High dimensional spatial entanglement with photon pairs

Filippo M. Miatto

A Thesis submitted for the degree of Doctor of Philosophy

Department of Physics

University of Strathclyde

August 2012

To my family

Abstract

This work is a theoretical investigation of the process of spontaneous parametric down-conversion, which consists in pumping a crystal with a laser to create entangled pairs of photons. In particular, it is an investigation of the properties of the entangled pairs that are created and on how their entanglement can be measured. We approach some of these questions both analytically and numerically.

The work is divided in three parts. The first is an introductory part, that consists of three chapters in which we set the framework and give the tools that are needed to understand the rest of the work. The second part consists of two chapters and it is concerned with the bi-photon states that are generated in down-conversion, in particular it approaches the problem of quantifying the amount of entanglement that those pairs carry and to understand on what parameters it depends in an analytical (chapter 4, based on the paper EPJD **66** (7) 183 (2012)) and a numerical way (chapter 5, based on the paper "Spatial Schmidt modes in parametric down-conversion", submitted to EPJD, in press (2012)). The numerical approach allows to overcome all the approximations previously employed. We will learn that the strength of the entanglement depends on quantities such as the size of the crystal, the size of the pump and the phase mismatch between the photons. The third part is written over three chapters and it is concerned with detecting the entangled pairs in a way that exposes the highest amount of entanglement possible. Various strategies will be analysed, ranging from projective measurements (chapter 6, based on the paper Physical Review A **83**, 033816 (2011), and chapter 7, based on the paper EPJD **66** (7), 178 (2012)), to the use phase masks (chapter 8, based on the paper New J. Phys. **14**, 073046 (2012) and on my master thesis [1]). We will learn that in certain situations, the most common approximations found in literature can be quite restrictive and we address the problem of understanding under what conditions they could or shouldn't be used. The thesis ends with chapter 9, which is a resumé of the work, and an outlook to future developments, which then blurs into the acknowledgments.

Contents

II Generation of high dimensional entanglement   39

iv

Part I

Background

Chapter 1

Introduction

1.1 Themes and structure

1.1.1 Down-conversion

The main subject of this thesis is spontaneous parametric down-conversion (SPDC). SPDC is a way of creating highly entangled systems. It consists in pumping a nonlinear crystal with a laser (details in chapter 3) to create pairs of photons. A crystal is called "nonlinear" depending on the symmetry properties of its lattice. In particular, it is called nonlinear if in order to express the polarisation density, one needs to use higher order susceptibilities, on top of the linear susceptibility. This is a consequence of the symmetry of the crystal lattice. The result of the nonlinearity of the crystal and the strong pumping light are terms in the hamiltonian which annihilate a pump photon and create down-converted photons [2].The first nonlinear term destroys one photon with momentum $\mathbf{k}$ and creates two photons (from now on called *signal* and *idler*, indicated with subscripts i and s) with momenta that satisfy momentum conservation $\mathbf{k}_i + \mathbf{k}_s = \mathbf{k}$. At the end of chapter 3 we will see more in detail how these terms arise. The impossibility of telling where and when the photon pair was created inside the crystal is deeply linked to the amount of entanglement. Down-conversion is a reliable way of creating entangled pairs of photons, and the rate at which they can be created

2

(up to hundreds of thousands per second in a narrow (~ 10 nm) frequency band [2]) makes it extremely interesting for practical implementations [3, 4], hence the interest, but we will see that apart from practical uses, down-conversion can provide quantum states that will help the study of quantum mechanics and entanglement themselves.

1.1.2 Orbital angular momentum

Orbital angular momentum is a very convenient quantum observable, in the sense that it has advantageous properties that make it a good choice for being an entangled quantity, especially if one is seeking high dimensional entanglement. The reason for this is that it is a spatial observable which relies on a compact support, i.e. the conjugate variable to OAM is an angle, which is a coordinate that lies on the circle [5]. As a consequence, eigenstates of orbital angular momentum can be manipulated with relative confidence even with optics of limited size, even though it is a spatial observable. Additionally, its spectrum is discrete and unbounded, therefore the entanglement in OAM can span in principle an infinite number of different modes [6, 7]. Another reason why OAM is convenient is that the azimuthal dependence, at the level of sheer mathematics, can be handled very straightforwardly; this gives a great advantage when manipulating the equations. We will explore the properties of OAM in details in chapter 3.

In addition to OAM we will consider also the radial degree of freedom, to make our treatment of the subject more complete, but mainly to make it possible to understand radial modes and having a chance to exploit them, which is something that is still not common practice. In fact, experimentally, addressing the radial modes is affected by technical difficulties due to their representation in real and momentum space as wave functions. However, it is an important step that ought to be made: as the equations of motion for optical states are separable in polar coordinates, the dimensionality of the Hilbert space that hosts the OAM part of the system is multiplied by the number of radial modes that can be used. This can lead to a realisable spatial entanglement made of hundreds or even thousands of joint modes [8].

1.1.3 Thesis structure

As the description of SPDC can become rather complicated, various different assumptions and approximations are made throughout the chapters. The aim of this section is to summarise them and to put them into a structured table, so that

the reader will have in mind where to find the chapters he or she is interested in. The various approximations of the SPDC state that will be used are presented for the first time in this section, in order to deliver a basic notion of which states will be used and how.

A fairly general wavefunction for the SPDC state comprises two terms[1]: one that states that the down-converted photons "originate from a pump photon" via a parametric process (i.e. that leaves the crystal invariant and therefore that conserves the properties of the pump photons) and one that states their phase matching relation. Accounting analytically for both terms is very hard, and we adopted some strategies to avoid getting stuck. These strategies are always addressing the phase matching function. In chapter 4 we replace the phase matching function, which is a sinc function, with a gaussian. This is known as the *gaussian approximation*. In chapter 5 we compare numerically the gaussian approximation and the sinc form of the phase matching, and we will see what are the consequences of employing the gaussian approximation. Another approximation that can be chosen is to remove the phase matching function entirely (i.e. setting it to 1), which is equivalent to stating that we have an infinitely thin crystal. This is known as the *thin crystal approximation*. In chapter 6 we use this approximation and in chapter 7 we remove it by restricting the projection basis, i.e. by keeping only those basis elements that allow a closed form of the projection amplitudes without using the thin crystal approximation or the gaussian approximation. Anticipating some of the conclusions of chapter 7, we will see that an infinitely thin crystal generates an infinite number of entangled modes, but the detection apparatus will set an upper bound to the detectable ones. In real life experiments we will have both effects at the same time: the source will produce a limited number of modes, and the detection apparatus will be able to see some of them. In chapter 7 we also give an intuitive way of calculating the strength of the entanglement that will be measured in an experiment.

Let's see now more precisely where these approximations fit. The thesis is separated into a background part, a generation part and a detection part. In the background part we set the scene by introducing the subject, by explaining how to approach this thesis (chapter 1), and by giving some fundamental tools on quantum mechanics (chapter 2) and optics (chapter 3). In the generation part we are concerned with understanding what the Schmidt modes are, i.e. how strong the entanglement is, without worrying about how to measure it: in chapter 4

---

[1]The most general description of the SPDC state is not used in this thesis. For instance we are not treating the temporal dependence of the state, nor azimuthal anisotropies, nor transverse phase matching.

we answer analytically using the gaussian approximation and in chapter 5 we remove the approximation, but we will answer numerically. Here ends the generation part. In the detection part we are concerned with understanding how much entanglement we would see if we were to detect the SPDC state under various circumstances. The detection part starts with chapter 6, where we project the SPDC state onto a doubly-complete family of joint Laguerre-Gauss modes and to do this we employ the thin crystal approximation. This allows us to find an analytical result, but crystals realistically are not infinitely thin, and this effect, although small, could be noticed. In chapter 7 we remove the thin crystal approximation, we restore the sinc phase matching (albeit without any phase mismatch) and we analytically calculate the projection amplitudes on a partial family of joint LG modes, i.e. for $p = 0$, where p is the radial mode number. Although this is an incomplete result, as it doesn't address projection on higher p modes, it gives us great insight on the detection of perfectly phase matched down-converted pairs. Very interestingly, we also provide a geometrical argument that reaches the same conclusions. Being a heuristic argument, we consider it an approximation and call it the "geometrical approximation". Chapter 8 is treating detection from a different perspective, as it presents a way of measuring the azimuthal part of the entanglement, using sector phase plates, a technique which allows to address spatial correlations without affecting the overall count rate.

The approximations used in this thesis, the consequent restrictions and the analytical (A) or numerical (N) results can be summarised in the following table:

Chapter	Approximation	Restriction	Result
4	gaussian	no mismatch	A
5	none	none	N
6	thin crystal	no mismatch, no thick crystals	A
7	$p = 0$/geometrical	no mismatch, no higher p modes	A

The first two (4 and 5) being concerned with the generation of entangled modes and the second two (6 and 7) with their detection.

1.2   A bit of history

In this section I will contextualise this thesis by giving a historical introduction to entanglement, to downconversion and to orbital angular momentum. I'd like to mention that the way I like to think about the history of entanglement is not as a research for a mean to be exploited for technological purposes, but as something

much deeper and honourable: a quest for something still indecipherable, yet in front of our eyes, as clear as the sun, but utterly obscure to our intellect. During the years, are we getting a hold of more and more pieces of the puzzle, or are we plunging into a bottomless cave, the end of which will always be out of reach for our limited brains?

1.2.1 Ancient history

In the beginning was Einstein. Or at least, it was him to strongly support the claim that there was a problem with quantum mechanics. In 1927, at the Solvay conference he pointed out: [9, 10] "The scattered wave moving towards P [the screen] does not show any preferred direction. If $|\psi|^2$ were simply regarded as the probability that at a certain point a given particle is found at a given time, it could happen that *the same* elementary process produces an action *in two or several* places of the screen. But the interpretation, according to which $|\psi|^2$ expresses the probability that *this* particle is found at a given point, assumes an entirely peculiar mechanism of action at a distance which prevents the wave continuously distributed in space from producing an action in *two* places on the screen."

What this implies is that locality[2], together with the predictions of quantum mechanics, together with the assumption that quantum mechanics is complete (meaning that ψ contains everything there can be about a system, in Einstein's example about the scattered matter wave) yields a contradiction, because if we call e_A the event of particle being detected at A, and e'_B the event of the particle being detected at B we have:

$$0 = P(e_A \ \& \ e'_B|\psi) = P(e_A|e'_B,\psi)P(e'_B|\psi) \tag{1.1}$$

$$= P(e_A|\psi)P(e'_B|\psi) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \tag{1.2}$$

The use of ψ comes from the completeness assumption, the first and second equality signs come from the statistics of quantum mechanics and the request for simultaneous detection, the third from locality, the fourth from quantum mechanics. If one wants to save locality and one cannot deny the way nature appears to work, the conclusion is that quantum mechanics is incomplete, i.e. there

---

[2]which can be thought of a property of the way probabilities work in quantum mechanics, in particular if A and B are two spacelike separated regions of spacetime, the probability of an event e happening in A, $P(e_A)$ is equal to the probability of e happening in A *given* that another event e' happened in B: $P(e_A) = P(e_A|e'_B)$

is something beyond ψ, in other words ψ is not what should be used in the series of equations above, so that the fourth equality wouldn't give $\frac{1}{2} \times \frac{1}{2}$.

The view of Einstein changed during the subsequent years, in particular he appears, in the context of the EPR paradox, to have stumbled across the concept of *steering* (the effect that *a choice* of measurement in A can result in the system in B being described by one of two (or many) states) in 1935: [11] "Now what is essential is exclusively that ψ_A and ψ_B are in general different from one another. I assert that this difference is incompatible with the hypothesis that the description is correlated one-to-one with the physical reality (the real state). After the collision, the real state of (AB) consists precisely of the real state of A and the real state of B, which two states have nothing to do with one another. *The real state of B thus cannot depend upon the kind of measurement I carry out on A*. But then for the same state of B there are two (in general arbitrarily many) equally justified ψ_B, which contradicts the hypothesis of a one-to-one or complete description of the real states."

In even simpler terms than the previous argument Einstein here realises that locality and the statistics of quantum mechanics imply that there can be more than one real states of affairs to be associated with a quantum system. Not only this is an argument for incompleteness, but it's an argument for epistemicity, i.e. for the property of states of reality not to be unique.

1.2.2 Medieval history

In 1964 John Bell wrote the historical paper "On the Einstein Podolsky Rosen paradox" [12]. In this paper he showed that a local theory cannot reproduce the predictions of quantum mechanics. The logic of his reasoning goes as follows: in the EPR state a measurement upon a subsystem makes it possible to predict the result of a parallel measurement on the other subsystem. If we accept the fact that spacelike separated events cannot influence each other, it seems that parallel measurement results should be locally predetermined[3]. But this leads to a contradiction with quantum statistics when we address non parallel measurements, therefore we cannot exclude that a measurement on one subsystem does not influence a measurement on the other anymore. The theorem applies to any local theory in which properties are assumed possessing a value *as such*, therefore it also applies to hidden variable theories, so that if one considers a hidden variable model, in which the measurement outcomes are elements of reality, one

---

[3]being predetermined does not imply being determinable, i.e. there is nothing in Bell's argument that is concerned with determinism

has to drop the locality assumption anyways, i.e. we must drop $P(x|y,\lambda) = P(x|\lambda)$ where x and y belong to spacelike separated regions X and Y and λ might be ψ (in which case we would not be dealing with a hidden variable theory) or a more general object. The derivation of Bell's inequalities (his 1971 derivation, to be precise) is rather straightforward. It is presented here for completeness, and because I like it.

Consider a measurement whose outcome can take two values, ± 1. This is a more general situation than we can naively think, as quantum measurements correspond to propositions about the system so they can be put in a form that, logically, implies a question with a yes/no answer, a simple example being "will this measurement give a result in this range?". Imagine having a system made of two parts A and B which interacted some time in the past, but design the experiment so that we have two instances of such measurement, performed in two *spatially separated* regions of spacetime. Call $A(a,\lambda)$ and $B(b,\lambda)$ the results of the measurements. Here a and b are the locally and freely chosen measurement settings and λ is the set of real values that characterise the system, may we or may we not have access to λ. We have made a locality assumption here: A does not depend on b and B does not depend on a. While it is natural to assume that the past interaction between the subsystems will correlate the measurement results in some way, any *additional* randomness of subsystems A and B has to be fully accounted for by something else, which we call λ, and consider it any sort of ontological entity, even just the wave function, interpreted in a realistic way. As the measurement can be repeated over and over we consider in full generality λ as a random variable, which therefore has some probability distribution $\rho(\lambda)$.

Define the expected value of the quantum correlation $E(a,b) := P(A = B) - P(A \neq B)$, which represents the difference between the probability that A will give the same result as B and the probability that A will give the opposite result of B. Clearly $E(a,b)$ is just a function of the measurement settings:

$$E(a,b) = \int_\Lambda A(a,\lambda)B(b,\lambda)\rho(\lambda)d\lambda \tag{1.3}$$

Now we play a bit with some algebra. We can consider two additional measurement settings a' and b' and write $E(a,b) + E(a',b)$ as

$$E(a,b) + E(a',b) = \int_\Lambda [A(a,\lambda)B(b,\lambda) + A(a',\lambda)B(b,\lambda)]\rho(\lambda)d\lambda \tag{1.4}$$

For simplicity of notation I will indicate $\rho(\lambda)$ with ρ, $A(a,\lambda)$ and $A(a',\lambda)$ with A and A' and similarly with B and B'. Now we make the reality assumption that is

key to the theorem: we add and subtract the same term $ABA'B'$, and this term contains values that must be all *possessed*, i.e. they must be *real*:

$$E(a,b) + E(a',b) = \int_\Lambda (AB + A'B + ABA'B' - ABA'B')\rho \, d\lambda \tag{1.5}$$

$$= \int_\Lambda AB[1 + A'B']\rho \, d\lambda + \int_\Lambda A'B[1 - AB']\rho \, d\lambda \tag{1.6}$$

The triangle inequality tells us that

$$|E(a,b) + E(a,b')| \leq \left| \int_\Lambda AB[1 + A'B']\rho \, d\lambda \right| + \left| \int_\Lambda A'B[1 - AB']\rho \, d\lambda \right| \tag{1.7}$$

Thanks to the absolute value, we can discard AB and $A'B$ in the integrals, because their value will be either $+1$ or -1. Moreover, the values in square bracket are either 0 or 2, so we can discard the absolute values. We are left with

$$|E(a,b) + E(a',b)| \leq \int_\Lambda [1 + A'B']\rho \, d\lambda + \int_\Lambda [1 - AB']\rho \, d\lambda \tag{1.8}$$

$$= 1 + \int_\Lambda A'B'\rho \, d\lambda + 1 - \int_\Lambda AB'\rho \, d\lambda \tag{1.9}$$

$$= 2 + E(a',b') - E(a,b') \tag{1.10}$$

which includes the CHSH inequality. As quantum mechanics can yield a value of $E(a,b) + E(a',b) - E(a',b') + E(a,b')$ which exceeds 2, this inequality is violated, therefore its premises must have something wrong, and our initial discussion applies. What I like about Bell's theorem, and that perhaps is often overlooked, is that it is not a result about local hidden variable theories, it is a result about *any* local theory that is interpreted realistically, i.e. any local theory in which the properties posses definite values, independently of wether we can or cannot know or predict them. This class of theories is much wider than just hidden variable theories and this is perhaps what makes Bell's theorem one of the most profound results in science.

1.2.3 Modern history

I think of a big gap between the sixties and the following couple of decades, because at some point technology allowed to drop the 'gedanken' from gedanken-experiment. The first convincing experiment in favour of Bell's theorem is described in the 1981 paper by Aspect, Grangier and Roger [13]. In their experiment they used an atomic source of photons, in particular a cascade in calcium atoms, which produced pairs of photons entangled in their polarisation. The suc-

cess of their experiment is considered a milestone in the history of science, for its significance: mankind had a first tangible proof that nature is really nonlocal, a formidable feature which, if you think of it, was never experienced ever before, by any human being, in all history. This is one of my favourite arguments in favour of our possible lack of the kantian categories needed to understand how entanglement works. Aspect's experiment was also a very stringent test of quantum mechanics, which proved to model reality astonishingly well.

Polarisation of photons is a quantity whose basis is made of two elements, therefore the Hilbert space of two such systems is four-dimensional. In contrast to this kind of source there are other kind of sources which emit states whose description is a state vector that needs many more basis vectors. This is why these sources are said to emit high-dimensional states. SPDC is amongst these sources, and a remarkable one of such kind, for its luminosity [14] and for the different types of quantities that are simultaneously entangled [2]. The high dimensionality of the Hilbert space is due to entanglement in spatial degrees of freedom, temporal degrees of freedom and polarisation degree of freedom at the same time [15]. The strength of the entanglement is extremely high and the consequences can be very interesting, we will talk about these all throughout the thesis, in particular we will talk about shared information and imperfect measurements in chapter 4.

1.2.4 Contemporary history

When I think of "contemporary history" I think of a way of approaching the topic of high dimensional entanglement and SPDC which follows a pattern which is comparable with the one that is currently used now in 2012. Looking back some years, we can track how high dimensional spatial entanglement, orbital angular momentum and SPDC all came together. We can start in 1985 with the work on the theory of parametric down-conversion by Hong and Mandel [16]. In their work, the most important result was that when two ideal photodetectors were appropriately located, the joint probability of two-photon detection can equal the single-photon detection probability, a feature that suggests the presence of entanglement.

In 1992, Les Allen, Marco Beijersbergen, Robert Spreeuw, and Han Woerdman [6] showed that Laguerre-Gaussian modes of light carry a definite amount of orbital angular momentum, and in 1995 the seminal work by Miles Padgett and Les Allen [7] on the Poynting vector of of Laguerre Gaussian modes started to clarify some fundamental properties of these modes of light.

SPDC was used as a reliable source of entangled photons (albeit still for polarisation) in 1995 by Kwiat et al. [14], while the spatial coherence properties of down-converted light was extensively studied in 1996 by Joobeur et al. [17]. In 1998 Carlos Monken et al. [18] showed transfer of pump properties to down-converted photons. Step by step, the scientific community was getting closer and closer to understanding high dimensional entanglement of orbital angular momentum in down-conversion.

It was 2001 (not at all long ago!) when Mair et al. from the Zeilinger group [19] discovered that entanglement in SPDC is much more structured than it was believed before, demonstrating entanglement between orbital angular momentum eigenstates. Since then, the research in this field has seen numerous advances, on the theoretical side by analysing the modal content of down-converted states, in particular we mention the seminal works by Torres, Alexandrescu and Torner (2003) on the spiral bandwidth (i.e. the profile of orbital angular momentum entanglement in down-converted states) [20], the famous work by Law and Eberly [21] where they address the Schmidt number and the difference between gaussian approximation and sinc form of the phase matching function.

The state of the art is held by a few groups. The Leiden group leads the way in the theory of phase masks, used in conjunction with measurements of entanglement in the orbital angular momentum space [22–25], while Miles Padgett's group in Glasgow University is on the cutting edge of entanglement measures utilising spatial light modulators to perform standard analysis of down-converted light, as well as more exotic measurements, such as the demonstration of entangled singularity knots [26] or violation of Bell's inequalities in a high dimensional space [27]. In chapter 8 we will see these two methods come together, namely the implementation of sector phase plates with spatial light modulators. The state of the art experimental research now revolves around pushing the limits of entanglement strength [8] and on utilising SPDC states for technological applications, while the state of the art theoretical research revolves around understanding how deeply entangled the SPDC state really is, by addressing both the time and the space parts of the state [28–33].

1.2.5 Future history

Many other groups in the world are currently investigating down-conversion sources also for their entangled temporal properties [28, 30, 31]. Putting together temporal and spatial entanglement leads to hyperentanglement (just entanglement, but spread across multiple types of degrees of freedom) [15], which is a

promising way of generalising high dimensional entanglement. There are still open questions, in particular the role of phase matching approximations, how far we can stretch them and the question of how the entangled eigenmodes depend on the phase matching function. Also, there isn't yet a full analytical description of the Schmidt modes with a nonzero phase mismatch parameter.

The future of research in down-conversion and entanglement will surely push to even higher figures for entangled modes, for control over the spatial shape of Schmidt modes, for an analytic theory of total rate of emission, optimised detection schemes and applications that are offered by having such an easy to manipulate, and entanglement-rich system.

Chapter 2

Entanglement

[...] note that the idea of an object bearing properties is encoded in the structure of our language itself: the subject-predicate form of simple sentences reflects the idea of properties adhering to things, and thereby our comonsense understanding of the question: "what is a thing?"

(Chris Isham, *Lectures on quantum theory*, 1995, p. 68)

In the following sections I will introduce a bit of quantum mechanics, with an emphasis on the concepts that are more important to understand the rest of the thesis, namely entangled states, orbital angular momentum, and entropic entanglement measures. I will start with some explanation of the concept of state in quantum mechanics (without entering the hazardous field of its interpretations and assuming an epistemological one), then I will introduce the concept of measurement in a fairly general way, and finally I will proceed to entanglement and entanglement measures.

2.1   Pure states and mixed states

The concept of state in quantum mechanics can be expressed in various different ways, depending on which kind of depth one is looking for. We can describe, in a fairly general way, a quantum state by using the *density operator*, $\hat{\rho}$. Such operator is a map between Hilbert spaces. In particular it's a linear endomorphism

of $\mathcal{H}$. In fact, it's a hermitian, trace 1, compact operator on a Hilbert space that maps vectors of a Hilbert space into vectors of a Hilbert space:

$$\hat{\rho} : \mathcal{H} \to \mathcal{H} \qquad (2.1)$$

All states are therefore vectors of the vector space $\mathcal{B}(\mathcal{H}) \subset \mathrm{Hom}(\mathcal{H}, \mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$[1]. The kernel of an operator $\hat{\rho}$ is a subspace of $\mathcal{H}$. We say that $\hat{\rho}$ represents a generalised state in the sense that it can account for a lack of knowledge. Dirac's "bra-ket" notation without outer product cannot include lack of knowledge, and is therefore limited to the description of states about which we know everything that there is to know, i.e. "pure states". In other words, a state is pure when everything about the system that can be known is taken into account. If a state is pure, we have

$$\hat{\rho}^2 = \hat{\rho} \qquad (2.2)$$

Which means that pure states are idempotent endomorphisms. On the other hand, if there is any attribute of the system which is unknown, the state presents itself as a mixed state and eq. (2.2) won't hold anymore. Such relation implies $\mathrm{Tr}(\hat{\rho}^2) = \mathrm{Tr}(\hat{\rho}) = 1$ for pure states. For mixed states one has $\mathrm{Tr}(\hat{\rho}^2) < \mathrm{Tr}(\hat{\rho}) = 1$ and we can finally understand that a mixed state $\hat{\rho}$ can be thought as a probability distribution over pure states $\hat{\rho}_n$ (where the summation can be replaced by an integration, were the pure states labelled by a continuous parameter):

$$\hat{\rho} = \sum_n p_n \hat{\rho}_n \qquad \sum_n p_n = 1 \qquad (2.3)$$

This form clearly indicates a certain lack of knowledge, as we are assigning probabilities to different pure states to describe the system, thus we know that the state really is described by only one of such pure states, but we are missing the information that is required to identify it.

On the other hand, a pure state contains all the possible knowledge about a state. The fact that Quantum Mechanics is a theory that can predict outcomes of measurements probabilistically and not deterministically has nothing to do with the completeness of the information that a pure state stores, but rather with the way nature works the way she works.

Using Dirac's notation, which defines vectors in a Hilbert space as $|\psi\rangle \in \mathcal{H}$, and linear functionals as $\langle \psi | \in \mathcal{H}^* : \mathcal{H} \to \mathbb{C}$, we can write a pure state by using

---

[1]Technically, they are nuclear operators, which is a subset of $\mathcal{B}(\mathcal{H})$

the outer product in the following way:

$$\hat{\psi} = |\psi\rangle\langle\psi|$$ (2.4)

If the pure state $|\psi\rangle$ is written in a complete basis of eigenkets of an operator $\hat{A}$, i.e. $|\psi\rangle = \sum_a \lambda_a |a\rangle$, then the pure state expressed with a density operator would be

$$\hat{\psi} = \sum_{a,a'} \lambda_a \lambda_{a'}^* |a\rangle\langle a'|$$ (2.5)

The density operator is called sometimes *density matrix* because the values $\lambda_a \lambda_{a'}$ can be arranged in a square matrix, and the probability normalisation condition $\sum_a |\lambda_a|^2 = 1$ is equivalent to requiring that the trace of the density matrix be equal to 1. Such requirement remains valid in the case of a mixed state, as it becomes $\sum_n p_n \sum_a |\lambda_{a,n}|^2 = \sum_n p_n = 1$.

2.2 Measurements

Measurements in Quantum Mechanics have always carried the burden of interpretation. After almost a century from the beginning of the Quantum Theory, there is still an ongoing debate. The so called *measurement problem* is a beautiful and utterly interesting subject, but it will be skipped in this introduction for the sake of clarity, as no discussion on the measurement problem that I know of can make one's mind less baffled than it was before.

Similarly to the way states were introduced, I will introduce quantum measurements in a general way. The measurement of an observable quantity of a state is a probabilistic operation that maps an initial state $\hat{\psi}$ onto an element of the set of possible final states $\{\hat{\sigma}_j\}$ with a finite probability $p_j = p_{\hat{\psi}\to\hat{\sigma}_j}$, with $\sum_j p_j = 1$. Such probabilities depend on the initial state and on the final state and they are given by the squared modulus of the inner product between these two states, according to the Born rule: the inner product is calculated through the trace in $\mathscr{B}(\mathscr{H})$

$$|\langle\hat{\psi},\hat{\sigma}\rangle|^2 = \text{Tr}(\hat{\psi}^\dagger\hat{\sigma})$$ (2.6)

We are allowed to do this because $\mathscr{B}(\mathscr{H})$ is a complete inner product space (it is a Hilbert space too), whose norm follows from its inner product (in this case the Hilbert-Schmidt norm). In case of pure states we obtain the familiar $p_{\hat{\psi}\to\hat{\sigma}} =$

$$\text{Tr}(\hat{\psi}^\dagger \hat{\sigma}) = |\langle \psi | \sigma \rangle|^2.$$

2.2.1 Projective measurements

If we consider a physical quantity A, which represents a property of a physical system, such as its position, its momentum, its energy and so on, we can associate to A a hermitian operator $\hat{A}$, which represents the property A in the quantum formalism. Given a system with states in a Hilbert space $\mathcal{H}$, it's possible to group together all the pure state vectors that satisfy the following functional equation:

$$\hat{A}|\psi\rangle = A_n|\psi\rangle \tag{2.7}$$

for some real (recall that $\hat{A}$ is hermitian) values A_n. We can then call $|\psi_n\rangle$ the vectors associated to the eigenvalue A_n, they may span a one-dimensional space or a multi-dimensional space. The set of eigenvalues $\{A_n\}$ is called *spectrum* of $\hat{A}$ and indicated by $\text{Sp}(\hat{A})$, and in general it can be discrete, continuous or mixed. We interpret the states $|\psi_n\rangle$ as those that possess the value A_n of the property A. The set $\{|\psi_n\rangle\}_{n\in\text{Sp}(\hat{A})}$ is the set of eigenstates of $\hat{A}$. A set of eigenstates is called *complete* if every one of them spans a one-dimensional space. If a set of eigenstates is not complete it is possible to use other observables B, C and so on until we have a (complete) set of eigenstates that span one-dimensional spaces and that all together generate $\mathcal{H}$. The choice of a complete set of eigenstates is not unique. From a set of eigenstates it is possible to build a set of projectors $P_n = |\psi_n\rangle\langle\psi_n|$, the action of which is to project any state $|\phi\rangle$ into its component in the subspace of $\mathcal{H}$ generated by $|\psi_n\rangle$. States which are not eigenstates, i.e. that are superpositions of eigenstates will not have a sharp value of the observable in question, and will *probabilistically* give one of the eigenvalues in the superposition with a probability that follows the Born rule, mentioned above. The state after the measurement will be a pure state and in particular the eigenstate relative to the eigenvalue just measured. The complete set of projectors has the following *four* properties:

1. $\hat{P}_n = \hat{P}_n^\dagger$ (hermiticity)

2. $\hat{P}_n \geq 0$ (positivity)

3. $\sum_n \hat{P}_n = \hat{\mathbb{1}}$ (completeness)

4. $\hat{P}_n \hat{P}_m = \delta_{mn} \hat{P}_n$ (orthogonality)

However, the last condition is not a fundamental property for an operator to represent a measurement. Probability Operator Measures (POM), also known as Positive Operator-Valued Measures (POVM) are a generalisation of projective measurements, but although they ought to be explained in a textbook on quantum mechanics, they are superfluous for the arguments in this thesis and their treatment will be skipped.

Therefore, a simple measurement of a property A on a state $|\phi\rangle$ can be described by a projective measurement in the following way. We can write $|\phi\rangle$ as a superposition of its components on the subspaces of eigenstates of $\hat{A}$, i.e. $|\phi\rangle = \sum_n P_n |\phi\rangle$, and then what happens (to the embarrassment of the physics community) is still a bit of a mystery, namely the state vector is randomly rotated to one of the eigenstates, although the randomness follows the probabilities given by the Born rule:

$$|\phi\rangle \rightarrow |\psi_n\rangle \qquad \text{with probability } |\langle \psi_n | \phi \rangle|^2 \qquad (2.8)$$

Many and many more pages could be devoted to the interpretation of the term *randomly* above and to the significance of the passage from the 'possible' (i.e. with probabilities given by the Born rule) to the 'actual' (i.e. to the actual final result and resulting state), but of course this is not the place for such pages.

2.3   Entanglement

2.3.1   The idea

Entanglement is perhaps the most dazzling property of quantum systems. It is what makes *quantum* fundamentally different from *classical*. The interest that it generates comes from deep philosophical questions and from practical uses for eventual future technologies. Entanglement is a hot topic in quantum foundations, quantum information, quantum computation, quantum optics, quantum control and metrology, engineered quantum systems, and recently one can hear biologists starting to mumble about entanglement [34].

