N-m;0,0\rangle + e^{ik\theta}|0,0;N-k,k\rangle + e^{i(N-k)\chi}e^{i\bar{\zeta}_k}|0,0;k,N-k\rangle\Big], \quad \text{(VII.37)}$$

where ζ_m and $\bar{\zeta}_k$ are two relative phases. This is by no means the only class of states, but we will restrict our discussion to this one for now. Observe that this is a superposition on the amplitude level, which allows destructive interference in the deposition rate in order to create dark spots on the substrate. Alternatively, we could have used the one-dimensional method [with states given by Eq. (VII.9)] in the x- and y-direction, but this cannot give interference effects between the modes a, b and c, d.

The phase-shifts φ and χ in the light beams b and d (see figure VII.4) result in respective displacements x and y of the interference pattern on the substrate. A phase-shift of 2π in a given direction will displace the pattern, say, N times. This means that the maxima are closer together, yielding an effective resolution equal to $\Delta x = \Delta y = \lambda/4N$. This happens in both the x- and the y-direction.

We proceed again as in section 2 by evaluating the N^{th} order moment $\hat{\delta}_N$ of the electric field operator [see Eq. (VII.7)]. On a substrate sensitive to N photons this gives the deposition rate $\Delta^{Nm'k'}_{Nmk} = \langle\psi^k_{Nm}|\delta_N|\psi^{k'}_{Nm'}\rangle$ [with $|\psi^k_{Nm}\rangle$ given by Eq. (VII.37)]:

$$\begin{aligned}
\Delta^{Nm'k'}_{Nmk} \propto\ & \binom{N}{m}\binom{N}{m'}\Big(e^{-im\varphi}e^{im'\varphi} + e^{-im\varphi}e^{i(N-m')\varphi}e^{i\zeta_{m'}} + e^{-i(N-m)\varphi}e^{im'\varphi}e^{-i\zeta_m} \\
& + e^{-i(N-m)\varphi}e^{i(N-m')\varphi}e^{-i(\zeta_m-\zeta_{m'})}\Big) \\
& + \binom{N}{m}\binom{N}{k'}\Big(e^{-im\varphi}e^{ik'\chi} + e^{-im\varphi}e^{i(N-k')\chi}e^{i\bar{\zeta}_{k'}} + e^{-i(N-m)\varphi}e^{ik'\chi}e^{-i\zeta_m} \\
& + e^{-i(N-m)\varphi}e^{i(N-k')\chi}e^{-i(\zeta_m-\bar{\zeta}_{k'})}\Big) \\
& + \binom{N}{k}\binom{N}{m'}\Big(e^{-ik\chi}e^{im'\varphi} + e^{-ik\chi}e^{i(N-m')\varphi}e^{i\zeta_{m'}} + e^{-i(N-k)\chi}e^{im'\varphi}e^{-i\bar{\zeta}_k} \\
& + e^{-i(N-k)\chi}e^{i(N-m')\varphi}e^{-i(\bar{\zeta}_k-\zeta_{m'})}\Big) \\
& + \binom{N}{k}\binom{N}{k'}\Big(e^{-ik\chi}e^{ik'\chi} + e^{-ik\chi}e^{i(N-k')\chi}e^{i\bar{\zeta}_{k'}} + e^{-i(N-k)\chi}e^{ik'\chi}e^{-i\bar{\zeta}_k} \\
& + e^{-i(N-k)\chi}e^{i(N-k')\chi}e^{-i(\bar{\zeta}_k-\bar{\zeta}_{k'})}\Big)\,. \qquad \text{(VII.38)}
\end{aligned}$$

For the special choice of $m' = m$ and $k' = k$ we have

$$\begin{aligned}\Delta^k_{Nm} \propto & \binom{N}{m}^2 (1+\cos[(N-2m)\varphi+\zeta_m]) + \binom{N}{k}^2 (1+\cos[(N-2k)\chi+\bar{\zeta}_k]) \\ & +4\binom{N}{m}\binom{N}{k}\cos\frac{1}{2}\left[N(\varphi-\chi)+(\zeta_m-\bar{\zeta}_k)\right] \\ & \times\cos\frac{1}{2}\left[(N-2m)\varphi-\zeta_m\right]\cos\frac{1}{2}\left[(N-2k)\chi-\bar{\zeta}_k\right] . \qquad \text{(VII.39)}\end{aligned}$$

We can again generalise this method and use superpositions of the states given in Eq. (VII.37). Note that there are now three numbers N, m and k which can be varied. Furthermore, as we have seen in the one-dimensional case, superpositions of different n do not give interference terms in the deposition rate.

Suppose we want to approximate a pattern $F(\varphi,\chi)$, with $\{\varphi,\chi\} \in [0,2\pi)$. This pattern can always be written in a Fourier expansion:

$$\begin{aligned}F(\varphi,\chi) = & \sum_{p,q=0}^{\infty} a_{pq}\cos p\varphi\cos q\chi + b_{pq}\cos p\varphi\sin q\chi\times \\ & c_{pq}\sin p\varphi\cos q\chi + d_{pq}\sin p\varphi\sin q\chi . \qquad \text{(VII.40)}\end{aligned}$$

with a_p, b_p, c_q and d_q real. In the previous section we showed that quantum lithography could approximate the Fourier series of a one-dimensional pattern up to a constant displacement. This relied on absence of interference between the terms with different photon numbers. The question is now whether we can do the same for patterns in *two* dimensions. Or alternatively, can general superpositions of the state $|\psi^k_{Nm}\rangle$ approximate the pattern $F(\varphi,\chi)$?

From Eq. (VII.38) it is not obvious that we can obtain the four trigonometric terms given by the Fourier expansion of Eq. (VII.40):

$$\Delta \propto \cos p\varphi\,\cos q\chi , \qquad \text{(VII.41)}$$

$$\Delta \propto \cos p\varphi\,\sin q\chi , \qquad \text{(VII.42)}$$

$$\Delta \propto \sin p\varphi\,\cos q\chi , \qquad \text{(VII.43)}$$

$$\Delta \propto \sin p\varphi\,\sin q\chi . \qquad \text{(VII.44)}$$

We can therefore not claim that two-dimensional quantum lithography can approximate arbitrary patterns in the sense of one-dimensional lithography. Only simple patterns like the one given in figure VII.5 can be inferred from Eq. (VII.38). In order to find the best fit to an arbitrary pattern one has to use a minimisation procedure.

For example, we calculate the total deposition rate due to the quantum state $|\Psi_N\rangle$, where

$$|\Psi_N\rangle = \sum_{m=0}^{\lfloor N/2\rfloor}\sum_{k=0}^{\lfloor N/2\rfloor} \alpha_{mk}|\psi^k_{Nm}\rangle . \qquad \text{(VII.45)}$$

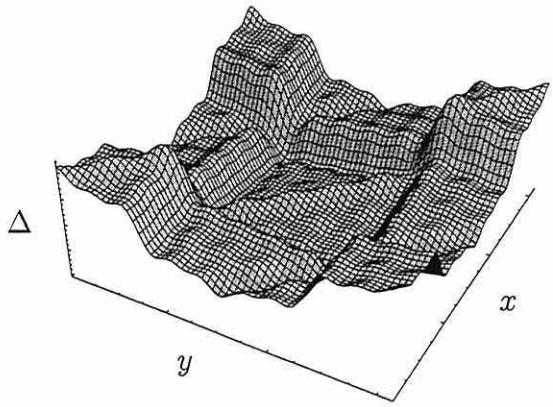

Figure VII.6: A simulation of a two-dimensional intensity pattern on an area λ^2, where λ denotes the wavelength of the used light. Here, I modelled a square area with sharp edges. The pattern was generated by a Fourier series of up to ten photons (see also figure VII.4 for the one-dimensional case).

Here, α_{mk} are complex coefficients. We now proceed by choosing a particular intensity pattern $F(\varphi, \chi)$ and optimising the coefficients α_{mk} for a chosen number of photons. The deposition rate due to the state $|\Psi_N\rangle$ is now

$$\Delta_N(\vec{\alpha}) = \sum_{m,m'=0}^{\lfloor N/2 \rfloor} \sum_{k,k'=0}^{\lfloor N/2 \rfloor} \alpha^*_{mk} \alpha_{m'k'} \Delta^{Nm'k'}_{Nmk} \,, \tag{VII.46}$$

with $\vec{\alpha} = (\alpha_{0,0}, \alpha_{0,1} \ldots, \alpha_{N/2,N/2})$. We again have to evaluate the $\vec{\alpha}$ and t which minimise

$$\int_0^{2\pi} \int_0^{2\pi} |F(\varphi, \chi) - \Delta_n(\vec{\alpha}) t|^2 \, d\varphi \, d\chi \,. \tag{VII.47}$$

The values of $\vec{\alpha}$ and t can again be found using a genetic algorithm.

5 Physical implementation

With current experimental capabilities, the physical implementation of quantum lithography is very challenging. In particular, there are two major issues to be dealt with before quantum lithography can become a mature technology. First of all, we not only need the ability to create the entangled photon states given by Eqs. (VII.9) and (VII.37), but we should also be able to create coherent

superpositions of these states. One possibility might be to use optical components like parametric down-converters. Contrary to the results of Ref. [100], we are not concerned with the usually large vacuum contribution of these processes, since the vacuum will not contribute to the spatial profile of the deposition [see Eqs. (VII.6) and (VII.7)].

Secondly, we need substrates which are sensitive to the higher moments of the electric field operator. When we want to use the pseudo-Fourier method, up to N photons for quantum lithography in one dimension, the substrate needs to be reasonably sensitive to all the higher moments up to N, the maximum photon number. Alternatively, we can use the superposition method for N photons when the substrate is sensitive to predominantly one higher moment corresponding to N photons. Generally, the method of lithography determines the requirements of the substrate.

There are also some considerations about the approximation of patterns. For example, we might not *need* arbitrary patterns. It might be the case that it is sufficient to have a set of patterns which can then be used to generate any desired circuit. This is analogous to having a universal set of logical gates, permitting any conceivable logical expression. In that case we only need to determine this elementary (universal) set of patterns.

Furthermore, we have to study whether the uniform background penalty exposure really presents a practical problem. One might argue that a sufficient difference between the maximum deposition rate and the uniform background penalty exposure is enough to accommodate lithography. This depends on the details of the substrate's reaction to the electro-magnetic field.

Before quantum lithography can be physically implemented and used in the production of nano circuits, these issues have to be addressed satisfactorily.

6 SUMMARY

In this chapter I have generalised the theory of quantum lithography as first outlined in Ref. [22]. In particular, I have shown how we can create arbitrary patterns in one dimension, albeit with a uniform background penalty exposure. We can also create some patterns in two dimensions, but we have no proof that this method can be extended to give arbitrary patterns.

For lithography in one dimension we distinguish two methods: the pseudo-Fourier method' and the superposition method. The pseudo-Fourier method is conceptually easier since it depends on Fourier analysis, but it also involves a finite amount of unwanted exposure of the substrate. More specifically, the deposition rate equals the pattern in its Fourier basis plus a term yielding unwanted background exposure. The superposition method gets around this problem and seems to give better results, but lacks the intuitive clarity of the Fourier method. Furthermore, we do not have a proof that this method can approximate arbitrary patterns.

Quantum lithography in two dimensions is more involved. Starting with a superposition of states, given by Eq. (VII.37), we found that we can indeed create two-dimensional patterns with sub-wavelength resolution, but we do not have a proof that we can create *arbitrary* patterns. Nevertheless, we might be able to create a certain set of elementary basis patterns.

There are several issues to be addressed in the future. First, we need to study the specific restrictions on the substrate and how we can physically realize them. Secondly, we need to create the various entangled states involved in the quantum lithography protocol.

Finally, G.S. Agarwal and R. Boyd have called to our attention that quantum lithography works also if the weak parametric down-converter source, described in Ref. [22] is replaced by a high-flux optical parametric amplifier [2]. The visibility saturates at 20% in the limit of large gain, but this is quite sufficient for some lithography purposes, as well as for 3D optical holography used for data storage.

Appendices

A

Complex vector spaces

In this appendix I review some properties of complex vector spaces, since quantum mechanics is defined in terms of a complex vector space.

1 Vector spaces

A vector space $\mathcal{V}$ consists of a set of vectors $\{\mathbf{v}_i\}$ on which two operations are defined:

Addition: for every $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ the vector $\mathbf{x} + \mathbf{y}$ is an element of $\mathcal{V}$.

Scalar multiplication: for every $\alpha \in \mathbb{C}$ and $\mathbf{x} \in \mathcal{V}$ there is a unique element $\alpha\mathbf{x}$ in $\mathcal{V}$.

Furthermore, for every (complex) vector space the following conditions hold:

1. For all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (commutativity of addition);
2. for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$, $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ (associativity of addition);
3. there exists an element 0 in $\mathcal{V}$ such that $\mathbf{x} + 0 = \mathbf{x}$ for every $\mathbf{x} \in \mathcal{V}$;
4. for each element $\mathbf{x}$ in $\mathcal{V}$ there exists an element $\mathbf{y}$ in $\mathcal{V}$ such that $\mathbf{x} + \mathbf{y} = 0$;
5. for each element $\mathbf{x}$ in $\mathcal{V}$, $1\mathbf{x} = \mathbf{x}$;
6. for each pair α and β in $\mathbb{C}$ and each $\mathbf{x} \in \mathcal{V}$ we have $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$;
7. for each $\alpha \in \mathbb{C}$ and each pair $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ we have $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$;
8. for each pair α and β in $\mathbb{C}$ and each $\mathbf{x} \in \mathcal{V}$ we have $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$.

On a vector space we can also define an inner product (sometimes called the scalar product, not to be confused with scalar multiplication). When two vectors in $\mathcal{V}$ are denoted by $\mathbf{x}$ and $\mathbf{y}$, their inner product is a (complex) number written as $(\mathbf{x}, \mathbf{y})$. For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$ and $\alpha \in \mathbb{C}$ the inner product obeys the following rules

1. $(\mathbf{x} + \mathbf{z}, \mathbf{y}) = (\mathbf{x}, \mathbf{y}) + (\mathbf{z}, \mathbf{y})$;

2. $(\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y})$;

3. $(\mathbf{x}, \mathbf{y})^* = (\mathbf{y}, \mathbf{x})$, where * denotes complex conjugation;

4. $(\mathbf{x}, \mathbf{x}) > 0$ if $\mathbf{x} \neq 0$.

A complex vector space with an inner product is called a *Hilbert space*. Note that we only defined algebraic rules for the inner product, the actual form of $(\mathbf{x}, \mathbf{y})$ depends on the representation.

In Dirac's bracket notation, the elements of the Hilbert space $\mathscr{H}$ are written as so-called *kets*: $|\psi\rangle, |\phi\rangle \in \mathscr{H}$. The adjoints of these kets are called *bras*: $\langle\psi|, \langle\phi|$. The inner product is given by $\langle\psi|\phi\rangle$. It necessarily obeys all the conditions given above. Two vectors $|\psi\rangle$ and $|\phi\rangle$ in Hilbert space are *orthogonal* if and only if their inner product vanishes: $\langle\psi|\phi\rangle = 0$.

In a Hilbert space $\mathscr{H}$ of dimension d we can construct a set of d orthogonal vectors, called an orthogonal basis of $\mathscr{H}$. When the vectors in the basis have unit length, i.e., if for every basis vector $|\psi_i\rangle$ we have $\langle\psi_i|\psi_i\rangle = 1$, then the basis is *orthonormal*.

Next, we define *linear operators* on $\mathscr{H}$. Consider a transformation $A : \mathscr{H} \to \mathscr{H}$. A is called a linear operator on $\mathscr{H}$ if for every $|\psi\rangle, |\phi\rangle \in \mathscr{H}$ and $\alpha \in \mathbb{C}$

1. $A(|\psi\rangle + |\phi\rangle) = A|\psi\rangle + A|\phi\rangle$;

2. $A(\alpha|\psi\rangle) = \alpha(A|\psi\rangle)$.

A linear operator transforms one vector in Hilbert space to another: $A|\psi\rangle = |\psi'\rangle$. In matrix notation A corresponds to a matrix, while kets correspond to column vectors and bras to row vectors.

Linear operators do not necessarily *commute*. That is, when we have two linear operators A and B on $\mathscr{H}$, their commutation relation $[A, B] = AB - BA$ is not necessarily zero. Non-zero commutation relations play an important rôle in quantum mechanics. For example, non-commuting operators lie at the heart of quantum cryptography.

Suppose a linear operator A obeys the following relation:

$$A|\psi\rangle = \alpha|\psi\rangle \ , \tag{A.1}$$

where $|\psi\rangle$ is a vector in $\mathscr{H}$ and α a complex number. This is called an eigenvalue equation for A, where α is the *eigenvalue* and $|\psi\rangle$ the corresponding *eigenvector*. When the dimension of $\mathscr{H}$ is d, every linear operator on $\mathscr{H}$ has d eigenvalue equations. An eigenvalue α might be d_α-fold degenerate, in which case α generates a d_α-dimensional *eigenspace*: there are d_α orthogonal vectors $|\psi_\alpha\rangle$ which obey Eq. (A.1), thus forming a basis for a d_α-dimensional subspace of $\mathscr{H}$.