Entanglement comes in multiple flavours, depending on which property *is* entangled. However, the mathematical framework of quantum mechanics allows one to treat all entangled systems in a general way. One of the various properties that characterises a quantum system is the dimensionality of the Hilbert space $\mathscr{H}$ that embeds its description. If the system is an entangled one, the dimensionality of $\mathscr{H}$ obtains a particular role, as we shall see in a moment. This will be a

fundamental concept in this thesis. Indeed, the content of the next chapters will orbit around the topic of *high dimensional* entanglement, where *high* refers to the dimensionality of the Hilbert space of an entangled system. At this point the reader might reply that the dimensionality of a Hilbert space could be trivially extended, simply by adding new orthogonal basis vectors. This trivial extension enlarges the dimensionality of $\mathscr{H}$ by working on the way we write the states: it is not this trivial extension that we are interested in! Instead, we are interested in the effective *minimal* dimensionality that a Hilbert space has to have in order to embed the allowed states of the system. Such dimensionality is known as the Schmidt number, and we aim at maximising it, i.e. (roughly speaking) we let the system obtain new states, so that consequently the minimal dimensionality of $\mathscr{H}$ has to increase, and with it, the strength of the entanglement. The Schmidt number is not the only way of measuring entanglement, there are many types of measure. In our case, however, of a bipartite system, things are much easier in this regard, and the Schmidt number is chosen for both its advantageous calculation and its intuitive significance of "number of entangled eigenstates".

Some of the reasons why high dimensional entanglement is a topic worth attention and time and efforts is that highly entangled systems have a behaviour that is favourable over systems with low dimensional entanglement when the application of a quantum technology is eventually realised. In particular, one finds that highly entangled states can tolerate imperfect measurements better than less entangled states [32], or that when they propagate in a turbulent atmosphere, they retain some properties better than less entangled states [35]. The main reason, though, to favour high dimensional entanglement is based on the fact that an n-dimensional vector space has a basis of n independent vectors, and this means that one can implement quantum algorithms that involve entangled states with up to n different possible joint states [3, 4]. The more joint states one can use, the more powerful and complex the algorithm can be, and this can be a great advantage for future technologies.

2.3.2 Entanglement strength

The strength of the entanglement depends on the process that one uses to produce an entangled state. As anticipated, the quantity that we will use to gauge the strength of the entanglement is called *Schmidt rank* or *Schmidt number* and we indicate it by K. The only possible value of the Schmidt number for a separable state (a non-entangled state) is 1, all numbers above 1 indicate entangled states. Schmidt numbers of less than 1 cannot exist, as one cannot have "less

than one state". This is a consequence of its definition for bipartite states:

$$K = \frac{\text{Tr}(\hat{\rho})^2}{\text{Tr}(\hat{\rho}^2)} \tag{2.9}$$

where $\hat{\rho}$ is *any* of the two reduced density matrices. The numerator can be set to 1 if the density matrices are normalised. The reason why $K \geq 1$ is that the denominator reaches the maximum of 1 if $\hat{\rho}$ is a normalised pure state; and considering that here $\hat{\rho}$ is a *reduced* density matrix, it follows that the minimum of $K = 1$ is reached if the reduced density matrix is a pure state, which happens only if the whole system is not entangled. To better understand what it means, we make use of a simple example.

Example: *Consider the state*

$$\sqrt{1-\varepsilon}\,|0,1\rangle + \sqrt{\varepsilon}\,|1,0\rangle$$

where ε is any number between 0 and 1. This state is made of two vectors, which lives in a four dimensional vector space, as it is spanned by $|0,0\rangle$, $|0,1\rangle$, $|1,0\rangle$ and $|1,1\rangle$. For $\varepsilon = 0$ and for $\varepsilon = 1$ the state is separable, while for all the other values of ε, the state is entangled. How can we calculate the Schmidt number? For this particular example the Schmidt number is $K = \left(|\sqrt{1-\varepsilon}|^4 + |\sqrt{\varepsilon}|^4\right)^{-1} = (2\varepsilon^2 - 2\varepsilon + 1)^{-1}$ which is plotted in figure 2.1 (For details on the definition of the Schmidt number we need to introduce the Schmidt decomposition, however one could briefly jump to section 2.3.4 for explicit calculations).

The minimal dimensionality for a Hilbert space to host a bipartite entangled state is 4, because a 1-dimensional Hilbert space cannot even hold two systems, and as we saw in the example, 2- and 3-dimensional spaces can only host separable states, which are not entangled. On the other hand, we can say that here is no fundamental upper limit. The only upper limit depends on the quality of the instruments used for preparing the state, and it depends on each different implementation. It is interesting to recall that different flavours of entanglement can be combined together (spatial, temporal, spin/polarisation) to yield hyperentanglement, in which the dimensionalities of the Hilbert spaces holding the spatial, temporal and spin representations multiply together to give a possibly extremely large Hilbert space and an extremely high entanglement.
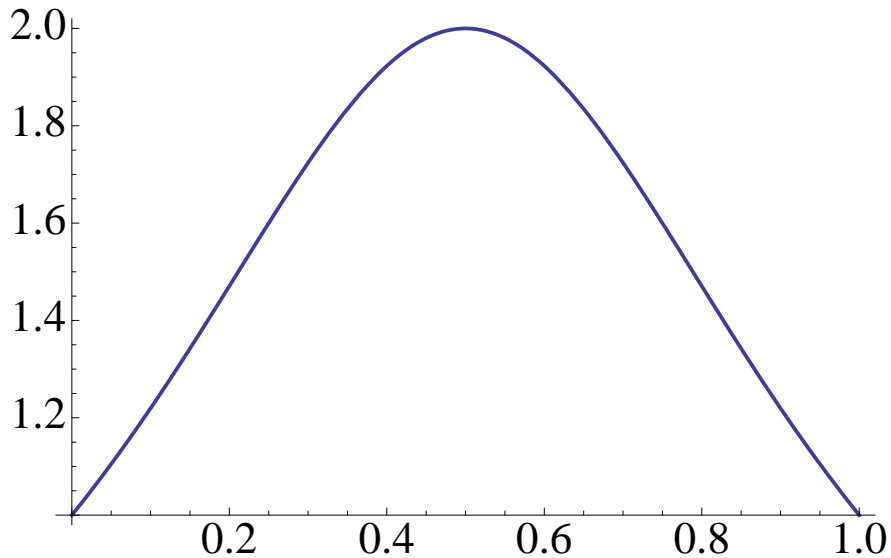
Figure 2.1: Schmidt number K as a function of ε for the state in the example

2.3.3 Bipartite entanglement and composition of Hilbert spaces

Entanglement is a nonclassical property of quantum states, that appears to us in the form of correlations. Correlations due to entanglement are stronger than correlations given by classical systems, because they persist even if the measurements are performed in different bases. This fact is the basis of Bell's theorem. The most simple form of entanglement happens between two quantum systems, and it is this class of entangled systems that will be at the core of the main topic of this thesis. Entanglement has a very close relation with separability, and the mathematical framework behind separability is embedded in the type of vector space operations that determine the unification of two systems. To explain this concept through an example, let's consider a system of two photons, and let's examine the polarisation degree of freedom. Call the complex Hilbert spaces of the polarisation states $\mathcal{H}_1$ and $\mathcal{H}_2$. To understand dimensionality in the correct way, I should specify the difference between complex dimension and real dimension: a complex space of dimension d has real dimension $2d$, this is simply a consequence of the fact that two real numbers are needed to specify a complex number. The state of polarisation can be considered as a qubit: it can take all the normalised complex combinations of two basis vectors. Therefore, the Hilbert space of a qubit is isomorphic to $\mathbb{C}^2$ and has *complex* dimension 2. However, to talk about the *state space* we need to fix normalisation and global phase, because it's not true that different points in $\mathbb{C}^2$ always correspond to different states, and instead we would like to have an injective map from $\mathcal{H}$ to the state space. After

we introduce these two constraints we have separated the qubit Hilbert space into equivalence classes, each representing a state. The state space can be represented as the surface of a sphere of *real* dimension 2, which is called Bloch sphere[2]. In fact, any pure state of the polarisation of a photon can be represented as a point on the surface of the Bloch sphere, which needs 2 coordinates to be specified. This was the description of *one* qubit, but what is the correct way of combining $\mathscr{H}_1$ and $\mathscr{H}_2$, in order to deal with a state of two qubits? Let's start by looking at what *not* to do, i.e. at what it is done for two classical states.
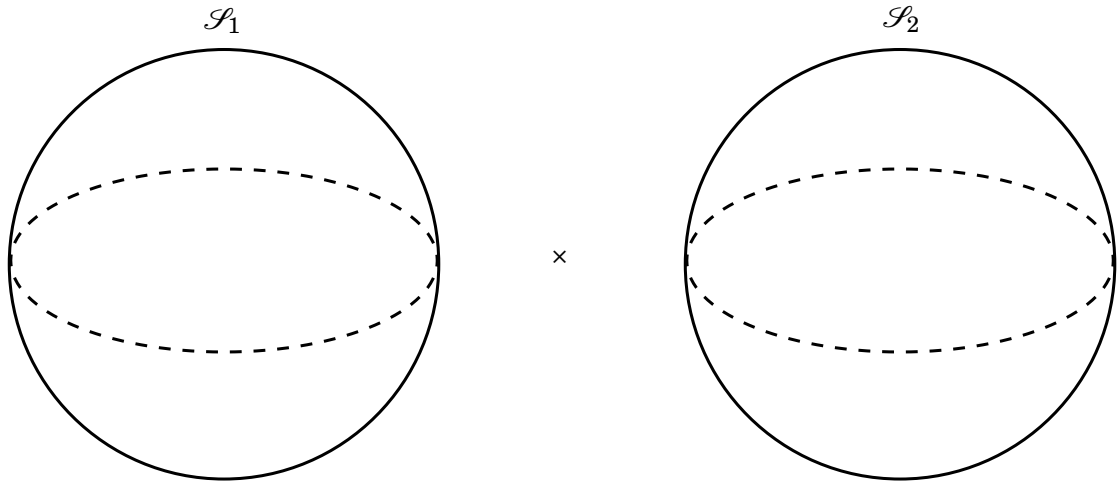
Figure 2.2: The space of separable states $\mathscr{S}_\times$. Although it consists of two Bloch spheres it is not the space that can host entangled states, because what is really needed is the tensor product of the Bloch spheres.

The cartesian product space $\mathscr{H}_\times = \mathscr{H}_1 \times \mathscr{H}_2$ is the space of pairs: each point of $\mathscr{H}_\times$ is a pair of vectors (h, g), where $h \in \mathscr{H}_1$ and $g \in \mathscr{H}_2$. Each point in $\mathscr{H}_\times$ is a separable state by definition (it is made by picking a state from each Hilbert space[3]) and therefore cannot represent an entangled state. However, this is the right way of combining two independent *classical* systems. This fact alone sets entangled states in a different class with respect to classical states.

How is it possible that $\mathscr{H}_\times$ has a vector space structure and yet cannot include entangled states? The answer is that its vector space structure is carried by the structure of the components, in a component-wise way: the operations that make $\mathscr{H}_\times$ a vector space are $(h, g) + (h', g') = (h + h', g + g')$ and the linear multiplication by scalars. It's straightforward to see that any linear combination

---

[2]The Bloch sphere for a qubit is homeomorphic to $U(2)/(U(1) \times U(1))$, where $U(2)$ is the group of transformations on the Hilbert space of the 2-level system, and the two $U(1)$ factors form the isotropy group of a 1-dimensional complex subspace of a state vector and of its complement. For a general n-level system the Bloch sphere is homeomorphic to $U(n)/(U(1) \times U(n-1))$.

[3]Formally, there is a unique projection map on the factors $\mathscr{H}_1$ and $\mathscr{H}_2$

of separable states would be mapped to another separable state, and clearly this is not what we want! So although all the states contained in $\mathscr{H}_\times$ are valid states, the cartesian product cannot be the correct operation, because we wouldn't be able to include entangled states. The correct operation is the tensor product $\mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2$, which can be understood as the cartesian product, but with some restrictions. The restrictions are that the operations that make $\mathscr{H}$ a vector space are $(h,g) + (h',g) = (h + h',g)$, $(h,g) + (h,g') = (h,g + g')$ and linear scalar multiplication. Note how these differ from the ones that make $\mathscr{H}_\times$ a vector space, in particular the combination $(h,g) + (h',g')$ is not mapped to any vector in the form (a,b). So separable states are still members of $\mathscr{H}_\times \subset \mathscr{H}$, but with the restricted operations there are states that cannot be written as a pair, and these are the entangled states. Notice that with the restricted operations $\mathscr{H}_\times$ is not a proper subspace of $\mathscr{H}$.

If we regard $\mathscr{H}_1$ and $\mathscr{H}_2$ as two copies of $\mathbb{C}^2$, $\mathscr{H}$ has complex dimension 4. After fixing global phase and normalisation, a state vector now belongs to the tensor product of two Bloch spheres, $\mathscr{S}$, and it needs six real numbers to be specified. Generalising this, a state of n qubits would live in a n-fold tensor product of $\mathbb{C}^2$, which has complex dimension 2^n, and if we take into account normalisation and global phase, it would have real dimension $2^{n+1} - 2$. generalising even further, a state of n m-level states would live in a n-fold tensor product of $\mathbb{C}^m$, which has complex dimension m^n, so the state space would have real dimension $2m^n - 2$.

Example: *A state $|H\rangle \otimes |V\rangle$, that lives in $\mathscr{H}_\times$, is not entangled, because it is separable. Likewise, the state $|H\rangle \otimes \left(\frac{|V\rangle + |H\rangle}{\sqrt{2}}\right)$ is not entangled, because it is separable. But outside $\mathscr{H}_\times$, we can find for instance states like $\frac{1}{\sqrt{2}}(|H\rangle \otimes |H\rangle + |V\rangle \otimes |V\rangle)$, which, in a 4-dimensional complex Hilbert space, is entangled. Such state lives outside $\mathscr{H}_\times$ because with the restricted vector space operations it cannot be identified with any vector in $\mathscr{H}_\times$. Such entangled state is the linear superposition of two of the four vectors that form a possible basis for the joint Hilbert space ($|H\rangle \otimes |H\rangle$, $|H\rangle \otimes |V\rangle$, $|V\rangle \otimes |H\rangle$, $|V\rangle \otimes |V\rangle$) and the squares of the coefficients sum to one. Thus the Schmidt number of this state is 2. The Schmidt number cannot be larger than the complex dimension of the smaller Hilbert space of the single subsystems.*

In this example, each photon polarisation lives in a copy of $\mathbb{C}^2$, therefore the Schmidt number cannot be larger than 2. This also means that the state in the example was maximally entangled. To show what happens if one tries to use more than two joint states, imagine we used all four states in the basis: $\frac{1}{2}(|H\rangle \otimes |H\rangle + |H\rangle \otimes |V\rangle + |V\rangle \otimes |H\rangle + |V\rangle \otimes |V\rangle)$. This state is separable, because

even with the restricted operations, it can be written as $\frac{1}{2}(|H\rangle + |V\rangle) \otimes (|H\rangle + |V\rangle)$ and thus it belongs to $\mathscr{H}_\times$. If we take the previous example to a higher level and consider $\mathscr{H}_1$ and $\mathscr{H}_2$ to have a much greater dimensionality and if we add some constraints on the type of entangled states that can be built (for instance, the constraint that the sum of the eigenvalues of the entangled observable of the pairs of states has to be 0) we obtain a framework which is very similar to the one that this thesis is built around. In down-conversion, the entangled systems are 2, so $n = 2$, and the number of possible states of each system is unbounded, so the dimensionality of the Hilbert space of the photon pairs is also unbounded. However, not all states are equally probable. Some of them are so unlikely to be measured that they can be safely ignored. How do we deal with this? The answer is given by the Schmidt decomposition.

2.3.4 Partial traces

After a bipartite entangled state has been written as a sum over joint states

$$\left|\psi\right\rangle = \sum_{n,n'} c_{n,n'} |a_n\rangle \otimes |b_{n'}\rangle, \tag{2.10}$$

it's possible to calculate what state we should assign to each individual system, which is equivalent to ask the question: "how would system A be described by an experimenter that didn't know that it is entangled with system B?" The answer to this question is found by projecting the part B of the system onto any possible state, i.e. taking the partial trace over B.

$$\mathrm{Tr}_\mathrm{B}\left|\psi\right\rangle\left\langle\psi\right| = \sum_k \langle b_k | \psi\rangle\langle\psi | b_k\rangle \tag{2.11}$$

$$= \sum_k \langle b_k | \left(\sum_{\substack{n,n' \\ m,m'}} c_{n,n'} c^*_{m,m'} |a_n\rangle \langle a_m| \otimes |b_{n'}\rangle \langle b_{m'}| \right) |b_k\rangle \tag{2.12}$$

$$= \sum_{m,n,k} c_{n,k} c^*_{m,k} |a_m\rangle \langle a_n| \tag{2.13}$$

The purity of the density matrix given by summing over the index k is inversely proportional to the entanglement of the initial state. Although it might seem obvious after we have calculated partial traces numerous times before, it is rather puzzling to think about: what it means is that system A was a statistical mixture of states all along, and the same is true for B, as the same calculation can be done on B. Even more puzzling is the fact that even though the states are

mixed, there are cases in which it is impossible to assign prior values to some observables (see GHZ paradox [36]), this is why reduced density matrices are referred to as improper mixtures. This leads to the obvious question: "how in the world can nature make the entanglement correlations work?" [37]. Nobody knows, yet. Similarly to the partial trace over system B, the partial trace over system A gives:

$$\text{Tr}_A \left|\psi\right\rangle\left\langle\psi\right| = \sum_{m',n',k} c_{k,n'} c^*_{k,m'} \left|b_{n'}\right\rangle\left\langle b_{m'}\right| \tag{2.14}$$

2.3.5 Schmidt decomposition

The following proof of the Schmidt decomposition follows Ekert and Knight's in [38]. Consider a state $\left|\psi\right\rangle$ in the space $\mathcal{H}_A \otimes \mathcal{H}_B$. The form of its density matrix $\hat{\psi}$ is the same as the density matrix in brackets in eq. 2.12:

$$\hat{\psi} = \sum_{\substack{n,n' \\ m,m'}} c_{n,n'} c^*_{m,m'} \left|a_n\right\rangle\left\langle a_m\right| \otimes \left|b_{n'}\right\rangle\left\langle b_{m'}\right| \tag{2.15}$$

We will show now that there is a basis $\{|u_i\rangle\}$ in $\mathcal{H}_A$ and a basis $\{|v_j\rangle\}$ in $\mathcal{H}_B$ such that the state $\left|\psi\right\rangle$ can be written as

$$\left|\psi\right\rangle = \sum_i \sqrt{\lambda_i} \left|u_i\right\rangle \otimes \left|v_i\right\rangle \tag{2.16}$$

or in terms of its density matrix,

$$\hat{\psi} = \sum_{ij} \sqrt{\lambda_i \lambda_j} \left|u_i\right\rangle\left\langle u_j\right| \otimes \left|v_i\right\rangle\left\langle v_j\right| \tag{2.17}$$

with the coefficients all real and normalised, i.e. such that λ_i is a discrete probability distribution.

Choose a basis $\{|u_i\rangle\}$ of $\mathcal{H}_A$ such that $\text{Tr}_B \hat{\psi} = \sum_{m,n,k} c_{n,k} c^*_{m,k} \left|u_m\right\rangle\left\langle u_n\right|$ is diagonal, i.e. $\sum_k c_{n,k} c^*_{m,k} = \lambda_n \delta_{m,n}$, where the λ_i coefficients are a probability distribution. We can now rewrite the initial state where $\lambda_i \neq 0$:

$$\hat{\psi} = \sum_{\substack{n,n' \\ m,m'}} \sqrt{\lambda_m}\sqrt{\lambda_n} \frac{c_{n,n'}}{\sqrt{\lambda_n}} \frac{c^*_{m,m'}}{\sqrt{\lambda_m}} \left|u_n\right\rangle\left\langle u_m\right| \otimes \left|b_{n'}\right\rangle\left\langle b_{m'}\right| \tag{2.18}$$

$$= \sum_{n,m} \sqrt{\lambda_m}\sqrt{\lambda_n} \left|u_n\right\rangle\left\langle u_m\right| \otimes \left|v_n\right\rangle\left\langle v_m\right| \tag{2.19}$$

where $\sum_{n'} \frac{c_{n,n'}}{\sqrt{\lambda_n}} \left|b_{n'}\right\rangle = \left|v_n\right\rangle$ and similarly for the sum over m'. The proof ends here.

Once we have written the state in this form it is interesting to address the probability distribution $\{\lambda_i\}$, as we have that the Schmidt number can be readily found from it: $K = \left(\sum_i \lambda_i^2\right)^{-1}$ for a normalised initial state.

2.3.6 Entanglement strength

The way to understand how much a state is entangled could be to address some metric distance from the space of separable states $\mathscr{H}_\times$, and in a way the Schmidt number is doing exactly this. The way the Schmidt decomposition works is by using a particular basis for the joint Hilbert space that allows to re-write the entangled state with the least number of basis elements. Calculating the Schmidt number then boils down to counting the minimum number of basis elements that are needed to write the state. If this number is greater than 1, the state is entangled. The calculation of the Schmidt number is less trivial than it might seem at first glance. In fact, the elements of the basis to use have to be custom built: there is no "one size fits all" basis that will do the job.

A few pages back, in equation (2.9) we defined K in terms of a reduced density operator. Instead, the definition $K = \left(\sum_i \lambda_i^2\right)^{-1}$, which is simpler, requires normalisation of the reduced density matrix, and requires a decomposition to be calculated first, so that the state can be written in the form (2.16). In chapter 4, this is precisely the path that we will take. The reason why K has this form is rather intuitive: what we are doing is a weighted sum of the basis elements, by assigning to each basis element a weight proportional to its probability. The result is the average probability of a basis element, $\sum_k \lambda_k^2$. The Schmidt number is then the total number of basis elements that a flat probability distribution with the same average would have to have, in order to be properly normalised. In other words, it's the width of a normalised flat distribution of the same average probability of the distribution $\{\lambda_i\}$. As it is using the least number of basis elements it automatically follows that the Schmidt basis is the basis that maximises the shared information between the entangled systems. It is interesting to notice that the Schmidt number shares some characteristics of entropic measures of entanglement (in chapter 4 we will see that its logarithm is the Rényi entropy of order 2), in fact it is a very good measure of the "compactness" of a distribution: suppose we have a distribution which is peaked in two distant places, the variance of such distribution fails to gauge the fact that, although the peaks are far they might be very narrow. An entropic measure overcomes this problem and what it really does is it measures how much the distribution gathers around some points, regardless of how distant they might be in the domain. The Schmidt

number shares this characteristic with entropic measures, which makes it a good tool to gauge an effective number of dimensions: a relabelling of their representation on one or multiple axes (to assign them a probability distribution) should leave their effective number invariant.

More details on the Schmidt number will be found in chapter 4, where a full Schmidt decomposition is analytically calculated.

Chapter 3

Orbital angular momentum and down-conversion

The topics in this thesis are mainly concerned with generation, propagation and detection of particular states of light. Although they will be treated from a quantum mechanical point of view, many are closely linked to topics in classical optics, such as paraxial modes, spatial light modulators, optical fibres and photon counters. To give the necessary tools to fully understand the next chapters, I will start by introducing paraxial field modes and in particular Laguerre-Gauss and Hermite-Gauss modes.

3.1 Paraxial field modes

Paraxial fields are fields which propagate mainly in one direction. In this thesis, paraxial field modes are fundamental to describe all the possible modes that a propagating beam can be decomposed into. There are various families of paraxial fields, depending on which coordinate system one chooses to solve the paraxial Helmholtz equation, which is the spatial part of the wave equation of a field which propagates along a preferred axis.

3.1.1 Helmholtz equation

Let's start by deriving the Helmholtz equation. We are describing a field of light and we are interested in the amplitude of the field (not on its polarisation), therefore we can start with the d'Alambert wave equation for a scalar field:

$$\nabla^2 \phi - \frac{1}{c^2} \ddot{\phi} = 0 \tag{3.1}$$

27

we assume that the solution is separable between time and space coordinates, i.e. $\phi(\mathbf{r}, t) = u(\mathbf{r})T(t)$. If we substitute this form of $\phi(\mathbf{r}, t)$ and simplify we obtain

$$\frac{\nabla^2 u}{u} = \frac{1}{c^2} \frac{\ddot{T}}{T} \tag{3.2}$$

The left-hand side depends only on $\mathbf{r}$ and the right-hand side only on t, so the whole equation holds if the two sides are equal to a constant value. We choose the value to be $-k^2$ for later convenience. Being interested in the spatial part we write

$$(\nabla^2 + k^2)u = 0 \tag{3.3}$$

which is known as Helmholtz equation, and depending on the coordinate system in which it is solved and on the boundary conditions, it will lead to common plane waves, but also Hankel functions, Bessel functions, Mathieu functions and other special functions.

3.1.2 Paraxial approximation

A sufficient condition for a field to carry a finite amount of energy is to be "mod-square integrable", i.e. to belong to an L^2 space. Fields that are solutions of the Helmholtz equation might not be L^2, in which case they would not physically realisable, or are only approximately realisable. A way to overcome this problem is to solve instead the paraxial approximation of the Helmholtz equation [39]. We start by writing the field in form $u(x, y, z) = E(x, y, z)e^{ikz}$. Substituting in eq. (3.3) we get

$$\nabla^2 E + 2ik \frac{\partial}{\partial z} E = 0$$

where e^{ikz} is factored out.

The paraxial approximation consists in assuming that the beam will vary slowly in the coordinates orthogonal to z. This justifies the form for $u(x, y, z)$ above, and leads to the approximation

$$\left| \frac{\partial^2 E}{\partial^2 z} \right| \ll \left| \frac{k \partial E}{\partial z} \right| \tag{3.4}$$

that allows us to neglect the term $\frac{\partial^2 E}{\partial^2 z}$ and write the paraxial approximation of the Helmholtz equation:

$$\left(\nabla_{\perp}^2 + 2ik\frac{\partial}{\partial z}\right)E = 0. \tag{3.5}$$

If cartesian coordinates are used, a family of solutions of this equation will be the Hermite-Gaussian (or Hermite-Gauss) modes. But of course other coordinate systems can be chosen and different solutions will arise.

3.1.3 Hermite-Gaussian modes

To show how Hermite-Gauss and Laguerre-Gauss modes are found, I follow the illuminating paper by Pampaloni and Enderlein in [40].

In general, a paraxial field in momentum space, written as a superposition of plane waves, is:

$$E = \int \frac{dk_x dk_y}{2\pi^2} A(k_x, k_y) e^{ik_x x + ik_y y + i(k - \frac{k_x^2 + k_y^2}{2k})z}, \tag{3.6}$$

where $A(k_x, k_y)$ is the amplitude of the plane wave with wave vector $\mathbf{k} = (k_x, k_y, k - \frac{k_x^2 + k_y^2}{2k})$. The third component has that form because by the paraxial approximation $k \simeq k_z$, so k_x and k_y are small with respect to k. This means that $k_z = \sqrt{k^2 - (k_x^2 + k_y^2)}$ can be approximated to the first order. If one seeks the beam that minimises the transversal extension while minimising the divergence one finds the fundamental gaussian mode E_0:

$$E_0 = \int \frac{dk_x dk_y}{2\pi^2} e^{-\frac{w_0^2}{4}(k_x^2 + k_y^2)} e^{ik_x x + ik_y y + i(k - \frac{k_x^2 + k_y^2}{2k})z}. \tag{3.7}$$

So for this mode, the amplitude function of plane waves follows a gaussian shape: $A(k_x, k_y) = e^{-\frac{w_0^2}{4}(k_x^2 + k_y^2)}$. The simplest way to find Hermite-Gaussian modes is to use the rectangular symmetry of cartesian coordinates and include the additional modulation $(ik_x)^m (ik_y)^n$ in the amplitude $A(k_x, k_y)$. It is easy to see that the same modulation can be obtained by introducing the following differential operator acting on the fundamental gaussian mode:

$$\frac{\partial^{m+n}}{\partial x^m \partial y^n} E_0. \tag{3.8}$$

If we use the closed gaussian solution to the integral in (3.7) we will encounter

an expression like

$$\frac{\partial^m}{\partial x^m} e^{-\alpha x^2},$$ (3.9)

and similarly for the y coordinate. We can therefore use the definition of Hermite polynomials

$$H_m(x) = (-1)^m e^{x^2} \frac{\partial^m}{\partial x^m} e^{-x^2}$$ (3.10)

to simplify the result and finally obtain a closed expression for the Elegant Hermite-Gaussian mode of order (m, n):

$$EHG_{m,n}(x,y,\zeta) = \frac{1}{w(\zeta)^{\frac{m+n}{2}+1}} H_m\left(\frac{x}{w_0\sqrt{1+i\zeta}}\right) H_n\left(\frac{y}{w_0\sqrt{1+i\zeta}}\right) e^{ikz - \frac{x^2+y^2}{w_0^2(1+i\zeta)} - i\tilde{\psi}_{m,n}}$$ (3.11)

where $\zeta = z/z_R$, with $z_R = \frac{\pi w_0^2}{\lambda}$ the Rayleigh length of a Gaussian beam of width w_0 and wavelength λ. We also have defined $w(\zeta) = w_0\sqrt{1+\zeta^2}$. $\tilde{\psi}_{m,n}$ is the modified Guoy phase $\tilde{\psi}_{m,n} = (1 + \frac{m+n}{2})\arctan\zeta$. However, it would be more convenient to have complex values only in the exponential, in order to distinguish between the field amplitude and its phase. To do this we use a modulation in the form $(ik_x + \frac{1}{u}\partial_{k_x})^m (ik_y + \frac{1}{u}\partial_{k_y})^n$ for a constant u that we are free to choose and decide to set to $-iw_0^2$ for later convenience. This modulation can be achieved by acting with the following differential operator on the fundamental gaussian mode:

$$f^{m+n} e^{\frac{x^2+y^2}{2w_0^2 f}} \frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial x^n} e^{-\frac{x^2+y^2}{2w_0^2 f}} E_0,$$ (3.12)

where $f = 1 - \frac{w_0^2(1+i\zeta)}{2iu}$. With this choice of modulation and constant u we obtain the standard Hermite-Gaussian modes:

$$HG_{m,n}(x,y) = \frac{1}{w(\zeta)} H_m\left(\frac{\sqrt{2}x}{w(\zeta)}\right) H_n\left(\frac{\sqrt{2}y}{w(\zeta)}\right) e^{ikz - \frac{x^2+y^2}{w_0^2(1+i\zeta)} - i\psi_{m,n}},$$ (3.13)

where the phase now is given by the exponential, which contains the usual Guoi phase $\psi_{m,n} = (1 + m + n)\arctan\zeta$.

30

3.1.4 Laguerre-Gaussian modes

In a very similar manner Laguerre-Gaussian (or Laguerre-Gauss) modes can be found, the difference being using polar coordinates instead of cartesian. The chosen modulation is separated into a radial and an angular contribution:

$$(k_x^2 + k_y^2)^{p+\frac{\ell}{2}} e^{i\ell \arctan\left(\frac{k_y}{k_x}\right)} = (k_x + ik_y)^p (k_x - ik_y)^{p+\ell} \tag{3.14}$$

Using the same ideas as before we find that the same modulation can be found with the following differential operator acting on the fundamental Gaussian mode:

$$\frac{\partial^p}{\partial(x + iy)^p} \frac{\partial^{p+\ell}}{\partial(x - iy)^{p+\ell}} E_0 \tag{3.15}$$

Using the definition of the generalised Laguerre polynomials and supposing that the Rayleigh range is big, we can write the Laguerre-Gaussian modes in the pupil plane:

$$LG_{\ell,p}(\rho, \phi) = \frac{1}{w_0^{|\ell|+p+1}} L_p^{|\ell|}\left(\frac{\rho^2}{w_0^2}\right) \rho^p e^{-\frac{\rho^2}{w_0^2}} e^{-i\ell\phi} \tag{3.16}$$

Also here a normalisation factor has to be imposed in an analogous way. These are the only two families of solutions of the paraxial Helmholtz wave equation that will be mentioned in the rest of this thesis.

These families are complete, which means that any paraxial field can be decomposed into a suitable superposition of Laguerre-Gauss modes or of Hermite-Gauss modes, in other words they form complete bases for the Hilbert space $L^2(\mathbb{R}^2)$ of almost-everywhere mod-square integrable functions on $\mathbb{R}^2$, amongst which elements one finds also all possible finite-energy paraxial fields.

3.2 Orbital angular momentum

Something remarkable happens when we switch to polar coordinates. In polar coordinates one coordinate is compact (the angle), as it is formally a coordinate in S^1, i.e. the circle. Just as position and momentum are conjugated variables, angle and angular momentum also are [4]. This topic can become very complicated very easily, for the sake of clarity I will keep it brief and intuitive. Since S^1 is circle, which is compact, we can calculate Fourier series on it, and in this way we can relate conjugate variables with each other. The Fourier transform

on S^1 yields a *discrete* set of integer values which span the integer numbers, and we can interpret them as the *orbital angular momentum* (OAM) of a photon if we consider the coordinate on the circle as the azimuthal coordinate of the wave function written in polar coordinates. [5].

The fact that the phase structure of a mode with a certain OAM does not depend on the radial coordinate means that if we choose to address the angle-OAM conjugated variables only (this request is very important for the point here), changing between the two representations with a Fourier lens, or an equivalent setup, can be done with the contribution of very high-order modes, because their *angular* features won't fall out of our lenses, mirrors and detectors (here is where the request of addressing exclusively the angular degree of freedom becomes fundamental). Even though the radial distribution of the amplitude of a mode is OAM-dependent (the radial position of the peak of the intensity grows like $\sqrt{|\ell| + 1}$ where ℓ is the OAM eigenvalue of the mode), if outer parts of the wave function fail to overlap with a detector, the detected mode is still the same eigenfunction of the OAM operator.

This suggests that radial modes are more difficult to handle, and in fact this is the drawback of using modes based on polar coordinates. There is a limitation on the order of radial eigenmodes that can be addressed by the experimental apparatus. In other words, the modes can be too big: if some outer part of the wave function fails to overlap with the detector, the detected mode corresponds to a superposition of different radial modes. As one gets close to the "limit", using smaller laser beams and larger optics helps, but eventually diffraction sets a hard bound.

To understand the meaning of OAM, we can visualise what transformations it generates. Let's start by considering rotations about the propagation axis of a vector field V, say the z axis. The generator of rotations about z is the z component of the total angular momentum operator J_z, this is a simple consequence of Stone's theorem. As the field that we rotate using J_z is a vector field, the rotation can be separated into two parts. Intuitively, the two parts are a spatial rotation of the field, which rotates about the z axis the origin of each vector, but that leaves the direction where they are pointing (their polarisation) fixed, and a inner rotation of the field, which rotates the polarisation of each vector, but that leaves their origin fixed. Now the question is: is this a fancy and useless decomposition or are there hermitian operators that generate such transformations? It turns out that the answer is positive, although there are some subtleties. The inner rotation is generated by S_z, the component of the spin along z, while the

spatial rotation is generated by the z component of the OAM, L_z. I would like to say a few words about why it's important to consider a particular axis, and this is one of the subtleties of OAM and rotations of massless vector fields. If the field V is a massless spin-1 field, like a photon field, although a spin 1 field should have three eigenstates of S_z, there are just two projections of the spin along the propagation axis, and this is because the photon has no mass, i.e. there is no reference frame in which the photon is at rest, so the spin is always parallel or antiparallel to the direction of propagation. This fact has some interesting consequences. For instance, the three components S_x, S_y and S_z of the spin commute, which means that they are not true angular momenta [41, 42]. The orbital part L_x, L_y and L_z also do not follow the commutation relations of an angular momentum. However, the sums $S_x + L_x$, $S_y + L_y$ and $S_z + L_z$ are true angular momenta, as they should be, being the components of the *total angular momentum*, which is a true angular momentum. As a consequence, we have that OAM and spin are well defined quantities only in the direction of propagation [41–43].

What are the characteristics of eigenstates of the OAM? To answer this question it is sufficient to have a look at the representation of the operator in spherical coordinates: $L_z = -i\partial/\partial\phi$ where ϕ is the azimuthal angle. L_z has this form because it is the conjugated quantity to the azimuthal angle and therefore generates angular rotations about the azimuthal axis, and it follows from the considerations, above, about the Fourier series on the circle. The fields that are eigenstates must therefore have an azimuthal dependence that is proportional to $\exp(i\ell\phi)$ for some integer value ℓ. As it is possible to see in eq. (3.16), Laguerre-Gauss modes are indeed eigenmodes of the OAM. The particular form of the phase turns out to be also rather suggestive.

3.3   Optical fibres

The detection of single photons requires photon counters. In many cases the photons are guided to the photon counters via optical fibres. The physics of optical fibres is a fairly wide area, with contributions from materials physics, engineering and optics [44]. Without covering unnecessary details, I will introduce optical fibres in connection with Part 3 of this thesis, i.e. that concerned with detection.

An optical fibre is a flexible fibre, usually made of silica, that acts as a waveguide for light. Fibres are characterised by the optical properties of their core and of the cladding that surrounds it. An important parameter that characterises the properties of optical fibres is the size of the core, which determines if a fibre
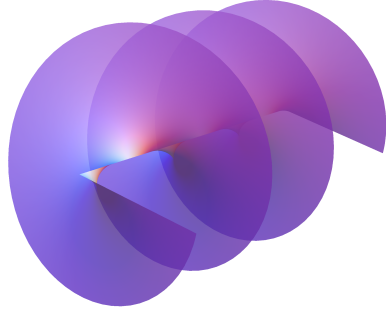
(a) A surface of constant phase of an LG beam with $\ell = 1$.
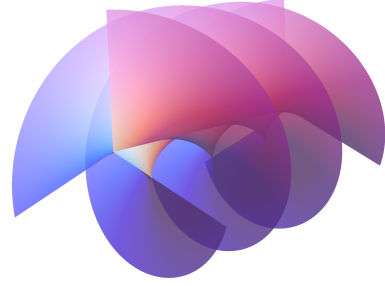
(b) A surface of constant phase of an LG beam with $\ell = 3$.

Figure 3.1: Two examples of surfaces of constant phase of LG beams

is Single Mode (SMF) or Multi Mode (MMF) fibre. A SMF lets only the fundamental gaussian mode propagate, while a MMF can carry multiple modes. This optical property of SM fibres becomes important when one wants to measure single photons in exactly one eigenstate, or as it will be explained in detail in the next chapters, by using a pair of fibres, when one wants to measure exactly one joint mode of an entangled pair of photons.

Light needs to be coupled into the fibre and from the end of the fibre to the photon counter. In order to do so one has to demagnify the beam to fit in the core of the fibre. Additionally, one has to make sure that the angle of incidence is appropriate, giving in total 6 degrees of freedom (3 spatial and 3 angular) to properly align a fibre and couple the signal in. Different setups can employ different alignment techniques, some of which can be automated. Bad alignment can be a major problem and a potential source of error, especially when counting is set up in coincidence. More details of particular implementations will be given in Part 3 of this thesis.

3.4 SPDC state and its approximations

3.4.1 General form of the interaction hamiltonian

As anticipated in chapter 1, the origin of the SPDC effect is due to the particular form of the polarisation density inside the crystal medium. The interaction hamiltonian that gives origin to the down-converted pairs can be written in this
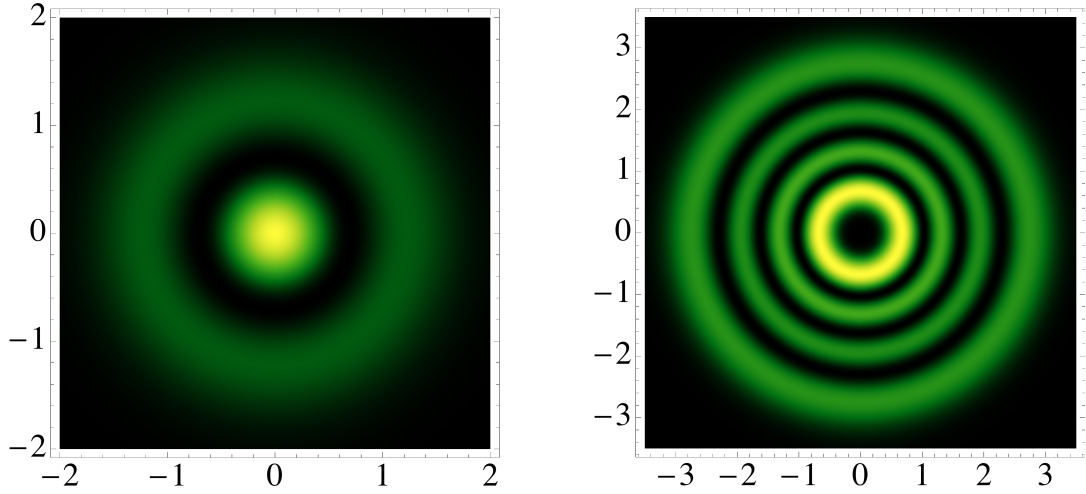
(a) Intensity structure of an LG beam with $p = 1$, $\ell = 0$ and gaussian width $w_0 = 1$. Notice that having no OAM gives no phase singularity at the center, and that $p = 1$ gives one dark ring.

(b) Intensity structure of an LG beam with $p = 3$, $\ell = 3$ and gaussian width $w_0 = 1$. Notice that $p = 3$ gives 3 dark rings.

Figure 3.2: Two examples of intensity structure of LG beams. The intensity scale follows this color bar: min max.

way (I'm using Einstein's notation for tensor sums):

$$H_I = \int_{L^3} d^3x P_k E_k = \int_{L^3} d^3x \chi_{ijk}^{(2)} V_i V_j E_k \tag{3.17}$$

Where $\chi_{ijk}^{(2)}$ is the second order susceptibility tensor; L^3 is the integration volume, i.e. the crystal; E is the field of the pump, V is the field of the generated signal-idler pairs. Being V a very weak field we can quantise it, while maintaining $\mathbf{E}$ a classical field, and obtain

$$\hat{H}_I = \int_{L^3} d^3x \sum_{\mathbf{k}',s'} \sum_{\mathbf{k}'',s''} \hat{a}_{\mathbf{k}',s'}^\dagger \varepsilon_{\mathbf{k}',s'}^* \hat{a}_{\mathbf{k}'',s''}^\dagger \varepsilon_{\mathbf{k}'',s''}^* \chi_{ijk}^{(2)} \mathbf{E}_k e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} \tag{3.18}$$

If $\chi_{ijk}^{(2)}$ does not depend on $\mathbf{r}$ we can integrate and obtain

$$\hat{H}_I = \sum_{\mathbf{k}',s'} \sum_{\mathbf{k}'',s''} \hat{a}_{\mathbf{k}',s'}^\dagger \varepsilon_{\mathbf{k}',s'}^* \hat{a}_{\mathbf{k}'',s''}^\dagger \varepsilon_{\mathbf{k}'',s''}^* \chi_{ijk}^{(2)} \mathbf{E}_k e^{i\omega t} \prod_{m=1}^3 \operatorname{sinc}\left(\frac{1}{2}(\mathbf{k} - \mathbf{k}' - \mathbf{k}'')_m L_m\right) \tag{3.19}$$

If we follow the paraxial prescription and concentrate on the z direction and if we don't use polarisation-dependent detectors, we can simplify the above expression

to obtain

$$\hat{H}_I = \sum_{\mathbf{k}'} \sum_{\mathbf{k}''} \hat{a}_{\mathbf{k}'}^\dagger \hat{a}_{\mathbf{k}''}^\dagger \chi_{ijk}^{(2)} \mathbf{E}_k e^{i\omega t} \mathrm{sinc}\left(\frac{1}{2}\Delta\mathbf{k}_z L\right) \qquad (3.20)$$

In realistic conditions the spatial part of the field $\mathbf{E}$ is a gaussian laser pump field, and interference filters are employed in order to work at degeneracy, i.e. so that signal and idler have the same frequency, to reach indistinguishability in the frequency domain. More general kinds of interaction hamiltonians can be achieved by engineering the susceptibility of the crystal, so that the integration would give a different function than the sinc. The $\chi^{(2)}$ susceptibility, given its small numerical value, yields the low efficiency of the down-conversion process (roughly one photon in 10^{12} is down-converted), so the SPDC states are to be interpreted as "conditional" on the success of the down-conversion process.

3.4.2 SPDC states used in this thesis

It is probably best to show the most general form of the SPDC state that will be used in this thesis and its approximations in quick succession, and then in each chapter we will justify and explain where exactly they come from. In this way the reader knows what he or she is going to read about, and it will be possible to concentrate in each chapter on just what is important for the time being. In chapter 4 we derive the gaussian approximation, in chapter 5 (which of all the chapters is the one that approaches the problem of Schmidt decomposition in the most general way) we give a very detailed justification of the general form of the state, in chapter 6 we carefully introduce the geometry of the beams that allows us to set up the projection integral onto LG modes, and during the calculations we will take the limit that generates the thin crystal approximation and finally in chapter 7 we will show how to set up and use the geometrical approximation.

The SPDC states that will be used throughout the rest of the thesis are:

Full SPDC state

$$\psi(\mathbf{q}_i,\mathbf{q}_s) = \mathcal{N} \underbrace{\exp\left(-\frac{w_p^2}{4}|\mathbf{q}_s+\mathbf{q}_i|^2\right)}_{\text{pump}} \overbrace{\mathrm{sinc}\left(\frac{L|\mathbf{q}_s-\mathbf{q}_i|}{4k_p}+\Phi\right)}^{\text{phase matching}} \qquad (3.21)$$

Gaussian approximation

$$\psi(\mathbf{q}_i, \mathbf{q}_s) = \mathscr{N} \underbrace{\exp\left(-\frac{w_p^2}{4}|\mathbf{q}_s + \mathbf{q}_i|^2\right)}_{\text{pump}} \overbrace{\exp\left(-\frac{L|\mathbf{q}_i - \mathbf{q}_s|^2}{4k_p}\right)}^{\text{phase matching}} \qquad (3.22)$$

Thin crystal approximation

$$\psi(\mathbf{q}_i, \mathbf{q}_s) = \mathscr{N} \underbrace{\exp\left(-\frac{w_p^2}{4}|\mathbf{q}_s + \mathbf{q}_i|^2\right)}_{\text{pump}} \qquad (3.23)$$

In the full SPDC state there is a constant term in the sinc function, Φ, which represents a constant phase mismatch between the down-converted photons. It can be achieved either by tilting the crystal, or by changing its temperature, depending on the type of crystal employed. Notice that in the Gaussian approximation we set to 0 the phase mismatch: $\Phi = 0$, so the state written under this approximation cannot account for a phase mismatch. A workaround is to use another approximation, known as "double gaussian" (or "supergaussian") approximation, but as we will see in chapter 5, it would be an unnecessary complication.

The argument of the phase matching function in these forms of the SPDC state is written making the assumption that the angle between the down-converted photons is not too large. This assumption is conceptually similar to the paraxial approximation. More about it in the next chapters.

To repeat once more what will be done in the next chapters, in chapter 4 we will use the state (3.22) and separate the signal and idler contributions using the smallest number of terms (i.e. we will perform a Schmidt decomposition). In chapter 5 we will do the same, this time numerically, for the state (3.21). In chapter 6 we will project the state (3.23) onto a complete family of LG modes (precisely, we will start with the state (3.21) and then take the limit $L \to 0$ in order to get past the last integration). In chapter 7 we will project the state (3.21) for $\Phi = 0$ onto a $p = 0$ family of LG modes.

As we would like chapters to be roughly self-contained, so that a reader wouldn't need to go back and forth between this introductory chapter and the chapter being read, we will state the relevant notions at the beginning of each chapter, especially if some particular notation is needed, instead of putting everything here. As anticipated, this decision is also motivated by the fact that different chapters need different parts of the description of the SPDC state and

in general none needs everything.

Part II

Generation of high dimensional entanglement

Chapter 4

Analytical Schmidt decomposition, Gaussian approximation

Ah, but if less is more, then just think of how much more "more" will be!

(Dr. Frasier Crane, *Frasier, 7.13*, 2000)

4.1   Introduction

This chapter is based on the paper "Cartesian and polar Schmidt bases for down-converted photons", EPJD **66** (7) 183 (2012) [32]. It consists in a detailed study of the two-photon state that is generated in SPDC. In order to study the entanglement between the two photons, an analytical calculation of the Schmidt decomposition is performed. The conditions under which this can be accomplished are to approximate the phase matching function with a gaussian function. This does not allow to include a phase mismatch, but in a collinear regime (the regime in which there is no phase mismatch) it constitutes a way of proceeding smoothly towards the Schmidt decomposition. One of the results that follows from having an analytical form of the Schmidt decomposition is a closed form of the shared information. We can then investigate how such quantity would change in case of imperfect measurements.

4.2 The SPDC state

In spontaneous parametric down-conversion (SPDC) two lower-frequency photons, commonly referred to as signal and idler, are generated when a pump field interacts with a nonlinear crystal [16]. The spatial structure of the down-converted biphotons depends both on the pump field and on the phase matching. For a gaussian pump beam it can be written as[18]

$$\psi(\mathbf{q}_i, \mathbf{q}_s) = \mathcal{N} \exp\left(-\frac{w_p^2}{4}|\mathbf{q}_s + \mathbf{q}_i|^2\right) \text{sinc}\left(\frac{L\Delta k_z}{2} + \Phi\right) \quad (4.1)$$

where $\mathbf{q}_{s,i}$ are the transverse components of the wave vectors $\mathbf{k}_{s,i}$ for the signal and idler fields, Δk_z is the longitudinal component of the wave mismatch between the pump wave vector $\mathbf{k}_p$ and the down-converted photons wave vectors $\Delta k_z = (\mathbf{k}_p - \mathbf{k}_i - \mathbf{k}_s)_z \simeq (|\mathbf{q}_i - \mathbf{q}_s|^2)/k_p$, where we have made use of the paraxial approximation. Φ is an additional phase mismatch which depends on the internal refractive indices (i.e. its value can be tuned by tilting the crystal or by changing its temperature). For the current analysis $\Phi = 0$.

Near the collinear phase matching regime the analysis can be simplified by using a gaussian approximation of the phase matching term [18, 21] so the state can be written in the form:

$$\psi(\mathbf{q}_i, \mathbf{q}_s) \propto \exp\left(-\frac{|\mathbf{q}_i + \mathbf{q}_s|^2}{\sigma^2}\right) \exp\left(-b^2|\mathbf{q}_i - \mathbf{q}_s|^2\right) \quad (4.2)$$

where b and σ depend on the pump waist, w_p, and wave number, k_p, and on the crystal length, L, in the following way

$$b = \frac{1}{2}\sqrt{\frac{L}{k_p}} \; ; \sigma = \frac{2}{w_p}. \quad (4.3)$$

If we scale the wavevectors $\mathbf{q}_{i,s}$ by a factor $\Gamma' = \sqrt{b/\sigma}$ such that $\mathbf{q} = \mathbf{q}'/\Gamma'$ then we can re-write (4.2) as

$$\psi(\mathbf{q}'_i, \mathbf{q}'_s) \propto \exp\left(-\frac{|\mathbf{q}'_i + \mathbf{q}'_s|^2}{b\sigma}\right) \exp\left(-b\sigma|\mathbf{q}'_i - \mathbf{q}'_s|^2\right) \quad (4.4)$$

where

$$b\sigma = \sqrt{\frac{L}{2z_r}} \quad (4.5)$$

41

with $z_r = \pi w_0^2/\lambda$ the Rayleigh range of the pump beam. Writing (4.2) in this form allows us to see the symmetry between the two parts of the wave function and will help explain some of the later results. In particular, we can see that the results will depend on the product $b\sigma$ and not b and σ independently.

4.3   Schmidt decomposition

As was anticipated in the first chapter, a useful, and experimentally convenient, measure of the entanglement is the Schmidt number, K. In order to calculate this we first need to calculate the Schmidt decomposition of the down-converted state, which is done by writing it in the form

$$|\Psi\rangle_{AB} = \sum_i \sqrt{\lambda_i}|\alpha_i\rangle_A|\beta_i\rangle_B, \qquad (4.6)$$

where $|\alpha_i\rangle_A, |\beta_i\rangle_B$ are the Schmidt modes, defined by the eigenvectors of the reduced density matrices, and the real and positive Schmidt coefficients, $\sqrt{\lambda_i}$, are the corresponding eigenvalues, with each of the factors in the normalized set $\{\lambda_i\}$ representing the probability of detecting the entangled state in the ith entangled Schmidt mode $|\alpha_i\rangle_A|\beta_i\rangle_B$. If all of the coefficients $\sqrt{\lambda_i}$ are different, then the Schmidt decomposition is unique. Whenever some of the coefficients $\sqrt{\lambda_i}$ are equal, one has a choice of infinitely many different Schmidt bases, as the degenerate vectors form a vector space for which a basis can be chosen at will and the Schmidt decompositions are connected by a unitary transformation.

The Schmidt decomposition provides insights into the nature of the bipartite entanglement by determining the natural set of biorthogonal mode pairs (or orthonormal bases) for the two systems [3, 21, 38, 45, 46] while the coefficients allow us to calculate the Schmidt number, K (i.e. the average number of modes in the state) and the entropy of entanglement. By knowing the Schmidt decomposition explicitly in HG and LG modes, one can easily implement an ideal detection basis.

As the SPDC state can be equivalently described using either HG or LG modes it is interesting to perform the Schmidt decomposition in the two corresponding coordinate systems. By approximating the biphoton state as a Gaussian in (4.2) it is possible to calculate the Schmidt decomposition in an analytical form for both. Their relationship can be shown by using the well-known relationship between HG and LG modes [47]. Note that in both cases, the Schmidt modes for the SPDC state will have the same form for both signal and idler, due

to symmetry requirements [46] and so we will obtain a decomposition of the form:

$$\psi(\mathbf{q}_i, \mathbf{q}_s) = \sum_{a,b} \sqrt{\lambda_{a,b}}\, u_{a,b}(\mathbf{q}_i) u^*_{a,b}(\mathbf{q}_s), \qquad (4.7)$$

where the functions $u_{a,b}$ depend on the coordinate system employed, and the labels a and b correspond to different degrees of freedom: in the cartesian case a and b will be replaced with m and n, in the polar case with ℓ and p. The m and n quantum numbers label the two transverse degrees of freedom, while the ℓ and p quantum numbers label the angular and radial degrees of freedom, respectively. The sum is calculated on two indices because we perform the decomposition in the two-dimensional plane perpendicular to the direction of emission.

4.3.1 Decomposition in cartesian coordinates

In a recent paper, Straupe *et al.* [48] reported a proof-of-principle experiment demonstrating that an appropriately chosen set of HG modes constitutes a Schmidt decomposition for transverse momentum states of biphotons generated by SPDC. For clarity and completeness we perform an equivalent Schmidt decomposition in cartesian coordinates before extending our analysis to polar coordinates in the next section and then demonstrating their equivalence.

The cartesian decomposition requires a separation of each of the variables $\mathbf{q}_i$ and $\mathbf{q}_s$ into a pair of orthogonal variables, q and $q_\perp$, so that the wave function assumes the form $\psi(\mathbf{q}_i, \mathbf{q}_s) \to \psi(q_i, q_s, q_{i\perp}, q_{s\perp})$. We define the cartesian basis of HG modes as

$$h_n(\Gamma q) = \frac{\sqrt{\Gamma}\, e^{-\Gamma^2 q^2/2} H_n(\Gamma q)}{(n!2^n \sqrt{\pi})^{1/2}} \qquad (4.8)$$

where $\Gamma = 2\sqrt{\frac{b}{\sigma}} = w_p \sqrt[4]{\frac{L}{2z_r}}$ is the width of the HG modes. If we express the wave function (4.2) in terms of this basis we obtain

$$\psi = (1 - \mu^2) \sum_{m,n \geq 0} \mu^{m+n} h_{mn}(q_i, q_{i\perp}) h_{mn}(q_s, q_{s\perp}), \qquad (4.9)$$

where

$$h_{mn}(x, y) = h_m(\Gamma x) h_n(\Gamma y) \qquad (4.10)$$

and

$$\mu = \left| \frac{b\sigma - 1}{b\sigma + 1} \right|. \qquad (4.11)$$

Expression (4.9) is the cartesian Schmidt form for the SPDC state with cor-

responding coefficients:

$$\sqrt{\lambda_{m,n}} = (1-\mu^2)\mu^{m+n} = \frac{4b\sigma}{(1+b\sigma)^2}\left|\frac{b\sigma-1}{b\sigma+1}\right|^{m+n}. \tag{4.12}$$

Note that these are exactly equivalent to those given in equation (5) of [48].

The way this result was obtained required the following mathematical result:

$$e^{-G(x^2+y^2-2\eta xy)} = \sqrt{1-|\mu|^2}\sum_{n=0}^{\infty}\mu^n h_n(\Gamma x)h_n(\Gamma y) \tag{4.13}$$

for $|\eta| < 1$, $G > 0$ and where

$$h_n(\Gamma w) = \frac{\sqrt{\Gamma}e^{-\Gamma^2 w^2/2}H_n(\Gamma w)}{(n!2^n\sqrt{\pi})^{1/2}},$$

where $H_n(v)$ are Hermite polynomials. A proof of this formula can be easily obtained with the use of generating functions for the Hermite polynomials.

In order to apply this result to the state (4.2) we set $G = b^2 + 1/\sigma^2$ and $\eta = (b^2\sigma^2-1)/(b^2\sigma^2+1)$. A bit of algebra gives the relations

$$\mu = \left|\frac{G\eta}{(G+\Gamma^2/2)}\right| = \left|\frac{b\sigma-1}{b\sigma+1}\right|;\Gamma = \sqrt{\frac{4b}{\sigma}}. \tag{4.14}$$

Equation (4.2) can thus be written in the form

$$\psi = \mathcal{N}e^{-G(q_i^2+q_s^2-2\eta q_i q_s)}e^{-G(q_{i\perp}^2+q_{s\perp}^2-2\eta q_{i\perp}q_{s\perp})}$$
$$= (1-\mu^2)\sum_{m,n}\mu^m\mu^n h_m(\Gamma q_i)h_m(\Gamma q_s)h_n(\Gamma q_{i\perp})h_n(\Gamma q_{s\perp}) \tag{4.15}$$

Let $h_{mn}(x,y) = h_m(\Gamma x)h_n(\Gamma y)$. Using the properties of Hermite polynomials one can verify that these functions form a complete orthonormal set for $\mathcal{L}^2(\mathbb{R}^2)$ and the result in eq. (4.9) follows.

4.3.2 Decomposition in polar coordinates

As LG modes are currently the preferred basis for many spatial entanglement experiments we also calculate the Schmidt decomposition in polar coordinates. The polar decomposition requires a separation of each of the variables $\mathbf{q}_i$ and $\mathbf{q}_s$ into a pair of polar variables, so that the wave function assumes the form $\psi(\mathbf{q}_i,\mathbf{q}_s) \rightarrow \psi(\rho_i,\rho_s,\varphi_i,\varphi_s)$, where ρ and φ are the radial and angular variables.

We take the LG modes to have the standard definition in momentum space, and we add the same scaling factor Γ as in the cartesian case:

$$LG_p^\ell(\Gamma\rho, \varphi) = \sqrt{\frac{\Gamma^2 p!}{\pi(p+|\ell|)!}} e^{-\frac{\Gamma^2\rho^2}{2}} (\Gamma\rho)^{|\ell|} L_p^{(|\ell|)} \left(\Gamma\rho^2\right) e^{i\ell\varphi} \qquad (4.16)$$

where $L_p^{(\ell)}$ are generalized Laguerre polynomials. One can express the wave function (4.2) in terms of LG modes, of width $\Gamma/\sqrt{2}$

$$\psi = (1-\mu^2) \sum_{\ell=-\infty}^{\infty} \sum_{p=0}^{\infty} \mu^{2p+|\ell|} LG_p^\ell(\Gamma\rho_i, \varphi_i) LG_p^{-\ell}(\Gamma\rho_s, \varphi_s). \qquad (4.17)$$

where $\Gamma = 2\sqrt{\frac{b}{\sigma}} = w_p \sqrt[4]{\frac{L}{2z_r}}$.

The above expression is the polar coordinate form for the Schmidt decomposition of the SPDC state, where μ is defined in (4.11) and the polar Schmidt coefficients are given by

$$\sqrt{\lambda_{\ell,p}} = (1-\mu^2)\mu^{2p+|\ell|} = \frac{4b\sigma}{(1+b\sigma)^2} \left|\frac{b\sigma-1}{b\sigma+1}\right|^{2p+|\ell|}, \qquad (4.18)$$

to be compared to (4.12).

Unlike for the cartesian decomposition, which was performed in one step because cartesian orthogonal degrees of freedom play the same role, for the polar decomposition it is necessary to separate the angular variables and the radial variables in a different way. The angular variables will be separated with the Fourier transform, while the radial variables will be separated with a variation of formula (4.13).

As the first step, we can rewrite the wave function (4.2) in polar coordinates and show that it is a function of the *difference* of the angular variables. This fact enforces the conservation of OAM and allows to write it as a sum over the Fourier components of the difference of the angular variables:

$$\psi = \mathcal{N} \exp\left[-\frac{1}{\sigma^2} \left(\rho_i^2 + \rho_s^2 + 2\rho_i\rho_s \cos(\varphi_i - \varphi_s)\right) \right.$$
$$\left. -b^2 \left(\rho_i^2 + \rho_s^2 - 2\rho_i\rho_s \cos(\varphi_i - \varphi_s)\right) \right] \qquad (4.19)$$

$$= \frac{1}{2\pi} \sum_\ell \sqrt{P_\ell} F_\ell(\rho_i, \rho_s) e^{i\ell(\varphi_i - \varphi_s)} \qquad (4.20)$$

where the sum runs over all integers. The Fourier components are easily found:

$$\sqrt{P_\ell}F_\ell(\rho_i,\rho_s) = \mathcal{N}e^{-\left(b^2+\frac{1}{\sigma^2}\right)\left(\rho_i^2+\rho_s^2\right)}I_{|\ell|}\left[2\left(b^2-\frac{1}{\sigma^2}\right)\rho_i\rho_s\right], \qquad (4.21)$$

where $I_\ell(\cdot)$ is the ℓth order modified Bessel function of the first kind.

The next step is to decompose each angular eigenfunction into a radial superposition of orthogonal modes. The mathematical result needed to proceed is:

$$\sum_{p=0}^{\infty}\mu^{2p}r_p^{(\ell)}(x)r_p^{(\ell)}(y) = \frac{|\mu|^{-|\ell|}}{1-\mu^2}e^{-\frac{x^2+y^2}{2}\frac{1+\mu^2}{1-\mu^2}}I_{|\ell|}\left(2xy\frac{|\mu|}{\mu^2-1}\right) \qquad (4.22)$$

a proof of which can be found in [49]. Here the $r_p^{(\ell)}$ functions are given by

$$r_p^{(\ell)}(x) = \sqrt{\frac{2p!}{(p+|\ell|)!}}e^{-\frac{x^2}{2}}x^{|\ell|}L_p^{(|\ell|)}(x^2) \qquad (4.23)$$

where $L_p^{(\ell)}$ are generalized Laguerre polynomials.

We apply the formula (4.22) to the functions in eq. (4.21). A bit of algebra yields the correct value of the parameter μ and the correct scaling, Γ, of the $r_p^{(\ell)}(\Gamma\rho)$ functions:

$$\mu^2 = \left(\frac{1-b\sigma}{1+b\sigma}\right)^2 ; \Gamma = \sqrt{\frac{4b}{\sigma}}. \qquad (4.24)$$

Notice that the values are analogous to the cartesian case.

Applying these results and normalizing the radial modes, the result in eq. (4.17) follows. It's interesting to note how easy it is to get rid of the pair of angular degrees of freedom. This is a consequence of the mathematical form of LG modes.

4.3.3   Equivalence of Schmidt bases

The expressions calculated above are equivalent descriptions of the entangled state and, just as it is possible to transform HG modes into LG modes [47] and vice versa [6], we are also able to convert between our two Schmidt bases. In fact, it is straightforward to convert the Schmidt decomposition in cartesian coordinates, equation (4.8), into that in polar coordinates, (4.16). The first step is to notice that the values of m and n that satisfy $m+n=N$ yield the same Schmidt coefficient $(1-\mu^2)\mu^N$; for the polar case this happens for all the values of ℓ and p that satisfy $|\ell|+2p=N$. The number N is called the mode order. We previ-

ously mentioned that whenever there was a degeneracy of a Schmidt coefficient, each degeneracy subspace could be used for writing a different Schmidt decomposition. The mode order labels the degeneracy subspaces, and this is what is needed for passing between the cartesian and the polar Schmidt bases. The relation between the cartesian basis (4.8) and the polar basis (4.16) is [47, 50]

$$LG_p^{(\ell)}(\rho) = \sum_{k=0}^{N} b_{p,k}^{(N)} h_{N-k,k}(q, q_\perp) \tag{4.25}$$

where the relation between cartesian and polar coordinates is the canonical one. Notice how the above formula takes place in a modal order subspace. The b coefficients are grouped by the mode order and they span the degeneracy subspace through the label k, and are given by

$$b_{p,k}^{(N)} = \frac{i^k(-1)^{p+k}}{2^{N/2}k!} \frac{d^k}{d\mu^k}[(1-\mu)^n(1+\mu)^m]_{t=0} \tag{4.26}$$

where $m + n = |\ell| + 2p = N$. This fact enforces the conservation of the mode order when changing Schmidt basis.