There also exists a property called the *trace* of an operator A. We write

$$\mathrm{Tr}A = \sum_i \langle\psi_i|A|\psi_i\rangle \ , \tag{A.2}$$

where $\{|\psi_i\rangle\}$ can be *any* complete orthonormal basis. The trace has the following properties:

1. if $A^\dagger = A$ then $\mathrm{Tr}A$ is real;
2. $\mathrm{Tr}(\alpha A) = \alpha \mathrm{Tr}A$;
3. $\mathrm{Tr}(A + B) = \mathrm{Tr}A + \mathrm{Tr}B$;
4. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, the cyclic property.

These properties are easily proved using the knowledge that the trace as defined in Eq. (A.2) is independent of the basis $\{|\psi_i\rangle\}$.

2 TENSOR PRODUCT SPACES

A tensor product of two operators A and B is defined as follows:

$$(A \otimes B)(|\psi_i\rangle_1 \otimes |\phi_j\rangle_2) = (A|\psi_i\rangle_1) \otimes (B|\phi_j\rangle_2) \,, \tag{A.3}$$

which is equivalent to $(A \otimes B)(C \otimes D) = (AC \otimes BD)$. In other words, every operator sticks to its own Hilbert space. It should be noted, however, that *not every operator on $\mathscr{H}_1 \otimes \mathscr{H}_2$ is of the form $A \otimes B$.* The fact that this is not the case is also of fundamental importance to quantum information theory, as we shall see in the remainder of this thesis (see also appendix B).

Other properties of the tensor product of operators are [85]

1. $A \otimes 0 = 0 \otimes B = 0$,
2. $\mathbb{1} \otimes \mathbb{1} = \mathbb{1}$,
3. $(A_1 + A_2) \otimes B = (A_1 \otimes B) + (A_2 \otimes B)$,
4. $\alpha A \otimes \beta B = \alpha\beta(A \otimes B)$,
5. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$,
6. $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$,
7. $\mathrm{Tr}(A \otimes B) = \mathrm{Tr}A \cdot \mathrm{Tr}B$.

For notational brevity the tensor product symbol $\otimes$ is often omitted, yielding, e.g., $|\psi\rangle \otimes |\phi\rangle = |\psi\rangle|\phi\rangle = |\psi, \phi\rangle$. When this abbreviated notation is used, one should always remember which state or operator is defined on which Hilbert space.

3 Projection operators

So far, we have only considered tensor products of Hilbert spaces. However, there is also an operation '$\oplus$', called the *direct sum* of two vector spaces. The direct sum of two vector spaces $\mathcal{W} = \mathcal{V}_1 \oplus \mathcal{V}_2$ is again a vector space, and $\mathcal{V}_1$ and $\mathcal{V}_2$ are called its *subspaces*. We can write a vector in $\mathcal{W}$ as $\mathbf{w} = (\mathbf{v}_1, \mathbf{v}_2)$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are vectors in the respective subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$. The subspaces are linear if

$$\alpha(\mathbf{u}_1, \mathbf{u}_2) + \beta(\mathbf{v}_1, \mathbf{v}_2) = (\alpha\mathbf{u}_1 + \beta\mathbf{v}_1, \alpha\mathbf{u}_2 + \beta\mathbf{v}_2) \,. \tag{A.4}$$

Here, I will only consider linear subspaces.

Suppose we have a Hilbert space which can be written as the direct sum of two subspaces. These are again Hilbert spaces:

$$\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2 \,. \tag{A.5}$$

A state $|\psi\rangle$ in $\mathscr{H}$ can then be written as

$$|\psi\rangle = \alpha|\psi_1\rangle + \beta|\psi_2\rangle \,, \tag{A.6}$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ are restricted to their respective subspaces and $|\alpha|^2 + |\beta|^2 = 1$. We can define an operator P_1 which yields

$$P_1|\psi\rangle = P_1\left(\alpha|\psi_1\rangle + \beta|\psi_2\rangle\right) = \alpha|\psi_1\rangle \,. \tag{A.7}$$

In other words, P_1 projects the state $|\psi\rangle$ onto the linear subspace $\mathscr{H}_1$. P_1 is said to be a *projection operator* or *projector* [85]. An operator is a projection operator if and only if

$$P^2 = P = P^\dagger \,. \tag{A.8}$$

Projection operators have the following properties:

1. two projection operators P and Q are called *orthogonal* projections if and only if $[P, Q] = 0$, they project onto linearly independent subspaces;
2. the sum of two (orthogonal) projectors is again a projector;
3. the sum over all orthogonal projectors in $\mathscr{H}$ is the identity operator $\mathbb{1}$;
4. the orthocomplement of a projector P in $\mathscr{H}$ is given by $\mathbb{1} - P$;
5. the eigenvalues of a projector are 1 and 0.

When viewed as measurement outcomes (see postulate 3), the eigenvalues of a projection operator indicate whether the state is in the subspace spanned by P or not.

It is easily checked that for any state $|\psi\rangle$ in $\mathscr{H}$ a projector on the subspace spanned by $|\psi\rangle$ can be written as

$$P_{|\psi\rangle} = |\psi\rangle\langle\psi| \,. \tag{A.9}$$

When $\{|\psi_i\rangle\}$ is a complete orthonormal basis of $\mathscr{H}$, the identity operator can then be written as

$$\mathbb{1} = \sum_i P_{|\psi_i\rangle} = \sum_i |\psi_i\rangle\langle\psi_i| \,. \tag{A.10}$$

This is called the *completeness relation.* An operator A with eigenvalues α_i whose eigenvectors are given by the basis $\{|\psi_i\rangle\}$ can then be written as

$$A = \sum_i \alpha_i |\psi_i\rangle\langle\psi_i| \,. \tag{A.11}$$

This is sometimes called the *spectral decomposition* of A. In general, when the eigenvectors of an operator A are *not* given by this basis, A can be written as

$$A = \sum_{ij} \alpha_{ij} |\psi_i\rangle\langle\psi_j| \,. \tag{A.12}$$

When A is Hermitian ($A^\dagger = A$), we have $\alpha_{ij} = \alpha^*_{ji}$.

B

States, Operators and Maps

In this appendix I summarise some background knowledge about states, operators and maps in the context of quantum mechanics. This knowledge is important for the understanding of the Peres-Horodecki partial transpose criterion for the separability of bi-partite density matrices, and it also lays the foundations for the definition of positive operator valued measures. For this appendix I am indebted to professor Rajiah Simon, who guided me through Hilbert space.

1 Single systems

Suppose we have a physical system which is described by a set of accessible states $\{|\phi_j\rangle\}$. The superposition principle and the linearity of quantum mechanics imply that this set spans a Hilbert space $\mathscr{H}$ of dimension d. This is a complex vector space with an orthonormal basis

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

We can define a set of linear operators $\{A_k\}$ on $\mathscr{H}$, the elements of which transform one state to another:

$$A : |\phi\rangle \longrightarrow |\phi'\rangle\,, \tag{B.1}$$

where $|\phi'\rangle$ is again a state in $\mathscr{H}$. Hermiticity of A ($A^\dagger = A$) implies that all the eigenvalues of A are real. When for all $|\phi\rangle$ we have $\langle\phi|A|\phi\rangle \geq 0$, A is *non-negative*. A non-negative operator with trace 1 ($\mathrm{Tr}[A] = 1$) is called a *density operator*, usually denoted by ρ. Note that all non-negative operators are also Hermitian, and have non-negative eigenvalues. A choice of a basis in $\mathscr{H}$ puts Hermitian non-negative operators into a one-to-one correspondence with Hermitian non-negative matrices.

The set of all linear operators $\{A_k\}$ in turn define a Hilbert space $\mathscr{H} \otimes \mathscr{H}$ of

dimension d^2, one orthonormal basis of which can be written as

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \tag{B.2}$$

We can now define an even higher set of objects called *maps*, denoted by $\{\mathcal{L}\}$, the elements of which linearly transform the set of linear operators into itself:

$$\mathcal{L} : A \longrightarrow A' = \mathcal{L}(A) . \tag{B.3}$$

These linear maps are sometimes called *super*-operators. They are operators on the Hilbert space $\mathscr{H} \otimes \mathscr{H}$, but we give them a different name to avoid confusion. A simple example of a map corresponds to a unitary transformation $U : A \rightarrow A' = U^\dagger A U$. The corresponding map $\mathcal{L}_U$ may be written as $\mathcal{L}_U = U^\dagger \otimes U$. The set of all maps thus constitutes a Hilbert space $\mathscr{H}^{\otimes 4}$ of dimension d^4. The concept of non-negative operators lead us to define *positive* maps.

Definition: a map $\mathcal{L}$ is called *positive* if for every non-negative operator A the operator $A' = \mathcal{L}(A)$ is again a non-negative operator.

When $\mathrm{Tr}[A'] = \mathrm{Tr}[A]$ for every A, the map $\mathcal{L}$ is a *trace-preserving* map. Trace-preserving positive maps are important in quantum mechanics, since they transform the set of density operators to itself. This property may lead one to expect that these maps correspond to physical processes or symmetries. It is an interesting aspect of quantum mechanics that *not all* positive maps can be associated with physical processes. This subtle fact becomes important when we consider composite systems.

2 Composite systems

Suppose we have *two* systems 1 and 2 with respective accessible states $\{|\phi_j^{(1)}\rangle\}$ and $\{|\psi_k^{(2)}\rangle\}$. These states span two Hilbert spaces $\mathscr{H}^{(1)}$ and $\mathscr{H}^{(2)}$ with dimensions d_1 and d_2 respectively. The accessible states of the composite system can be written on the basis of the tensor product of the states $|\phi_j^{(1)}\rangle \otimes |\psi_k^{(2)}\rangle$, generating a Hilbert space $\mathscr{H}^{(1)} \otimes \mathscr{H}^{(2)}$ of dimension $d_1 \times d_2$. Similarly, the set of linear operators $\{A_j\}$ on $\mathscr{H}^{(1)} \otimes \mathscr{H}^{(2)}$ generates a Hilbert space of dimension $(d_1 \times d_2)^2$, and the set of maps generates a Hilbert space of dimension $(d_1 \times d_2)^4$.

Consider a map $\mathcal{L}_1$, defined for subsystem 1. When this map is positive (and trace-preserving) it transforms density operators of the subsystem to density operators. When system 1 is part of a composite system $1 + 2$, we want to know when a positive map of system 1 (leaving system 2 unchanged) would transform

a density operator defined on the composite system again into a density operator. In other words, we ask when the *extended* map $\mathcal{L}_{12} = \mathcal{L}_1 \otimes \mathbb{1}_2$, with $\mathbb{1}_2$ the identity map of system 2, is again positive.

Definition: a map $\mathcal{L}_1$ is called *completely positive* if all its extensions are positive.

There exist maps which are positive, but not completely positive. One such map is the transpose. Take, for example, the singlet state of a two-level bi-partite system (written in the computational basis):

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|0,1\rangle - |1,0\rangle \right) . \tag{B.4}$$

The density operator of this state can be written as

$$\rho = |\Psi\rangle\langle\Psi| = \frac{1}{2} \left(|0,1\rangle\langle 0,1| - |0,1\rangle\langle 1,0| - |1,0\rangle\langle 0,1| + |1,0\rangle\langle 1,0| \right) . \tag{B.5}$$

The transpose of a general density operator for a single system in this notation is given by

$$\begin{aligned} T : a|0\rangle\langle 0| + b|0\rangle\langle 1| + c|1\rangle\langle 0| + d|1\rangle\langle 1| \\ \longrightarrow \quad a|0\rangle\langle 0| + b|1\rangle\langle 0| + c|0\rangle\langle 1| + d|1\rangle\langle 1| , \end{aligned} \tag{B.6}$$

that is, we *exchange* the entries of the bras and kets. This is a positive map. The extended transpose (or *partial* transpose) on a compound system $PT = T_1 \otimes \mathbb{1}_2$, however, is *not* positive. To see this, apply the extended transpose to the density operator given in Eq. (B.5), we obtain

$$PT : \rho \longrightarrow \rho' = \frac{1}{2} \left(|0,1\rangle\langle 0,1| - |0,0\rangle\langle 1,1| - |1,1\rangle\langle 0,0| + |1,0\rangle\langle 1,0| \right) . \tag{B.7}$$

If the eigenvalues of ρ' are non-negative, ρ' is again a density operator. In order to find the eigenvalues of this operator we write ρ' in matrix representation on the computational basis:

$$\rho' = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} . \tag{B.8}$$

It is easily found that this matrix has eigenvalues 1 (with multiplicity 3) and -1. Therefore, ρ' is not a density operator, and T, although positive, is not a *completely* positive map.

I will now present an important class of completely positive maps. Consider the general map

$$\mathcal{L} : A \longrightarrow A' = \mathcal{L}(A) , \tag{B.9}$$

with A and A' linear operators on the system Hilbert space. An important special case of such a map is given by

$$\mathcal{L} : A \longrightarrow A' = \sum_k \lambda_k B_k A B_k^\dagger \,, \tag{B.10}$$

where the B_k's are again linear operators. In particular, we can define a family of such maps, as given by Eq. (IV.11). When $\lambda_k \geq 0$ for all k, the map in Eq. (B.10) is again a positive map. To prove this statement, note that

$$\langle\phi|B_k A B_k^\dagger|\phi\rangle = \langle B_k^\dagger\phi|A|B_k^\dagger\phi\rangle \equiv \langle\phi'|A|\phi'\rangle \geq 0 \tag{B.11}$$

for all B_k and $|\phi\rangle$ if A is non-negative. We then have

$$\langle\phi|A'|\phi\rangle = \sum_k \lambda_k \langle\phi|B_k A B_k^\dagger|\phi\rangle = \sum_k \lambda_k \langle\phi_k'|A|\phi_k'\rangle \,. \tag{B.12}$$

The right-hand side of this equation is positive for all B_k's, $\lambda_k \geq 0$'s and non-negative operators A. Hence A' is a non-negative operator and $\mathcal{L}$ is positive.

Furthermore, when $\lambda_k \geq 0$ such an $\mathcal{L}$ is *completely* positive. To prove this statement, let $\mathcal{L}$ be a map on system 1 (henceforth denoted by $\mathcal{L}_1$) and consider a second system 2. Recall that $\mathcal{L}_1$ is completely positive if all its extensions $\mathcal{L}_1 \otimes \mathbb{1}_2$ are positive. I will now show that this is the case.

Define the extension $\mathcal{L}_{12} = \mathcal{L}_1 \otimes \mathbb{1}_2$. System 2 can have arbitrary dimension, and may itself be composite. We thus have to show that $\mathcal{L}_{12}$ is positive. Let A_{12} be a non-negative operator on the composite system $1+2$:

$$A_{12} = \sum_{j,k;l,m} a_{jk,lm} |\phi_j\rangle_1 |\psi_k\rangle_2 \langle\psi_m|_1 \langle\phi_l| \,, \tag{B.13}$$

with $a_{jk,lm} = a^*_{lm,jk}$, and define the operator B_i on system 1 as

$$B_i = \sum_{p,q} b^i_{pq} |\phi_p\rangle\langle\phi_q| \,. \tag{B.14}$$

The map $\mathcal{L}_1 \otimes \mathbb{1}_2$ is then given by the transformation

$$\begin{aligned}
A_{12} \to A'_{12} &= \sum_i \lambda_i (B_i \otimes \mathbb{1}_2) A_{12} (B_i \otimes \mathbb{1}_2)^\dagger \\
&= \sum_i \lambda_i \sum_{p,q;r,s} \sum_{j,k;l,m} a_{jk,lm} b^i_{pq} |\phi_p\rangle_1 \langle\phi_q|\phi_j\rangle_1 |\psi_k\rangle_2 \langle\psi_m|_1 \langle\phi_l|\phi_r\rangle_1 \langle\phi_s| {b^i_{sr}}^* \\
&= \sum_i \lambda_i \sum_{p,s} \sum_{j,k;l,m} a_{jk,lm} b_{pj} |\phi_p\rangle_1 |\psi_k\rangle_2 \langle\psi_m|_1 \langle\phi_s| {b^i_{sl}}^* \\
&= \sum_i \lambda_i \sum_{j,k;l,m} a_{jk,lm} |\phi'_j\rangle_1 |\psi_k\rangle_2 \langle\psi_m|_1 \langle\phi'_l| \,,
\end{aligned} \tag{B.15}$$

where I defined $|\phi'_j\rangle = \sum_p b_{pj}|\phi_p\rangle$. Since this is a convex sum over non-negative operators, the resulting operator is again non-negative and $\mathcal{L}$ is completely positive. This completes the proof.

The fact that positive but not completely positive maps on a subsystem do not necessarily transform density operators on the composite system to density operators can be exploited to detect (or witness) quantum entanglement. This is the subject of the next section.

3 Partial transpose criterion

In quantum information theory, it is important to know whether a composite system, characterised by a density operator ρ is separable or not. One way to test this is to use the Peres-Horodecki partial transpose criterion [134, 87].