4.4 Analysis of the entanglement

We now look at two ways of assessing the amount of entanglement in the down-converted state. The first is the Schmidt number, defined in (2.9), the second is the Rényi entropy.

4.4.1 Schmidt number

The entanglement of a state can be quantified by the probability distribution of the modes it contains. Intuitively, a state is more entangled whenever this probability distribution is more 'spread out'. A particularly important measure of entanglement is the Schmidt number, K, which corresponds to the number of significant modes in the Schmidt decomposition [38, 46]. We remind that for states in the form (4.7), it is defined as

$$K = \frac{\text{Tr}[\hat{\rho}]^2}{\text{Tr}[\hat{\rho}^2]} \equiv \frac{1}{\sum_{a,b} \lambda_{a,b}^2} \tag{4.27}$$

where $\hat{\rho}$ is the reduced state formed by tracing over one part of a pure bipartite state and $\sqrt{\lambda_{a,b}}$ are the Schmidt coefficients that appear in the Schmidt decom-

position of the bipartite state. One can immediately see that the eigenvalues of the reduced state are just the square of the Schmidt coefficients. A state will be separable when $K = 1$ and entangled if $K > 1$. Applying this to the Schmidt decompositions calculated earlier, gives

$$K = \left[\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\lambda_{m,n}^2\right]^{-1} = \left[\sum_{\ell=-\infty}^{\infty}\sum_{p=0}^{\infty}\lambda_{\ell,p}^2\right]^{-1}$$
$$= \frac{1}{4}\left(b\sigma + \frac{1}{b\sigma}\right)^2. \tag{4.28}$$

This result agrees with previous calculations of K [21] and, as is clear from eq. (4.28) that is independent of the Schmidt basis used, as it should be. This means that one has the freedom to choose the basis that best matches the experimental conditions with no consequence on the dimensionality of the Hilbert space that is spanned by the detection basis.

The effect of the experimental parameters can be seen more clearly if, as in (4.5), we write $b\sigma = w_p\sqrt{L/k_p} = \sqrt{L/2z_r}$ where L is the crystal thickness, w_p and k_p are the width and wave number, respectively, of the pump and z_r is its Rayleigh range. Note that $K = 1$ (which means that the state is not entangled) whenever $b\sigma = 1$, which corresponds to choosing experimental parameters such that the crystal length is twice the Rayleigh range ($L = 2z_r$).

4.4.2 Rényi entropy

An alternative approach to quantifying the entanglement of an SPDC state is to calculate its entropy. The most famous entropic function is the Shannon entropy, which appears in information theory and statistical mechanics [51, 52]. A more general measure, however, is the Rényi entropy, which is obtained by neglecting the grouping property of entropy [53, 54]. For a probability distribution $\{p_k\}$ the Réyni entropy of order α is defined as

$$H_\alpha(\{p_k\}) = \frac{1}{1-\alpha}\log_2\left(\sum_k p_k^\alpha\right), \ \alpha > 0. \tag{4.29}$$

Convergence of the series above has to be ensured when dealing with values of α between 0 and 1. Note that when $\alpha \to 1$ one regains the Shannon entropy.

A simple calculation shows that the Rényi entropy of the Schmidt coefficients

in either eq. (4.12) or (4.18) is

$$H_\alpha(b\sigma) = \frac{2}{\alpha - 1} \log_2 \frac{(4b\sigma)^\alpha}{|1 + b\sigma|^{2\alpha} - |1 - b\sigma|^{2\alpha}}. \tag{4.30}$$

Using equation (4.28) we can rewrite the parameter $b\sigma$ in terms of the Schmidt number, K, as $b\sigma = \sqrt{K} - \sqrt{K-1}$ and hence find the Rényi entropy in terms of the Schmidt number. Replacing $b\sigma$ with $\sqrt{K} - \sqrt{K-1}$ in (4.30) gives the Rényi entropy as a function of K, which can be approximated by

$$H_\alpha(K) \simeq \log_2 K - f(\alpha) \tag{4.31}$$

where $f(\alpha) = 2 - \log_2(\alpha^2)/(\alpha - 1)$. For a discussion of the quality of this approximation, see the next subsection. This tells us that, to a good approximation, valid for sufficiently high K, different orders of the Rényi entropy differ by a constant value. Note that $f(2) = 0$ and thus $H_2 = \log_2 K$, in fact the Schmidt number is related to the Rényi entropy of order 2 by $K = 2^{H_2}$, and H_2 can be interpreted as the number of bits needed to express K.

4.4.3 Von Neumann entropy

For quantum systems we can write the quantum Rényi entropy as

$$H_\alpha[\hat{\rho}] = \frac{1}{1 - \alpha} \log_2 \left[\mathrm{Tr}\left(\hat{\rho}^\alpha \right) \right], \ \alpha > 0, \tag{4.32}$$

where $\hat{\rho}$ is the reduced density matrix [55]. An important special case of the Rényi entropy is when one takes the limit $\alpha \to 1$ in which case (4.32) reduces to the von Neumann entropy of the reduced state [3, 45]

$$S[\hat{\rho}] = -\mathrm{Tr}[\hat{\rho} \log_2 \hat{\rho}], \tag{4.33}$$

as it should be, the von Neumann entropy being the quantum version of the Shannon entropy. The entropy of a reduced state is known as either the index of correlation [56] or the entanglement entropy and the importance of the entanglement entropy stems from the fact that it quantifies the number of entangled bits (or ebits) within the state [3, 45, 57]. This is very important in case of qubits: means that if one has n copies of an entangled qubit, with entanglement entropy $0 < S < 1$, then one can asymptotically convert this collection to approximately nS maximally entangled states, for instance using Bennett's concentration algorithm [57, 58]. It has been shown that the maximum amount of

shared information that two parties can extract from an entangled pure state is given by entanglement entropy of their state [56, 59–61]. Our results thus enable us to determine the maximum amount of shared bits per photon pair that two parties can extract from SPDC states. Using our results, one can also see how changing the parameters of the pump or the crystal affects the amount of shared information.

Taking the limit $\alpha \to 1$ in (4.31), shows that the entanglement entropy for the SPDC state can be approximated by a logarithmic relation:

$$S(K) = \lim_{\alpha \to 1} H_\alpha \simeq 1 + \log_2(K). \tag{4.34}$$

Such a relation is an approximation that holds well for large enough values of K. In fact, a Taylor expansion of the non approximated von Neumann entropy for large K yields

$$S(K) = \underbrace{\frac{2}{\log 2}}_{\sim 0.9} - 2 + \log_2 K - \frac{1}{K \log 8} + O\left(\frac{1}{K}\right)^2 \tag{4.35}$$

Without considering such a power expansion, it is not immediately obvious how much the relation between Schmidt number and von Neumann entropy differs from a purely logarithmic approximation. As an example of the failure of the approximation (4.34), note that in the regime where the state is not entangled, i.e. if there is only one joint mode (and so $K = 1$) the amount of quantum correlation in the state has to be 0 (dashed line in Fig. 4.1) and not 1 (solid line in Fig. 4.1). Common experimental conditions where one seeks high dimensional entanglement are in the range of $b\sigma \ll 1$, where the relation (4.34) is accurate. However, experiments with a tightly focussed pump, or a long crystal, may fall in the region closer to $b\sigma \sim 1$, where it fails. A plot of the von Neuman entropy and of its approximation are given in figure 4.1.

An interesting feature of figure 4.1 is that the results are symmetric under the substitution $b\sigma \to 1/b\sigma$. This can be explained by reference to equation (4.22) which describes the correlations in the two conjugate planes $(\mathbf{q}'_i + \mathbf{q}'_s)$ and $(\mathbf{q}'_i - \mathbf{q}'_s)$. Interchanging $b\sigma$ and $1/b\sigma$ corresponds to 'squeezing' in one plane but 'expanding' in the other: the product of the two is constant. Physically this means that an experiment with a crystal of length L and Rayleigh range z_r is equivalent to an experiment with crystal length az_r and Rayleigh range of L/a. We expect this symmetry to be no longer exact outside the approximation of the gaussian phase matching function.
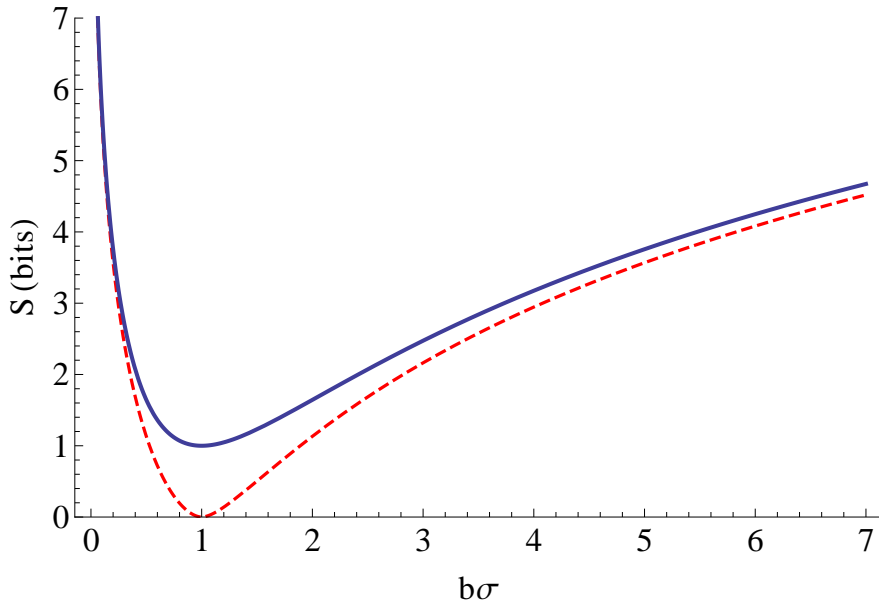
Figure 4.1: A plot of the shared bits of information per photon pair, given in eq. (4.34), plotted against $b\sigma = \sqrt{L/2z_r}$ (solid line) and of the non approximated Von Neumann entropy (dashed line). Note that they match for large values of K, that is for small or very large values of $b\sigma$.

As we already mentioned, the value of the shared bits reaches its minimum for $b\sigma = 1$. In this regime the state is separable, i.e. not entangled. In fact, for such a value of $b\sigma$, the Schmidt number K is also 1, which means that the state can be written using only one Schmidt mode, which is separable by definition. We will now assume a large enough Schmidt number to safely use the definition (4.34) of the entanglement entropy.

The logarithmic dependence of $S(K)$, for large K, has two important consequences. Both are due to the small value of the slope of the logarithm curve for large values of the argument. The first is that if we can prepare an SPDC state with a large number of modes, and thus large Schmidt number, any further increase in the number of modes will provide only a modest increase in both the entanglement and the number of shared bits that one can extract. The second consequence is concerned with the non-ideal detection of the entangled state and is discussed in the following subsection.

It is interesting to compare this result with the work in ref. [29], in particular the isotropic case, but also [28], where a relation between spectral entanglement and a control parameter similar to (4.28) is found. In particular we note that in their assessment of the spectral-temporal part of the down-converted state, the entanglement strength has its minimum at a value considerably larger than zero, meaning that spectral entanglement is never small. In our case, however,

the state is spatially separable when $b\sigma = 1$. In both cases, the control parameters depend upon the characteristics of the crystal and of the pump beam and this could have an implication on the extent of hyperentanglement.

4.4.4 Non-ideal detection

There can be many sources of non-ideal detection. These can range from defects in the measuring apparatus (which give rise to cross-talk between channels, accidental coincidences, dark counts, etc.) to turbulence (that can affect the propagation of the states), to non-ideal choices of the optical elements in the setup (which determine a mismatch between the Schmidt modes and the detection modes and therefore impairs the ability to detect high order modes). We concentrate, in particular, on cases in which the number of modes that a measurement apparatus can detect is less than the number of modes that the source is producing. This type of experimental inaccuracy is fundamentally different from the others, which have been studied for instance in [62, 63], the difference being that in the previous work it was always assumed that a detector could have access to all the modes that are produced by the source. Detection of entangled states by projection onto modes that match the Schmidt modes, such as the ones given by eqs. (4.12) and (4.18), will yield the maximum amount of shared bits. However, if the detection basis does not exactly match the Schmidt basis, the effective number of Schmidt modes that are measured, K_{eff}, will be less than K. Provided the state is highly entangled (i.e. large K), the logarithmic relationship between the number of shared bits and the Schmidt number, given in equation (4.34), means that even if the fraction of entangled modes that are detected, $\eta = K_{\text{eff}}/K$, is small, this need not be too detrimental to the fraction of shared bits $S(\eta K)/S(K)$ that one can extract. This result may seem counter-intuitive, however, the key point to note is that information is measured by entropy, not by the number of modes [51, 53]. For example, the number of different messages that one could encode using 4 modes can be described using 2 binary digits, while 8 modes require 3 binary digits, i.e. the information increases by one bit every time the number of different messages doubles.

To illustrate this idea consider the following example. Suppose that one can generate a state with a large Schmidt number, K, but that imperfections in the detection of the modes means that the number of effective modes that can be accessed is only $K/2$, i.e. $\eta = 1/2$. The number of shared bits will then be $S(K/2) = S(K) - 1$: every time K is halved, one shared bit is lost. As we see in figure 4.2, the reduction in the entropy will be negligible for large enough values of K.
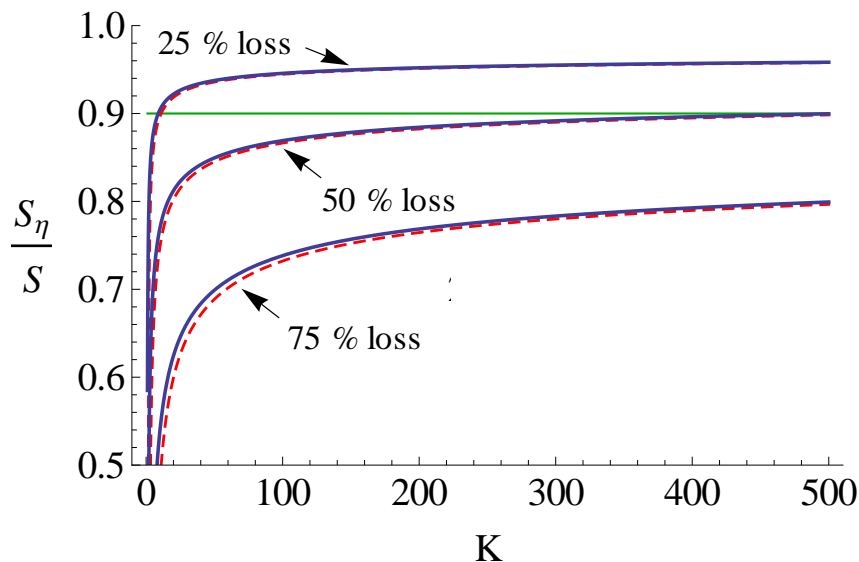
Figure 4.2: A plot of the fraction of shared bits per photon pair as a function of K for three types of measurements that yield different amounts of loss of joint modes. The solid lines are derived from the logarithmic approximation (4.34), the dashed lines are derived from the non-approximated Von Neumann entropy.

If one is interested in determining the experimental parameters needed to retain a certain amount of shared bits, in Fig. 4.3 we recast Fig. 4.2 in terms of $b\sigma$. To give some realistic numbers, even if the detection basis allows only half of the modes to be detected, it is still possible to retain more than 90% of the shared bits, if states with more than ~ 500 entangled modes are used, which corresponds to $b\sigma \sim 0.02$. Values of hundreds of entangled modes can be produced within the limitations of experimental equipment [64].

4.5   Conclusions

We have seen how an analytical Schmidt decompositions for the biphoton state produced using SPDC can be derived in both cartesian and polar coordinates for cases in which the sinc phase matching term can be approximated with a gaussian. The resultant modes exhibit either the orthogonal characteristics of Hermite-Gauss modes of width Γ or the angular and radial characteristics of Laguerre-Gauss modes of width $\Gamma/\sqrt{2}$, respectively, and can be shown to be equivalent using the transformation relations between HG and LG modes. An entropic analysis of these different derivations confirms that the strength of the spatial entanglement is independent of the choice of coordinates and gives the freedom
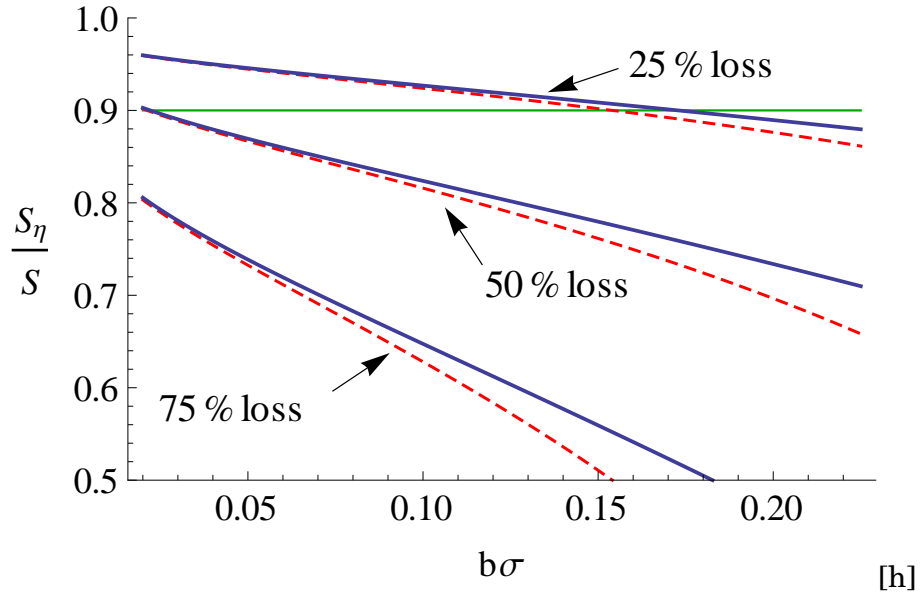
Figure 4.3: A plot of the fraction of shared bits per photon pair as a function of $b\sigma$ for three types of measurements that yield different amounts of loss of joint modes. The solid lines are derived from the logarithmic approximation (4.34), the dashed lines are derived from the non-approximated Von Neumann entropy.

to choose the Schmidt basis that is most appropriate for given experimental conditions. We have seen that an insight into the strength of the entanglement can be gained by calculating the Schmidt number or the Rényi and von Neumann entropies for the SPDC state. We saw under what conditions it is safe to use a logarithmic relation between the number of Schmidt modes and the maximum number of bits per photon that one can extract. Such logarithmic relation demonstrates that, for highly entangled states, the information loss due to non-ideal measurements can be only a small fraction of the maximum information. We finally showed, given any non-ideal measurement, which will allow to detect a smaller number of entangled modes K_{eff}, what is the Schimdt number K that is needed in order to retain a given fraction of shared bits, and from this what is the experimental parameter $b\sigma$ that one should look for.

Chapter 5

Numerical Schmidt decomposition

5.1 Introduction

This chapter is based on the paper "Spatial Schmidt modes in parametric down-conversion", submitted to EPJD, in press (2012) [33]. In this chapter we will see how a numerical approach can yield more insight on the entanglement between down-converted pairs. In particular, the issue of being able to take into account a phase mismatch between the photons, when calculating entanglement strength and detection probabilities, has not yet found a full resolution. Approximations have been made, and with these, very interesting analytical results have been found [32, 48]. However, approximations naturally restrict the domain of validity of the results. The most common approximation used in SPDC theory (replacing a sinc function with a Gaussian function) is discussed in this chapter. Our key challenge is how to decompose (with numerical tools) the two-photon field $A(\mathbf{q}_1, \mathbf{q}_2)$ into its Schmidt modes $u_i(\mathbf{q}_1)$ and $v_i(\mathbf{q_2})$:

$$A(\mathbf{q}_1, \mathbf{q}_2) = \sum_i \sqrt{\lambda_i} u_i(\mathbf{q}_1) v_i(\mathbf{q_2}), \qquad (5.1)$$

without the aid of the gaussian phase matching approximation. We will look at a numerical analysis of the entanglement strength, of the Schmidt modes, and of the detection amplitudes without the usual approximations. The first part is a discussion of the physical system under study and a presentation of the meaning of the Schmidt decomposition. In the second part we will analyse the numerical results, and study some interesting features of the structure of the entanglement.

5.2 Theory

5.2.1 Generated two-photon field

We consider the generation of entangled photon pairs via the nonlinear optical process of SPDC, where single pump photons which propagate through a nonlinear crystal occasionally split in two photons of lower energy. We operate in the quasi-monochromatic frequency-degenerate limit, by combining a narrow-band cw pump laser at frequency ω_p with narrow-band spectral filters at the degenerate frequency $\omega_1 = \omega_2 = \omega_p/2$. When the pump field is spatially coherent and the spectral filtering is sufficiently narrow, the generated two-photon state (ignoring polarisation and therefore describing both type-I and type-II down-conversion) is pure and has the general form [4]

$$|\Psi\rangle = \int\int A(\mathbf{q}_1, \mathbf{q}_2)\hat{a}^\dagger(\mathbf{q}_1)\hat{a}^\dagger(\mathbf{q}_2)\,d\mathbf{q}_1 d\mathbf{q}_2|0\rangle, \tag{5.2}$$

where $\hat{a}^\dagger(\mathbf{q}_i)$ is the creation operator of a plane wave with transverse momentum $\mathbf{q}_i$ and i is the photon label.

Formally, the two photon amplitude $A(\mathbf{q}_1, \mathbf{q}_2)$ is just the SPDC wavefunction. Notice that we are using the transverse coordinates in momentum space. In chapter 4, as we set up in detail the geometry of the beams, we will start from a more general amplitude which is a function of the complete $\mathbf{k}$ vectors (not just their transverse component $\mathbf{q}$), and then show how the transverse part is indeed the one that matters most to us.

One of the factors of the two-photon amplitude $A(\mathbf{q}_1, \mathbf{q}_2)$ in eq. (5.2), is the phase matching function, which describes the efficiency of the down-conversion process, as a function of transverse k-vectors. The longitudinal phase matching function in the two-photon amplitude is a sinc function, being the Fourier transform of the step function that represents the uniform amplitude of creating the photon pair along the crystal. It is usually approximated with a Gaussian function if no or very little phase mismatch (indicated by Φ) is present, as it was done in chapter 2. However, the description that derives from such approximation cannot account for an arbitrary phase mismatch. Moreover, even in the regime where $\Phi = 0$, the difference in shape of the approximated amplitude from its complete form has an effect in the analysis of the down converted state, as it will be shown.

The the two-photon amplitude in the far field, or in momentum space, presents a ringed pattern, due to the sinc function. If the phase is matched ($\Phi = 0$) the

sinc gives a bright spot in the centre and secondary rings around it. Most of the intensity of the field is concentrated in the central area, and it is for this reason that in the case of perfect phase matching it is possible to roughly approximate the phase matching function with a gaussian function (see below for more details). We would like to analyse also situations in which $\Phi \neq 0$. In this case the argument of the sinc function is modified by a constant value, see eq. (5.4). The effect in the amplitude for negative and decreasing Φ is to lose the central spot and obtain a more divergent field, in fact every secondary ring obtains a larger radius the lower the value of Φ. On the other hand, for positive and increasing Φ the effect is to obtain an overall weaker field, less divergent, with a central spot intensity that depends on the value of Φ. We will refer to the effect of tuning Φ as "opening and closing the rings".

Sinc phase matching

The complete form of the two photon amplitude is [21]

$$A(\mathbf{q}_1, \mathbf{q}_2) \propto \mathscr{E}_p(\mathbf{q}_1 + \mathbf{q}_2)\mathrm{sinc}\left(\frac{1}{2}\Delta k_z L\right), \tag{5.3}$$

where $\mathscr{E}_p(\mathbf{q_1} + \mathbf{q_2})$ is the transverse momentum profile of the pump beam ($\mathbf{q}_p = \mathbf{q}_1 + \mathbf{q}_2$). The function $\mathrm{sinc}\left(\frac{1}{2}\Delta k_z L\right)$ quantifies the influence of phase matching, where L is the crystal length and where the projected wave-vector mismatch $\Delta k_z = \Delta k_z(\mathbf{q}_1, \mathbf{q}_2)$ is a function of $\mathbf{q}_1$ and $\mathbf{q}_2$.

We make the general Eq. (5.3) more specific by inserting a (rotationally-symmetric) Gaussian pump profile. Other, more general, types of pump profiles can be treated analytically in the limit for a short crystal [65]. We also assume non-critical phase matching, to remove any linear dependence of Δk_z on $\mathbf{q}_i$, and perform a Taylor expansion of $\Delta k_z(\mathbf{q}_1, \mathbf{q}_2)$, neglecting a small term that scales with the difference between the refractive indices at the pump and SPDC wavelength multiplied by $|\mathbf{q}_1 + \mathbf{q}_2|^2$. Following the notation of Law and Eberly [21], we thus rewrite Eq. (5.3)

$$A(\mathbf{q}_1, \mathbf{q}_2) \propto \exp\left(-|\mathbf{q}_1 + \mathbf{q}_2|^2/\sigma^2\right) \times \mathrm{sinc}\left(b^2|\mathbf{q}_1 - \mathbf{q}_2|^2 + \Phi\right), \tag{5.4}$$

where $\sigma = 2/w_p$, for a Gaussian pump profile $\mathscr{E}_p(\mathbf{x}) \propto \exp(-|\mathbf{x}|^2/w_p^2)$, and $b^2 = L/(4k_p)$, where $k_p = n\omega_p/c$ is the momentum of a pump photon in the crystal [18], and Φ is the collinear phase mismatch. Note that σ and b have the exact same meaning as in the previous chapter, but now they belong to the general SPDC state.

Although this challenge has been tackled mathematically by Law and Eberly [21], several aspects and physical implications of the Schmidt decomposition remained undiscussed and are still a topic of active research [32, 48]. The work that is summarised in this chapter aims to be as complete as possible in this discussion.

Gaussian phase matching

As we saw previously, to obtain analytic solutions, many authors [18, 21, 32] replace the sinc-type phase matching function at $\Phi = 0$ by its Gaussian approximation, using the substitution $\mathrm{sinc}\left(b^2|\mathbf{q}_1 - \mathbf{q}_2|^2\right) \approx \exp\left(-b^2|\mathbf{q}_1 - \mathbf{q}_2|^2\right)$. The Gaussian approximation, as explained in chapter 2, allows an exact Schmidt decomposition with Hermite-Gaussian or Laguerre-Gaussian eigenmodes [21, 32, 48, 66]. The Schmidt number of this decomposition, i.e., the effective number of modes that participate in the modal decomposition, is [21]

$$K = \frac{1}{4}\left(b\sigma + \frac{1}{b\sigma}\right)^2 . \tag{5.5}$$

Most experiments operate in the weak-focusing limit (large w_p, long Rayleigh range), where $b\sigma \ll 1$ and $K \approx 1/(2b\sigma)^2$. The generic beam waist of all Hermite-Gauss (HG) or Laguerre-Gauss (LG) modes, defined by the fundamental mode $u_{0,0}(\mathbf{q}) = \exp(-q^2/q_0^2)$, is $q_0 = \sqrt{\sigma/(2b)}$ being the geometric mean between the width of the pump profile and the Gaussian phase matching function.

The eigenmodes of the Gaussian Schmidt decomposition are not unique (because of the degeneracy of all the eigenmodes belonging to a fixed mode order) and, for example, can equally well be taken as the Hermite-Gaussian or Laguerre-Gaussian set of modes.

The numerical results presented in section 3 suggest that the Gaussian expansion can be improved somewhat by writing $\mathrm{sinc}\left(b^2|\mathbf{q}_1 - \mathbf{q}_2|^2\right) \approx \exp\left(-\alpha^2 b^2|\mathbf{q}_1 - \mathbf{q}_2|^2\right)$, where α is a (scaling) constant chosen such that both functions satisfy some common criterium. This modified scaling results in a replacement of $b \to \alpha b$ in all expressions that originate from the Gaussian approximation. We postpone the discussion on the rescaling factor α at the end of the chapter, as additional concepts have to be introduced first. In Figure 5.1 we plot the sinc phase matching (solid green line), comparing it to its Gaussian (red dashed) and supergaussian (or "double gaussian", blue dashed) approximations. Notice that the Gaussian approximation $\exp(-q^2)$ scales very rapidly, thus missing out high-q modes. The sinc function $\mathrm{sinc}(q^2)$, instead, has a $1/q^2$ overall scaling, letting high-q modes

have an effect on the Schmidt decomposition. The supergaussian $\exp(-bq^4)$ with $b = 0.193$ is a better approximation than the Gaussian, as it captures the central scaling of the sinc much better [67, 68]. This is due to the Taylor expansion of the supergaussian having the same powers as the sinc. However, there are two main problems associated with the supergaussian approximation. The first is that although it models the central scaling for a wider range of q values, it misses the $1/q^2$ overall scaling too, failing to carry high-q modes, and it does not display any ring structure. The second problem is that analytically it is not an advantage: for perfect phase mismatch it is possible to find angular Schmidt modes in terms of Bessel functions of two variables, which are defined in [69, 70], but as the literature on generalised Bessel functions is poorer than the literature on single variable Bessel functions, it is perhaps an unprofitable effort to try to find the full analytical Schmidt decomposition using the supergaussian approximation, and as we will show in section 3.3, a heuristic rescaling of the Gaussian approximation is sufficient to express the Schmidt number correctly for a perfect phase mismatch.
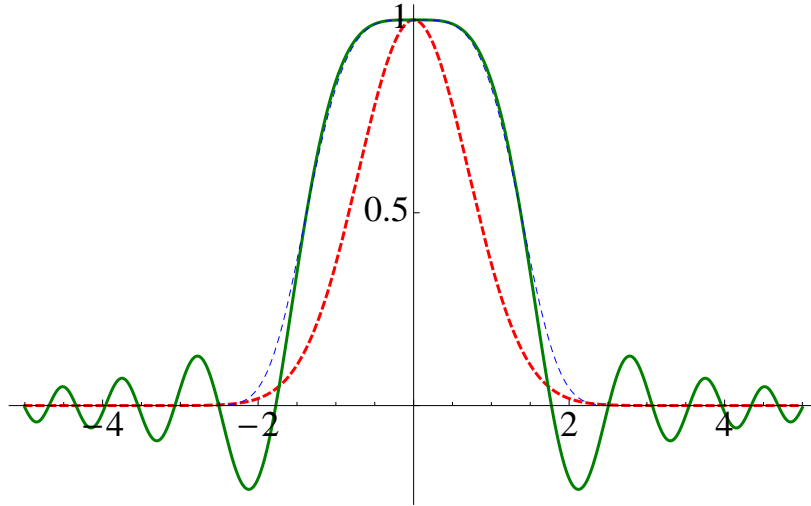

Figure 5.1: (colour online) Comparison between the three different phase matching functions. The solid line is the full sinc form $\text{sinc}(q^2)$, the dashed lines are the gaussian (red, narrower) $\exp(-q^2)$ and optimised supergaussian (blue) approximation $\exp(-0.193q^4)$. The similarity between the Gaussian and sinc functions is questionable, as it is the similarity between supergaussan and sinc for the overall scaling and for the oscillations at $q \gtrsim 2$.

5.2.2 Schmidt decomposition and Entanglement

For the arguments in this chapter it is sufficient to recall that the Schmidt decomposition of an entangled state is its representation in the joint basis that maximises the mutual information. Notice that in Eq. (5.1) the joint modes are the product of one mode per subsystem (hence each Schmidt mode is separable). The decomposition, in fact, also specifies the weights $\sqrt{\lambda_i}$ of single joint modes $u_i(\mathbf{q}_1)v_i(\mathbf{q}_2)$, the square of which is interpreted as the normalised probability (via $\sum_\ell \lambda_\ell = 1$) of detecting the i-th joint mode, were we to measure in the Schmidt basis. The more the distribution of the probabilities $\{\lambda_i\}$ is "spread out" the more entangled the state is. This leads to the natural choice of choosing entropic measures to quantify entanglement. In our case we choose the Schmidt number $K = 1/\sum_i \lambda_i^2$, where the probabilities are normalised: $\sum_i \lambda_i = 1$.

The Schmidt decomposition of the state (5.2) can be thought as made of two stages, because we are working with two pairs of coordinates: an angular coordinate per subsystem and a radial coordinate per subsystem. In the first stage we separate the angular variables, in the second stage we separate the radial variables. The decomposition will have therefore two sums, and the weights and joint modes will have two indices [32].

We start with the two-photon amplitude $A(\mathbf{q}_1,\mathbf{q}_2)$ and we write it in polar coordinates (which we call $q_{1,2}$ and $\theta_{1,2}$). The rotational symmetry of $A(\mathbf{q}_1,\mathbf{q}_2)$ limits the angular dependance to a function of $(\theta_1 - \theta_2)$ only. The orthogonality of the azimuthal functions $e^{i\ell(\theta_1-\theta_2)}$ allows for an easy extraction of the quantity

$$\sqrt{P_\ell} F_\ell(q_1,q_2) = \int_0^{2\pi} A(\mathbf{q}_1,\mathbf{q}_2) e^{-i\ell(\theta_1-\theta_2)} d(\theta_1 - \theta_2). \tag{5.6}$$

This means that with a Fourier transform we can write the two photon amplitude in a joint basis of the OAM:

$$A(\mathbf{q}_1,\mathbf{q}_2) = \frac{1}{2\pi} \sum_\ell \sqrt{P_\ell} F_\ell(q_1,q_2) e^{i\ell\theta_1} e^{-i\ell\theta_2}, \tag{5.7}$$

where P_ℓ is the probability of measuring a pair of photons with OAM ℓ and $-\ell$, were we to measure in the joint OAM basis, and the radial amplitude is normalised by imposing $\int |F_\ell(q_1,q_2)|^2 q_1 q_2 \, dq_1 dq_2 = 1$. In this way we have separated the two angular variables, and each joint OAM mode is labelled by an integer number ℓ. We can therefore proceed to the second stage and consider the functions $\sqrt{P_\ell} F_\ell(q_1,q_2)$ and separate the radial variables. In order to make this separation easy and truly one-dimensional, we follow the reasoning of Law and

Eberly [21] by introducing a prefactor $\sqrt{q_1 q_2}$, to account for the Jacobian of the transformation $\mathbf{q} \to (q, \theta)$, and separating the radial function as

$$\sqrt{q_1 q_2} \sqrt{P_\ell} F_\ell(q_1, q_2) = \sum_{p \geq 0} \sqrt{\lambda_{\ell,p}} \phi_{\ell,p}(q_1) \phi_{-\ell,p}(q_2). \tag{5.8}$$

Symmetry considerations also lead to the separated functions to be equal, i.e. we use $\phi(q_1)$ as well as $\phi(q_2)$. The prefactor $\sqrt{q_1 q_2}$ is needed to facilitate the simple normalisation of the eigenfunctions ($\int \phi_{\ell,p}(q) \phi_{l,p'}(q) dq = \delta_{p,p'}$) that is required for the numerical decomposition that we discuss later on. For convenience, we will also define $u_{\ell,p}(q) \equiv \phi_{\ell,p}(q)/\sqrt{q}$ such that $\int u_{\ell,p}(q) u_{l,p'}(q) q \, dq = \delta_{p,p'}$.

Note that we had to introduce the non negative, integer label p. Such quantum number labels radial eigenmodes. The numerical analysis will lead to some more insight on this quantum number, as we will show in section 5.3. After putting these two results together we obtain the full two-dimensional Schmidt form of the two-photon amplitude

$$A(\mathbf{q}_1, \mathbf{q}_2) = \sum_{\ell=-\infty}^{\infty} \sum_{p=0}^{\infty} \sqrt{\lambda_{\ell,p}} u_{\ell,p}(\mathbf{q}_1) u_{-\ell,p}(\mathbf{q}_2). \tag{5.9}$$

Note that this form is analogous to Eq. (5.1) with $u_i(\mathbf{q}) \equiv u_{\ell,p}(q) e^{i\ell\theta}$.

The radial shape $\phi_{\ell,p}(q)$ of these profiles depends strongly on the collinear phase mismatch Φ; it resembles the well-known Laguerre-Gaussian (ℓ, p)-profiles only for $\Phi \approx 0$, but deviates strongly from these profiles at $\Phi \neq 0$ (see below). As anticipated, the effective number of spatial modes (which can be thought of as the strength of the entanglement) is given by the Schmidt number

$$K = \left(\sum_{\ell,p} \lambda_{\ell,p}^2 \right)^{-1} \tag{5.10}$$

where the probabilities are normalised: $\sum_{\ell,p} \lambda_{\ell,p} = 1$.

A simple scaling argument shows that the outcome of the decomposition depends only on two dimensionless parameters. The first parameter is the product $b\sigma = \sqrt{\frac{L}{w_p^2 k_p}} = \sqrt{L_R/2}$ where we define L_R to be the crystal thickness normalised to the Rayleigh range of the pump beam, $L_R \equiv L/z_R$. The second parameter is the collinear phase mismatch Φ.

Entanglement in orbital angular momentum

The rotational symmetry of the considered geometry and the associated conservation of OAM makes the azimuthal component of the entanglement easier to address and more fundamental than its radial counterpart. Hence, most experiments have concentrated on the OAM part of the spatial entanglement. In this context we are interested in the probability P_ℓ (which appears in Eq. (5.7)) of detecting a pair of photons with OAM, respectively, ℓ and $-\ell$. Such probability is related to the weights $\lambda_{\ell,p}$ via $P_\ell = \sum_p \lambda_{\ell,p}$. P_ℓ is called the *spiral weight*, its full distribution $\{P_\ell\}$ is called the *spiral spectrum*, and the width of this distribution is called the *spiral bandwidth* [20]. The dimensionality of the generated OAM entanglement is defined by the azimuthal Schmidt number

$$K_{\mathrm{az}} = \left(\sum_\ell P_\ell^2 \right)^{-1}. \tag{5.11}$$

In case of high-dimensional spatial entanglement ($K \gg 1$) and $\Phi \approx 0$, the relation between both forms of entanglement is $K_{\mathrm{az}} \approx 2\sqrt{K}$, as the effective range of the ℓ labels is typically $\approx 4\times$ as large as that of the p labels. This statement is exact for the Gaussian approximation of the two-photon field, where [66]

$$\lambda_{\ell,p} = C_1 \exp(-(2p + |\ell|)/C_2), \tag{5.12}$$

where C_1 and C_2 are two constants. However, it does not apply to the full expression for $\Phi \neq 0$. Our numerical approach allows us to analyse what happens also in such regime.

5.2.3   Optical etendue

The Schmidt decomposition of the two-photon field is equivalent to a coherent mode decomposition of the coherence function of the reduced one-photon state. The reduced one-photon operator $\hat{\rho}^{(1)}$, obtained by tracing $|\Psi\rangle\langle\Psi|$ over all possible states of the second photon, is an incoherent (and weighted) mixture of all of all Schmidt states. The extent to which the reduced density operator is mixed reflects the degree of entanglement for the pure two-photon state $|\psi\rangle$. The Schmidt number K can thus be estimated by comparing the one-photon coherence of the SPDC emission with that of more standard optical sources.

At sufficiently large $K \gg 1$, the reduced one-photon field is quasi-homogeneous, i.e., the reduced one-photon density matrix $\rho^{(1)}(\mathbf{q}_1, \mathbf{q}_1')$ factorizes in a product of

an intensity function of sum coordinate $(\mathbf{q}_1 + \mathbf{q}_1')$ and a coherence function of difference coordinate $(\mathbf{q}_1 - \mathbf{q}_1')$. The Schmidt number K of the quantum state then reduces to the normalised optical etendue $N \equiv A\Omega/\lambda^2$ of the one-photon field, where A is the effective area of the source and Ω is its effective (solid) opening angle [64, 66]. The optical etendue quantifies the effective number of transverse modes contained in a partially-coherent beam. The relation $K \approx N$ becomes exact if the source is quasi-homogeneous and if the effective area A and opening angle Ω are defined in a convenient way [64].

The relation between K and N yields an easy and intuitive interpretation of the Schmidt dimension. For this, we convert the opening angle of the source Ω into a transverse coherence width w_{coh} in the source plane, which we conveniently define as

$$w_{\mathrm{coh}} \equiv 4b = \sqrt{\frac{L\lambda_0}{\pi n}}, \tag{5.13}$$

where L is the crystal length, λ_0 is the emission wavelength in vacuum and n is the refractive index. The transverse coherence width w_{coh} of the one-photon field is similar to the width of the two-photon coherence function $V(\vec{x}_1 - \vec{x}_2)$ defined in refs. [64, 71] as the Fourier-transform of the sinc-type phase matching function. With the above definition of w_{coh}, the Law-and-Eberly expression for the Schmidt number at $b\sigma \ll 1$ reduces to the logical form

$$K \approx \left(\frac{1}{2b\sigma}\right)^2 = \left(\frac{w_p}{w_{\mathrm{coh}}}\right)^2, \tag{5.14}$$

which shows that the Schmidt number K simply counts the number of "independent" coherent regions in the source. The above definition of w_{coh} also provides for an easy rewrite of the results of the Gaussian expansion discussed in subsection (2.1.2). The width w_0 of the fundamental Gaussian Schmidt mode $u_{0,0}(\vec{q}) \propto \exp{-|\vec{q}|^2/q_0^2}$ in real space is $w_0 = 2/q_0 = \sqrt{w_p w_{\mathrm{coh}}}$, being the geometric mean between the pump waist and the transverse coherence length of the generated field. The modal amplitude of the Gaussian $u_{\ell,p}$ mode is

$$\sqrt{\lambda_{\ell,p}} \approx \exp{[-(2p + |\ell|)/K_{1D}]}, \tag{5.15}$$

where $K_{1D} = w_p/w_{\mathrm{coh}}$ is the one-dimensional equivalent of the Schmidt number.

5.2.4 Projective measurement of OAM entanglement

In this subsection, we will briefly compare two experimental techniques that have been developed to characterise in particular the OAM contents of two-photon sources. We distinguish between measurements with bucket detectors, which record the complete field, and measurements with single-mode detectors, which record projected components of the field.

A Schmidt analysis of the two-photon field with bucket detectors, which by definition have no spatial sensitivity, requires two-photon interference before measurement. More specifically, one should measure the visibility of the famous two-photon (Hong-Ou-Mandel) dip [72] as a function of the relative alignment of the interfering beams. Measurements with one beam rotated with respect to the other provide complete and unbiased information on the azimuthal part of the entanglement [73, 74], as the visibility of the observed two-photon interference is

$$V(\Delta\tilde{\theta}) \propto \int\int A^*(q_1,q_2;\Delta\theta)A(q_1,q_2;\Delta\theta+\Delta\tilde{\theta})q_1q_2dq_1dq_2$$

$$= \sum_{\ell=-\infty}^{\infty} P_\ell \exp(i\ell\Delta\tilde{\theta}), \tag{5.16}$$

where $\Delta\tilde{\theta}$ is the rotation angle between the interfering two-photon fields, which is twice the rotation angle applied in a single arm of the interferometer. This angular dependence originates from the interference between two-photon fields with either the ℓ or the $-\ell$ photon in arm 1 and the other photon in arm 2, i.e. the azimuthal terms both contribute of the same amount. Equation (5.16) is based on the entanglement between these contributions and the natural symmetry $P_{-\ell} = P_\ell$.

A Schmidt analysis of the two-photon field with single-mode detectors, each comprising a single-mode fiber positioned in front of a detector, requires some form of mode transformation before projection. This mode transformation is typically performed with (a set of) fixed phase plates or an (adjustable) spatial light modulator (SLM). The combined transformation-projection technique has been used successfully for the analysis of the OAM contents of the two-photon field [64]. The generated two-photon field is then typically projected onto two modes of the form $\phi_d(q)\exp(if(\theta))$, where $\phi_d(q)$ is the (image of the) mode profile of the fiber and $\exp(if(\theta))$ is the phase profile imposed by the phase-transforming element, which typically has the standard form $\exp(i\ell\theta)$. The rotational symmetry of the generated field $A(\mathbf{q}_1,\mathbf{q}_2)$, which manifests itself in the generation of photon pairs with opposite OAM only, allows one to derive simple expressions for

the projected two-photon field. It is easy to show that the maximum information contained in these projections, for any combination of azimuthal phase profiles $f(\theta)$, is contained in the function

$$
\begin{aligned}
g_{\text{proj}}(\Delta\theta) &= \int\int A(\mathbf{q}_1,\mathbf{q}_2) u_d^*(q_1) u_d^*(q_2) q_1 q_2 \, dq_1 \, dq_2 \\
&\propto \sum_{\ell=-\infty}^{\infty} \sum_{p=0}^{\infty} C_{\ell,p} \sqrt{\lambda_{\ell,p}} \exp(i\ell\Delta\theta),
\end{aligned}
\tag{5.17}
$$

where $\Delta\theta \equiv \theta_1 - \theta_2$. The projection coefficients

$$
C_{\ell,p} = \left| \int u_{\ell,p}(q) u_d(q) q \, dq \right|^2 \le 1,
\tag{5.18}
$$

quantify the spatial overlap between each generated Schmidt mode and detector mode. They also quantify the bias imposed by the mode projections.

As an example, let us consider the projection on a detector mode $u_d(q)$ that is much wider than a range of lower-order $u_{\ell,p}(q)$ modes. Under these conditions, the projection coefficients will have a strong bias towards the $p = 0$ modes, as all $u_{\ell,p}(q)$ with $p \ge 1$ modes exhibit at least one oscillation (in the radial direction) making $C_{\ell,p\neq0} \ll C_{\ell,p=0}$. Hence, projection experiments are mainly sensitive to the $p = 0$ components of the two-photon field. A calculation of the projection-imposed bias on the ℓ modes is more complicated. A rigorous analysis, in terms of Lerch transcendent functions, is presented in [75], which is discussed in chapter 7.

5.3 Numerical results

In this section we will present the numerical results for the general Schmidt decomposition of Eq. (5.4) into Eq. (5.9). The topics that it allows us to address are the distribution of the modal weights $\lambda_{\ell,p}$ and their separation over the azimuthal and radial degrees of freedom, the shape of the Schmidt modes $u_{\ell,p}(q)$, and the influence of the collinear phase mismatch Φ.

The numerical analysis is relatively straightforward. We start from the generated two-photon field $A(\mathbf{q}_1,\mathbf{q}_2)$ defined by Eq. (5.4) and perform an OAM decomposition into radial-only function $\sqrt{P_\ell}F_\ell(q_1,q_2)$ using Eq. (5.6). The analytical integration is performed using an approximation that works well for $b\sigma \lesssim 0.2$, which is the most relevant experimental situation as it corresponds to a weakly focused pump. Employing the analytical integration of the angular variable yields analogous final results in much less time, as the numerical

methods consist only in a diagonalisation of matrices. The prescribed radial integration is performed on a discrete equidistant grid of q-values. After multiplication by a factor $\sqrt{q_1 q_2}$, to account for the Jacobian of the transformation, we obtain the scaled $N \times N$ matrix F with a dimension set by the required accuracy (N is typically in the order of the hundreds). The Schmidt decomposition in Eq. (5.8) of $\sqrt{P_\ell} F_\ell(q_1, q_2)$ is equivalent to the diagonalisation of the matrix F. To avoid potential complications associated with the two-particle nature of the $\phi(q_1)\phi(q_2)$ phase space, we first multiply F by its transpose $F^\dagger$, in order to obtain the equivalent of the reduced one-photon density matrix, and diagonalize the matrix $F F^\dagger = F^2$. The diagonalisation of the matrix F^2 yields three matrices: $F^2 = MEM^T$, two of which represent the modes (M) and one of which is diagonal (E) and represents the eigenvalues. The resulting eigenvalues are $\lambda_{\ell,p}$ (at fixed ℓ); the resulting eigenmodes are $\phi_{\ell,p} = \sqrt{q} u_{\ell,p}(q)$.

In the following subsection we present the weights in the matrix E and we show what are the consequences of having a nonzero phase mismatch. Similarly, we also present the modes in the matrix M, and we show the consequences of having a nonzero phase mismatch.

5.3.1 Schmidt weights

As a typical example we will concentrate on the case $b\sigma = 0.05$, which for instance corresponds to a 2 mm thick PPKTP pumped by a $w_p = 230$ μm diameter pump laser at a pump wavelength of 413 nm.

Phase mismatch $\Phi = 0$

Figure 5.2 is a false-colour plot of the probabilities $\lambda_{\ell,p}$ of the modal decomposition for $b\sigma = 0.05$ and $\Phi = 0$. The Schmidt numbers deduced from this calculation are $K \approx 231$ and $K_{\mathrm{az}} \approx 32$, respectively. These numbers satisfy the approximate relation $K_{\mathrm{az}} \approx 2\sqrt{K}$ associated with the triangular mode spectrum in figure 5.2. They are, however, considerably larger than the values of $K \approx 100$ and $K_{\mathrm{az}} \approx 20$ expected from a simple Gaussian expansion based on the approximation $\mathrm{sinc}(q^2) \approx \exp(-q^2)$. The reason for this difference, as anticipated, is the fact that the Gaussian approximation is missing high-q transverse modes.

Phase mismatch $\Phi < 0$

We consider now geometries with non-perfect phase matching and in particular the case $\Phi < 0$ for which the SPDC rings have opened up. The choice of $\Phi = -4$
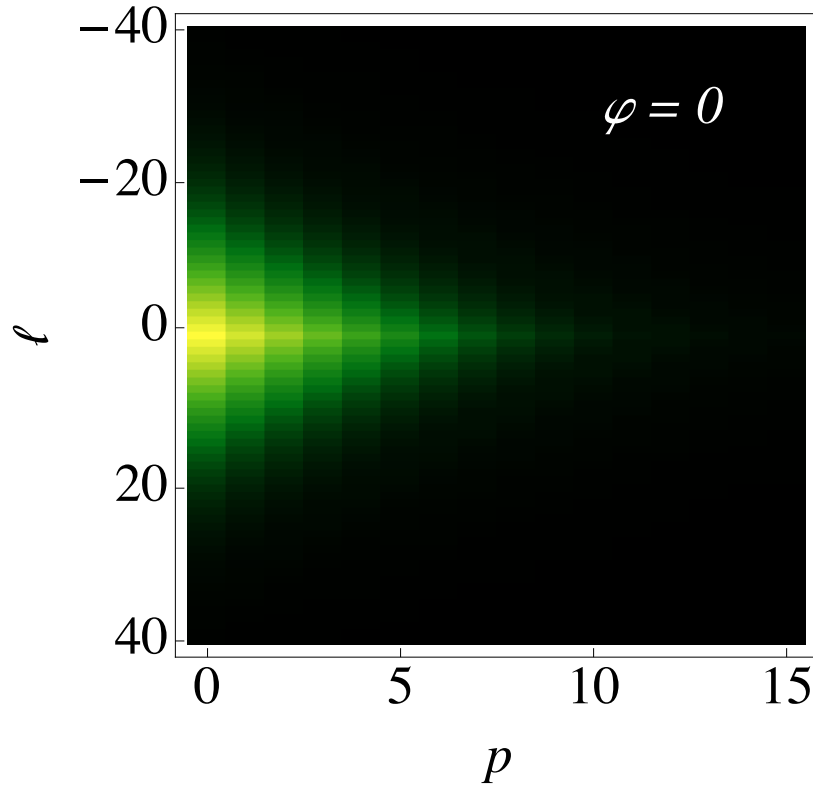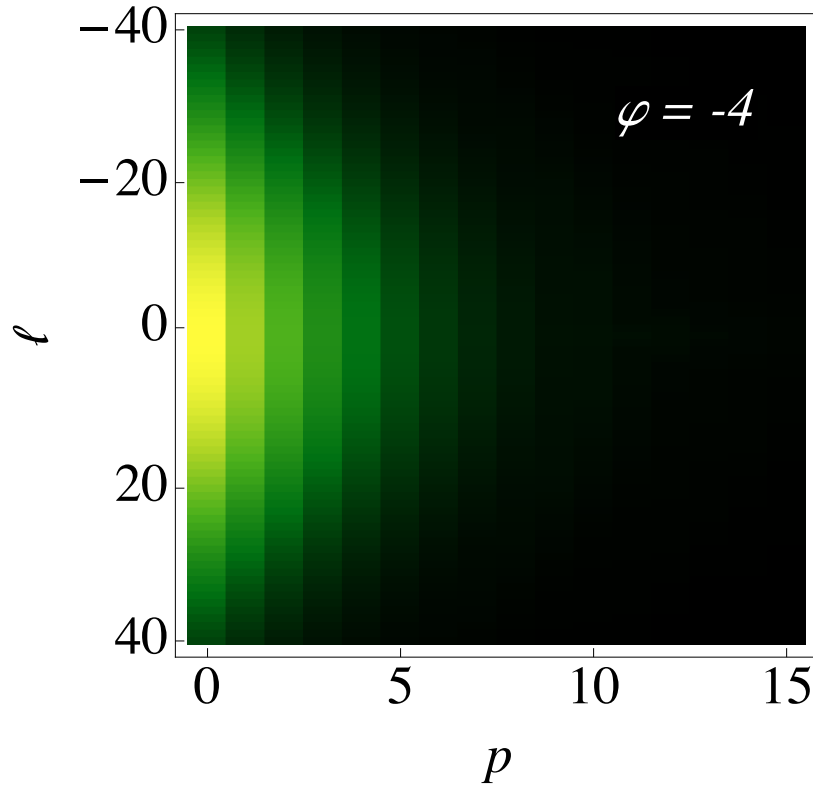
Figure 5.2: False colour plot of the modal spectrum $\lambda_{\ell,p}$ (with p horizontal and ℓ vertical) for $b\sigma = 0.05$ and $\Phi = 0$. Note the typical triangular form and the $\approx 4\times$ larger range of relevant ℓ values as compared to p-values. The intensity scale follows this colour bar: min ▮▮▮ max.

is a compromise between having a low enough value to show the different distribution of Schmidt weights and avoiding an excessive computational complexity.

Figure 5.3 is a false colour plot of the probability $\lambda_{\ell,p}$ of the modal decomposition calculated for $b\sigma = 0.05$ and $\Phi = -4$. On account of phase mismatch, the full 2-dimensional Schmidt number has increased by a factor $\approx 2\times$ to $K \approx 425$. This increase coincides with a similar increase in the space angle and the angular integrated output of the SPDC source. Most noticeably, the spread in OAM values is now much larger then in the $\Phi = 0$ case, at the expense of the spread in p values. The azimuthal Schmidt number has increased from its prior value of 32 at $\Phi = 0$ to $K_{\text{az}} \approx 72$.

Phase mismatch $\Phi > 0$

For the sake of completeness we consider now the case of positive phase mismatch, $\Phi > 0$. Similarly to the case of negative phase mismatch, the choice of

Figure 5.3: False colour plot of modal spectrum $\lambda_{\ell,p}$ (with p horizontal and ℓ vertical) for $b\sigma = 0.05$ and $\Phi = -4$. Note the elongated form in the ℓ direction and the reduced width in p. The intensity scale follows this colour bar: min ▬▬▬▬ max.

$\Phi = 4$ is a compromise between having a high enough value to show the different distribution of Schmidt weights and avoiding an excessive computational complexity. In this case the SPDC rings have closed, and the efficiency of the down-conversion process is considerably lower, normally by two orders of magnitude, which degrades the S/N ratio by a large extent.

Figure 5.4 is a false colour plot of the eigenvalue $\lambda_{\ell,p}$ of the modal decomposition calculated for $b\sigma = 0.05$ and $\Phi = 4$. As compared to the case of negative detuning, we find that the width in the p direction has increased at the expense of the width in the ℓ-direction. In this example we calculated values of $K \approx 625$ and $K_{az} \approx 57$.

The optical etendue argument presented in Sec. (5.2.3) allows one to explain why the spatial entanglement gets concentrated in the azimuthal degree of free-dom, instead of the radial one, for $\Phi < 0$ and why this concentration hardly occurs for $\Phi > 0$. Our geometric argument is based on the notion that the transverse coherence width in the azimuthal and radial directions are comparable, being
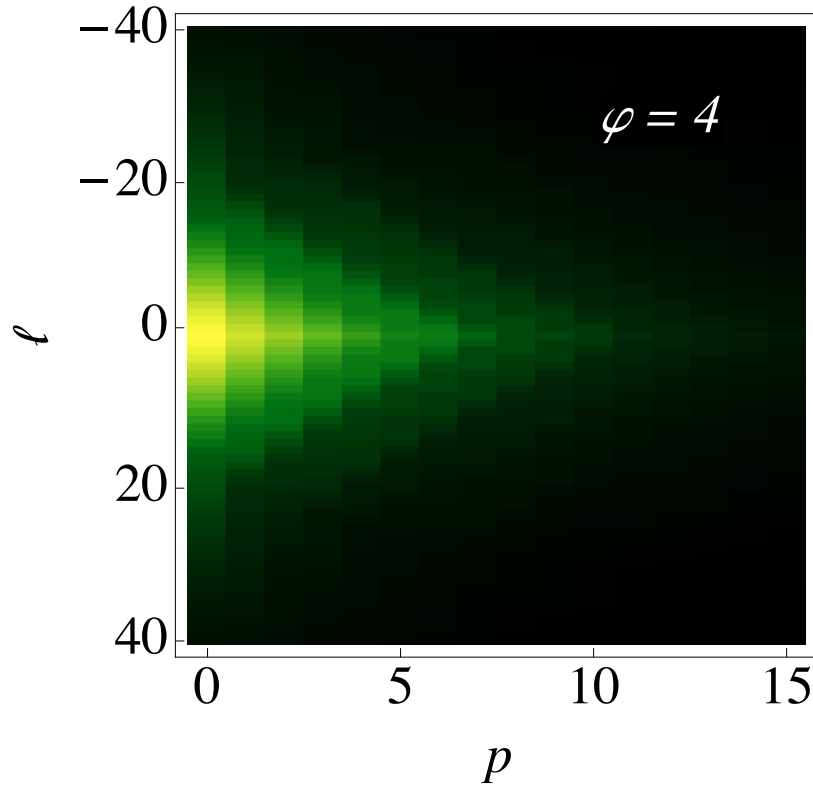
Figure 5.4: False colour plot of modal spectrum $\lambda_{\ell,p}$ (with p horizontal and ℓ vertical) for $b\sigma = 0.05$ and $\Phi = 4$. Note the strongly elongated form in the p direction and the reduced width in ℓ as compared to the case $\Phi = -4$. The intensity scale follows this colour bar: min ▮▮▮ max.

Fourier related to the size of the pump beam. At negative detuning $\Phi < 0$, where the fundamental SPDC ring opens up, the circumference of this SPDC increases with phase mismatch while the radial thickness of the ring decrease accordingly. As a result, the number of spatial modes that fit in the azimuthal (ℓ) direction, increase while the number of modes in the radial (p) direction decrease. The total area of the open SPDC ring at $\Phi \ll -1$ is approximately twice as large as the area of central emission at $\Phi = 0$, making the 2-dimensional K also about two times larger. At positive detuning $\Phi > 0$ the fundamental SPDC ring disappears and only weak secondary rings remain. As these rings are numerous and have similar intensities, the number of spatial modes in the Schmidt decomposition can be quite large and is more evenly distributed over ℓ and p mode numbers. However, we stress once again that if $\Phi > 0$ the down-conversion process is inhibited, as the phase matching function is a measure of efficiency, so even though the single pairs are more entangled, their number can decrease considerably.

Next we single out the OAM part of the entanglement for the considered

cases $\Phi = 0, -4$ and 4. We do so by summing the modal weights $\lambda_{\ell,p}$, depicted in figures 5.2, 5.3 and 5.4, over its radial quantum number p. What one then obtains are three different probability distributions that show different behaviour over the same range of ℓ values. Figure 5.5 shows the modal weights $(\sum_p \lambda_{\ell,p})$ of the OAM modes. Note that the behaviour changes from Lorentzian-like to Gaussian-like as the phase mismatch goes from positive to negative values [74]. We stress that the case of positive detuning is experimentally unfavorable on account of the limited brightness of a source where the fundamental SPDC ring has closed/disappeared.
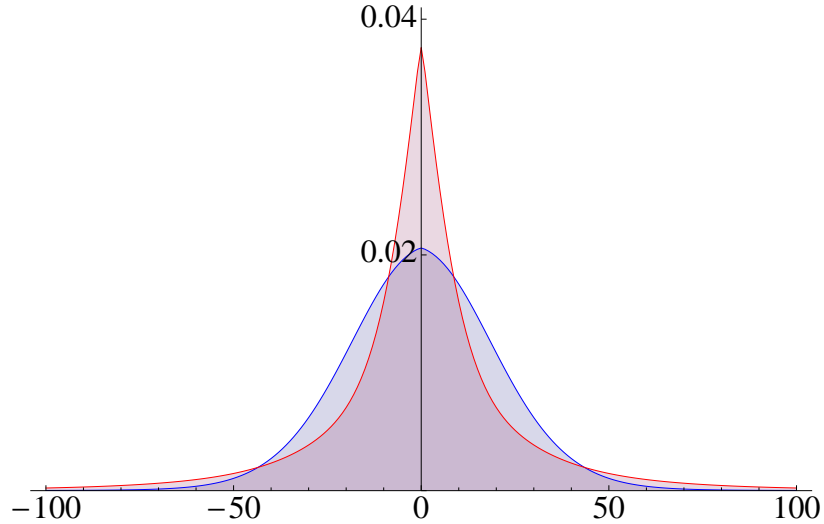
Figure 5.5: Probability distributions of OAM modes (summing over p). The Lorentzian-like distribution (red) is for positive mismatch ($\Phi = 4$) and the Gaussian-like distribution (blue) is for negative phase mismatch ($\Phi = -4$).

5.3.2 Schmidt modes

In this subsection we will present a series of numerical results that show the radial profile of the Schmidt modes in momentum space. The profile of the modes is influenced by the presence of a phase mismatch, and as anticipated the modes resemble the LG modes only at $\Phi = 0$.

Phase mismatch $\Phi = 0$

Figure 5.6 shows the result of the singular value decomposition for $b\sigma = 0.05$. What is represented is a portion of the matrix M corresponding to a value of the OAM of 10. This choice gives a good example of the radial displacement in k-space between Schmidt modes that we observe. Such behaviour will become clear

below, where we compare modes from a perfect phase matched case to modes from a negatively detuned phase mismatch. Each horizontal line represents the intensity profile $|\phi_{\ell,p}(q)|^2$ of a radial mode $\phi_{l,p}(q)$ with fixed ℓ and a different radial quantum number p. The first three modes in figure 5.6 are plotted in figure 5.7.

We would like to give an interpretation of the behaviour of the Schmidt modes in figure 5.6. Notice that each eigenmode resembles a common eigenmode from the family of Hermite-Gauss modes, only centred at specific distances from the origin. There is a one to one correspondence between a family of modes and a region between zeros in the function $\sqrt{P_\ell}F_\ell(q_1, q_2)$. Such function represents the joint amplitude density of the photons having radial momentum q_1 and q_2, respectively (examples are given in figures 5.12 and 5.13). Therefore it's also closely related to the radial sinc profile of the down-converted beam. Such correspondence might be explained by considering the size of the coherence length compared to the size of the nonzero regions in $\sqrt{P_\ell}F_\ell(q_1, q_2)$: if the coherence length is smaller than the thickness of a ring, there can be no strong coherence between photons in different rings, and any wave function in a single ring can be written in terms of a set of fundamental modes of fixed OAM and variable p.