A density operator ρ of two systems $1+2$ is separable if and only if it can be written as

$$\rho = \sum_k p_k \rho_k^{(1)} \otimes \rho_k^{(2)} , \tag{B.16}$$

with $p_k > 0$ and $\sum_k p_k = 1$. The density operator $\rho_k^{(j)}$ is defined on system $j = 1, 2$. Consider again the transpose of an operator A:

$$T : A \longrightarrow A^T . \tag{B.17}$$

As we have seen, this is a trace-preserving positive, but not completely positive map. We extended this map to the *partial transpose* $PT \equiv T_1 \otimes \mathbb{1}_2$. The partial transpose is not positive on the composite system.

Under the partial transpose, the separable density operator from Eq. (B.16) will transform according to

$$PT : \rho \longrightarrow \rho' = \sum_k p_k \left(\rho_k^{(1)}\right)^T \otimes \rho_k^{(2)} . \tag{B.18}$$

However, T is positive and $(\rho_k^{(1)})^T$ is again a density operator. Therefore ρ' is another (separable) density operator. Now look at the eigenvalues of ρ'. Clearly, if ρ is separable, then ρ' has positive eigenvalues. Therefore, if ρ' has one or more *negative* eigenvalues, the original density operator ρ must have been entangled. This is the Peres-Horodecki partial transpose criterion. Clearly, it is only a necessary condition for separability.

It has been proved [87] that for the Hilbert spaces $\mathscr{H}_2 \otimes \mathscr{H}_2$ and $\mathscr{H}_2 \otimes \mathscr{H}_3$ the partial transpose criterion is both necessary and sufficient. In other words, ρ is separable *if and only if* the eigenvalues of its partial transpose ρ' are positive. For higher dimensional Hilbert spaces this is no longer true. In that case there can exist density operators which are *not* separable, but for which the eigenvalues of ρ' are non-negative. Such states are said to exhibit *bound* entanglement [88, 89]. It is generally believed that this form of entanglement cannot be purified.

4 PROJECTION OPERATOR VALUED MEASURES

Let us now return to the case of a single system and the states, operators and maps defined on it. Consider a projection operator P defined by

$$P^\dagger = P \qquad \text{and} \qquad P^2 = P \,. \tag{B.19}$$

In terms of the states $\{|\phi_j\rangle\}$ this operator can be written as

$$P_{|\phi\rangle} = |\phi\rangle\langle\phi| \,. \tag{B.20}$$

Suppose we have a set of projection operators $\{P_{|\mu\rangle}\}$, with the states $|\mu\rangle$ not necessarily orthogonal. We can define a *generalised* projection operator $\hat{E}_\nu$ as a weighted measure over this set:

$$\hat{E}_\nu = \sum_\mu \lambda_\mu^\nu P_{|\mu\rangle} \,, \tag{B.21}$$

with $\lambda_\mu^\nu > 0$ and

$$\sum_\nu \hat{E}_\nu = \mathbb{1} \,. \tag{B.22}$$

The operator E_k is called a *projection operator valued measure* or POVM for short [107].

This can be generalised further by observing that

$$\lambda_\mu^\nu P_{|\mu\rangle} = \alpha_\mu^\nu |\mu\rangle\langle\nu|\nu\rangle\langle\mu|(\alpha_\mu^\nu)^* \,, \tag{B.23}$$

with $|\nu\rangle \in \{|\mu\rangle\}$ and $|\alpha_\mu^\nu|^2 = \lambda_\mu^\nu$. When we define the operator $\mathcal{A}$:

$$\mathcal{A} \equiv \sum_{\mu,\nu} \alpha_\mu^\nu |\mu\rangle\langle\nu| \tag{B.24}$$

we can write the POVM as

$$E_\nu = \sum_\mu \mathcal{A}_{\mu\nu}\mathcal{A}_{\mu\nu}^\dagger \,. \tag{B.25}$$

C

Elementary Group Theory

In this appendix I give some background theory of Lie groups. I am indebted to the book by De Wit and Smith [176], which gives a good and concise exposition of the subject. Further Lie group theory in particle physics is presented in Halzen and Martin [77]. For group theory in quantum mechanics, see also Chaichian and Hagedorn [43]. For a formal treatment of Lie groups, see Gilmore [69].

A set G is called a *group* when G satisfies the following requirements:

- There exists a multiplication rule $(\cdot)$ such that for every two elements g_1 and g_2 of the group, their product $g_1 \cdot g_2$ is again an element of the group;
- the multiplication rule is associative, i.e., $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ for all $g_1, g_2, g_3 \in \mathsf{G}$;
- there exists an element $e \in \mathsf{G}$, called the unit element, for which the product $e \cdot g = g \cdot e = g$, with g any element of G;
- for every $g \in \mathsf{G}$ there exists an element $g^{-1} \in \mathsf{G}$, called the inverse element of g, such that $g \cdot g^{-1} = g^{-1} \cdot g = e$.

When $g_1 \cdot g_2 = g_2 \cdot g_1$, the group is called *Abelian.* In other words, the elements of G *commute.* A subset H of G is called a *subgroup* of G if the group requirements above hold for H. This is written as $\mathsf{H} \subset \mathsf{G}$.

1 Lie groups

If a group G has a finite number of elements, this number is called the *order* of G. For finite groups, see e.g., Serre [150]. When the group has an infinite number of elements, the group can be either continuous or discontinuous. In the context of this thesis I am mostly interested in continuous groups. If the elements of a continuous group G depend analytically on a (finite) set of parameters $(g = g(\vec{\xi}))$, we speak of a *Lie group*[1]. The *dimension* of the Lie group is given by the number of independent parameters: if $\vec{\xi} = (\xi_1, \ldots, \xi_N)$, we have $\dim \mathsf{G} = N$. The N-dimensional space generated by the parameters is called *parameter space.*

[1] After the Norwegian mathematician Marius Sophus Lie (1842–1899).

Let G be a one-dimensional Lie group with elements $g(\xi)$. We can always choose the parametrisation such that [176]

$$g(\xi)g(\xi') = g(\xi + \xi') \tag{C.1}$$

with

$$g(0) \equiv e \qquad \text{and} \qquad (g(\xi))^{-1} = g(-\xi) \; . \tag{C.2}$$

If we interpret the group elements as operators (acting on other group elements), the unit element is the identity operator: $e = \mathbb{1}$. In the neighbourhood of the identity, a group element can thus be written as an expansion

$$g(\xi) = g(0) + \xi T + O(\xi^2) \; , \tag{C.3}$$

where T is some operator.

Since G is a continuous group, we can write the transformation from the identity [written as $g(0)$] to $g(\xi)$ in terms of n small steps $g(\xi/n)$:

$$g(\xi) = [g(\xi/n)]^n \; . \tag{C.4}$$

If I now take the limit of $n \to \infty$ the higher-order terms vanish and we obtain

$$g(\xi) = \lim_{n\to\infty} \left[g(0) + \frac{\xi T}{n} + \ldots \right]^n = \lim_{n\to\infty} \left[\mathbb{1} + \frac{\xi T}{n} \right]^n \equiv \exp\left[\xi T\right] \; . \tag{C.5}$$

The operator T is said to be the *generator* of the group G because it generates the elements of G. This can be generalised immediately to N-dimensional Lie groups, yielding

$$g(\vec{\xi}) = \exp\left[\sum_{i=1}^{N} \xi_i T_i \right] \; , \tag{C.6}$$

where the ξ_i's are the independent parameters of the group and the T_i's the generators. There are as many different generators as there are parameters.

In terms of the generators, the group multiplication can be written as

$$g(\vec{\xi}) \cdot g(\vec{\xi'}) = \exp\left[\sum_{i=1}^{N} \xi_i T_i \right] \exp\left[\sum_{i=1}^{N} \xi_i' T_i \right] \; . \tag{C.7}$$

The right-hand side can be expressed as the argument of a single exponent by means of the Baker-Campbell-Hausdorff formula:

$$\exp\left[\sum_{i=1}^{N} \xi_i T_i \right] \exp\left[\sum_{i=1}^{N} \xi_i' T_i \right] = \exp\left[\sum_{i=1}^{N} (\xi_i + \xi_i') T_i + \frac{1}{2} \sum_{i,j=1}^{N} \xi_i \xi_j' [T_i, T_j] + \ldots \right] \; , \tag{C.8}$$

where $[T_i, T_j]$ denotes the commutator between T_i and T_j, and the dots indicate a series of terms with higher-order commutators of the T's (like, for example $[T_i, [T_j, T_k]]$). This series does not necessarily terminate.

I started with the condition that G is a group, which implies that the right-hand side of Eq. (C.8) is again an element of G and thus can be written as $e^{\sum_{i=1}^{N} \xi_i'' T_i}$. In turn, this means that the generators are closed under commutation:

$$[T_i, T_j] = \sum_{k=1}^{N} c_{ij}^k T_k \,, \tag{C.9}$$

with the (complex) numbers c_{ij}^k the so-called *structure constants*. To see that this equation must hold, suppose that the argument of the right-hand side of Eq. (C.8) *does not* imply Eq. (C.9). There is then a commutator $[T_l, T_m]$ which cannot be written as a sum over the generators: $[T_l, T_m] = X$. Repeated commutators should then cancel X, because the right-hand side of Eq. (C.8) is a group element. This can only happen when repeated commutators yield X. By definition, X is then a member of the set of generators. This contradicts our assumption.

The structure constants c_{ij}^k define a so-called Lie *algebra*. They obey the Jacobi identity for structure constants:

$$\sum_k c_{ij}^k c_{kl}^m + c_{jl}^k c_{ki}^m + c_{li}^k c_{kj}^m = 0 \,. \tag{C.10}$$

This is easily proved using the Jacobi identity for any three operators A, B and C:

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0 \,. \tag{C.11}$$

2 Representations

When we have a set of matrices M_i with $i = 1, \ldots, N$, and the commutation relations between these matrices are given by

$$[M_i, M_j] = \sum_{k=1}^{N} c_{ij}^k M_k \,, \tag{C.12}$$

then this set of matrices is said to form a *representation* of the Lie algebra defined in Eq. (C.9). When these matrices are multiplied by θ_i and exponentiated, they define a representation of the group G, denoted by $D(\mathsf{G})$:

$$D(g(\vec{\theta})) = \exp\left[\sum_{i=1}^{N} \theta_i M_i\right] \,. \tag{C.13}$$

The matrices M_i form a *basis* of the representation. For N matrices the representation is said to be N-dimensional. If there exists a non-trivial subspace of $\mathcal{V}$ spanned by the basis $\{M_i\}$ (i.e., a subspace other than $\mathbf{0}$ and $\mathcal{V}$ itself) which is invariant under the group transformations, the representation is called *reducible*. If no such invariant subspace exists, the representation is *irreducible*. The theory of representations is important for many applications in physics.

3 Examples of Lie groups

One of the most important Lie groups in quantum mechanics must be the group of 2×2 unitary matrices. This group is called $SU(2)$. The corresponding Lie algebra $su(2)$ is given by three generators ($i, j, k \in \{x, y, z\}$):

$$[J_i, J_j] = \sum_k i\epsilon_{ij}^k J_k \,, \tag{C.14}$$

where ϵ_{ij}^k are the entries of the Levi-Civita tensor of rank three (entries with even permutations of the indices are 1, odd permutations give -1, and repeated indices give 0). The generators are given by $J_i = \frac{1}{2}\sigma_i$, with

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \tag{C.15}$$

the so-called Pauli matrices. Representations of this group are used in the description of angular momentum, spin and iso-spin, as well as in quantum optics (see appendix D for a relation between Lie algebras and optical devices). If the parameters are given by ξ_x, ξ_y and ξ_z, a general $SU(2)$ group element in the fundamental (two-dimensional) representation can be written as ($\xi \equiv \sqrt{\xi_x^2 + \xi_y^2 + \xi_z^2}$)

$$g(\vec{\xi}) = \cos\tfrac{1}{2}\xi \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{i \sin\frac{1}{2}\xi}{\xi} \begin{pmatrix} \xi_z & \xi_x - i\xi_y \\ \xi_x + i\xi_y & -\xi_z \end{pmatrix} . \tag{C.16}$$

Since the group elements depend periodically on ξ, we have $\xi = \sqrt{\xi_x^2 + \xi_y^2 + \xi_z^2} \leq 2\pi$, which means that the parameter space of $SU(2)$ is *compact*; it can be restricted to a sphere with radius 2π.

The group $SU(2)$ is closely related to the group $SO(3)$, the group of orthogonal 3×3 matrices, better known as the *rotation group* in three dimensions.

We know that a rotation [in $SO(3)$] over 2π is equal to the identity. However, when $\xi = 2\pi$ in Eq. (C.16), we see that the $SU(2)$ group element is equal to $-\mathbb{1}$. This behaviour is the reason why $SU(2)$, rather than $SO(3)$, is used to describe particles with spin. After all, spin $\frac{1}{2}$ particles need a rotation over 4π in order to return to their original state. The group $SU(2)$ is called the *covering group* of $SO(3)$.

Another Lie group which is important in the context of this thesis is the group $SU(1,1)$. Its Lie algebra is given by

$$[J_x, J_y] = -iJ_z \ , \qquad [J_y, J_z] = iJ_x \ , \qquad [J_z, J_x] = iJ_y \ , \tag{C.17}$$

and the elements of the fundamental (two-dimensional) representation are generated by the matrices $J_i = \frac{1}{2}\rho_i$, with

$$\rho_x = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \rho_y = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad \rho_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{C.18}$$

The group elements can be written in the fundamental representation as

$$g(\vec{\zeta}) = \cosh \tfrac{1}{2}\zeta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sinh \tfrac{1}{2}\zeta \begin{pmatrix} 0 & \zeta_1 - i\zeta_2 \\ \zeta_1 + i\zeta_2 & 0 \end{pmatrix} , \tag{C.19}$$

with $\vec{\zeta} = (\zeta_1, \zeta_2, \zeta_3)$ and $\zeta \equiv \sqrt{\zeta_3^2 - \zeta_2^2 - \zeta_1^2}$. There is no periodicity in ζ, and the parameter space is therefore not compact. In quantum optics, this group is associated with *squeezing*.

D

Bilinear and Quadratic Forms

In many problems in quantum optics we are faced with a unitary evolution U due to a Hermitian operator $\mathcal{H}$ (with $U = \exp[-it\mathcal{H}/\hbar]$), generally an interaction Hamiltonian. For computational simplicity we often wish that this evolution is in normal ordered form. This form (or, consequently, the corresponding Baker-Campbell-Hausdorff formula) is usually very complicated, if it exists at all. In this appendix I present two important classes of operators for which the normal ordered form of the unitary evolution can be derived.

I first define the so-called *bilinear* and *quadratic forms* for the creation and annihilation operators. We will present the normal ordering for evolutions generated by the Hermitian operators which can be written in terms of these bilinear and quadratic forms.

Suppose we have two vectors $\vec{x} = (x_1, \ldots, x_n)$ and $\vec{y} = (y_1, \ldots, y_n)$. The scalar product between these vectors is denoted by $(\vec{x}, \vec{y})$. Furthermore, let C be an $n \times n$ matrix. With C we associate the quadratic form [59]

$$\phi(\vec{x}, \vec{x}) = (\vec{x}, C\vec{x}) \ , \tag{D.1}$$

and the bilinear form

$$\phi(\vec{x}, \vec{y}) = (\vec{x}, C\vec{y}) \ . \tag{D.2}$$

We have made no assumptions about the nature of the vector components, and it is possible to define bilinear and quadratic forms in terms of creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$. These operators obey the well-known commutation relations

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij} \quad \text{and} \quad [\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0 \ . \tag{D.3}$$

The quadratic form now reads $(\vec{a}, C\vec{a})$ or $(\vec{a}^\dagger, C\vec{a}^\dagger)$, with $\vec{a} = (\hat{a}_1, \ldots, \hat{a}_n)$ and $\vec{a}^\dagger = (\hat{a}_1^\dagger, \ldots, \hat{a}_n^\dagger)$. The bilinear form can be chosen many ways (according to Eq. (D.2)), but for our present purposes I write it as $(\vec{a}^\dagger, C\vec{a})$.

In quantum optics, the bilinear and quadratic form of creation and annihilation operators occurs very often. Take, for instance, the interaction Hamiltonian for the beam-splitter in modes a_1 and a_2:

$$\mathcal{H}_I = \kappa \, \hat{a}_1^\dagger \hat{a}_2 + \kappa^* \, \hat{a}_2^\dagger \hat{a}_1 \ , \tag{D.4}$$

where κ is a coupling constant. This Hamiltonian is a bilinear form which may be written as

$$(\vec{a}^\dagger, C\vec{a}) = (\hat{a}_1^\dagger, \hat{a}_2^\dagger) \begin{pmatrix} 0 & \kappa \\ \kappa^* & 0 \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} . \tag{D.5}$$

Another example is the interaction Hamiltonian due to parametric down-conversion in two modes a_1 and a_2. This is a sum of two quadratic forms (one for the creation and one for the annihilation operators):

$$\mathcal{H}_I = \nu\, \hat{a}_1^\dagger \hat{a}_2^\dagger + \nu^*\, \hat{a}_2 \hat{a}_1 , \tag{D.6}$$

where ν is a coupling constant. In symmetric form, the quadratic form of the creation operators reads

$$(\vec{a}^\dagger, C\vec{a}^\dagger) = \frac{1}{2}(\hat{a}_1^\dagger, \hat{a}_2^\dagger) \begin{pmatrix} 0 & \nu \\ \nu & 0 \end{pmatrix} \begin{pmatrix} \hat{a}_1^\dagger \\ \hat{a}_2^\dagger \end{pmatrix} . \tag{D.7}$$

Usually, these interaction Hamiltonians are exponentiated to generate the unitary evolution of a system, and studying the behaviour of the bilinear and quadratic forms might simplify our computational task. In particular, we would like to find the normal ordered form of $\exp[-it\mathcal{H}_I/\hbar]$, where $\mathcal{H}_I$ is given by Eq. (D.4) or Eq. (D.6).