We find the organisation of the eigenmodes in groups quite remarkable. Each of the groups looks like a local complete family of HG-like modes, giving the impression of the existence of an additional symmetry and therefore of a quantum number that could be specified to identify each group. This can be explained if the matrix F^2 is taken into account and in particular the shape of the lobes that form along the diagonal, as in figure 5.12 and 5.13. They can be thought of as potential wells, which have to be filled by eigenmodes. For each of them the eigenmodes build in local families because from the Taylor expansion of the neighbourhood of the bottom of any potential well we can infer that the HG modes form a natural set of local modes for each well. The potential walls between neighbouring wells could be tunneled by the modes, if their size were large enough to reach over to the next well. In practice, the size of each mode is small enough and the families are well confined within individual lobes of the sinc-fucntion if the pump beam is not too focussed. It would be interesting to look at modes with high values of p, in which case they could be large enough for a tunneling to take place. I plan to do so in a future work.
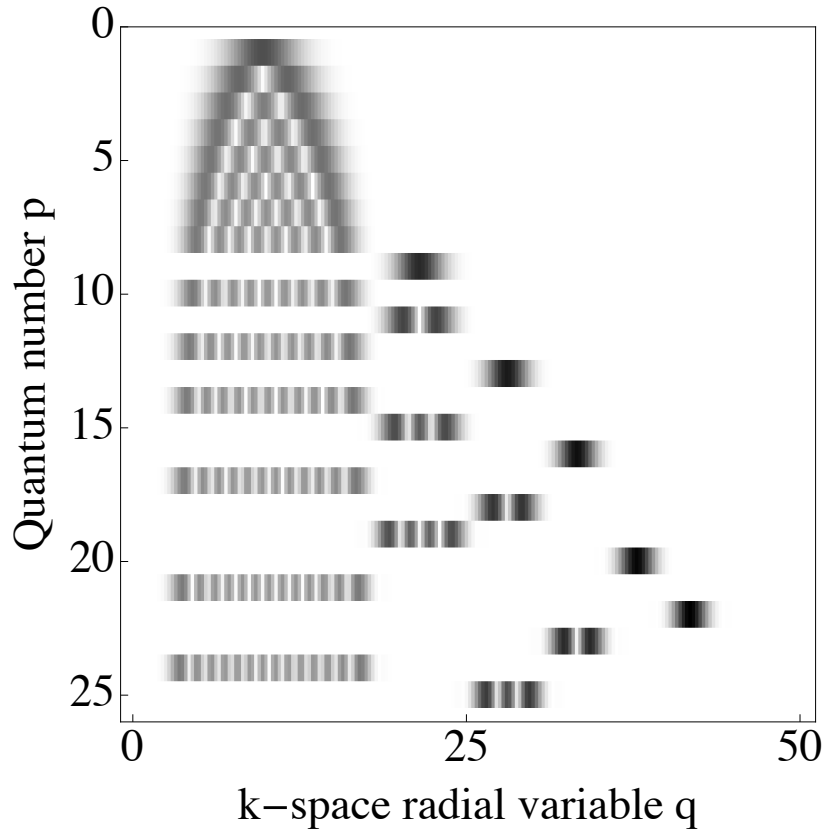
Figure 5.6: A part of the matrix M for $\ell = 10$ is plotted in a black and white array plot. The intensity profile $|\phi_{l,p}(q)|^2$ of a single radial eigenmode is represented in each horizontal series. The radial quantum number p labels the different series. They are organised from top to bottom in order of contribution to the total wave function. In this example the radial field that was decomposed was the one in figure 5.12. The intensity scale follows this colour bar: min [colour bar] max.

Phase mismatch $\Phi \neq 0$

Figure 5.8 shows the calculated radial profile $|\phi_{\ell,p}(q)|^2$ of the eigenmodes at $\ell = 10$ for the detuned case $\Phi = -4$. At $\Phi \ll 0$, where the SPDC ring has opened, even the fundamental $\ell = p = 0$ Schmidt mode is ring-shaped and has a wide region of zero intensity around the central axis. As with the zero mismatch example, we also plot the first three modes of figure 5.8 in figure 5.9. Note that although the nonzero phase mismatch, the shape of the first three eigenmodes is remarkably similar to the one in figure 5.7, the difference being just an offset with respect to the origin.

This happens because in the two cases the radial functions that are being decomposed are remarkably similar. The modes firstly fill the largest ring, until the difference between their combined intensity and the intensity of the main
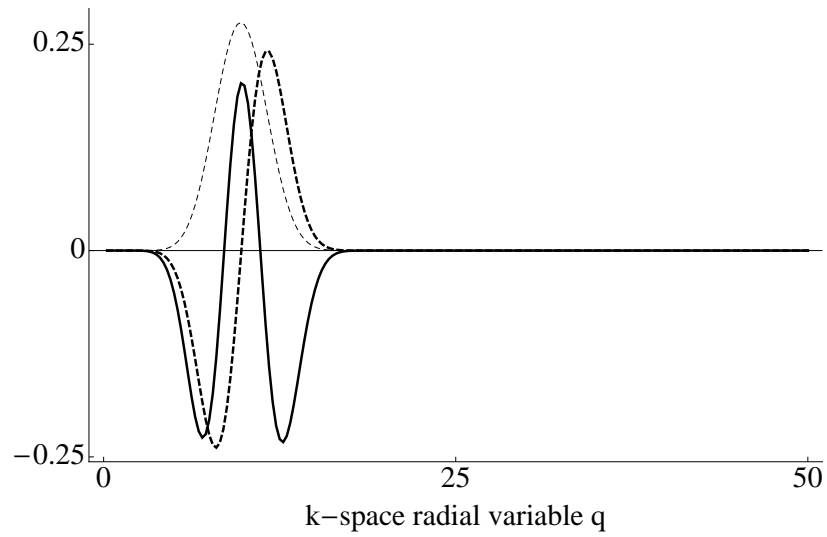
Figure 5.7: Amplitudes in the form $\phi_{\ell,p}(q)$ of the first three p-modes in figure 5.6
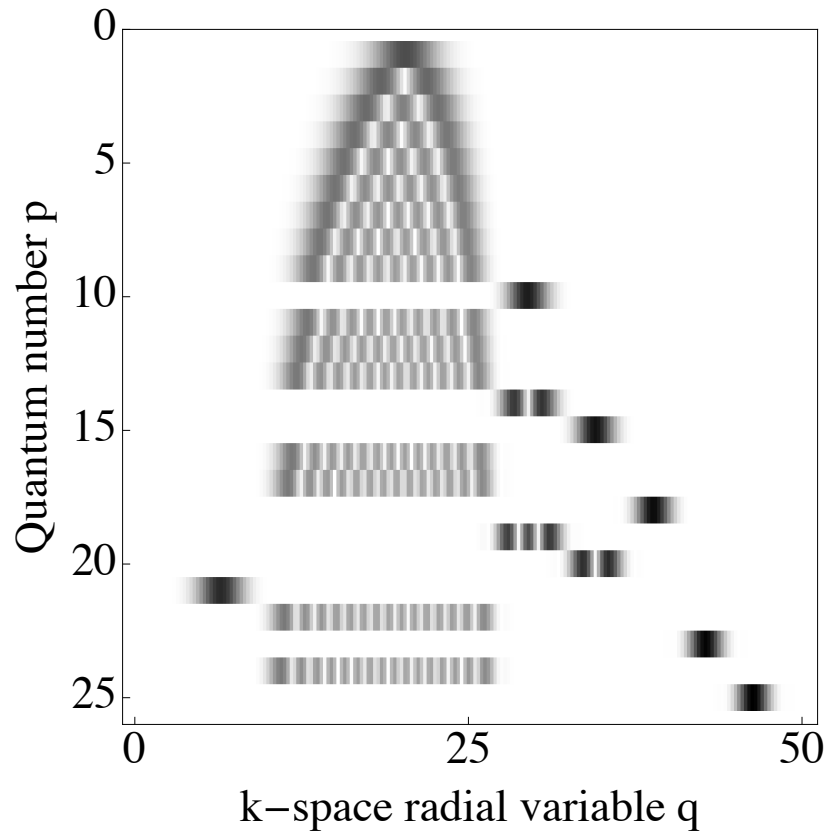

Figure 5.8: A part of the matrix M for $\ell = 0$ in case of phase mismatch $\Phi = -4$. In this example the radial field that was decomposed was the one in figure 5.13. The intensity scale follows this colour bar: min ▭ max.

ring is small enough to be less than the intensity of secondary rings. When this happens we see that the subsequent modes fill different rings, and jump discretely between them. This is in accordance with the explanation that we gave at zero phase mismatch: the modes are compact enough to be well contained within individual rings.
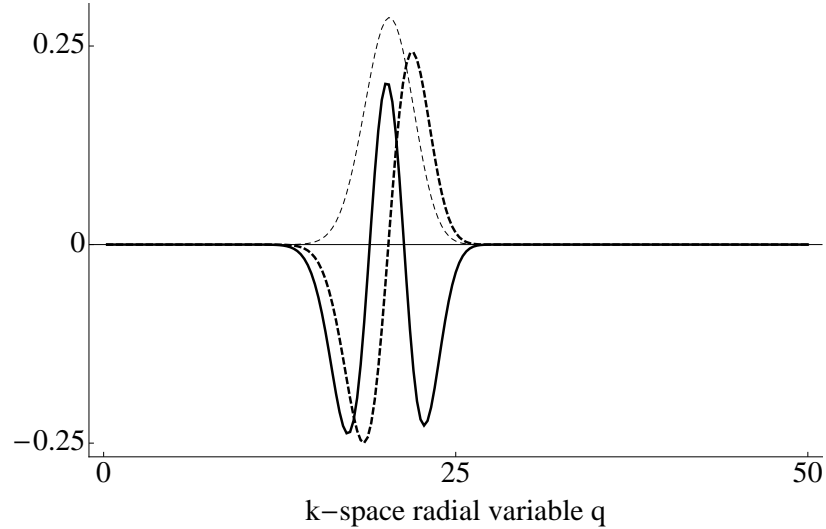

Figure 5.9: Amplitudes in the form $\phi_{\ell,p}(q)$ of the first three p-modes in figure 5.8

5.3.3 Heuristic rescaling of Gaussian approximation

In this section we conclude the argument, anticipated in the previous sections, on the scaling factor α. Such scaling factor would allow us to describe more accurately, for $\Phi = 0$, the Schmidt number when approximating the sinc phase matching with a Gaussian phase matching. Below, we present two arguments for why we think it is better to replace the function $\text{sinc}(b^2 q^2)$ by $\exp(-\alpha^2 b^2 q^2)$ rather than by $\exp(-b^2 q^2)$. There are experimental results that support our claims, for instance see [76].

Figure 5.10 depicts the numerically-calculated Schmidt number K and its azimuthal counterpart K_{az} as functions of $b\sigma$ for the exact sinc phase matching. In order to numerically calculate these curves we had to abandon momentarily the faster approach used to find the Schmidt modes, as $b\sigma$ in figure 5.10 is not restricted to small values. The Gaussian approximation of the phase matching function results in expressions that are symmetric under the exchange $b\sigma \leftrightarrow \frac{1}{b\sigma}$. The same argument cannot be applied to the sinc phase matching, as it's possible to see from figure 5.10: the graph is symmetric with respect to a point between

74

1 and 2, meaning that there could be an extra factor α such that the symmetry now reads $b\sigma\alpha \leftrightarrow \frac{1}{b\sigma\alpha}$, representing a reflection around the point $1/\alpha$. If it were so, the factor α should have a value around $\alpha \simeq 0.85$. However, the introduction of α alone is not enough to obtain the correct K values. There are two ways to proceed, either to choose a simple modification, like $b\sigma \rightarrow b\sigma\alpha$, or to add also a second multiplicative factor β. We find that the relation

$$K = \frac{\beta}{4}\left(\frac{1}{b\sigma\alpha} + b\sigma\alpha\right)^2 \qquad (5.19)$$

fits extremely well the values in figure 5.10 for $\alpha = 0.85$ and $\beta = 1.65$. However, if one is only interested in the regime $b\sigma \lesssim 1$, which experimentally is the most relevant, there is no need to introduce β and the relation $K \approx 1/(2b\sigma\alpha')^2$ works well for $\alpha' \equiv \alpha/\sqrt{\beta} \approx 0.65$, as it is possible to see in figure 5.11. The α-modification is a consistent choice for replacing the sinc phase matching with a Gaussian phase matching, as it is implemented by modifying the width of the Gaussian and not by modifying the relation between K and the product $b\sigma$ ad hoc, as in (5.19).

The second argument relays on a common criterion to choose a sensible value for α without recurring to ad hoc arguments. The common criterion is equal width at $1/e$ from the peak intensity [17]. This criterion yields again a value of $\alpha \approx 0.65$.
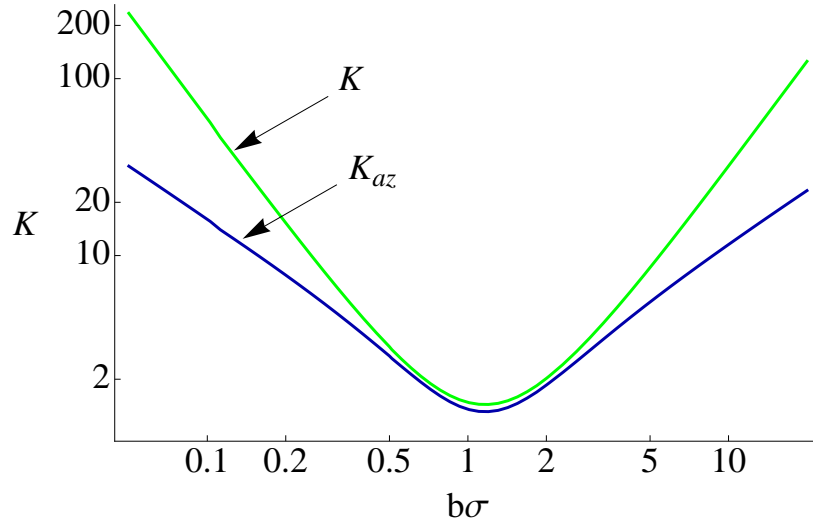

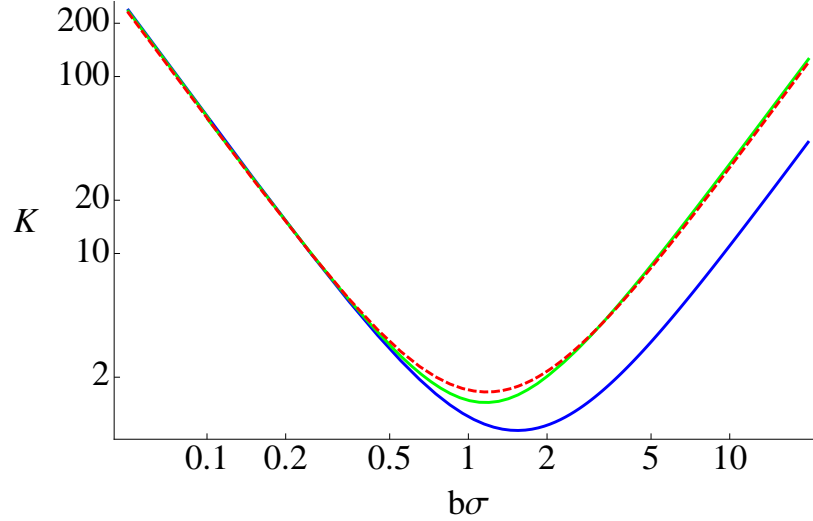Figure 5.10: K and K_{az} as functions of $b\sigma$ for $\Phi = 0$.

Figure 5.11: K in green and the two different modifications of the law (5.5). The dashed red line is the (α, β)-modification, the solid blue line is the α-only modification. The fits obviously match for $b\sigma \lesssim 1$.

5.4 Conclusions

We saw how a fast numerical approach to perform a numerical Schmidt decomposition without approximating the phase matching function, and for any phase mismatch, gave us insight on the entanglement between SPDC pairs. Thanks to the numerical approach the limitations of the Gaussian approximation were investigated and a scaling $b \to b\alpha$, for $\alpha \simeq 0.65$ was proposed. Such scaling allows the Gaussian approximation with no phase mismatch to reproduce the results of the more complete sinc phase matching. The etendue argument was successfully verified by calculating K and K_{az} for different values of the phase mismatch.

The Schmidt modes that were found exhibit numerous features, the most remarkable being the grouping in local families of HG-like modes and the fact that they don't span more than one lobe on the diagonal of the matrices of joint radial probability density. We gave an interpretation of these features by using the meaning of coherence length and the way of naturally assigning eigenmodes to potential wells. In fact, the other sets of HG-like radial modes are simply shifted by different amounts into other rings of the sinc-type SPDC emission. These other sets have similar HG-like radial structures but lower modal weights. The rather strict separation in discrete sets suggests that we might want to use the set number as third quantum number, i.e. to split the radial quantum number p in two quantum numbers instead.
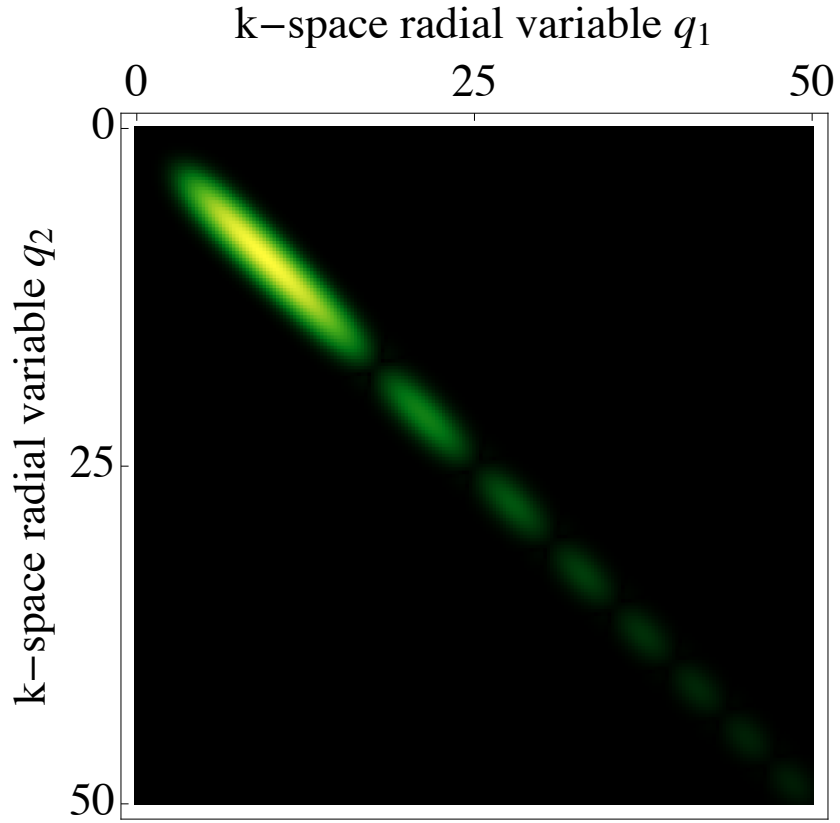
Figure 5.12: Joint radial probability density of generating a pair of photons with OAM $\ell = 10$ and radial momentum q_1 and q_2. This is the matrix F for $\Phi = 0$ used in the examples. The intensity scale follows this colour bar: min ▮▮▮ max.

5.5 Appendix

Here we report more in detail the function that was used to fill the matrix F. To avoid confusion between the azimuthal coordinate φ and the phase mismatch, here we indicate the latter with Φ. We start with the following two-photon wave function

$$\Psi(\mathbf{q}_1, \mathbf{q}_2) = \mathcal{N} e^{-\frac{w_0^2}{4}|\mathbf{q}_1 + \mathbf{q}_2|^2} \mathrm{sinc}\left(\Phi + \frac{L}{4k_p}|\mathbf{q}_1 - \mathbf{q}_2|^2\right) \tag{5.20}$$

k−space radial variable q_1
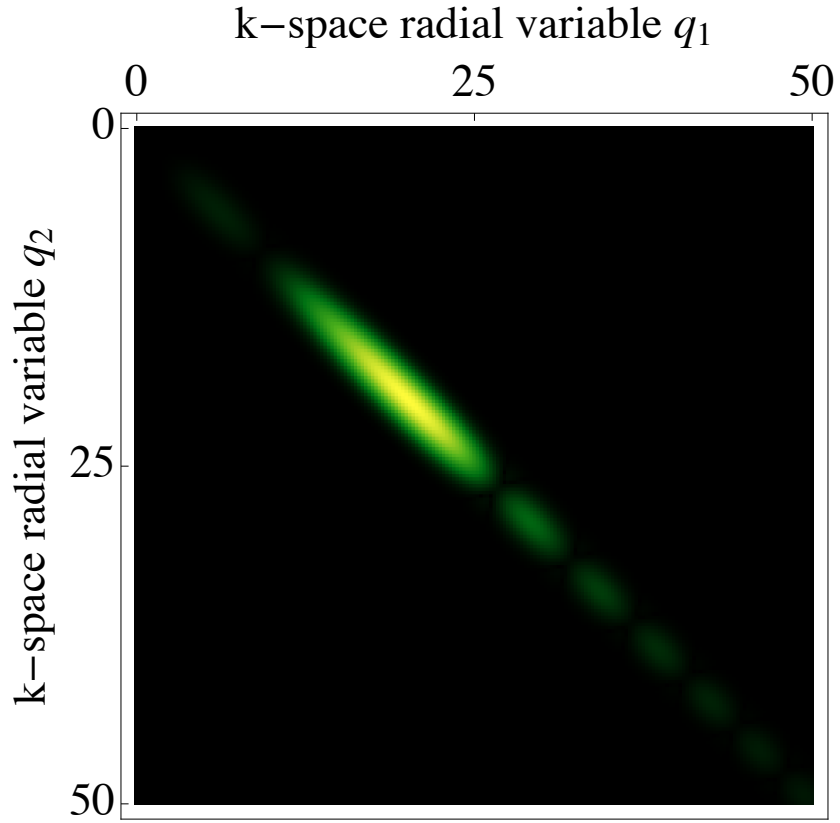
k−space radial variable q_2

Figure 5.13: Joint radial probability density of generating a pair of photons with OAM $\ell = 10$ and radial momentum q_1 and q_2. This is the matrix F for $\Phi = -4$ used in the examples. The intensity scale follows this colour bar: min ▮ max.

which can be written in cylindrical coordinates:

$$\Psi(q_1,\varphi_1,q_2,\varphi_2) =$$

$$= \mathcal{N} \exp\left(-\frac{w_0^2}{4}(q_1^2 + q_2^2) - \underbrace{\frac{w_0^2}{2} q_1 q_2 \cos(\varphi_1 - \varphi_2)}_{\alpha} \right)$$

$$\times \operatorname{sinc}\left(\underbrace{\Phi + \frac{L}{4k_p}(q_1^2 + q_2^2)}_{\beta} - \underbrace{\frac{L}{2k_p} q_1 q_2 \cos(\varphi_1 - \varphi_2)}_{\gamma} \right) \tag{5.21}$$

$$= \mathcal{N}' \exp(-\alpha \cos(\varphi_1 - \varphi_2)) \operatorname{sinc}(\beta - \gamma \cos(\varphi_1 - \varphi_2)) \tag{5.22}$$

where $\mathcal{N}' = \mathcal{N} \exp(-\frac{w_0^2}{4}(q_1^2 + q_2^2))$

Similarly to what was done for the exp phase matching, we begin the evaluation of the Fourier transform with $\exp(-i\ell(\varphi_1 - \varphi_2))$ as the Fourier kernel, which

will give the eigenfunctions of a state of OAM $\pm \ell$.

$$\sqrt{P_\ell} F(q_1, q_2) = 2\pi \mathcal{N}' \underbrace{\int_0^{2\pi} e^{-\alpha \cos \varphi} \mathrm{sinc}(\beta - \gamma \cos \varphi) e^{-i\ell\varphi} d\varphi}_{H_\ell} \qquad (5.23)$$

We have

$$H_\ell = \int_0^{2\pi} e^{-\alpha \cos(\varphi)} \mathrm{sinc}(\beta - \gamma \cos(\varphi)) e^{-i\ell\varphi} d\varphi \qquad (5.24)$$

$$= \frac{1}{L} \int_{-L/2}^{L/2} e^{\frac{2i\beta}{L}t} \int_0^{2\pi} e^{-(\alpha + \frac{2i\gamma}{L}t)\cos(\varphi) - i\ell\varphi} d\varphi dt \qquad (5.25)$$

$$= \frac{2\pi}{L} \int_{-L/2}^{L/2} e^{\frac{2i\beta}{L}t} I_{|\ell|}\left(-\alpha - \frac{2i\gamma}{L}t\right) dt \qquad (5.26)$$

$$= \pi \int_{-1}^{1} e^{i\beta t} I_{|\ell|}(-\alpha - i\gamma t) dt \qquad (5.27)$$

This last form of the integral can be approximated if the product $b\sigma$ can be considered sufficiently small. The amount of precision sought restricts the maximum value of L_R, but at $L_R < 0.1$ (which corresponds to $b\sigma \lesssim 0.2$) the maximum error is less than 1%. The approximating function of $e^{-\frac{w_0^2}{4}(q_1^2 + q_2^2)} H_\ell$ is found to be

$$f(q_1, q_2) = 2\pi e^{-(q_1^2 + q_2^2)} I_\ell(2q_1 q_2) \mathrm{sinc}\left[\sqrt{2L_R}\left(q_1^2 + q_2^2\right) + \Phi\right] \qquad (5.28)$$

A matrix F is filled with entries $(i,j) = s\sqrt{ij} f(is, js)$, where s is the step, until the size N of the matrix is reached. Then $F^2 = FF^\dagger$ is used in the singular value decomposition. The resulting eigenvalues are $\lambda_{\ell,p}$ (at fixed ℓ); the resulting eigenmode are $\phi_{\ell,p}(q) = \sqrt{q} u_{\ell,p}(q)$.

At sufficiently large x, the modified Bessel function can be expanded as $I_\ell(x) \approx e^x / \sqrt{2\pi x}$. By combining this expansion with the mentioned multiplication by $\sqrt{q_1 q_2}$, we find that the matrix F has a similar appearance as the original two-photon field of Eq.(4), but with the roles of $(q_1 - q_2)$ and $(q_1 + q_2)$ interchanged and a removal of the vector character. After Taylor expansion of the sinc-function, which works best in the considered limit $L_R << 1$ for $\Phi < 0$, the resulting bi-exponential function yields a set of Hermite-Gaussian radial Schmidt modes centered around a displaced maximum.

Part III

Detection of high dimensional entanglement

Chapter 6

Projection on LG modes, thin crystal approximation

All this theory is fine, but is useless unless you know how to apply it.

(John Baez, *Lectures on classical mechanics*, 2005, p. 27)

6.1   Introduction

This chapter is based on the paper "Full characterisation of the spiral bandwidth of entangled biphotons", Physical Review A **83**, 033816 (2011) [77]. In this work we investigated the detection of a down-converted state generated using a Gaussian pump beam via projection onto Laguerre-Gaussian modes. We derived an expression for the simultaneous correlations in the orbital angular momentum modes and radial modes of the down-converted beams. We showed that, with the usual paraxial, collinear and Gaussian phase matching approximations, a fully analytic expression for the correlations can be derived in the limit for a short crystal. For comparison, in chapter 7 we will see how the thin crystal approximation can be analytically overcome, at the expense of projecting only onto the $p = 0$ subset of the LG modes, but with the advantage of using the full phase matching term.

6.2 Coincidence amplitudes on pairs of LG modes

6.2.1 Geometry and notation

We now describe the geometry of the beams, so that setting up the projection integral will be straightforward. A typical SPDC setup consists of a continuous-wave Gaussian pump beam propagating in the z direction, as shown in figure 6.1, incident on a short (typically 1–5mm) nonlinear crystal of length L. This produces two highly-correlated, lower-frequency photons. Energy is conserved in this process so that $\omega_p = \omega_s + \omega_i$, where the subscripts p, s, i refer to the pump, signal and idler, respectively. The photons are emitted at angles $\vartheta_{s,i}$ to the direction of propagation of the pump, $\hat{\mathbf{z}}$, and the components $\mathbf{q}_{s,i}$ (perpendicular to the z axis) of their $\mathbf{k}_{s,i}$ vectors are at angles φ_s and φ_i to the x axis. This means that φ is an azimuthal coordinate and this will play a central role in the description of the orbital angular momentum. The $\mathbf{q}$ components can be decomposed, on the planes perpendicular to the z axis, as

$$\mathbf{q_{s,i}} = \begin{pmatrix} \rho_{s,i} \cos\varphi_{s,i} \\ \rho_{s,i} \sin\varphi_{s,i} \\ 0 \end{pmatrix}, \tag{6.1}$$

where the radial variable ρ extends outwards from the z axis. The magnitude of the wave-vector inside the medium is $\omega n/c$ and for type-I phase matching (eoo) the extra-ordinary and ordinary refractive indices are $n_e(\omega_p)$ and $n_o(\omega_{s,i})$. Typical values, for example for β-barium borate (BBO), are: at $\lambda = 1064$nm, $n_o = 1.65, n_e = 1.54$ and at $\lambda = 532$nm, $n_o = 1.67, n_e = 1.55$ [78].

At the output of the nonlinear crystal the two-photon state in the wave-vector domain is given by [20, 79]

$$|\psi\rangle = \int\!\!\int d\mathbf{k}_s d\mathbf{k}_i A(\mathbf{k}_s, \mathbf{k}_i)\hat{a}_s^\dagger(\mathbf{k}_s)\hat{a}_i^\dagger(\mathbf{k}_i)|0\rangle, \tag{6.2}$$

where $A(\mathbf{k}_s, \mathbf{k}_i)$ describes the mode function of the pump and the phase matching conditions (as seen in more detail in chapter 3), $|0\rangle$ is the multimode vacuum state and $\hat{a}_s^\dagger(\mathbf{k}_s), \hat{a}_i^\dagger(\mathbf{k}_i)$ are creation operators for the signal and idler modes with wave vectors $\mathbf{k}_s, \mathbf{k}_i$, respectively.

Photon pairs generated by parametric down-conversion are entangled in arbitrary superpositions of OAM modes and we aim to complete this description by including the radial behaviour that can be experimentally probed by detecting the state in joint pairs of LG modes. A natural way to do this is to directly
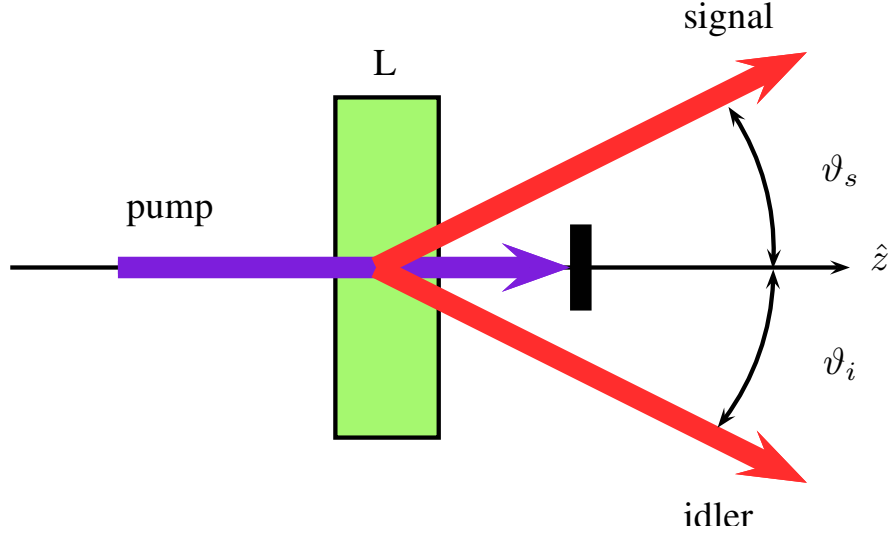
Figure 6.1: Sketch of the SPCD process. A gaussian pump propagating in the z direction is incident on a short nonlinear crystal. Signal and idler photons are produced at angles ϑ_s and ϑ_i to the pump direction.

describe the down-converted photons in terms of the LG modes, LG_p^ℓ. As usual, ℓ corresponds to the angular momentum carried by the mode, $\ell\hbar$, and describes the helical structure of the wave front around a wave front singularity and p is the number of radial zero crossings.

6.2.2   Calculation of the coincidence amplitudes

The coincidence probability for finding one signal photon in a given LG mode characterized by the ℓ_s and p_s numbers and one idler photon in a given LG mode characterized by the ℓ_i and p_i numbers is $P_{p_s,p_i}^{\ell_s,\ell_i} = \left|C_{p_s,p_i}^{\ell_s,\ell_i}\right|^2$, where the coincidence amplitudes $C_{p_s,p_i}^{\ell_s,\ell_i}$ are given by the overlap integral

$$
\begin{aligned}
C_{p_s,p_i}^{\ell_s,\ell_i} &= \langle LG_i, LG_s | \psi_{SPDC} \rangle \\
&= \int\int d^3k_s d^3k_i\, A(\mathbf{k}_s,\mathbf{k}_i) \left[LG_{p_s}^{\ell_s}(\mathbf{k}_s)\right]^* \left[LG_{p_i}^{\ell_i}(\mathbf{k}_i)\right]^*
\end{aligned}
\tag{6.3}
$$

The pump profile and phase matching condition were shown in previous chapters, but in those discussions their form was simplified. Here we need to set up a projection integral onto LG modes, and it is the best occasion to specify why those simplifications are enough for us. Most of what follows comes from geometrical considerations about the shape of the beams and separation of transverse and longitudinal quantities. We therefore start from a rather general spatial

amplitude [79]:

$$A(\mathbf{k}_s, \mathbf{k}_i) = \int d^3 k_p \tilde{E}_p(\mathbf{k}_p) \xi(\mathbf{k}_p - \mathbf{k}_s - \mathbf{k}_i) \delta(\omega - \omega_s - \omega_i), \qquad (6.4)$$

where the δ term enforces energy conservation and the ξ term arises due to phase matching. We can simplify this if, instead of the wave vector $\mathbf{k}$, we use its transverse component $\mathbf{q}$ and the corresponding frequency ω. As we have assumed a monochromatic gaussian pump of frequency $\omega_p = \omega_i + \omega_s$ we can write

$$\tilde{E}_p(\mathbf{k}_p) = \tilde{E}_p(q) \delta(\omega - \omega_p) \qquad (6.5)$$

where $\tilde{E}_p(q)$ is the Fourier transform of

$$E_p(r, \phi) = \sqrt{\frac{2}{\pi}} \frac{1}{w_p} \exp\left(\frac{-r^2}{w_p^2}\right), \qquad (6.6)$$

the spatial distribution of the pump at the input face of the crystal.