In the next two sections I will establish relations between the bilinear and quadratic forms and the Lie algebras of $SU(2)$ and $SU(1,1)$ respectively. The two resulting theorems place restrictions on the matrix C in the bilinear and quadratic forms.

1 Bilinear Forms and $SU(2)$

In quantum optics, linear unitary operations like beam-splitters, half- and quarter-wave plates, phase-shifters, polarisation rotations, etc. all preserve the number of photons. When we write these operations as $U = \exp[-it\mathcal{H}/\hbar]$, with $\mathcal{H}$ some Hermitian operator (an interaction Hamiltonian), it is clear that every term in $\mathcal{H}$ should be a product of an equal number of creation and annihilation operators (i.e., for every photon which is created, another will be destroyed). The lowest order interaction Hamiltonian which satisfies this requirement has a bilinear form (see, e.g., Eq. (D.4)). Furthermore, the resulting unitary operations form representations of the group $SU(2)$, and I therefore study the relation between bilinear forms and $SU(2)$.

Define $K_-(\Lambda)$ to be

$$K_-(\Lambda) = \sum_{ij} \hat{a}_i^\dagger \lambda_{ij} \hat{a}_j \equiv (\vec{a}^\dagger, \Lambda\vec{a}) , \tag{D.8}$$

a bilinear form. Let $K_+(\Lambda)$ be the adjoint of $K_-(\Lambda)$:

$$K_+(\Lambda) = K_-^\dagger(\Lambda) = (\vec{a}^\dagger\Lambda, \vec{a}) \; . \tag{D.9}$$

I can now define a third operator $K_0(\Lambda)$ in such a way that the three operators generate an $su(2)$ algebra[1]:

$$K_0 = -\frac{1}{2}[K_-, K_+] \; . \tag{D.10}$$

Using Eqs. (D.8) and (D.9), and normal ordering the commutator in Eq. (D.10) yields

$$K_0(\Lambda) = -\frac{1}{2}\sum_{ij} \hat{a}_i^\dagger [\Lambda, \Lambda^\dagger]_{ij} \hat{a}_j = K_0^\dagger(\Lambda) \; . \tag{D.11}$$

With $M \equiv [\Lambda, \Lambda^\dagger]$ this can be written as

$$K_0(M) = -\frac{1}{2}\sum_{ij} \hat{a}_i^\dagger \mu_{ij} \hat{a}_j = -\frac{1}{2}(\vec{a}^\dagger, M\vec{a}) \; . \tag{D.12}$$

The last commutation relation which, together with Eq. (D.10) and $K_- = K_+^\dagger$, $K_0 = K_0^\dagger$ constitutes the $su(2)$ algebra is

$$[K_0, K_\pm] = \pm K_\pm \; . \tag{D.13}$$

This equation places a constraint on the allowed matrices Λ (all the other commutation relations so far have not placed any restrictions on the form of Λ). Using Eqs. (D.8), (D.9) and (D.11) yields

$$\begin{aligned} [K_0, K_-] = -K_- \quad &\Leftrightarrow \quad \Lambda = \frac{1}{2}[M, \Lambda] \\ [K_0, K_+] = K_+ \quad &\Leftrightarrow \quad \Lambda^\dagger = \frac{1}{2}[\Lambda^\dagger, M] \; . \end{aligned} \tag{D.14}$$

This can be summarised in the following theorem:

Theorem 1 Consider a bilinear operator of the form $K(\Lambda) = (\vec{a}^\dagger, \Lambda\vec{a})$. K and $K^\dagger$ define a third operator $K_0 = -\frac{1}{2}[K, K^\dagger]$. These operators are generators of an $su(2)$ algebra if and only if $\Lambda = \frac{1}{2}[M, \Lambda]$, with $M = [\Lambda, \Lambda^\dagger]$.

Suppose that the interaction Hamiltonian can be written as

$$\mathcal{H}_I = \kappa \; K_+(\Lambda) + \kappa^* \; K_-(\Lambda) \; , \tag{D.15}$$

[1] Traditionally, the *group* is denoted by capital letters, e.g., $SU(2)$, whereas the corresponding *algebra* is written with lowercase letters, e.g., $su(2)$.

with κ again a coupling constant. Note that this is now a *sum* of two bilinear forms, unlike in Eq. (D.5).· Since K_+ and K_- generate an $su(2)$ algebra, we know what the normal ordering for the unitary evolution U associated with this interaction Hamiltonian is ($\tau = -i\kappa t/\hbar$ and $\hat{\tau}$ is the unit vector in the direction of τ) [164]:

$$\begin{aligned} U(\tau) &= \exp[\tau K_+(\Lambda) - \tau^* K_-(\Lambda)] \\ &= \exp[\hat{\tau}\tan|\tau|K_+]\exp[-2\ln\cos|\tau|K_0]\exp[-\hat{\tau}^*\tan|\tau|K_-] \ . \end{aligned} \tag{D.16}$$

This is the Baker-Campbell-Hausdorff formula for $su(2)$ [69].

2 Quadratic Forms and $SU(1,1)$

In quantum optics, squeezers and down-converters are described by interaction Hamiltonians which are quadratic in the creation and annihilation operators. These Hamiltonians generate unitary transformations which do not conserve the photon number. In particular, these transformations can be viewed as photon sources.

Write the unitary evolution of these sources as $U = \exp[-it\mathcal{H}/\hbar]$, with $\mathcal{H}$ a Hermitian operator (the interaction Hamiltonian). Here, every term in $\mathcal{H}$ is proportional to a product of either two creation or two annihilation operators, i.e., $\mathcal{H}$ is proportional to a sum of quadratic forms. Subsequently we know that squeezing and parametric down-conversion are representations of the group $SU(1,1)$, and I therefore study the relation between this group and quadratic forms.

Let $L_-(\Lambda)$ be a quadratic form:

$$L_-(\Lambda) = \frac{1}{2}\sum_{ij}\hat{a}_i\lambda_{ij}\hat{a}_j \equiv \frac{1}{2}(\vec{a}, \Lambda\vec{a}) \ . \tag{D.17}$$

Since the annihilation operators commute, it is clear that Λ can always be chosen symmetric. The adjoint of $L_-(\Lambda)$ is given by

$$L_+(\Lambda) = L_-^\dagger(\Lambda) = \frac{1}{2}\sum_{ij}\hat{a}_i^\dagger\lambda_{ij}^*\hat{a}_j^\dagger \equiv \frac{1}{2}(\vec{a}^\dagger\Lambda^\dagger, \vec{a}^\dagger) \ . \tag{D.18}$$

When we want to construct an $su(1,1)$ algebra with these operators we need to show that there exists an operator $L_0(\Lambda)$ which satisfies the commutation relations

$$[L_-, L_+] = 2L_0 \quad \text{and} \quad [L_0, L_\pm] = \pm L_\pm \ . \tag{D.19}$$

The first relation in Eq. (D.19) defines $L_0(\Lambda)$:

$$L_0(\Lambda) = \frac{1}{2}\left\{(\Lambda\vec{a})^\dagger\cdot(\Lambda\vec{a}) + \frac{1}{2}\mathrm{Tr}(\Lambda^\dagger\Lambda)\right\} = L_0^\dagger(\Lambda) \ . \tag{D.20}$$

The second relation in Eq. (D.19) places a constraint on Λ:

$$\begin{aligned} [L_0, L_+] = L_+ &\Leftrightarrow \Lambda^\dagger = \Lambda^\dagger \Lambda \Lambda^\dagger \\ [L_0, L_-] = -L_- &\Leftrightarrow \Lambda = \Lambda \Lambda^\dagger \Lambda \, . \end{aligned} \tag{D.21}$$

The matrix Λ is unitary if it is invertible.

I can now formulate these results in terms of a theorem:

Theorem 2 Consider a quadratic operator of the form $L(\Lambda) = (\vec{a}, \Lambda \vec{a})$. L and $L^\dagger$ define a third operator $L_0 = \frac{1}{2}[L, L^\dagger]$. These operators are generators of an $su(1,1)$ algebra if and only if $\Lambda^\dagger = \Lambda^\dagger \Lambda \Lambda^\dagger$. Such a Λ is unitary if and only if it is invertible.

Suppose that the interaction Hamiltonian can be written as

$$\mathcal{H}_I = \nu \, L_+(\Lambda) + \nu^* \, L_-(\Lambda) \, , \tag{D.22}$$

with ν the coupling constant. Since L_+ and L_- generate an $su(1,1)$ algebra, we know what the normal ordering for the unitary evolution U associated with this interaction Hamiltonian is ($\tau = -i\nu t/\hbar$) [164]:

$$\begin{aligned} U(\tau) &= \exp[\tau L_+(\Lambda) - \tau^* L_-(\Lambda)] \\ &= \exp[\hat{\tau} \tanh |\tau| L_+] \exp[-2 \ln \cosh |\tau| L_0] \exp[-\hat{\tau}^* \tanh |\tau| L_-] \, . \end{aligned} \tag{D.23}$$

This is the Baker-Campbell-Hausdorff formula for $su(1,1)$.

E

TRANSFORMATION PROPERTIES OF MAXIMAL ENTANGLEMENT

In this appendix I will show[1] that any maximally entangled state can be transformed into any other maximally entangled state by means of a unitary transformation on only one of the subsystems. I will treat this in a formal way by considering an arbitrary maximally entangled state of two N-level systems in the Schmidt decomposition:

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} e^{i\phi_j} |n_j, m_j\rangle \ , \tag{E.1}$$

that is, a state with equal amplitudes on all possible branches. There always exist two orthonormal bases $\{|n_i\rangle\}$ and $\{|m_i\rangle\}$ such that Eq. (E.1) can be written this way, by virtue of the definition for maximal entanglement. Clearly, we can obtain *any* maximally entangled state by applying the (bi-local) unitary transformation $U_1 \otimes U_2$. That is, each maximally entangled state can be transformed into any other by a pair of local unitary transformations on each of the subsystems. We will now show that any two maximally entangled states $|\psi\rangle$ and $|\psi'\rangle$ are connected by a local unitary transformation on one subsystem alone:

$$|\psi\rangle = U \otimes \mathbb{1} |\psi'\rangle = \mathbb{1} \otimes U' |\psi'\rangle \ . \tag{E.2}$$

First, I will prove that any transformation $U_1 \otimes U_2$ on a *particular* maximally entangled state $|\phi\rangle$ can be written as $V \otimes \mathbb{1}|\phi\rangle$, where $V = U_1 U_2^T$. To this end I will give the proofs for two theorems. Take the special maximally entangled state

$$|\phi\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} |n_j, m_j\rangle \ . \tag{E.3}$$

Theorem 1: For any state $|\phi\rangle$ given by Eq. (E.3) and any unitary operator U we have

$$U \otimes U^* |\phi\rangle = |\phi\rangle \ . \tag{E.4}$$

[1] This is not new material, it is included here for reasons of completeness.

Proof: Using the completeness relation

$$\sum_{k=1}^{N} |n_k\rangle\langle n_k| = \mathbb{1} \tag{E.5}$$

on both subsystems we have

$$U \otimes U^*|\phi\rangle = \sum_{k,l=1}^{N} |n_k, m_l\rangle\langle n_k, m_l| \, (U \otimes \mathbb{1}) \, (\mathbb{1} \otimes U^*) \, |\phi\rangle \; . \tag{E.6}$$

By writing out $|\phi\rangle$ explicitly according to Eq. (E.3) we obtain

$$\begin{aligned} U \otimes U^*|\phi\rangle &= \frac{1}{\sqrt{N}} \sum_{j,k,l=1}^{N} U_{kj} U^*_{lj} |n_k, m_l\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{k,l=1}^{N} (UU^\dagger)_{kl} |n_k, m_l\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{k,l=1}^{N} \delta_{kl} |n_k, m_l\rangle \; , \end{aligned} \tag{E.7}$$

which is just $|\phi\rangle$. □

Theorem 2: Every unitary transformation $U_1 \otimes U_2$ acting on the state $|\phi\rangle$ given by Eq. (E.3) is equivalent to a transformation $V \otimes \mathbb{1}$ acting on $|\phi\rangle$, where $V = U_1 U_2^T$.

Proof: The equality

$$\left(\mathbb{1} \otimes U^T\right) (U \otimes U^*) \, |\phi\rangle = (U \otimes \mathbb{1}) \, |\phi\rangle \; , \tag{E.8}$$

together with theorem 1 immediately gives us

$$\begin{aligned} U \otimes \mathbb{1}|\phi\rangle &= \mathbb{1} \otimes U^T|\phi\rangle \qquad \text{and} \\ U^T \otimes \mathbb{1}|\phi\rangle &= \mathbb{1} \otimes U|\phi\rangle \; . \end{aligned} \tag{E.9}$$

From Eq. (E.9) we obtain

$$\begin{aligned} U_1 \otimes U_2|\phi\rangle &= (U_1 \otimes \mathbb{1}) \, (\mathbb{1} \otimes U_2) \, |\phi\rangle \\ &= (U_1 \otimes \mathbb{1}) \left(U_2^T \otimes \mathbb{1}\right) |\phi\rangle \\ &= U_1 U_2^T \otimes \mathbb{1}|\phi\rangle \; . \end{aligned} \tag{E.10}$$

Similarly,

$$U_1 \otimes U_2|\phi\rangle = \mathbb{1} \otimes U_1^T U_2|\phi\rangle \ . \tag{E.11}$$

We therefore obtain that $U_1 \otimes U_2|\phi\rangle$ is equal to $V \otimes \mathbb{1}|\phi\rangle$ with $V = U_1 U_2^T$, and similarly that it is equal to $\mathbb{1} \otimes V'|\phi\rangle$ with $V' = U_1^T U_2$. □

Since every maximally entangled state can be obtained by applying $U_1 \otimes U_2$ to $|\phi\rangle$, two maximally entangled states $|\psi\rangle$ and $|\psi'\rangle$ can be transformed into any other by choosing

$$\begin{aligned} |\psi\rangle &= U \otimes \mathbb{1} \ |\phi\rangle \\ |\psi'\rangle &= V \otimes \mathbb{1} \ |\phi\rangle \ , \end{aligned} \tag{E.12}$$

which gives

$$|\psi\rangle = UV^\dagger \otimes \mathbb{1} \ |\psi'\rangle \ . \tag{E.13}$$

Thus each maximally entangled two-system state can be obtained from any other by means of a local unitary transformation on one subsystem alone.

F

Statistical Distance

In this appendix I review the concept of the statistical distance. It is first and foremost a concept form *classical* probability theory, which has been extended to quantum theory by Wootters [177], Hilgevoord and Uffink [84] and Braunstein and Caves [29].

Suppose we have a vase containing red, blue and green marbles in some proportion. We can draw a marble from the vase and register its colour. When the proportion of red, blue and green marbles is known, we can predict that we will draw a red marble with some probability p_{red}. In this situation the probability quantifies our uncertainty of *prediction*.

Alternatively, we might be in a different situation where the proportion of red, blue and green marbles is *not* known. When we draw a marble, it gives us extra knowledge which can be used to estimate the proportion of marbles. We do not know for certain what the proportion is until we have drawn all the marbles from the vase, but every new draw will yield extra information about the proportion. The number of drawn red, blue and green marbles estimates the probability distribution of drawing red, blue and green marbles, and the uncertainty after a number of draws is the uncertainty of *inference*. There are therefore two kinds of uncertainty: one associated with the prediction of the outcome of a stochastic process and one associated with the inference of a probability distribution based on a set of outcomes [84].

The *statistical distance* quantifies the distinguishability of two probability distributions, and is therefore closely related to the uncertainty of inference. Since it is a distance, it obeys the four well-known requirements [39]:

1. A distance $s(x, y)$ between two points x and y is positive;
2. $s(x, y) = 0$ if and only if $x = y$;
3. the distance is symmetric: $s(x, y) = s(y, x)$;
4. the distance obeys the triangle inequality: $s(x, z) \leq s(x, y) + s(y, z)$.

The points x, y and z are elements of some (continuous) space. In the case of the statistical distance, these points are probability distributions, which are elements of the so-called *probability simplex* (see figure F.1).