For a crystal of finite thickness, L, in the longitudinal direction and transverse length much larger than the pump beam size, the phase matching condition is [79]

$$\begin{aligned} \xi(\mathbf{k}_p - \mathbf{k}_s - \mathbf{k}_i) &= \delta(q_p - q_s - q_i) \\ &\times \sqrt{\frac{L}{\pi k_p}} \operatorname{sinc}\left(\frac{L \Delta k_z}{2}\right) \exp\left(\frac{-i L \Delta k_z}{2}\right) \end{aligned}$$

where $\Delta k_z = k_{p,z} - k_{s,z} - k_{i,z}$ and $k_z = \sqrt{n^2(\omega)\omega^2/c^2 - q^2}$ is the longitudinal component of the wave vector, $\mathbf{k}$, with transverse component q, angular frequency ω and refractive index $n(\omega)$.

If the angle between signal and idler beams is small enough that the z-component of the momentum vector ($\sqrt{k^2 - q^2}$) can be approximated by $k - q^2/2k$, where $q = |\mathbf{q}|$, we can write the phase matching function as

$$A(\mathbf{k}_s, \mathbf{k}_i) = \tilde{E}(q_s + q_i) \sqrt{\frac{2L}{\pi^2 k_p}} \operatorname{sinc}\left(\frac{L \Delta k}{2}\right) e^{-i\frac{L\Delta k}{2}} \qquad (6.7)$$

where

$$\Delta k = \frac{|\mathbf{q}_i - \mathbf{q}_s|^2}{2 k_p}. \qquad (6.8)$$

Explicitly:

$$A(\mathbf{q}_s, \mathbf{q}_i) = \overbrace{\frac{w_p}{\sqrt{2\pi}} e^{-\frac{w_p^2}{4}|\mathbf{q}_s + \mathbf{q}_i|^2}}^{\text{Pump}} \times$$

$$\times \underbrace{\sqrt{\frac{2L}{\pi^2 k_p}} \operatorname{sinc}\left(\frac{L|\mathbf{q}_i - \mathbf{q}_s|^2}{4k_p}\right) e^{-i\frac{L|\mathbf{q}_i - \mathbf{q}_s|^2}{4k_p}}}_{\text{Phase Matching}} \tag{6.9}$$

We now calculate the projection amplitudes onto a pair of LG modes in a cylindrical coordinate system, so we re-express (6.3) as:

$$\begin{aligned}
C_{p_s,p_i}^{\ell_s,\ell_i} &\propto \int_0^\infty\int_0^\infty\int_0^{2\pi}\int_0^{2\pi} A(\rho_i, \rho_s, \varphi_i, \varphi_s)\left[LG_{p_s}^{\ell_s}(\rho_s, \varphi_s)\right]^* \\
&\times \left[LG_{p_i}^{\ell_i}(\rho_i, \varphi_i)\right]^* \rho_i \rho_s \, d\rho_i d\rho_s d\varphi_i d\varphi_s,
\end{aligned} \tag{6.10}$$

where ρ and φ are the modulus and azimuthal angle, respectively, of the transverse component $\mathbf{q}$ of the wave vector and the normalized LG modes in k-space are given by

$$\begin{aligned}
LG_p^\ell(\rho, \varphi) &= \sqrt{\frac{w^2 p!}{2\pi(p + |\ell|)!}}\left(\frac{\rho w}{\sqrt{2}}\right)^{|\ell|} \exp\left(\frac{-\rho^2 w^2}{4}\right) \\
&\times (-1)^p \, L_p^{|\ell|}\left(\frac{\rho^2 w^2}{2}\right)\exp\left[i\ell\left(\varphi + \frac{\pi}{2}\right)\right].
\end{aligned} \tag{6.11}$$

Here w is the beam waist (we have assumed $z = 0$) and $L_p^{|\ell|}(\cdot)$ is an associated Laguerre polynomial.

The pump and phase matching functions in (6.9) can be written in cylindrical coordinates by performing the substitution

$$|\mathbf{q}_i \pm \mathbf{q}_s|^2 = \rho_i^2 + \rho_s^2 \pm 2\rho_i\rho_s\cos(\varphi_i - \varphi_s) \tag{6.12}$$

to obtain

$$\begin{aligned}
A(\rho_i, \rho_s, \varphi_i, \varphi_s) &= \frac{w_p}{\sqrt{2\pi}} e^{-\frac{w_p^2}{4}(\rho_i^2 + \rho_s^2 + 2\rho_i\rho_s\cos(\varphi_i - \varphi_s))}\sqrt{\frac{2L}{\pi^2 k_p}} \\
&\operatorname{sinc}\left(L\frac{\rho_i^2 + \rho_s^2 - 2\rho_i\rho_s\cos(\varphi_i - \varphi_s)}{4k_p}\right) e^{-iL\frac{\rho_i^2 + \rho_s^2 - 2\rho_i\rho_s\cos(\varphi_i - \varphi_s)}{4k_p}}
\end{aligned} \tag{6.13}$$

It is then straightforward to see that $A(\rho_i, \rho_s, \varphi_i, \varphi_s)$ depends on the radial coordinates in the momentum space, ρ_i and ρ_s, and on the *difference* between the

azimuthal angles, $\varphi_i - \varphi_s$. This allows the function to be written as a superposition of plane waves with phase $\exp(i\ell(\varphi_i - \varphi_s))$:

$$A(\rho_i, \rho_s, \varphi_i - \varphi_s) = \sum_{\ell=-\infty}^{\infty} f_\ell(\rho_i, \rho_s) e^{i\ell(\varphi_i - \varphi_s)}. \tag{6.14}$$

Using this in (6.10), the angular integral becomes

$$\sum_{\ell=-\infty}^{\infty} f_\ell(\rho_i, \rho_s) \int_0^{2\pi} \int_0^{2\pi} e^{i\ell_s \varphi_s} e^{i\ell_i \varphi_i} e^{i\ell(\varphi_i - \varphi_s)} d\varphi_i \, d\varphi_s \propto \delta_{\ell, -\ell_i} \delta_{\ell, \ell_s}, \tag{6.15}$$

which clearly enforces the angular momentum conservation, $\ell_i + \ell_s = 0$.

We also re-write the sinc function as the inverse Fourier transform of the step function:

$$\mathrm{sinc}\left(\frac{|\mathbf{q}_i - \mathbf{q}_s|^2 L}{4k_p}\right) = \frac{1}{L} \int_{-L/2}^{L/2} dt \exp\left(-\frac{i|\mathbf{q}_i - \mathbf{q}_s|^2 t}{2k_p}\right). \tag{6.16}$$

In this way we calculate

$$
\begin{aligned}
C_{p_i,p_s}^{\ell,-\ell} \quad &\propto K_{p_i,p_s}^{|\ell|} \int_{-L/2}^{L/2} dt \, \frac{B^{|\ell|}(1 - \frac{4I}{T})^{p_s}(1 - \frac{4S}{T})^{p_i}}{T^{|\ell|+1}} \\
&\times \quad {}_2F_1\left[\begin{matrix} -p_i, -p_s \\ -p_i - p_s - |\ell| \end{matrix}; \frac{T(T - 4I - 4S + 4)}{(T - 4S)(A - 4I)}\right]
\end{aligned} \tag{6.17}
$$

where the combinatorial coefficient $K_{p_i,p_s}^{|\ell|}$ is given by:

$$K_{p_i,p_s}^{|\ell|} = \frac{(p_i + p_s + |\ell|)!}{\sqrt{p_i! p_s! (p_s + |\ell|)!(p_i + |\ell|)!}}. \tag{6.18}$$

B, I, S and T are functions of the dummy variable t:

$$
\begin{aligned}
B \quad &= -\left(\frac{2t}{w_i w_s k_p} + \frac{L}{w_i w_s k_p} + \frac{i w_p^2}{w_i w_s}\right) \\
T \quad &= 4IS + B^2 \\
I \quad &= \frac{w_p^2}{2w_i^2} + \frac{1}{2} + \frac{it}{w_i^2 k_p} + \frac{iL}{2w_i^2 k_p} \\
S \quad &= \frac{w_p^2}{2w_s^2} + \frac{1}{2} + \frac{it}{w_s^2 k_p} + \frac{iL}{2w_s^2 k_p}
\end{aligned} \tag{6.19}
$$

and ${}_2F_1$ is the Gauss hypergeometric function. Although this integral is too complicated to be calculated analytically, it is simple enough to be evaluated numerically. However, this is not what we would like to do, as we are interested in an analytical result.

As we show now, rather conveniently, in the limit of a thin crystal we can solve the integral analytically (because the integration limits depend on the crystal length). This gives

$$C_{p_i,p_s}^{\ell,-\ell} \propto K_{p_i,p_s}^{|\ell|} \frac{(1-\gamma_i^2+\gamma_s^2)^{p_i}(1+\gamma_i^2-\gamma_s^2)^{p_s}(-2\gamma_i\gamma_s)^{|\ell|}}{(1+\gamma_i^2+\gamma_s^2)^{p_i+p_s+|\ell|}}$$
$$\times \, {}_2F_1\left[\begin{array}{c} -p_i,-p_s \\ -p_i-p_s-|\ell| \end{array}; \frac{1-(\gamma_i^2+\gamma_s^2)^2}{1-(\gamma_i^2-\gamma_s^2)^2}\right] \qquad (6.20)$$

where γ_i and γ_s are the ratios w_p/w_i and w_p/w_s, so that the two γ factors are the inverse signal and idler widths normalized to the pump width. This means that every amplitude is invariant under the joint scaling of signal, idler and pump widths. Of course, the same result can be achieved by starting directly with the state (3.23), i.e. with the thin crystal approximation of the SPDC state.

Comparing (6.20) with the numerical evaluation of (6.17) we find excellent agreement. The exact conditions in which such agreement is met will be explored in the next chapter. We will see that in normal experimental conditions these results are valid, but it is important to have a feeling for when such conditions are not met anymore.

6.3 Interpretation of the projection amplitudes

The advantage of an analytical result over an integral one, such as (6.17), is that it allows us to see more easily the role that each parameter plays in determining the state of the down-converted photon and the resulting ℓ distribution, or spiral bandwidth [20]. Full knowledge of this allows the quality of the entangled state to be determined and also allows us to compare the width of the distribution of ℓ modes of different states. From (6.20) it is clear that the distribution of modes is determined by the ratio of the signal and idler widths to the pump width. We can see this more clearly, and also investigate more the effect of taking the radial modes into account, if we consider some special cases.

6.3.1 Comparison with published results

Note that for the specific case of $p_s = p_i = 0$ eq.(6.20) reduces to

$$|C_{0,0}^{\ell,-\ell}|^2 \propto \left(\frac{2\gamma_i\gamma_s}{1+\gamma_i^2+\gamma_s^2}\right)^{2|\ell|}. \qquad (6.21)$$
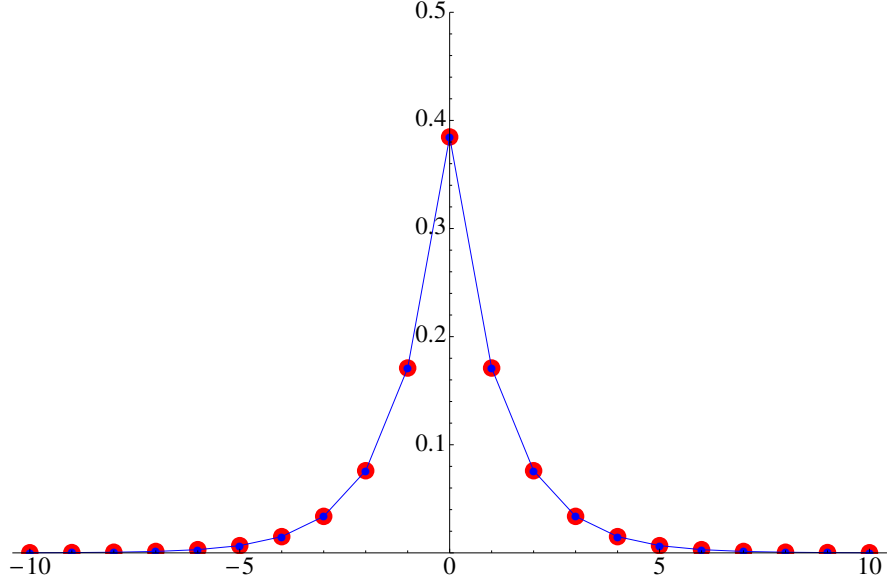
Figure 6.2: Normalised spiral bandwidth for $\gamma_s = \gamma_i = 1$ and $p_s = p_i = 0$. Results from eq. (10) in [20] are in red (large dots). The blue line is only a guide for the eye.

If we further choose the signal and idler widths to be equal, which is a common experimental condition, so that $\gamma_i = \gamma_s = \gamma$, then this simplifies to

$$|C_{0,0}^{\ell,-\ell}|^2 \propto \left(\frac{2\gamma^2}{1+2\gamma^2}\right)^{2|\ell|}. \tag{6.22}$$

Note that this result has been calculated in the thin crystal approximation, whereas the same result (i.e. with $p_s = p_i = 0$ and $\gamma_i = \gamma_s = \gamma$) without the thin crystal approximation is described in the next chapter. This agrees with the result found in [20] as depicted in figure 6.2:

Although the form of eq.(6.22) is very simple, it is remarkably precise when it is compared to the numerical evaluation of eq.(6.17), where the signal and idler sizes are equal and if the eigenstates of $p_i = p_s = 0$ have been selected.

6.3.2 Correlation between p_i and p_s

It is possible to have an extended view of the structure of the entangled system with an array plot that shows the contribution $P_{p_s,p_i}^{\ell,-\ell} = |C_{p_s,p_i}^{\ell,-\ell}|^2$ for each pair of modes.
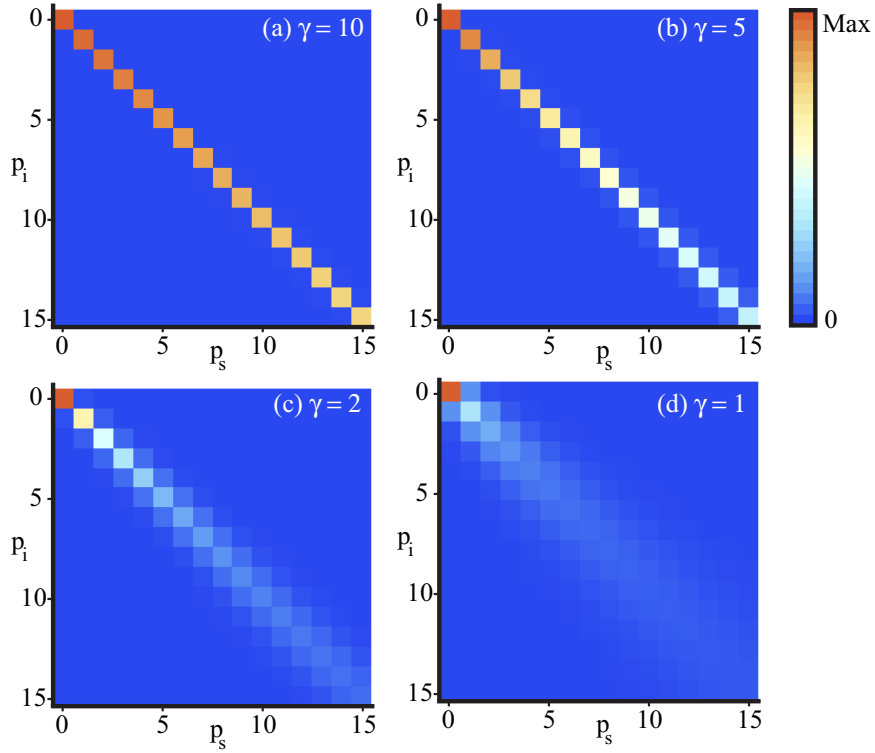
Figure 6.3: $P_{p_s,p_i}^{0,0}$ for $0 < p_i < 15$ and $0 < p_s < 15$, for different pump sizes.

Effect of the pump width on the correlation between p_i and p_s when $\ell = 0$ and signal and idler have the same size.

We show in figure 6.3 the probabilities of projecting onto modes in the form $|p_i, 0\rangle \otimes |p_s, 0\rangle$. Since $\gamma_i = \gamma_s$ in this case, we will omit the subscript and simply write γ. On the axes we are scanning the discrete values of p_i and p_s. In the limit of an infinite pump width p_i and p_s are delta correlated. Notice how the correlation between p_i and p_s breaks down as the pump beam size approaches that of signal and idler. (This corresponds to $\gamma \to 1$).

In an experiment it is not possible to use a pump with arbitrarily large beam waist because the crystal size places an effective upper limit. It is also not possible to shrink the size of signal and idler enough so that $\gamma \gg 1$ (although we will see in chapter 5 that it is in this direction that it is interesting to move). Therefore, the experimental scenario will be closer to the bottom right graph rather than the top left. This means that cross correlations between eigenstates of different p are to be expected. This relates closely with what we anticipated in chapter 1, namely that orthogonality between radial modes is much harder to achieve than orthogonality between OAM modes.
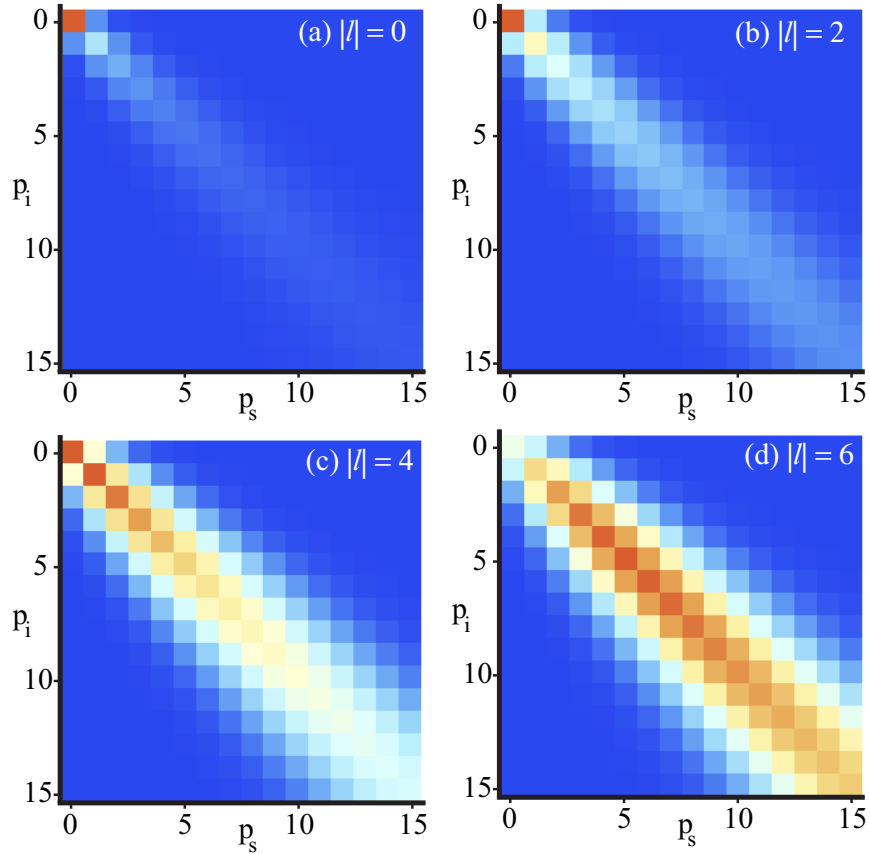
Figure 6.4: Correlations between $0 < p_i < 15$ and $0 < p_s < 15$, for different $|\ell|$ values and a pump of the same size as signal and idler.

Effect of ℓ on the correlation between p_i and p_s.

In the previous section we described the p correlations between states of OAM with $\ell = 0$. We now consider the case $\ell \neq 0$. It is worth recalling that (in general) the probability of selecting a joint state of a given OAM, $|\ell|$, decreases as the value of $|\ell|$ increases. For this reason, if we select a specific value of ℓ, as is done with SLMs, the average rate of coincidence counts will be lower than the rate measured after selecting states with a lower value of $|\ell|$. If the value of the selected $|\ell|$ is too large, the rate could drop too much to obtain a meaningful signal to noise ratio and the effects could be masked by noise. This is why we show the effect only up to the value $|\ell| = 6$. The reason why the maximum probability of detecting modes of similar p shifts to a higher value of p for states of a higher OAM can be found by considering what happens to the product of the pump mode with specific signal and idler modes. In this case the key idea is that when we calculate the overlap integral between the pump and modes with a given ℓ, there is an optimal combination of p_s and p_i that maximizes the overlap. The higher the value of ℓ, the higher the optimal value of the combination of radial indices.
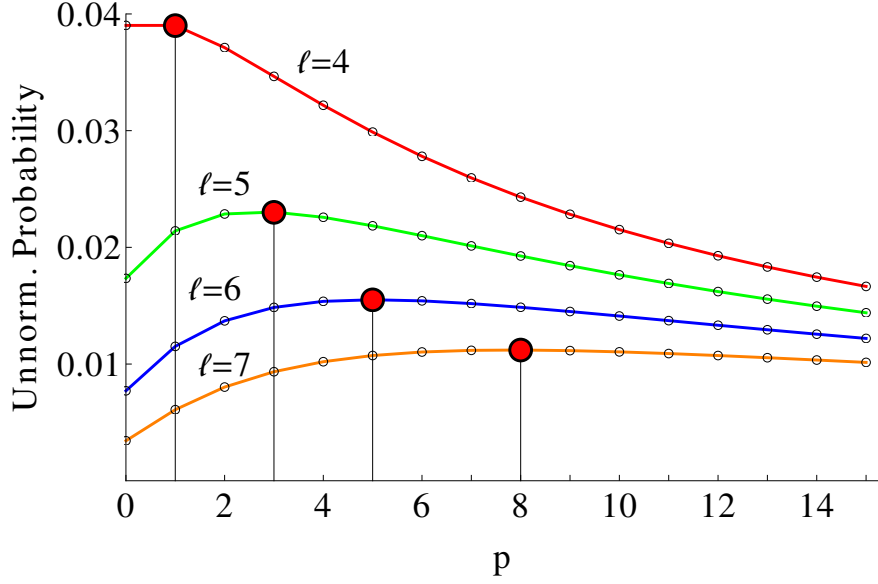
Figure 6.5: The overlap between LG_0^0 and $LG_p^\ell LG_p^{-\ell}$ is best achieved at an optimal value of p (indicated by red disks), once $|\ell|$ is fixed.

To show how this happens, consider the product of the fundamental pump mode with two $p_i = p_s$ modes (i.e. the integrand of (6.10) in the short crystal limit). If we calculate the (unnormalised) overlap integral, although the effect of increasing the radial indices moves the inner rings towards the origin of the coordinates (where the maximum value of the gaussian mode of the pump is), the radius of the main ring reduces and consequently the overlap reduces. Thus, as the number of rings increases, an optimal value of the overlap integral is reached and subsequently it drops, which is exactly what happens along the diagonals of the graphs in figure 6.4, which are also shown in figure 6.5.

Effect of signal-idler size mismatch on the correlation between p_i and p_s.

In the analysis of the correlations between states of different p index we assumed signal and idler fields to have the same size ($\gamma_i = \gamma_s$). Although detection systems are generally set to detect fields of the same size to maximise coincidences, it is worth having a look at what could in principle happen with detection modes of different size. When the sizes of signal and idler beams begin to differ, the main correlation "direction" (namely the main diagonal in the correlation matrix) between modes of p_i and p_s shifts towards one or the other axis, depending which of the beams is larger, thus yielding the highest correlations between modes of

different radial indices rather than between modes of equal radial indices. Figure 6.6 is a group of four graphs, each featuring the probabilities of detecting an state in the form $|p_i, 0\rangle \otimes |p_s, 0\rangle$, where the size of the idler beam increases to twice the size of the signal beam, which remains of the same size of the pump.
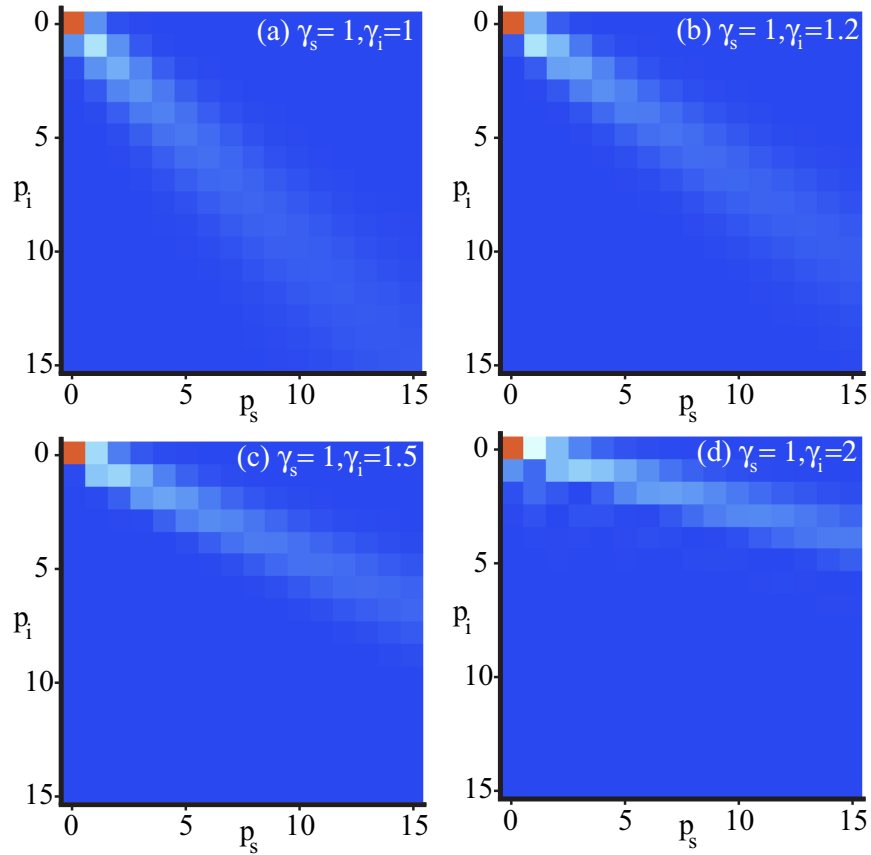

Figure 6.6: These graphs show that the effect of a width mismatch yields a higher probability of finding a state with different radial indices. If, instead, the value of γ_s were larger than γ_i the graphs would be mirrored with respect to the leading diagonal.

It is clear from the graphs that we can increase the detection probability of modes of very different values of p simply by changing the relative size of signal and idler modes. The consequence of selecting a different ℓ mode is analogous to what is shown in figure 6.4.

As we are able to calculate the amplitudes for any values of the beams sizes, it is interesting to see what the effects of the beams sizes are on the spiral bandwidth. We will discuss this in the next section and give multiple examples of the effect of varying the beams sizes.

6.3.3   Effect of modes widths on the spiral bandwidth.

We now want to see the effects of the beams sizes on the spiral bandwidth (SB). In this context, the SB is the collection of detection probabilities in the form $P_{p_s,p_i}^{\ell,-\ell}$, where the radial indices p_i and p_s have been fixed and we scan the ℓ eigenvalues in an interval (in our case of ± 20) around $\ell = 0$. Unlike the previous graphs, each graph now *does not* represent the structure of the SPDC state, but rather a collection of "slices" of the total modal content, each slice representing the SB, i.e. each graph features only the ℓ eigenvalues.

Each group of graphs will highlight a particular feature of the SB determined by the radial indices. Note that the normalization has to be performed on each individual horizontal line and not on the graph as a whole, because the beam size is just a parameter of the detection basis. We stress that each value in the graphs corresponds to the integral over the whole area where the pump overlaps with the detection mode.

We will divide the effects into four families, each one corresponding to equal or different radial indices and equal or different signal-idler sizes. In this way we can cover all the experimentally interesting effects on the SB.

Equal radial indices - Equal signal-idler sizes.

This case is of particular importance because now the ratio between pump and signal-idler beams size is the only feature that influences the detection of the SPDC state. It is a quantity that is relatively easy to manipulate, for example by choosing an initial pump beam size and then magnifying or de-magnifying the signal and idler beams. We acknowledge that it can be hard to achieve a particular magnification, because of limiting apertures in the setup, especially between the crystal and the detection instrumentation. Such finite apertures will give a limit on the maximum size of the beams, and also losses if this maximum size is exceeded.

The graphs on the left-hand side of figure 6.7 show the SB (horizontally) for different values of the ratio of the pump width to the signal-idler width, i.e. γ (which changes on the vertical axis). The positions where the bandwidths on the right are taken are marked on the graphs at positions $\gamma_s = 0.5$ red, $\gamma_s = 2$ green, $\gamma_s = 4$ blue. Note that the blue bandwidth on the right is indicated in yellow ($\gamma_s = 4$) on the left. The same goes for all the graphs in this chapter with analogous structure.

A larger pump width (or smaller signal and idler widths) increases the width of the SB. Also, measuring on a basis with larger radial indices gives a larger SB.
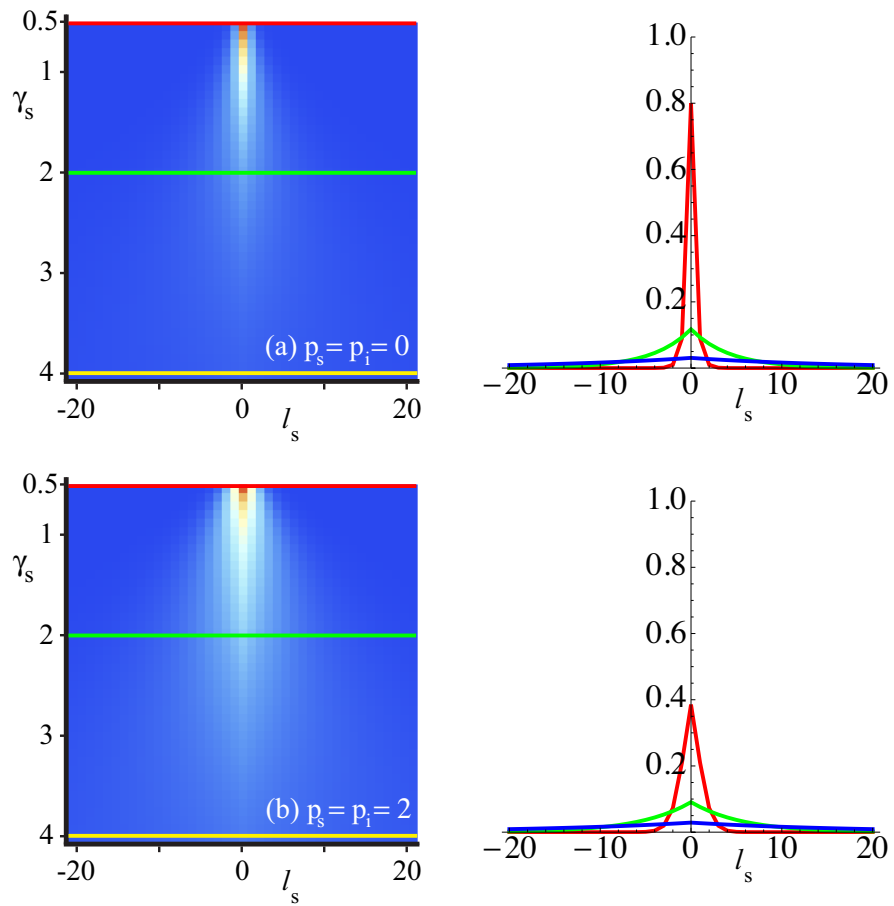
Figure 6.7: Dependence of the SB on the pump width and radial number.