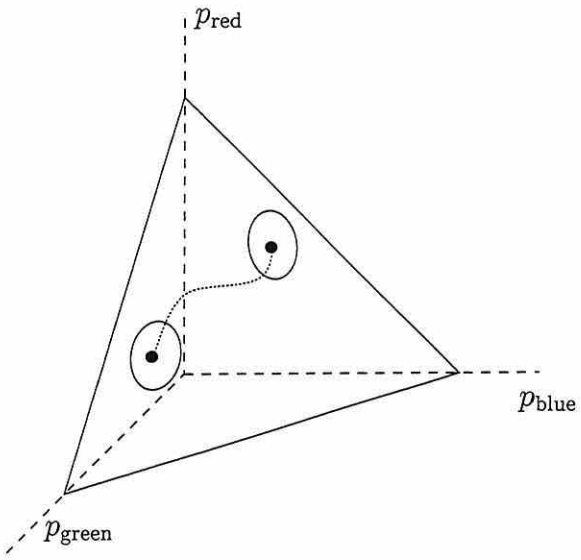

Figure F.1: The probability simplex corresponding to three possible outcomes 'red', 'green' and 'blue'. The two dots correspond to normalised probability distributions. Their uncertainty regions after N trials is depicted by the circle around the dots. The distance between the two distributions is the shortest path in the simplex, measured in units of the typical statistical fluctuation.

The distance function in a space (in this case the simplex) is defined by the so-called *metric*. The metric g is a real symmetric matrix which obeys $\sum_k g_{jk} \cdot g^{kl} = \delta^l_j$, where δ^l_j is the Kronecker delta. Furthermore, it transforms *covariant* vectors x_j to *contravariant* vectors x^j, distinguished by lower and upper indices respectively[1]:

$$x^j = \sum_k g^{jk} x_k \qquad \text{and} \qquad x_j = \sum_k g_{jk} x^k \,. \tag{F.1}$$

The contraction $\sum_j x_j y^j$ yields a scalar (which is invariant under all transformations).

In general, an (incremental) distance ds on the simplex separating points p^j and $p^j + dp^j$ can be written as a quadratic form

$$ds^2 = \sum_{jk} g_{jk} dp^j dp^k \,. \tag{F.2}$$

This is a scalar which is invariant under all coordinate transformations. The dp^j are the components of the incremental tangent vector $\vec{dp}$ along the shortest path between the two probability distributions in the probability simplex. We now aim to find the metric g of the simplex.

[1]To avoid confusion, I will *not* use Einstein's summation convention.

To this end, we define the *dual* A to $\vec{dp}$, i.e., every component A_j is paired with the component dp^j:

$$\langle A \rangle \equiv \sum_j A_j p^j \ , \tag{F.3}$$

where $\langle A \rangle$ can be interpreted as the mean value of A. In order to find the metric, we look at the two-point correlation function of A:

$$\langle A^2 \rangle = \sum_{jk} A_j A_k g^{jk} = \sum_j A_j^2 p^j \ . \tag{F.4}$$

The last equality is obtained by using Eq. (F.3). From this we immediately obtain the contravariant form of the metric:

$$g^{jk} = p^j \delta^{jk} \ , \tag{F.5}$$

with δ^{jk} the Kronecker delta. Since $\sum_k g_{jk} \cdot g^{kl} = \delta_j^l$, the covariant metric is $g_{jk} = \delta_{jk}/p^j$ and the statistical distance becomes

$$ds^2 = \sum_j \frac{dp^j dp^j}{p^j} \equiv \sum_j \frac{(dp^j)^2}{p^j} \ . \tag{F.6}$$

This is the incremental statistical distance used in chapter II. When we make the substitution $p^j = r_j^2$, we find the Euclidean distance $ds^2 = 4 \sum_j dr_j^2$.

Note that ds^2 tends to infinity when one of the probabilities p_j equals zero. This is expected since a probability distribution $\vec{p}^{(1)}$ with $p_j = 0$ is perfectly distinguishable from a distribution $\vec{p}^{(2)}$ with $p_j \neq 0$: one outcome corresponding to p_j will immediately tell us that we have the probability distribution $\vec{p}^{(2)}$.

In order to find the statistical distance between two well separated probability distributions, we have to integrate Eq. (F.6). Following Wootters [177], we find

$$s(\vec{p}^{(1)}, \vec{p}^{(2)}) = \arccos\left(\sum_j \sqrt{p_j^{(1)} p_j^{(2)}} \right) \ . \tag{F.7}$$

In other words, the statistical distance is the *angle* between two vectors with coordinates $\sqrt{p_j^{(1)}}$ and $\sqrt{p_j^{(2)}}$. Wootters [177] proved that this distance measure is the *only* Riemannian distance measure in a Hilbert space, which is invariant under all transformations. It should be noted that we only assumed classical probability theory in our derivation, which makes the appearance of probability amplitudes even more surprising.

An alternative way to arrive at the statistical distance is by using the Gaussian distribution for the observed frequencies f_j in a large number (N) of trials [142, 29]:

$$\rho(f_1, \dots) \propto \exp\left[-\frac{N}{2} \sum_j \frac{(f_j - p_j)^2}{p_j} \right] \ . \tag{F.8}$$

Two probability distributions $p^{(1)}$ and $p^{(1)}$ can then be distinguished if and only if the Gaussian function $\exp\left[-\frac{N}{2}\sum_j \frac{(p_j^{(1)}-p_j^{(2)})^2}{p_j^{(1)}}\right]$ is small. In other words, if $p_j^{(1)} - p_j^{(2)} \equiv dp_j$ we need

$$\frac{N}{2}\sum_j \frac{dp_j^2}{p_j} \gtrsim 1\,, \tag{F.9}$$

or

$$ds \gtrsim \frac{1}{\sqrt{N}}\,, \tag{F.10}$$

which is consistent with Eq. (III.35).

G

Multi-Dimensional Hermite Polynomials

In this appendix I will give the background of multi-dimensional Hermite polynomials. Early introductions to the subject were presented by P. Appell and J. Kampé de Fériet [3], and in the Bateman Manuscript Project [59]. M.M. Mizrahi [121] and M. Klauderer [96] further developed the mathematical theory, and in the context of quantum optics multi-dimensional Hermite polynomials have been applied by V.V. Dodonov, V.I. Man'ko, O.V. Man'ko, V.V. Semjonov, A. Vourdas and R.M. Weiner [52, 53, 54, 169].

1 Ordinary Hermite Polynomials

First, let me revisit the case of the ordinary Hermite polynomials, which are known to physicists as (part of) the eigenfunctions of the linear harmonic oscillator in quantum mechanics (see, for example Merzbacher [119]).

The definition of the Hermite polynomials can be obtained by the construction of a so-called *generating function* $G(x, s)$:

$$G(x,s) = e^{x^2-(s-x)^2} = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!} s^n \ . \tag{G.1}$$

The last equality will give rise to our definition of the Hermite polynomials $H_n(x)$. In order to arrive at this definition we use Taylors expansion:

$$f(x+s) = \left[1 + s\frac{d}{dx} + \frac{s^2}{2!}\left(\frac{d}{dx}\right)^2 + \cdots\right] f(x) = e^{s\frac{d}{dx}} f(x) \ , \tag{G.2}$$

where $\frac{d}{dx}$ denotes the derivative taken with respect to x. The second equality collects the derivatives in the exponential function $\exp[s\frac{d}{dx}]$. Using this relation we write the generating function as

$$G(x,s) = e^{x^2} e^{-s\frac{d}{dx}} e^{-x^2} \ . \tag{G.3}$$

By expanding the exponential $\exp[-s\frac{d}{dx}]$ and comparing with Eq. (G.1) we obtain the definition of the Hermite polynomials $H_n(x)$:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \ . \tag{G.4}$$

Every $H_n(x)$ is a polynomial with n real roots and traditionally normalised in such a way that the leading term x^n has pre-factor 2^n.

There are several relations connecting Hermite polynomials. For instance, the Hermite polynomials obey the orthogonality relation:

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-x^2} |H_n(x)|^2 \, dx &= 2^n n! \sqrt{\pi} \, . \\ \int_{-\infty}^{+\infty} e^{-x^2} H_n(x) H_m(x) \, dx &= 0 \qquad \text{for} \quad n \neq m \, . \end{aligned} \tag{G.5}$$

This relation ensures that the eigenfunctions of the harmonic oscillator are orthonormal.

Furthermore, there are two types of recursion relations connecting Hermite polynomials of different order. From the generating function in Eq. (G.1) it is relatively straightforward to derive the recursion relations

$$\frac{d}{dx} H_n(x) = 2n H_{n-1}(x) \, , \tag{G.6}$$

$$H_{n+1}(x) - 2x H_n(x) + 2n H_{n-1}(x) = 0 \, . \tag{G.7}$$

Combining these two relations yields a second-order homogeneous differential equation called the *Hermite equation*:

$$\frac{d^2}{dx^2} H_n(x) - 2x \frac{d}{dx} H_n(x) + 2n H_n(x) = 0 \, . \tag{G.8}$$

2 Real Multi-Dimensional Hermite Polynomials

The ordinary Hermite polynomials are functions of one variable x. The obvious way to generalise this is taking x to be a vector $\vec{x} = (x_1 \ldots, x_N)$ in an N-dimensional vector space. The generating function of the multi-dimensional Hermite polynomial (henceforth called MDHP) then has to change accordingly: $G(x, s) \to G(\vec{x}, \vec{s})$.

However, rather than replacing s^2 by $(\vec{s}, \vec{s})$ and sx by $(\vec{s}, \vec{x})$ (where we denote the inner product of two vectors $\vec{a}$ and $\vec{b}$ by $(\vec{a}, \vec{b})$), we take the generating function to be [59]

$$G_A(\vec{x}, \vec{s}) = \exp\left[(\vec{s}, A\vec{x}) - \frac{1}{2}(\vec{s}, A\vec{s}) \right] \, , \tag{G.9}$$

where A is a positive definite $N \times N$ matrix, called the *defining matrix*. We can always choose A symmetric. The reason we choose this generating function is that we now also include cross-terms $s_i s_j$ and $s_i x_j$. Without these cross-terms the generalisation would be trivial. When we define $\vec{n}$ as an N-tuple $(n_1, \ldots, n_N)$

with n_i a non-negative integer, $G_A(\vec{x},\vec{s})$ generates the multi-dimensional Hermite polynomials:

$$G_A(\vec{x},\vec{s}) = \sum_{\vec{n}} \frac{s_1^{n_1}}{n_1!}\cdots\frac{s_N^{n_N}}{n_N!} H_{\vec{n}}^A(\vec{x}) \; . \tag{G.10}$$

In this equation, $\sum_{\vec{n}}$ means the sum over all possible N-tuples $\vec{n}$.

The generating function leads to the following definition of the real multi-dimensional Hermite polynomial:

$$H_{\vec{n}}^A(\vec{x}) = (-1)^{\sum_i n_i} e^{\frac{1}{2}(\vec{x},A\vec{x})} \frac{\partial^{\sum_i n_i}}{\partial x_1^{n_1}\cdots\partial x_N^{n_N}} e^{-\frac{1}{2}(\vec{x},A\vec{x})} \; . \tag{G.11}$$

This definition is derived analogous to the one-dimensional case, which was presented above.

3 REDUCTION THEOREM

In order to simplify the derivation of the orthogonality and recursion relations for the real MDHP's, I derived a *Reduction Theorem*:

Reduction Theorem: For any real N-dimensional generating function $G_A(\vec{x},\vec{s})$ with positive definite defining matrix A there exists a linear transformation T which transforms G_A into a product of $N-M$ generating functions of one-dimensional Hermite polynomials:

$$G_A(\vec{x},\vec{s}) \longrightarrow_T \prod_{i=1}^{N-M} G(z_i,v_i) \; , \tag{G.12}$$

where M is the number of zero eigenvalues of A.

Proof: This theorem is proved by explicit construction of T. The transformation has two parts: an orthogonal transformation and a rescaling. The $N\times N$ matrix A is real and symmetric. It can therefore be diagonalised by an orthogonal matrix O [68]:

$$A = O^T\Lambda O \tag{G.13}$$

and

$$(\vec{s},A\vec{x}) = (\vec{s},O\Lambda O^T\vec{x}) = (O^T\vec{s},\Lambda O^T\vec{x}) \equiv (\vec{u},\Lambda\vec{y}) \; . \tag{G.14}$$

This last term can be written as $\sum_i \lambda_i u_i y_i$. The generating function of a real MDHP then transforms as

$$G_A(\vec{x},\vec{s}) \longrightarrow_O G_\Lambda(\vec{y},\vec{u}) = \exp\left[\sum_i \lambda_i\left(u_i y_i - \frac{1}{2}u_i^2\right)\right] \; . \tag{G.15}$$

We can now rescale the transformed coordinates $\vec{y}$ and $\vec{u}$:

$$v_i = \sqrt{\frac{\lambda_i}{2}} u_i \qquad \text{and} \qquad z_i = \sqrt{\frac{\lambda_i}{2}} y_i \; . \tag{G.16}$$

This *rescaled* transformation of the generating function of a real MDHP then gives

$$G_A(\vec{x}, \vec{s}) \;\longrightarrow_T\; G(\vec{z}, \vec{v}) = \prod_{i=1}^{N-M} \exp\left[2 z_i v_i - v_i^2\right] \; , \tag{G.17}$$

where M is the number of zero eigenvalues of A. This is the transformation T whose existence we had to prove. □

Since the new variables in Eq. (G.12) are linearly independent, the reduced generating function trivially generates the ordinary Hermite polynomials. We can now derive the orthogonality relation of the real MDHP's.

4 Orthogonality relation

The Reduction Theorem yields the diagonalised form:

$$\begin{aligned} &\prod_{i=1}^{N-M} \int_{-\infty}^{+\infty} e^{-z_i^2} |H_{n_i}^A(z_i)|^2 \, dz_i = \sqrt{\pi^{N-M}} \prod_{i=1}^{N-M} 2^{n_i} n_i! \\ &\prod_{i=1}^{N-M} \int_{-\infty}^{+\infty} e^{-z_i^2} H_{n_i}^A(z_i) H_{m_i}^A(z_i) \, dz_i = 0 \qquad \text{if } \exists \; n_i \neq m_i \; . \end{aligned} \tag{G.18}$$

Since the Jacobian J of an orthogonal transformation O is equal to 1, I omit it here. We can now transform Eq. (G.18) back to the non-diagonalised case. This yields

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\vec{x}, A\vec{x})} H_{\vec{n}}^A(\vec{x}) H_{\vec{m}}^A(\vec{x}) \, d\vec{x} = 0 \; . \tag{G.19}$$

The orthogonality relations for the real MDHP's are then:

$$\begin{aligned} &\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\vec{x}, A\vec{x})} |H_{\vec{n}}^A(\vec{x})|^2 \, d\vec{x} = \pi^{N/2} \prod_{i=1}^{N} 2^{n_i} n_i! \\ &\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\vec{x}, A\vec{x})} H_{\vec{n}}^A(\vec{x}) H_{\vec{m}}^A(\vec{x}) \, d\vec{x} = 0 \qquad \text{for } \vec{n} \neq \vec{m} \; . \end{aligned} \tag{G.20}$$

5 RECURSION RELATIONS

There are two classes of recursion relations for the real multi-dimensional Hermite polynomials. First, we present the differential recursion relations, which form a generalisation of Eq. (G.6). Subsequently, we present the type of recursion relations which form a generalisation of Eq. (G.7).

In the generalised form of the Hermite polynomials, we wish to evaluate the derivative $\partial_{x_i} H^A_{\vec{n}}(\vec{x})$. We proceed again from the generating function $G(\vec{x},\vec{s})$:

$$\frac{\partial}{\partial x_i} G(\vec{x},\vec{s}) = \frac{\partial}{\partial x_i} e^{-\frac{1}{2}(\vec{s},A\vec{s})+(\vec{s},A\vec{x})} = \sum_{j=1}^{N} A_{ij} s_j G(\vec{x},\vec{s}) \; . \tag{G.21}$$

Furthermore, from Eq. (G.10) we obtain

$$\frac{\partial}{\partial x_i} G(\vec{x},\vec{s}) = \sum_{\vec{n}} \frac{s_1^{n_1}}{n_1!} \cdots \frac{s_N^{n_N}}{n_N!} \frac{\partial H^A_{\vec{n}}(\vec{x})}{\partial x_i} \; . \tag{G.22}$$

Expanding the right-hand side of Eq. (G.21) into Hermite polynomials and equating it with the right-hand side in Eq. (G.22) yields

$$\sum_{\vec{n}} \frac{s_1^{n_1}}{n_1!} \cdots \frac{s_N^{n_N}}{n_N!} \frac{\partial H^A_{\vec{n}}(\vec{x})}{\partial x_i} = \sum_{\vec{n}} \sum_{j=1}^{N} A_{ij} n_j \frac{s_1^{n_1}}{n_1!} \cdots \frac{s_N^{n_N}}{n_N!} H^A_{\vec{n}-e_j}(\vec{x}) \; , \tag{G.23}$$

where $\vec{n} - e_j$ denotes the vector $\vec{n}$ with n_j replaced by $n_j - 1$. Comparing the terms with equal powers in n_k yields the generalised differential recursion relation

$$\frac{\partial H^A_{\vec{n}}(\vec{x})}{\partial x_i} = \sum_{j=1}^{N} A_{ij} n_j H^A_{\vec{n}-e_j}(\vec{x}) \; . \tag{G.24}$$

This relation is easily generalised for multiple derivatives on $H^A_{\vec{n}}(\vec{x})$.

The second recursion relation is given by

$$H^A_{\vec{n}+e_i}(\vec{x}) - \sum_{j=1}^{N} A_{ij} x_j H^A_{\vec{n}}(\vec{x}) + \sum_{j=1}^{N} A_{ij} n_j H^A_{\vec{n}-e_j}(\vec{x}) = 0 \; . \tag{G.25}$$

This relation can be proved by taking the derivative to x_i and using the recursion relation (G.24).

H

Mathematica Code for Teleportation Modelling

In this appendix, I present the MATHEMATICA code I used to derive the results of chapter VI.

```
Krondelta [j_, k_] := If[ j == k, Return [ 1 ], Return [ 0 ] ]

f [ ] := 1
f [0] := 0
f [y___, 1, x___ ] := f[y,x]

f[x___, y_ + z_, w___] := f[x, y, w] + f[x, z, w]

(*
   Let's define annihilation ops for polarisation k = "x" or "y"
   on mode j = 1, 2, 3, 4, or "a", "u", "v", "d" as a[k,j]
   and the creation ops as ad[k,j]
*)

f[ x___, n_  a[k_ ,j_ ], w___ ] := n f [ x,  a[k, j ], w ]
f[ x___, n_ ad[k_ ,j_ ], w___ ] := n f [ x, ad[k, j ], w ]

(*
   The normal ordering rule preserves commutator algebra.
*)

normOrder := f[ x___, a[k_, j_ ], ad[kk_, jj_ ], w___ ] :>
             f[ x, ad[kk, jj ], a[k, j ], w ] +
             Krondelta [ k, kk ] Krondelta [ j, jj ] f[x, w]

f[ x___, n_. f[y___], w___] := n f[ x, y, w ]

(*
   To what order do we expand the exponential.
*)

expandExp1 := myExp1[ x__ ] :> g[ x ] + g[ x, x ]/2

expandExp2 := myExp2[ x__ ] :> g[ 1 ] + g[ x ]

cutEnd := { g [ x_ ] :> ExpandAll[ x ],
```

```
    g [ y___, x_] :> f[ g [ y ], ExpandAll [ x ] ] }

(* Calculate the creation and annihilation results *)

numReduce := { f[ x___, ad[k_, ll_], aa__ ket[k_, ll_, n_ ] ] :>
                        Sqrt[n+1] f[ x, aa ket[ k, ll, n+1 ] ],
               f[ x___, a[k_, ll_ ], aa__ ket[k_, ll_, n_ ] ] :>
                        Sqrt[n] f[ x, aa ket[ k, ll, n-1 ] ] ,

(* and the adjoint  *)

               f[ bra[k_, ll_, n_ ] aa__, ad[k_, ll_], x___ ] :>
                        Sqrt[n] f[ aa bra[ k, ll, n-1 ], x ],
               f[ bra[k_, ll_, n_ ] aa__, a[k_, ll_ ], x___ ] :>
                        Sqrt[n+1] f[ aa bra[ k, ll, n+1 ], x ]
     }

(* The concise Form *)

consiseForm :=  f[ket["x", h3_, o_] ket["y", h3_, p_]] *
                f[bra["x", bh3_, bo_] bra["y", bh3_, bp_]] :>
       Infix[{" | ",o,", ",p," > < ",bo,", ",bp," |"}, "" ]

(*
   The beam-splitter has five entries: the first two are the input modes,
   the third and the fourth are the respective output modes and the fifth
   entry gives the beam-splitters coefficient.
*)

beamSplitter[ am_, bm_, cm_, dm_, eta_ ] :=
  { ad[ "x", am ] :> Sqrt[eta] ad[ "x", cm ] + Sqrt[1 - eta] ad[ "x", dm ],
     a[ "x", am ] :> Sqrt[eta]  a[ "x", cm ] + Sqrt[1 - eta]  a[ "x", dm ],
    ad[ "x", bm ] :> Sqrt[1 - eta] ad[ "x", cm ] - Sqrt[eta] ad[ "x", dm ],
     a[ "x", bm ] :> Sqrt[1 - eta]  a[ "x", cm ] - Sqrt[eta]  a[ "x", dm ],
    ad[ "y", am ] :> Sqrt[eta] ad[ "y", cm ] + Sqrt[1 - eta] ad[ "y", dm ],
     a[ "y", am ] :> Sqrt[eta]  a[ "y", cm ] + Sqrt[1 - eta]  a[ "y", dm ],
    ad[ "y", bm ] :> Sqrt[1 - eta] ad[ "y", cm ] - Sqrt[eta] ad[ "y", dm ],
     a[ "y", bm ] :> Sqrt[1 - eta]  a[ "y", cm ] - Sqrt[eta]  a[ "y", dm ],
     ket[ "x", am, n_ ] :> ket[ "x", cm, n ],
     ket[ "y", am, n_ ] :> ket[ "y", cm, n ],
     ket[ "x", bm, n_ ] :> ket[ "x", dm, n ],
     ket[ "y", bm, n_ ] :> ket[ "y", dm, n ]
  }

(*
   The polarisation filter performs a rotation over an angle theta (the
   second entry) on mode am (the first entry). The two directions of
   polarisation are called "x" and "y".
*)

polarizeFilter[ am_, theta_, aam_ ] := {
```

```
    ad["x", am] :>   Cos[theta] ad["x", aam ] + Sin[theta] ad["y", aam ],
     a["x", am] :>   Cos[theta]  a["x", aam ] + Sin[theta]  a["y", aam ],
    ad["y", am] :> - Sin[theta] ad["x", aam ] + Cos[theta] ad["y", aam ],
     a["y", am] :> - Sin[theta]  a["x", aam ] + Cos[theta]  a["y", aam ],
    ket[ "x", am, n_ ] :> ket[ "x", aam, n ],
    ket[ "y", am, m_ ] :> ket[ "y", aam, m ]
  }

(* The takeAdjoint rule changes kets into bras.  *)

takeAdjoint := ket[ k_, l_, n_ ] :> bra[ k, l, n ]

(*
   Polarisation insensitive detector. It assumes detectors cannot distinguish
   between a pulse containing one or more photons. The POVM acts on mode l and
   loss is Sqrt[1-efficiency^2]. The perfect detector therefore corresponds to
   loss=0. Loss is the AMPLITUDE loss.
*)

povMeasure[ b_, loss_ ] := f[ z__ ket[ "x", b, n_] ket[ "y", b, m_ ]] :>
  Sqrt[1-loss^( 2(n + m) )] f[z ket[ "x", b, n ] ket[ "y", b, m ]]


povAngle[ mode_, loss_ ] := {
   f[ z__ ket["x", mode, n_ ]] :> loss^n f[ z ket["x", mode, n]],
   f[ z__ ket["y", mode, m_ ]] :> loss^m f[ z ket["y", mode, m]]
    }

povHit[pol_, mode_, loss_ ] :=
  f[ z__ ket[pol, mode, m_ ]] :> Sqrt[1-loss^(2m)] f[ z ket[pol, mode, m]]

povMiss[pol_, mode_, loss_ ] :=
  f[ z__ ket[pol, mode, m_ ]] :> loss^m f[ z ket[pol, mode, m]]

(* partialTrace[ mode_ ] takes the partial trace of mode "mode".  *)

partialTrace[ mode_ ] := {
f[l__ bra[ "y", mode, n_ ] u__ ] f[w__ ket[ "y", mode, m_ ] v__ ] :>
  f[l u] f[w v] Krondelta[ n, m ],
f[l__ bra[ "x", mode, n_ ] u__ ] f[w__ ket[ "x", mode, m_ ] v__ ] :>
f[l u] f[w v] Krondelta[ n, m ]
 }

xTrace[ mode_ ] := {
f[l__ bra[ "x", mode, n_ ] u__ ] f[w__ ket[ "x", mode, m_ ] v__ ] :>
f[l u] f[w v] Krondelta[ n, m ]
 }

yTrace[ mode_ ] := {
f[l__ bra[ "y", mode, n_ ] u__ ] f[w__ ket[ "y", mode, m_ ] v__ ] :>
f[l u] f[w v] Krondelta[ n, m ]
```

```
 }

(*
   In the procedure "myCalc", ketval is assigned the function corresponding
   to the unitary transformation of creating EPR-pairs. After expansion to a
   certain order the beam-splitter on modes "b" and "c" is applied, the
   polarisation rotation on mode "a" is performed and the creation operators
   are calculated. Then the Hermitian conjugate is computed. With this we
   can define the density operator (densval). But first we apply the POVM on
   "u" and "v". After expanding the density operator we take the partial
   traces of "a", "u" and "v", which gives us the output mode "d".
*)

(*
   In the line "ketval = f[ myExp [ ...", tau corresponds to the normal
   ordered function tau/|tau| tanh(|tau|), and NOT the tau due to the
   Hamiltonian.
*)

myCalc := Block[ {ketval, braval, densval, myval},
  ketval = f [ myExp1 [ tau f [ ad[ "x", "a"], ad [ "y", "b"] ] -
                  tau f [ ad[ "y", "a"], ad [ "x", "b"] ] ],
       myExp2 [ tau f [ ad[ "x", "c"], ad [ "y", "d"] ] -
                  tau f [ ad[ "y", "c"], ad [ "x", "d"] ] ],
                  ket[ "x", "a", 0 ] ket [ "x", "b", 0] *
                  ket[ "y", "a", 0 ] ket [ "y", "b", 0] *
                  ket[ "x", "c", 0 ] ket [ "x", "d", 0] *
                  ket[ "y", "c", 0 ] ket [ "y", "d", 0] ];

  ketval = ketval /. beamSplitter[ "b", "c", "u", "v", 1/2 ];
  ketval = ketval /. polarizeFilter["a", theta, "a1" ];
  Print[ "<< beamSplitters and polarizeFilter >>" ];

  ketval = ketval /. expandExp1;
  ketval = ExpandAll[ ketval /. cutEnd ];
  ketval = ketval //. numReduce;
  ketval = ExpandAll[ ketval /. cutEnd ];
  ketval = ketval //. numReduce;
  Print[ "<< expandExp1 >>" ];

  ketval = ketval /. expandExp2;
  ketval = ExpandAll[ ketval /. cutEnd ];
  ketval = ketval //. numReduce;
  Print[ "<< expandExp2 >>" ];

  Save[ "ketval.m", ketval ];

  ketval = ketval /. povMeasure[ "u", loss1 ];
  ketval = ketval /. povMeasure[ "v", loss2 ];

  ketval = ketval /. povHit [ "x", "a1", loss3];
```

```
  ketval = ketval /. povMiss [ "y", "a1", loss3];

  Print[ "<< POVM's >>" ];

  ketval = ExpandAll[ ketval ];

  Print[ "<< expansion >>" ];

  braval = ketval /. takeAdjoint;

  ketlen = Length [ ketval ];
  bralen = Length [ braval ];
  Print[ "<< ketlen = ", ketlen, " >>" ];

  densval = Sum [
       Print [ N[ 100. jj / ketlen, 3] , " %" ];
       Sum [
        myval = ketval [[jj]] braval [[kk]];
myval = myval /. {tau^6 -> 0};
        myval = myval //. partialTrace[ "u" ];
        myval = myval //. partialTrace[ "v" ];
myval = myval //. partialTrace[ "a1"],
          { kk, 1, bralen }
           ],
          { jj, 1, ketlen }
                ];

  Save[ "densval4cas1.m", densval ];

  densval = Collect [ densval, tau ];
  densval = densval /. consiseForm;
  densval = Simplify[ densval, TimeConstraint -> Infinity ]
]
```

I

Genetic Algorithms

In this appendix I review genetic algorithms. In the first section I present the basics behind these algorithms [120], and in the second section I describe the so-called *differential evolution*-approach by Price and Storn [137]. It was this method I used in chapter VII. The fortran code of this application is given in the last section.

1 Genetic algorithms

Genetic algorithms can be used to find the best solution to a given problem. As the name already suggests, it is based on 'natural selection' over several generations of a 'population' of solutions to the problem. It works as follows.

Suppose we have a problem with a set of possible solutions. This set generally spans a high dimensional solution space. For instance, when the solutions to a particular problem are given by $x(\theta; a, b, c) = a\cos\theta + b\sin\theta + c$ (with θ its variable and a, b and c constants), the solution space is a three-dimensional space spanned by the vectors (a, b, c). In addition, we have a selection criterion which gives us a measure of the 'fitness' of a solution. For example, we might define the fitness of a solution as a distance measure between a function y and a solution x. The smaller this distance, the fitter the solution. One such fitness measure for the example above may be given by

$$f(a, b, c) = \int_0^{2\pi} |y(\theta) - x(\theta; a, b, c)|^2 d\theta \,. \tag{I.1}$$

In the genetic representation, the numbers a, b and c are the *genes* of a particular solution.

In any genetic algorithm, we first select a population of n solutions

$$P(t_0) = \{x_1^{t_0}, \ldots, x_n^{t_0}\} \,, \tag{I.2}$$

where P is the population (taken at the initial time t_0) and $x_i^{t_0}$ a candidate solution to the problem. The solutions are evaluated using the fitness measure, yielding a measure set $F(t_0) = \{f_1^{t_0}, \ldots, f_n^{t_0}\}$, where $f_i^{t_0}$ is a number associated with the fitness of solution $x_i^{t_0}$.

Depending on the details of our problem, we are looking for the smallest or the largest number $f_i^{t_0}$. Suppose better fitness means a smaller f, then we choose the

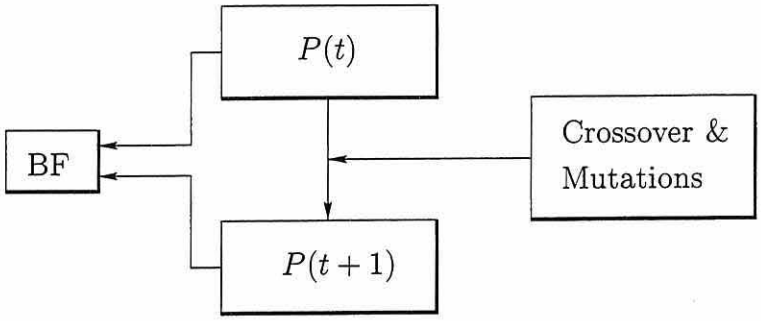

Figure I.1: Flowchart for genetic algorithms. At time t the fitness of the members of a population $P(t)$ is evaluated according to some criterion. The best fitting member (BF) of $P(t)$ is recorded. Subsequently, a new population (the next generation) $P(t+1)$ is formed from $P(t)$. In addition, crossover and mutations diversify the next generation. This generation is again tested for the best fitting member, which is recorded as the fittest if it defeats the previous fittest.

solution $x_k^{t_0}$ corresponding to the smallest $f_k^{t_0}$ in our measure set. Let $x_k^{t_0}$ be the best solution for this population. It will be stored in the memory. This memory slot is reserved for the best solution, and it will be updated if some solution x_l^t from a later generation outperforms $x_k^{t_0}$.

The next step is the crucial step of genetic algorithms. The old generation, the population $P(t_0)$, will now determine the next generation $P(t)$ of solutions. Low fitness solutions from $P(t_0)$, however, will be discarded: evolution has destined them to die. Thus only the fittest individuals from a population will generate a new population: they are making babies.

Just as in the offspring of real populations, the children inherit the traits of their parents. But they differ too. In biology, organisms often produce genetically different offspring by using *crossover*, in which genes of the parents are mixed. In a population of solutions to a mathematical problem, we can also introduce crossover. For example, when two solutions $x_1^{t_0}$ and $x_2^{t_0}$ are determined by the vectors

$$x_1^{t_0} = (a_1, b_1, c_1, d_1, e_1) \qquad \text{and} \qquad x_2^{t_0} = (a_2, b_2, c_2, d_2, e_2)\ ,$$

crossover can produce a child-solution $x_j^t = (a_1, b_2, c_2, d_1, e_1)$, where 'genes' a, d and e are taken from parent 1 and 'genes' b and c from parent 2.

A second mechanism for inducing changes in subsequent generations is *mutation*. In real life, background radioactivity or free radicals induce changes in the DNA structure which will affect future generations. Most of the time these changes are a setback and will be eliminated in the next generation, but once in a while it increases an individual's fitness.

Using crossover and mutation, a new population is formed. This is the next generation. The population size is usually held constant, but this is not nec-

essary. The new population is again evaluated, yielding a measure set $F(t) = \{f_1^t, \ldots, f_n^t\}$. The fittest solution of this generation is chosen, and will replace the previous one if it is better.

Next, the process of offspring generation using crossover and mutation is repeated to generate a population $P(t+1)$, the members of which are again tested for their fitness, and so on. This process terminates after a given number of generations (see also figure I.1).

Since less fit members of the parent population do not make children, the genetic algorithm does not conduct a random search in the solution space. Also, since the best overall solution is recorded, it is not necessarily a member of the final population. For example, looking for artistic and scientific traits in the human population, The genetic algorithm would probably select Leonardo da Vinci, even though he died in 1519.

In short, an optimisation algorithm is a *genetic algorithm* if it meets the following criteria:

1. The problem must allow a genetic representation for potential solutions. For example, a vector has a genetic representation, in which the entries correspond to genes. As we have seen, crossover exchanges these genes.

2. An initial population has to be created and a mechanism for producing the next generation must be given.

3. A fitness measure has to be defined in order to guide the evolution. It plays the role of the environment in the sense that it induces 'natural selection'.

4. The algorithm needs a crossover and mutation mechanisms to allow the generations to evolve.

5. Finally, the algorithm needs parameters like population size, number of generations, mutation probabilities, etc.

In the next section, I will take a closer look at differential evolution.

2 Differential evolution

The main difference between genetic algorithms and differential evolution lies in the parent-child relationship. In genetic algorithms described in the previous section, two[1] parents pass their genes on to a child by means of *uniform* crossover. This means that all parents have equal probability to pass on their genes to their children (note the distinction with the *unequal* probability for members of a population of having offspring at all).

[1] Or more: why let biology restrict this mathematical protocol?

In differential evolution, however, a fitter parent has a higher probability of passing on its genes to the child. The child is thus more closely related to its fitter parent, and is likely to have a good fitness rating. This accommodates a more directed evolution, in which successful branches are biased [137].

3 FORTRAN CODE FOR LITHOGRAPHY

I used a genetic algorithm to optimise one-dimensional quantum lithography used in the creation of a trench function. I have omitted the fitness function because it is quite lengthy. It can easily be generated using MATHEMATICA.

```
      program genetic
c     uses a GENETIC search algorithm
      implicit real*8 (a-h, o-z)
      integer time
      real RAN
      external time, RAN

c     number of parameters to fit
      parameter (n = 21)
c     maximum number of generations
      parameter ( gen_max = 1000 )
c     population size
      parameter ( NP = n*10 )
c     scaling mutation parameter
      parameter ( Fscale = 0.5 )
c     recombination parameter
      parameter ( CR = 0.1 )
      real*8 x1(n, NP), x2(n, NP), trial(n), cost(NP), psmallest(n)
c
c     set the random seed
      iseed = time()
c     iseed = 950015448
c
c     initialization
      do 800 i = 1, NP
        do 700 j = 1, n-1
          trial(j) = 2.0d0*RAN(iseed) - 1.0d0
          x1(j,i) = trial(j)
700     continue
c     initialise the exposure time parameter
        trial(21) = 1.0d-2*RAN(iseed)
        x1(21,i) = trial(21)
        cost(i) = f( n, trial )
c       write(*,*) i, trial, cost(i)
800   continue
c     initialise 'smallest'
      jsmallest = 1
      smallest = cost(1)*10
```

```
c       halt after 'gen_max' generations
        do 2000 jgen = 1, gen_max
c       loop through the population
          do 1800 i = 1, NP
c       mutate and recombine.
c       randomly generate three *different* vectors from each other and 'i'
1001        ia = 1.0 + NP*RAN(iseed)
            if ( ia.eq.i ) goto 1001
1002        ib = 1.0 + NP*RAN(iseed)
            if ( (ib.eq.i) .or. (ib.eq.ia) ) goto 1002
1003        ic = 1.0 + NP*RAN(iseed)
            if ( (ic.eq.i) .or. (ic.eq.ia) .or. (ic.eq.ib) ) goto 1003
c       randomly pick the first parameter
            j = 1.0 + RAN(iseed)*n
c       load n parameters into trial; perform n - 1 binomial trials
            do 1300 k = 1, n
              if ( (RAN(iseed).le.CR) .or. (k.eq.n) ) then
c       source for 'trial(j)' is a random vector plus weighted differential..
                trial(j) = x1(j,ic)+Fscale*( x1(j,ia) - x1(j,ib) )
              else
c       ... or the trial parameter comes from 'x1(j,i)' itself.
                trial(j) = x1(j,i)
              end if
c       get the next 'j' modulo n
              j = j + 1
              if ( j.gt.n ) j = 1
c       last parameter 'k=n' comes from noisy random vector.
1300        continue
c       evaluate/select.
c       score this trial
            score = f ( n, trial )
            if ( score.le.cost(i) ) then
              do 1400 j = 1, n
c       move trial to secondary vector (for next generation) ..
                x2(j,i) = trial(j)
1400          continue
              cost(i) = score
            else
              do 1450 j = 1, n
c       ... or place the old population member there
                x2(j,i) = x1(j,i)
1450          continue
            end if
1800      continue

c       end of population, swap arrays; move x2 onto x1 for next round
          do 1500 i = 1, NP
            do 1490 j = 1, n
              x1(j,i) = x2(j,i)
1490        continue
```

```
1500      continue

c       keep a record of progress so far
          do 1900 j = 1, NP
            if ( cost(j).lt. smallest ) then
              smallest = cost(j)
              jsmallest = j
              do 1600 kkk = 1, n
                psmallest(kkk) = x1(kkk,jsmallest)
1600         continue
             end if
c            write(*,*) jgen, j, cost(j)
1900        continue
c       display the progress each generation
          write(*,*) " gen", jgen, "     score=", float(smallest)

2000      continue
          xnorm = 0.0d0
          do 2050 i = 1, n
            xnorm = xnorm + psmallest(i)**2
2050      continue
          xnorm = dsqrt(xnorm)
          do 2100 i = 1, n
            psmallest(i) = psmallest(i) / xnorm
2100      continue
          write(*,*) " parameters:"

          stop
          end

          real*8  function f( n, trial )
          implicit real*8 (a-h, o-z)
          integer n
          real*8 trial(n)
          xnorm = 0.0d0
          do 100 i = 1, n-1
            xnorm = xnorm + trial(i)**2
100       continue
          xnorm = dsqrt ( xnorm )
c         renormalise the trials
          do 200 i = 1, n-1
            trial(i) = trial(i) / xnorm
200       continue

c         f = fitness function to be minimised
```

BIBLIOGRAPHY

[1] A. Acín, A. Adrianov, L. Costa, E. Jané, J.I. Latorre and R. Tarrach, *Generalized Schmidt decomposition and classification of three-quantum-bit states*, Phys. Rev. Lett. **85**, 1560 (2000).

[2] G.S. Agarwal and R. Boyd, private communication.

[3] P. Appell and J. Kampé de Fériet, *Fonctions Hypergéométriques et Hypersphériques*, Gauthier-Villars, Paris (1926).

[4] G.M. D'Ariano, C. Macchiavello and M.G.A. Paris, *Detection of the density matrix through optical homodyne tomography without filtered back projection*, Phys. Rev. A **50**, 4298 (1994).

[5] G.M. D'Ariano, L. Maccone, M.G.A. Paris and M.F. Sacci, *Optical Fock-state synthesizer*, Phys. Rev. A **61**, 053817 (2000).

[6] M. Artoni, U.P. Ortiz and J.L. Birman, *Photocount distribution of two-mode squeezed states*, Phys. Rev. A **43**, 3954 (1991).

[7] A. Aspect, P. Grangier and G. Roger, *Experimental tests of realistic local theories via Bell's theorem*, Phys. Rev. Lett. **47**, 460 (1981).

[8] A. Aspect, P. Grangier and G. Roger, *Experimental realisation of Einstein-Podolski-Rosen-Bohm gedankenexperiment: a new violation of Bell's inequalities*, Phys. Rev. Lett. **48**, 1804 (1982).

[9] V. Bargmann, Comm. Pure App. Math. **14**, 187 (1961).

[10] S.M. Barnett and P.L. Knight, *Thermofield analysis of squeezing and statistical mixtures in quantum optics* J. Opt. Soc. Am. B **2**, 467 (1985).

[11] J.S. Bell, *On the Einstein-Podolski-Rosen paradox*, Phys. **1**, 195 (1964); also in '*Speakable and unspeakable in quantum mechanics*', Cambridge University Press, Cambridge (1987).

[12] C.H. Bennett and G. Brassard, Proc. IEEE Int. Conf. Comp., IEEE New York (1984).

[13] C.H. Bennett and S.J. Wiesner, *Communication via one- and two-particle operators on Einstein-Podolski-Rosen states*, Phys. Rev. Lett. **69**, 2881 (1992).

[14] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W.K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolski-Rosen channels*, Phys. Rev. Lett. **70**, 1895 (1993).

[15] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin and W.K. Wootters, *Purification of noisy entanglement and faithful teleportation via noisy channels*, Phys. Rev. Lett. **76**, 722 (1996).

[16] C.H. Bennett, H.J. Bernstein, S. Popescu and B. Schumacher, *Concentrating partial entanglement by local operations*, Phys. Rev. A **53**, 2046 (1996).

[17] G. Björk, L.L. Sánchez Soto and J. Söderholm, *Entangled-state lithography: tailoring any pattern with a single state*, quant-ph/0011075 (2000).

[18] N. Bohr, *On the constitution of atoms and molecules*, Phil. Mag. **26**, 132 (1913).

[19] D. Boschi, F. De Martini and G. DiGiuseppe, *Test of the violation of local realism in quantum mechanics without Bell inequalities*, Phys. Lett. A **228**, 208 (1997).

[20] D. Boschi, S. Branca, F. De Martini, L. Hardy and S. Popescu, *Experimental realization of teleporting an unknown quantum state via dual classical and Einstein-Podolski-Rosen channels*, Phys. Rev. Lett. **80**, 1121 (1998).

[21] S. Bose, V. Vedral and P.L. Knight, *Purification via entanglement swapping and conserved entanglement*, Phys. Rev. A **60**, 194 (1999).

[22] A.N. Boto, P. Kok, D.S. Abrams, S.L. Braunstein, C.P. Williams and J.P. Dowling, *Quantum interferometric optical lithography: exploiting entanglement to beat the diffraction limit*, Phys. Rev. Lett. **85**, 2733 (2000).

[23] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter and A. Zeilinger, *Experimental quantum teleportation*, Nature **390**, 575 (1997).

[24] D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter, M. Zukowski and A. Zeilinger, *Reply to comment 'a posteriori teleportation'*, Nature **394**, 841 (1998).

[25] D. Bouwmeester, J.-W. Pan, H. Weinfurter and A. Zeilinger, *High-fidelity teleportation of independent qubits*, J. Mod. Opt. **47**, 279, Special Issue on the Physics of Quantum Information (2000).

[26] D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter and A. Zeilinger, *Observation of three-photon Greenberger-Horne-Zeilinger entanglement*, Phys. Rev. Lett. **82**, 1345 (1999).

[27] D. Bouwmeester, A. Ekert and A. Zeilinger (Eds.), *The physics of quantum information*, Springer Verlag (2000).

[28] B.H. Bransden and C.J. Joachain, *Introduction to quantum mechanics*, Longman, New York (1989).

[29] S.L. Braunstein and C.M. Caves, *Statistical distance and the geometry of quantum states*, Phys. Rev. Lett. **72**, 3439 (1994).

[30] S.L. Braunstein and A. Mann, *Measurement of the Bell operator and quantum teleportation*, Phys. Rev. A **51**, R1727 (1995).

[31] S.L. Braunstein, *Quantum teleportation without irreversible detection*, Phys. Rev. A **53**, 1900 (1996).

[32] S.L. Braunstein and H.J. Kimble, *A posteriori teleportation*, Nature **394**, 840 (1998).

[33] S.L. Braunstein and H.J. Kimble, *Teleportation of continuous quantum variables*, Phys. Rev. Lett. **80**, 869 (1998).

[34] S.L. Braunstein, *Quantum computation*, tutorial. In: *Quantum Computation: Where Do We Want to Go Tomorrow?*, S.L. Braunstein (Ed.), Wiley-VCH, Weinheim (1999).

[35] S.L. Braunstein, *Squeezing as an irreducible resource*, quant-ph/9904002 (1999).

[36] S.L. Braunstein, H.-K. Lo (Eds.) and P. Kok (Ass. Ed.), *Experimental proposals for quantum computation*, Fort. Phys. **48**, 9-11 (2000); *Scalable quantum computers; paving the way to realization*, S.L. Braunstein, H.-K. Lo (Eds.) and P. Kok (Ass. Ed.) Wiley-VCH (forthcoming).

[37] H.R. Brown and R. Harré (Eds.), *Philosophical foundations of quantum field theory*, Clarendon Press, Oxford (1988).

[38] S.R.J. Brück, S.H. Zaidi, X. Chen and Z. Zhang, *Interferometric lithography - from periodic arrays to arbitrary patterns*, Microelectron. Eng. **42**, 145 (1998).

[39] G. Buskes and A. van Rooij, *Topological spaces*, Undergraduate texts in mathematics, Springer Verlag (1997).

[40] C.M. Caves, *Quantum linits on noise in linear-amplifiers*, Phys. Rev. D **26**, 1817 (1982).

[41] C.M. Caves and B.L. Schumaker, *New formalism for 2-photon quantum optics. 1. quadrature phases and squeezed states*, Phys. Rev. A **31**, 3068 (1985).

[42] C.M. Caves, C. Zhu, G.J. Milburn and W. Schleich, *Photon statistics of two-mode squeezed states and interference in four-dimensional phase space*, Phys. Rev. A **43**, 3854 (1991).

[43] M. Chaichian and R. Hagedorn, *Symmetries in quantum mechanics*, Graduate student series in physics, IoP publishing, Bristol (1998).

[44] J.F. Clauser, M.A. Horne, A. Shimony and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[45] C. Cohen-Tannoudji, B. Diu and F. Laloë, *Quantum mechanics*, Vols. I and II, Wiley, New York (1977).

[46] M. Dakna, J. Clausen, L. Knöll and D.-G. Welsch, *Generation of arbitrary quantum states of travelling fields*, Phys. Rev. A **59**, 1658 (1999).

[47] D. Deutsch, Proc. R. Soc. London A **400**, 97 (1985).

[48] D. Deutsch and R. Jozsa, Proc. R. Soc. London A **439**, 553 (1992).

[49] D. Dieks, *Communication by EPR devices*, Phys. Lett. **92**A, 271 (1982).

[50] G. DiGiuseppe, F. DeMartini and D. Boschi, *Experimental test of the violation of local realism in quantum mechanics without Bell inequalities*, Phys. Rev. A **56**, 176 (1997).

[51] P.A.M. Dirac, *The principles of quantum mechanics*, Oxford University Press (1930).

[52] V.V. Dodonov, V.I. Man'ko and V.V. Semjonov, *The density matrix of the canonically transformed multidimensional Hamiltonian in the Fock basis*, Il Nuovo Cimento, **83**, 145 (1984).

[53] V.V. Dodonov, O.V. Man'ko and V.I. Man'ko, *Multidimensional Hermite polynomials and photon distribution for polymode mixed light*, Phys. Rev. A **50**, 813 (1994).

[54] V.V. Dodonov, O.V. Man'ko, V.I. Man'ko and P.G. Polynkin, *Linear optical transformator of the photon distribution function*, SPIE **2799**, 230 (1996).

[55] W. Dür, G. Vidal and J.I. Cirac, *Three qubits can be entangled in two inequivalent ways*, Phys. Rev. A **62**, 062315 (2000).

[56] A. Einstein, B. Podolski and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777 (1935).

[57] A.K. Ekert, *Quantum cryptography based on Bell's theorem*, Phys. Rev. Lett. **67**, 661 (1991).

[58] A.K. Ekert, *Distributed Quantum Computation over Noisy Channels* quant-ph/9803017 (1998).

[59] A. Erdély (ed.), *Bateman Manuscript Project*, McGraw-Hill, New York (1953).

[60] R.A. Fisher, M.M. Nieto and V.D. Sandberg, *Impossibility of naively generalizing squeezed coherent states*, Phys. Rev. D **29**, 1107 (1984).

[61] R.P. Feynman, *Lectures on Physics*, Volume III, Addison Wesley (1965).

[62] E.J.S. Fonseca, C.H. Monken and S. Pádua, *Measurement of the de Broglie wavelength of a multiphoton wave packet*, Phys. Rev. Lett. **82**, 2868 (1999).

[63] C.A. Fuchs, PhD thesis, University of New Mexico (1996).

[64] C.A. Fuchs and A. Peres, *Quantum-state disturbance versus information gain: Uncertainty relations for quantum information*, Phys. Rev. A **53**, 2038 (1996).

[65] C.A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu and A. Peres, *Optimal eavesdropping in quantum cryptography. I. Information bound and optimal strategy*, Phys. Rev. A **56**, 1163 (1997).

[66] C.A. Fuchs and A. Peres, *Quantum theory needs no 'interpretation'*, Physics Today, March 70 (2000).

[67] A. Furusawa, J.L. Sørensen, S.L. Braunstein, C.A. Fuchs, H.J. Kimble and E.S. Polzik, *Unconditional quantum teleportation*, Science **282**, 706 (1998).

[68] F.R. Gantmacher, The theory of matrices, Vols. I and II, Chelsea Publishing Company, New York (1959).

[69] R. Gilmore, *Lie groups, Lie algebras, and some of their applications*, Krieger Publishing Company, Florida (1994).

[70] M. Göppert-Mayer, Ann. Phys. **5**, 273 (1931).

[71] D.M. Greenberger, M.A. Horne and A. Zeilinger, *Going beyond Bell's theorem*, in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, M. Kafatos (Ed.), 69–72, Kluwer Academic Publishers (1989).

[72] D.M. Greenberger, M.A. Horne and A. Zeilinger, *Nonlocality of a single photon?*, Phys. Rev. Lett. **75**, 2064 (1995).

[73] D.M. Greenberger, M.A. Horne and A. Zeilinger, *Tangled concepts about entangled states*, in *Quantum interferometry: proceedings of an Adriatico workshop, Triëste*, F. DeMartini, G. Denardo and Y. Shih (Eds.), VCH Publishing, Weinheim (1996).

[74] L. Grover, Proc. 28 Ann. ACM Symp. on Th. Comp., ACM Press New York, 212 (1996).

[75] L. Grover, *Quantum Telecomputation*, quant-ph/9704012 (1997).

[76] R. Haag, *Local quantum physics*, Springer Verlag (1991).

[77] F. Halzen and A.D. Martin, *Quarks & Leptons: an introductory course in modern particle physics*, John Wiley and sons, New York (1984).

[78] L. Hardy, *Non-locality for two particles without inequalities for almost all entangled states*, Phys. Rev. Lett. **71**, 1665 (1993).

[79] L. Hardy, *Nonlocality of a single photon revisited*, Phys. Rev. Lett. **73**, 2279 (1994).

[80] L. Hardy, Phys. Rev. Lett. **75**, 2065 (1995).

[81] G. Harel, G. Kurizki, J.K. McIver and E. Coutsias, *Optimized preparation of quantum states by conditional measurements*, Phys. Rev. A **53**, 4534 (1996).

[82] J.L. Heilbron, *J.J. Thomson and the Bohr atom*, Physics Today, April 303 (1977).

[83] C.W. Helstrom, *Quantum detection and estimation theory*, Academic Press, New York (1976).

[84] J. Hilgevoord and J. Uffink, *Uncertainty in prediction and in inference*, Foundations of Physics, **21**, 323 (1991).

[85] J. Hilgevoord, *Foundations of quantum mechanics*, Utrecht University syllabus (in Dutch) (1993).

[86] A.S. Holevo, *Probabilistic and statistical aspects of quantum theory*, North Holland, Amsterdam (1982).

[87] M. Horodecki, P. Horodecki and R. Horodecki, *Separability of mixed states: Necessary and sufficient conditions*, Phys. Lett. A. **223**, 1 (1996).

[88] P. Horodecki, *Separability criterion and inseparable mixed states with positive partial transposition*, Phys. Lett. A. **232**, 333 (1997).

[89] M. Horodecki, P. Horodecki and R. Horodecki, *Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature?*, Phys. Rev. Lett. **80**, 5239 (1998).

[90] J. Jacobson, G. Björk, I. Chuang and Y. Yamamoto, *Photonic De Broglie waves*, Phys. Rev. Lett. **74**, 4835 (1995).

[91] J. Javanainen and P.L. Gould, *Linear intensity dependence of a two-photon transition rate*, Phys. Rev. A **41**, 5088 (1990).

[92] R. Jozsa, D.S. Abrams, J.P. Dowling and C.P. Williams, *Quantum clock synchronization based on shared prior entanglement*, Phys. Rev. Lett. **85**, 2010 (2000).

[93] J. Kim, S. Takeuchi, Y. Yamamoto and H. H. Hogue, *Multiphoton detection using visible light photon counter*, App. Phys. Lett. **74**, 902 (1999).

[94] Y.H. Kim, S.P. Kulik and Y. Shih, *High-intensity pulsed source of spacetime and polarization double-entangled photon pairs*, Phys. Rev. A **62**, 011802 (2000).

[95] Y.H. Kim, S.P. Kulik and Y. Shih, *Quantum teleportation with a complete Bell state measurement*, quant-ph/0010046 (2000).

[96] M. Klauderer, *Modes in n-dimensional first order systems*, J. Math. Phys. **34**, 4221 (1993).

[97] D.N. Klyshko, Sov. J. Quantum Electron. **7**, 591 (1977).

[98] E. Knill, R. Laflamme and G.J. Milburn, *Efficient linear optics quantum computation*, quant-ph/0006088 (2000).

[99] P. Kok and S.L. Braunstein, *Postselected versus non-postselected quantum teleportation using parametric down-conversion*, Phys. Rev. A **61**, 42304 (2000).

[100] P. Kok and S.L. Braunstein, *Limitations on the creation of maximal entanglement*, Phys. Rev. A, **62**, 064301 (2000).

[101] P. Kok and S.L. Braunstein, *Entanglement swapping as event-ready entanglement preparation*, Fort. Phys. **48**, 5-7, 553 (2000).

[102] P. Kok and S.L. Braunstein, *Event-ready entanglement*, in *Relativistic Quantum Measurement and Decoherence*, H.-P. Breuer and F. Petruccione (Eds.), Lecture Notes in Physics, pp. 15–29, Springer-Verlag (2000).

[103] P. Kok and S.L. Braunstein, *Detection devices in entanglement-based state preparation*, to appear in Phys. Rev. A (2000).

[104] P. Kok, A.N. Boto, D.S. Abrams, C.P. Williams, S.L. Braunstein and J.P. Dowling, *Quantum lithography*, in: *Quantum, Communication, Measurement and Computing 3*, O. Hirota and P. Tombesi (Eds.) Kluwer Academic/Plenum Publishers (forthcoming).

[105] P. Kok, A.N. Boto, D.S. Abrams, C.P. Williams, S.L. Braunstein and J.P. Dowling, *Quantum interferometric optical lithography: towards arbitrary two-dimensional patterns*, quant-ph/0011088 (2000).

[106] P. Kok and S.L. Braunstein, *Multi-dimensional Hermite polynomials in quantum optics*, quant-ph/0011114 (2000).

[107] K. Kraus, *States, effects and operations: fundamental notions of quantum theory*, Springer Berlin (1983).

[108] P.G. Kwiat, A.M. Steinberg, R.Y. Chiao, P.H. Eberhard and M.D. Petroff, *Absolute efficiency and time-response measurement of single-photon detectors*, Applied Optics **33**, 1844 (1994).

[109] P.G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A.V. Segienko and Y. Shih, *New high-intensity source of polarization-entangled photon pairs*, Phys. Rev. Lett. **75**, 4337 (1995).

[110] P.G. Kwiat and R. Hughes, private communication (1998).

[111] P.G. Kwiat, E. Waks, A.G. White, I. Appelbaum and P.H. Eberhard, *Ultra-bright source of polarization-entangled photons*, Phys. Rev. A **60**, R773 (1999).

[112] R. Loudon, *The quantum theory of light*, Oxford Science publications, Oxford (1983).

[113] R. Loudon and P.L. Knight, *Squeezed light*, J. Mod. Opt. **34**, 709 (1987).

[114] N. Lütkenhaus, J. Calsamiglia and K-A. Suominen, *Bell measurements for teleportation*, Phys. Rev. A. **59**, 3295 (1999).

[115] C.A. Mack, *Trends in optical lithography*, Opt. Phot. News **7**, 29 (1996).

[116] M. Mansuripur and R. Liang, *Projection photolithography*, Opt. Phot. News **11**, 36 (2000).

[117] S. Massar and S. Popescu, *Optimal extraction of information from finite quantum ensembles*, Phys. Rev. Lett. **74**, 1259 (1995).

[118] S. Massar and S. Popescu, *Amount of information obtained by a quantum measurement*, Phys. Rev. A **61**, 062303 (2000).

[119] E. Merzbacher, *Quantum Mechanics*, third ed., Wiley New York (1998).

[120] Z. Michalewicz, *Genetic algorithms + data structures = evolution programs*, Artificial Intelligence Series, Springer Verlag (1992).

[121] M.M. Mizrahi, *Generalized Hermite polynomials*, J. Comp. App. Math. **I**, 273 (1975).

[122] J. von Neumann, *Mathematical foundations of quantum mechanics*, Princeton University Press (1955); originally published in German in 1932.

[123] T.D. Newton and E.P. Wigner, , Rev. Mod. Phys. **21**, 400 (1949).

[124] M.A. Nielsen, E. Knill and R. Laflamme, *Complete quantum teleportation by nuclear magnetic resonance*, Nature **396**, 52 (1998).

[125] M. Oberparleiter and H. Weinfurter, *Cavity-enhanced generation of polarization-entangled photon pairs*, Opt. Comm. **183**, 133 (2000).

[126] J-W. Pan, D. Bouwmeester, H. Weinfurter and A. Zeilinger, *Experimental entanglement swapping: Entangling photons that never interacted*, Phys. Rev. Lett. **80**, 3891 (1998).

[127] J-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter and A. Zeilinger, *Experimental test of quantum non-locality in three-photon Greenberger-Horne-Zeilinger entanglement*, Nature **403**, 515 (2000).

[128] M.G.A. Paris, M.B. Plenio, S. Bose, D. Jonathan and G.M. D'Ariano, *Optical Bell measurement by Fock filtering*, Phys. Lett. A **273**, 153 (2000).

[129] H. Paul, P. Törmä, T. Kiss and I. Jex, *Photon chopping: new way to measure the quantum state of light*, Phys. Rev. Lett. **76**, 2464 (1996).

[130] M. Pavičić: *Event-ready entanglement preparation*, in *Quantum Interferometry*, F. De Martini, G. Denardo and Y. Shih (Eds.), VCH Publishing Division I, New York (1996).

[131] A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers, Dordrecht (1995).

[132] A. Peres, *Higher order Schmidt decompositions*, Phys. Lett. A **220**, 16 (1995).

[133] A. Peres, *Nonlocal effects in Fock space*, Phys. Rev. Lett. **74**, 4571 (1995); Erratum Phys. Rev. Lett. **76**, 2005 (1996).

[134] A. Peres, Separability criterion for density matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[135] J. Perina jr., B.E.A. Saleh and M.C. Teich, *Multiphoton absorption cross section and virtual-state spectroscopy for the entangled n-photon state*, Phys. Rev. A **57**, 3972 (1998).

[136] S. Popescu, *Bell's inequalities versus teleportation: what is nonlocality*, Phys. Rev. Lett. **72**, 797 (1994).

[137] K. Price and R. Storn, *Differential evolution*, *Dr. Dobb's Journal*, April, p. 18ff, (1997).

[138] J.G. Rarity and P.R. Tapster, *2-Color photons and nonlocality in 4th-order interference*, Phys. Rev. A **41**, 5139 (1990).

[139] Lord Rayleigh, Phil. Mag. **8**, 261 (1879).

[140] M. Reck, A. Zeilinger, H.J. Bernstein and P. Bertani, *Experimental realization of any discrete unitary operation*, Phys. Rev. Lett. **73**, 58 (1994).

[141] M. Redhead, *Incompleteness, nonlocality and realism*, Clarendon Press, Oxford (1987).

[142] F. Reif, *Fundamentals of statistical and thermal physics*, McGraw-Hill international editions, Singapore (1965).

[143] M.H. Rubin, *Entanglement and state preparation*, Phys. Rev. A **61**, 022311 (2000).

[144] L.H. Ryder, *Quantum field theory*, Cambridge University Press, Cambridge (1996).

[145] E. Santos, *Nonlocality of a single photon, comment*, Phys. Rev. Lett. **68**, 894 (1992).

[146] G. Schrade, V.M. Akulin, V.I. Man'ko and W.P. Schleich, *Photon statistics of a two-mode squeezed vacuum*, Phys. Rev. A **48**, 2398 (1993).

[147] B.L. Schumaker and C.M. Caves, *New formalism for 2-photon quantum optics. 2. mathematical foundation and compact notation*, Phys. Rev. A **31**, 3093 (1985).

[148] M.O. Scully and M.S. Zubairy, *Quantum Optics*, Cambridge University Press (1997).

[149] M.O. Scully, B.-G. Englert and C.J. Bednar, *Two-photon scheme for detecting the Bell basis using atomic coherence*, Phys. Rev. Lett. **83**, 4433 (1999).

[150] J.-P. Serre, *Représentations linéaires des groupes finis*, Collection Méthodes, Hermann Paris (1967).

[151] C.E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[152] B.-S. Shi, Y.-K. Jiang and G.-C. Guo, *Optimal entanglement purification via entanglement swapping*, Phys. Rev. A **62**, 054301 (2000).

[153] Y.H. Shih and C.O. Alley, *New type of Einstein-Podolski-Rosen-Bohm experiment using pairs of light quanta produced by optical parametric down conversion*, Phys. Rev. Lett. **61**, 2921 (1988).

[154] Y.H. Shih, A.V. Sergienko, M.H. Rubin, T.E. Kiess and C.O. Alley, *2-Photon entanglement in type II parametric down-conversion*, Phys. Rev. A **50**, 23 (1994).

[155] P.W. Shor, *Scheme for reducing decoherence in quantum computer memory*, Phys. Rev. A **52**, R2493 (1995).

[156] P.W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, S.I.A.M. J. Comp. **26**, 1484 (1997).

[157] D.T. Smithey, M. Beck, M.G. Raymer and A. Faridani, *Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum*, Phys. Rev. Lett. **70** 1244 (1993).

[158] S. Song, C.M. Caves and B. Yurke, *Generation of superpositions of classically distinguishable quantum states from optical back-action evasion*, Phys. Rev. A **41**, R5261 (1990).

[159] A.M. Steane, *Error correcting codes in quantum theory*, Phys. Rev. Lett. **77**, 793 (1995).

[160] S.M. Tan, D.F. Walls and M.J. Collett, *Nonlocality of a single photon*, Phys. Rev. Lett. **66**, 252 (1991).

[161] S.M. Tan, D.F. Walls and M.J. Collett, *Nonlocality of a single photon, reply*, Phys. Rev. Lett. **68**, 895 (1992).

[162] S. Takeuchi, J. Kim, Y. Yamamoto and H. H. Hogue, *Development of a high-quantum-efficiency single-photon counting system*, App. Phys. Lett. **74**, 1063 (1999).

[163] A. Trifonov, T. Tsegaye, G. Björk, J. Söderholm, E. Goobar, M. Atatüre and A.V. Sergienko, *Experimental demonstration of the relative phase operator*, J. Opt. B **2**, 105 (2000).

[164] R.D. Truax, *Baker-Campbell-Hausdorff relations and unitarity of SU(2) and SU(1,1) squeeze operators*, Phys. Rev. D **31**, 1988 (1985).

[165] L. Vaidman, *Teleportation of quantum states*, Phys. Rev. A **49**, 1473 (1994).

[166] L. Vaidman, *Nonlocality of a single photon revisited again*, Phys. Rev. Lett. **75**, 2063 (1995).

[167] L. Vaidman and N. Yoran, *Methods for reliable teleportation*, Phys. Rev. A **59**, 116 (1999).

[168] K. Vogel, V.M. Akulin and W.P. Schleich, *Quantum state engineering of the radiation field*, Phys. Rev. Lett. **71**, 1816 (1993).

[169] A. Vourdas and R.M. Weiner, *Photon-counting distribution in squeezed states*, Phys. Rev. A **36**, R5866 (1987).

[170] D.F. Walls, *Squeezed states of light*, Nature **306**, 141 (1983).

[171] D.F. Walls and G.J. Milburn, *Quantum optics*, Springer Verlag, Berlin (1994).

[172] H. Weinfurter, *Experimental Bell-state analysis*, Europhys. Lett. **25**, 559 (1994).

[173] H. Weinfurter, private communication (1998).

[174] R.F. Werner, *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, Phys. Rev. A **40**, 4277 (1989).

[175] H.M. Wiseman and G. J. Milburn, *Quantum theory of field-quadrature measurements*, Phys. Rev. A **47**, 642 (1993).

[176] B. de Wit and J. Smith, *Field theory in particle physics*, Vol. 1, North-Holland (1986).

[177] W.K. Wootters, *Statistical distance and Hilbert space*, Phys. Rev. D **23**, 375 (1981).

[178] W.K. Wootters and W. H. Zurek, *A single quantum cannot be cloned*, Nature **299**, 802 (1982).

[179] E. Yablonovich and R.B. Vrijen, *Optical projection lithography at half the Rayleigh resolution limit by two-photon exposure*, Opt. Eng. **38**, 334 (1999).

[180] H.P. Yuen, *Two-photon coherent states of the radiation field*, Phys. Rev. A **13**, 2226 (1976).

[181] H.P. Yuen and J.H. Shapiro, IEEE Trans. Inf. Theory **26**, 78 (1980).

[182] A. Zeilinger, M.A. Horne, H. Weinfurter and M. Żukowski, *Three-particle entanglements from two entangled pairs*, Phys. Rev. Lett. **78**, 3031 (1997).

[183] M. Żukowski, A. Zeilinger, M.A. Horne and A.K. Ekert, *"Event-ready-detectors" Bell experiment via entanglement swapping*, Phys. Rev. Lett. **71**, 4287 (1993).

[184] M. Żukowski and D. Kaszlikowski, *Entanglement swapping with PDC sources*, Acta Phys. Slov. **49**, 621 (1999).

INDEX