In fact, modes with more rings still give a significant overlap with the gaussian pump for larger values of ℓ than modes with less rings.

Equal radial indices - Different signal-idler sizes.

A rather different and interesting effect is achieved when the size of signal and idler differ by some amount. figure 6.8 shows similar graphs to figure 6.7, but here the signal and idler beams now have different widths.
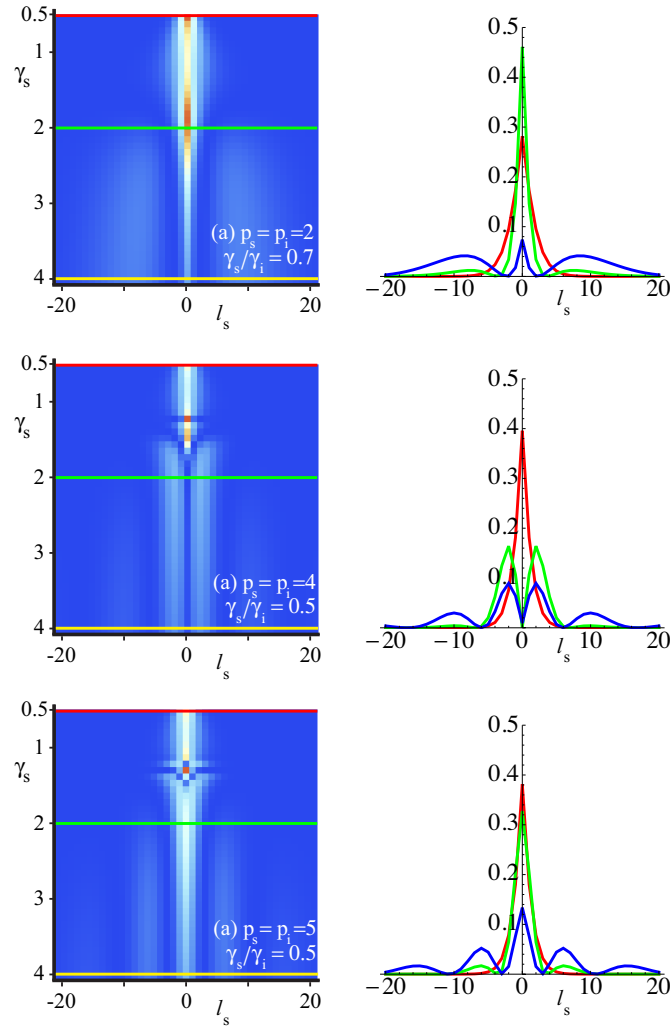

Figure 6.8: These graphs show how the difference in the width of signal and idler influences the SB.

A physical explanation of the multiple-branched SBs in figure 6.8 is given by the fact that the joint detection mode is made out of two modes with different widths (therefore it will consist of concentric rings of positive and negative value) and by the consequence of increasing ℓ: the rings in the modes move further away from where the pump mode is concentrated, but since they have a different

width their product will change shape, becoming alternately mostly negative or mostly positive. The higher the width difference, the higher number of times this process takes place in the same amount of ℓ values.

The consequence is that the value of the overlap with the pump mode (which is always positive) increases and decreases alternately as higher and higher ℓ modes are chosen, because the mostly positive or mostly negative part of the joint detection mode, since they are moving away from the center of the beam, will cease to have a substantial overlap.

The number of branches in the graphs in figure 6.8 depends on the relative dimensions of signal and idler, and also on the number of rings in the joint detection mode, as mostly positive and negative parts can co-exist in the same detection mode over many rings. If the pump is too small (top of the graphs on the left-hand side, and red SB in the graphs on the right-hand side) the effect can't be noticed because the gaussian pump always stays mostly inside the first ring (so it doesn't matter now many there are outside the first one, or their values). However, for a larger value of the pump (bottom of the graphs), it initially overlaps with many rings, whose overlap can then be lost, in the way explained, while we consider higher and higher values of $|\ell|$. The blue SB in the right-hand side graphs portrays this effect.

Different radial indices - Equal signal-idler sizes.

We now consider the case of different values of the radial indices p_i and p_s. Since the probability of detecting modes of different radial index depends on the ratios of the beam sizes, there are choices of widths that may significantly increase the probability of detecting some particular output modes, as seen in figure 6.6.

In figure 6.9 we show the effect of increasing the pump width while maintaining the signal and idler at the same size, but with different radial indices. As long as signal and idler beams retain the same width, the number of branches will be always two, independent of the values of p_i and p_s. This behavior is due to the fact that the beams have the same width, therefore, as higher and higher values of ℓ are reached, the rings in the joint detection mode will be moved away from the pump "rigidly", i.e. maintaining the same shape and thus delivering a smooth decay of the overlap.

Different radial indices - Different signal-idler sizes.

We showed that a multiple-branched SB can be obtained when signal and idler beams have different sizes, now we show the same effect while also varying p_i
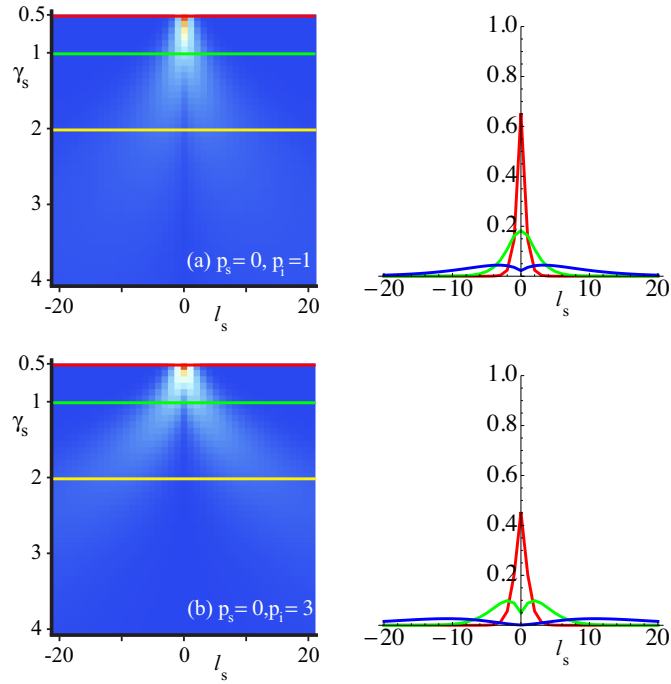
Figure 6.9: Joint effect of a signal and an idler mode with different radial indices.

and p_s. To give a more complete description of the possible SBs we will now consider the *difference* between γ_i and γ_s, rather than the ratio. In the first four graphs the ratios of the pump width with signal and idler widths have a difference (i.e. $\gamma_i - \gamma_s$) of 0.5, 0.75, 1 and 2. The pump width indicated on the left axis of the graphs is relative to the signal width.

The physical explanation for the shapes of the SBs is similar to the one supplied in the case of equal radial indices and different signal-idler sizes. The effect is dependent on both the relative size of signal and idler and on the number of rings in the modes. The fact that the number of branches depends on the smaller of p_i and p_s clarifies what is stated above, namely that it all depends on how the rings in the modes overlap: if the joint detection mode consists of two modes with $p_s + 1$ and $p_i + 1$ rings, the rings will overlap more or less effectively depending on both the relative size and the intensity of the modes. Therefore, it's the mode with less rings that counts. When ℓ is scanned, the two modes, having a different width, will shift the rings away from the center at different rates, giving alternating mostly positive or mostly negative groups of rings. Since the total number of rings depends on the lowest of the radial indices, the mode with less rings will determine the maximum number of branches in the SB, as is shown in figure 6.10.
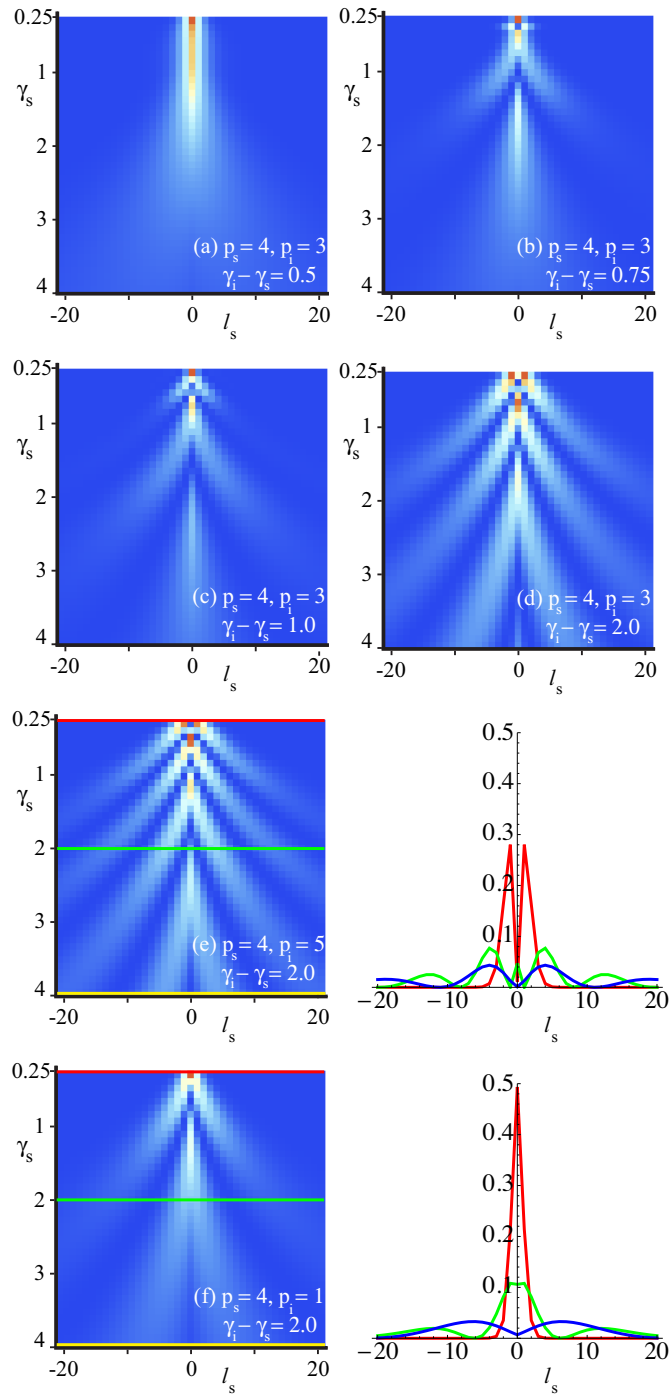
97

Figure 6.10: Joint effect of a signal and an idler mode with different radial indices and different widths.

6.4   Conclusion

We have shown the overlap of the SPDC state produced with a Gaussian pump beam with a set of LG modes and investigated the resulting correlations in the OAM and radial momentum. The results show excellent agreement with previous works on SPDC [20] and extend them to a much more general case, albeit at the cost of considering thin crystals only, which is equivalent to considering a source producing an infinite number of Schmidt modes. We will see in the next chapter what "thin crystal" really means and where the various bounds come into play.

Chapter 7

Projection of the full state onto pure OAM modes

I respect conscious guessing, because it comes from the best human qualities: courage and modesty.

(Imre Lakatos, *Proofs and Refutations*, 1976, p. 30)

7.1 Introduction

This chapter is based on the paper "Bounds and optimisation of orbital angular momentum bandwidths within parametric down-conversion systems", EPJD **66** (7) (2012) [75]. The measurement of high-dimensional entangled states of orbital angular momentum prepared by spontaneous parametric down-conversion can be considered in two separate stages: a generation stage and a detection stage. Given a certain number of generated modes, the number of measured modes is determined by the measurement apparatus. We derive a simple relationship between the generation and detection parameters and the number of measured entangled modes. We overcome the limitation of the Gaussian approximation of the phase matching function while still obtaining analytical results, at the expense of addressing only the $p = 0$ subspace.

The OAM is conserved in the down-conversion process and hence for a Gaussian ($\ell = 0$) pump, the OAM of the signal and idler fields are perfectly anticorrelated. There are also correlations on the radial direction, but these will not

concern us in this chapter. Our central concern will be the number of entangled lowest order ($p = 0$) Laguerre-Gaussian modes generated in a down-conversion experiment. The typical setup that we consider is a type-I or type-II, degenerate SPDC setup. We work in the regime of undepleted pump and we neglect eventual anisotropies of the down-converted beams.

We find that, for any given set of generation parameters (pump waist w_p, wavelength λ, crystal length L) the detection apparatus can be prepared in a way that maximises the measured number of entangled modes and that two important parameters are γ, the ratio of the width of the pump beam to the width of the detection modes, and L_R, the length of the crystal normalised to the Rayleigh range of the pump beam:

$$\gamma_{s,i} = \frac{w_p}{w_{s,i}} \quad \text{and} \quad L_R = \frac{L}{z_R}, \tag{7.1}$$

where the Rayleigh range, as previously stated, is $z_R = \frac{\pi w_p^2}{\lambda}$. We assume that the signal and idler modes have the same width so that $w_s = w_i$ and $\gamma_s = \gamma_i = \gamma$.

The precise calculation of $w_{s,i}$ depends upon the details of the detection system. Our analysis can be applied if the back-projected fundamental detection mode size, $w_{i,s}$, is approximately ℓ-independent over the range of OAM of interest, and if the modes with $p \neq 0$ couple only weakly with the fundamental mode of the fibre that carries the signal to the coincidence counter. We can investigate the L_R dependence of the OAM bandwidth, while recognising that many experiments operate in a regime where $L_R \ll 1$ [15, 19, 27, 80, 81]. In the short crystal limit and near to collinearity the familiar sinc phase can be dropped [82]. One can then obtain an analytical form for the down-converted state (chapter 6 of this thesis and [77, 83]) and its extension to non-Gaussian pump beams [65]. Our aim in this chapter is to go beyond these existing analyses and to explore regimes in which the sinc phase matching term becomes significant, which leads to the exact analytical expression (7.6) and to the characterisation of the detection parameters. We present both an analytical treatment and also a simple geometrical argument for our results.

The second section of the chapter specifies the definitions of the various bandwidths which are used. The third section contains the analytical approach to calculate the projection amplitudes. The fourth section contains the geometrical approach to calculate a simple formula that gives the measurement bandwidth. The fifth section contains the interpretation of the results and the conclusions.

7.2 Definition of bandwidths

For a distribution of probabilities, in our case for the OAM of the signal and idler photon in SPDC, we can define a number of statistical measures. For high-dimensional entanglement we require as many modes as possible to contribute to the state and, moreover, for these to contribute strongly, that is to have a significant probability. As before, we will use the Schmidt number [21, 25]:

$$K(\{p_i\}) := \frac{1}{\sum_i \lambda_i^2},\qquad(7.2)$$

where the probabilities $\{\lambda_i\}$ are, in our case, those for each of the joint OAM modes. The measure K gives the effective number of contributing modes and hence the effective dimensionality of the system. In experiments, it is typical to quote the full-width at half maximum as the measure of the bandwidth (FWHM) so as to include only modes that are well above the noise floor. FWHM should not be confused with K. For simple, symmetrical and single-peaked probability distributions, the FWHM provides a convenient measure of the bandwidth, but for more complicated distributions, the Schmidt number is a more accurate measure. The precise relationship between the Schmidt number and the FWHM depends upon the detailed shape of the distribution but typical of our systems is that the K exceeds the FWHM, see figure 7.1. For a distribution like this we can define an effective range of modes contributing to the state ranging from $\ell_{\max}$ to $\ell_{\min} = -\ell_{\max}$ such that $K = 1 + 2|\ell_{\max}|$.

The *generation bandwidth* is the effective number of entangled modes generated in the SPDC process. As it does not depend on the detection apparatus, it is a function only of the crystal length and of the size of the pump beam, combined into the quantity L_R, defined in eq. (7.1). This bandwidth can be thought of as the dimensionality of the entanglement in OAM and can be calculated through the Schmidt decomposition of the SPDC state [32].

The *measurement bandwidth* represents the number of modes that a detector will measure in an experiment and depends on both the generated modes and on the overlap of these with the detection modes. In doing so, we need to consider the optics used to image the light onto the detectors and any restriction arising from this, such as a restriction to $p = 0$ Laguerre-Gaussian modes. The overlap between the generated modes and the back-projected detection modes needs to be maintained both in the image plane and in the far field plane of the crystal: a setup with high overlap in the image plane may still suffer from low overlap in the far field or vice versa and this would translate into a decreased modal

sensitivity. This overlap requirement has a central role in the derivation of eq. (7.10), which is based on the argument that the angular spread of a generated mode cannot exceed the natural spread of the down-conversion cone. In the next sections we will define an image plane bandwidth and a far field bandwidth and, as we shall show, there is a natural way of combining the two. This geometrical result is strongly supported by the more complicated analytic result, which we evaluate numerically for a comparison in figure 7.3.

7.3 Analytical treatment

A direct calculation of the measurement bandwidth needs to consider the overlap between the SPDC state and a pair of joint detection modes [20, 77]. This yields a series of complex measurement amplitudes $\{C_\ell\}$ where ℓ labels each value of the OAM that was measured. The measured Schmidt number (or the measurement bandwidth) is therefore given by the measure K applied to the set of projection probabilities

$$K(\{P_\ell\}), \quad \text{where} \quad P_\ell = |C_\ell|^2. \tag{7.3}$$

We seek to evaluate this quantity for a Gaussian pump laser, taking full account of the sinc phase-matching term. In this way we extend the regime of validity of earlier calculations.

We consider the measurement modes for the signal and idler fields to be a pair of Laguerre-Gaussian modes. The LG modes are characterised by two integers ℓ and p and a real positive number w, which represent the OAM quantum number, the radial quantum number and the Gaussian modal width, respectively. For simplicity, we set $p = 0$, which limits our analysis to modes with a single bright ring in the transverse plane. Many of our experiments are designed to detect $p = 0$ modes with a higher efficiency, moreover, than higher-order modes. We note however, that modes with non-zero p are produced in the SPDC process [77] and, indeed, it is these that makes it possible to observe entanglement of three-dimensional vortex knots in SPDC [26].

The SPDC wave function $\psi(\mathbf{q}_i, \mathbf{q}_s)$, in momentum space, is written in the following way, where the subscripts s and i refer to signal and idler modes [20]:

$$\psi(\mathbf{q}_i, \mathbf{q}_s) = N e^{-\frac{w_p^2}{4}|\mathbf{q}_i + \mathbf{q}_s|^2} \operatorname{sinc}\left(\frac{L}{4 k_p}|\mathbf{q}_i - \mathbf{q}_s|^2\right). \tag{7.4}$$

Here $\mathbf{q}$ is the transverse component of the momentum vector $\mathbf{k}$, w_p is the pump width, L is the crystal thickness, k_p is the wave vector of the pump. The first term corresponds to the transverse wavevector components of the pump, while the second term represents the phase-matching imposed on the down-conversion process by the nonlinear crystal.

We consider each detection mode to be an LG mode with radial quantum number $p = 0$. In polar coordinates (ρ, φ) in momentum space it has the form

$$LG_\ell(\rho, \varphi) = \sqrt{\frac{w^2}{2\pi|\ell|!}} \left(\frac{\rho w}{\sqrt{2}}\right)^{|\ell|} e^{-\frac{\rho^2 w^2}{4}} e^{i\ell\varphi}. \tag{7.5}$$

The projection amplitude is therefore calculated by evaluating the overlap integral of ψ with two LG modes of opposite OAM (because of angular momentum conservation) [19, 20, 84]. The result is found to be (from eq. (6.20) for $p_i = p_s = 0$)

$$C_\ell^{L_R, \gamma} = \frac{\mathcal{N}}{L_R} \left(\frac{2\gamma^2}{1 + 2\gamma^2}\right)^{|\ell|} \left[\xi^{|\ell|+1} \Phi_\ell^{L_R, \gamma} - \Phi_\ell^{0, \gamma}\right]. \tag{7.6}$$

We note that the first term in brackets corresponds to that obtained previously [77], specialised to equal signal and idler widths and $p = 0$ modes. Here the function $\Phi_\ell^{L_R, \gamma}$ is the Lerch transcendent function of order $(1, |\ell| + 1)$ and argument $-2\gamma^2\xi$ [85]:

$$\Phi_\ell^{L_R, \gamma} = \Phi(-2\gamma^2\xi, 1, |\ell| + 1), \qquad \xi = \frac{i + L_R}{i - 2\gamma^2 L_R}. \tag{7.7}$$

Note that $\xi = 1$ for $L_R = 0$.

Once L_R and γ are specified, the amplitudes $C_\ell^{L_R, \gamma}$ are to be used in eq. (7.3), in order to calculate the measurement bandwidth. The dependence of the projection amplitudes on a transcendent function makes further analytical calculation difficult, and a numerical approach has to be employed. However, as the tails of the distribution of projection probabilities have a slow decay and therefore an effect on the width even at high $|\ell|$, the numerical approach is slow, if an accurate result is sought.

In figure 7.1 we give the probabilities for the angular momentum values ℓ for $L_R = 0.001$ and $\gamma = 2$. In this parameter range existing analytical expressions provide an excellent approximation [20, 77].
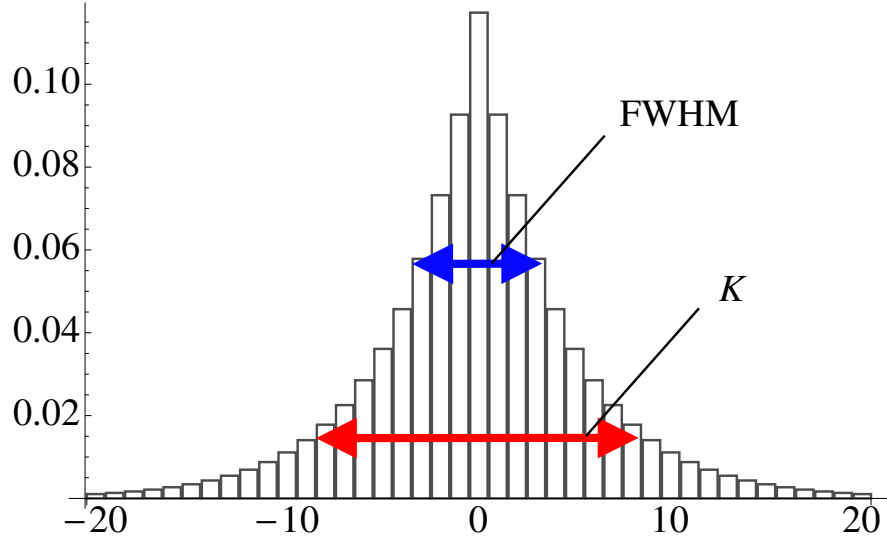
Figure 7.1: An example of a distribution of $|C_\ell^{L_R,\gamma}|^2$ for $L_R = 0.001$ and $\gamma = 2$. The FWHM and the measurement bandwidth K are shown in blue and red, respectively.

7.4 Geometrical argument

In this section we find an upper (and therefore lower) bound for the generated OAM values, and for the measured OAM values. The measurement bandwidth that we calculate from such bounds matches the analytic result of the previous section and therefore allows to avoid calculating numerically the distribution of projection probabilities.
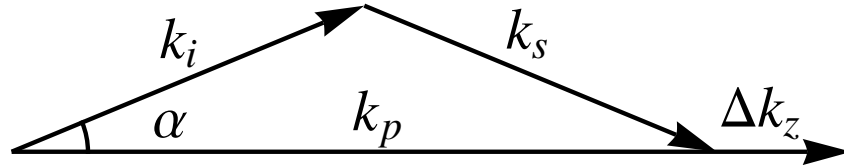

Figure 7.2: The relation between α and Δk_z sets a natural upper bound to α for near-collinear emission.

The phase-matching efficiency of the down-conversion process depends upon the axial mismatch Δk_z between wave vectors of the pump, signal and idler fields, and it is given by $\text{sinc}^2(L\Delta k_z/2)$. When optimised for degenerate, near-collinear phase-matching, the signal and idler output is obtained over a narrow range of angles, α, for which $L\Delta k_z \lesssim \pi$. With reference to figure 7.2, for small α

(which corresponds to being near to collinearity) we can write

$$\Delta k_z \simeq \frac{\alpha^2 k_p}{2}.$$ (7.8)

It follows, therefore, that the allowed values of α are bounded from above:

$$\alpha^2 \lesssim \frac{2\pi}{k_p L}.$$ (7.9)

For Laguerre-Gaussian modes, in the paraxial regime, we can define an effective local wavevector associated with the gradient of the phase. The helical form of the wavefronts gives rise to an angular spreading of these such that at a distance r from the mode axis, the angular spread is $\beta \simeq \ell/kr$ [7], which can be interpreted as the local spreading angle from the optical axis. The natural restriction on α imposed by the phase matching therefore sets a limit $\beta \lesssim \alpha$ on the efficiency of production of the OAM carrying beams, imposing a restriction on the generated OAM bandwidth. Such restriction is a natural consequence of the fact that a generated mode cannot be more divergent than the down-conversion cone. The relation $\beta \lesssim \alpha$, using the definitions and bounds given above for β and α, can be rewritten as

$$\ell \lesssim r\sqrt{\frac{\pi k_p}{2L}},$$ (7.10)

where we have made the approximation that $k_{s,i} \approx k_p/2$. This relation is the starting point to calculate the generation bandwidth and for the analysis in the far field of the image plane of the crystal.

7.4.1 Generation bandwidth

The beam size can be no bigger than that of the pump beam, i.e. $r \lesssim w_p$. Applying this bound to eq. (7.10) we obtain an upper bound for the generated OAM value:

$$\ell_{\text{gen}} \lesssim w_p\sqrt{\frac{\pi k_p}{2L}} = \sqrt{\frac{\pi}{L_R}}.$$ (7.11)

It follows, therefore, that the generation bandwidth is

$$K_{\text{gen}} = 1 + 2\sqrt{\frac{\pi}{L_R}}.$$ (7.12)

This number represents the effective number of entangled OAM modes generated by the source.

7.4.2 Image plane bandwidth

As anticipated in section 2, to calculate the measurement bandwidth we need to consider the overlap of the generated field with the detection modes in the image plane of the crystal and in its far field. Intuitively, a detection system which has a good overlap in the image plane, but that detects light that only comes from a narrow spread of directions would restrict the measured bandwidth. A similar restriction would occur for one that has a good overlap with the typical incoming angles of LG beams, but that has a poor overlap with the intensity in the image plane. It is clear that in order to optimise a detection system, both these quantities have to be taken into account.

To calculate the overlap in the image plane it suffices to note that a $p = 0$ Laguerre-Gaussian mode with OAM number ℓ and width w has its maximum intensity at a radius

$$r = w\sqrt{\frac{\ell}{2}}. \tag{7.13}$$

For efficient conversion of pump to signal and idler we require that the pump, single and idler beams should all overlap, giving a restriction on the maximum size of the down-converted beams ($r_{s,i} \lesssim w_p$) and hence an upper bound to the value of OAM in the plane of the crystal corresponding to

$$r_{s,i} = w_{s,i}\sqrt{\frac{\ell}{2}} \lesssim w_p. \tag{7.14}$$

In terms of γ, this gives an upper bound of the value of the OAM in the plane of the crystal:

$$\ell_{\text{ip}} \lesssim 2\gamma^2 \tag{7.15}$$

and hence an image plane bandwidth

$$K_{\text{ip}} = 1 + 4\gamma^2. \tag{7.16}$$

7.4.3 Far field bandwidth

It is clear that in the far field of the plane of the crystal, instead of a real space argument, we need to use the angular relationship $\beta \lesssim \alpha$, expressed in (7.10), where we apply the restriction for the maximum width of the detection modes given in (7.14):

$$\ell \lesssim w_{s,i} \sqrt{\frac{\ell}{2}} \sqrt{\frac{\pi k_p}{2L}}. \tag{7.17}$$

From which, replacing $w_{s,i}$ with w_p/γ, we obtain an upper bound of the value of the OAM in the far field of the plane of the crystal:

$$\ell_{\mathrm{FF}} \lesssim \frac{\pi}{2\gamma^2 L_R} \tag{7.18}$$

and therefore a far field bandwidth

$$K_{\mathrm{FF}} = 1 + \frac{\pi}{\gamma^2 L_R}. \tag{7.19}$$

7.4.4 Measurement bandwidth

If K_{ip} and K_{FF} are very different from each other, the resulting measurement bandwidth is given by the smaller of the two. For cases where the bandwidths are similar it is sensible to combine them. The convolution of two normal distributions of widths k and k' gives a normal distribution of width $(k^{-2} + k'^{-2})^{-1/2}$. Similarly, we can get an estimate of the total measurement bandwidth by considering the convolution of two normal distributions of widths K_{ip} and K_{FF}. The bandwidth of the resulting distribution is

$$\begin{aligned} K &= \left(K_{\mathrm{ip}}^{-2} + K_{\mathrm{FF}}^{-2} \right)^{-1/2} \\ &= \left(\left(1 + 4\gamma^2 \right)^{-2} + \left(1 + \frac{\pi}{\gamma^2 L_R} \right)^{-2} \right)^{-1/2}. \end{aligned} \tag{7.20}$$

7.5 Analysis of the results

For a comparison between the analytic and geometric arguments, we calculate the width of the distribution given by the modulus squared of the coefficients in (7.6) and compare it to (7.20). In figure 7.3 we plot the two bandwidths as functions of L_R for $\gamma = 3$, $\gamma = 5$ and $\gamma = 7$. The solid curves (red online) represent the measurement bandwidth calculated from the numerical evaluation of the ana-
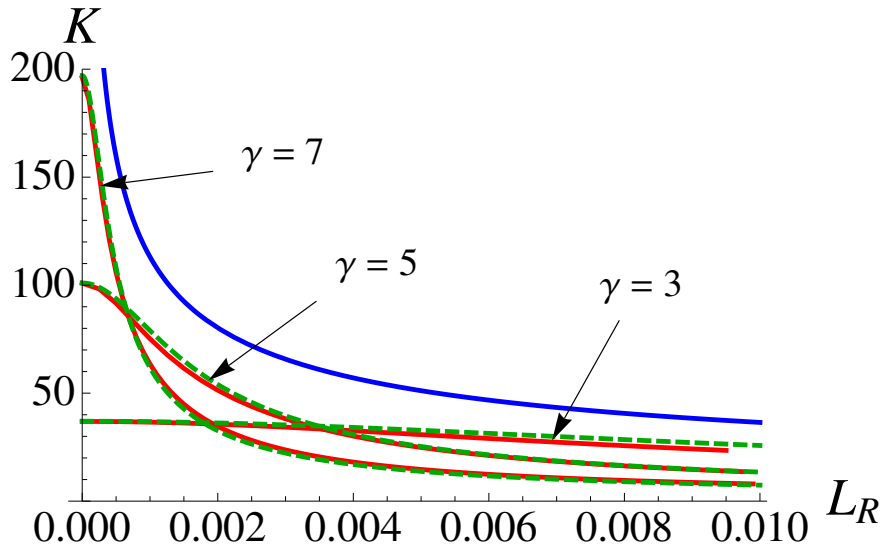
Figure 7.3: The blue line (uppermost) is the generation bandwidth defined in (7.12), the green curves (dashed) are calculated from our analytical treatment , and the red curves (solid) are the result of our geometrical argument.

lytical model. The dashed curves (green online) are the same bandwidths calculated with our geometrical argument. The uppermost solid line (blue online) is the generation bandwidth. Note that to achieve high dimensional entanglement the crystal length should be a small fraction of the Rayleigh range.

We see that the geometrical argument is in excellent agreement with the numerical evaluation of our analytical result. The effect of increasing γ yields a higher measurement bandwidth for very small values of L_R, but for large enough values of γ and for fixed L_R, the measurement bandwidth eventually drops. Therefore it reaches a maximum value for a particular crystal length. Under all conditions the measurement bandwidth never reaches that of the generation bandwidth, because we are restricting the measurement to modes with $p = 0$. Note, however, that the full generation bandwidth does not arise explicitly from additional values of the OAM but rather from entanglement in the radial quantum number p.

Differentiation of eq. (7.20) with respect to the crystal length gives an estimate of the value of γ corresponding to the highest measurement bandwidth for a given L_R. In this way we find

$$\gamma_{\text{opt}} \approx \sqrt[4]{\frac{\pi}{4L_R}}. \tag{7.21}$$

It is worth noting that for such value of γ we have that $K_{\text{ip}} = K_{\text{FF}} = K_{\text{gen}}$, where
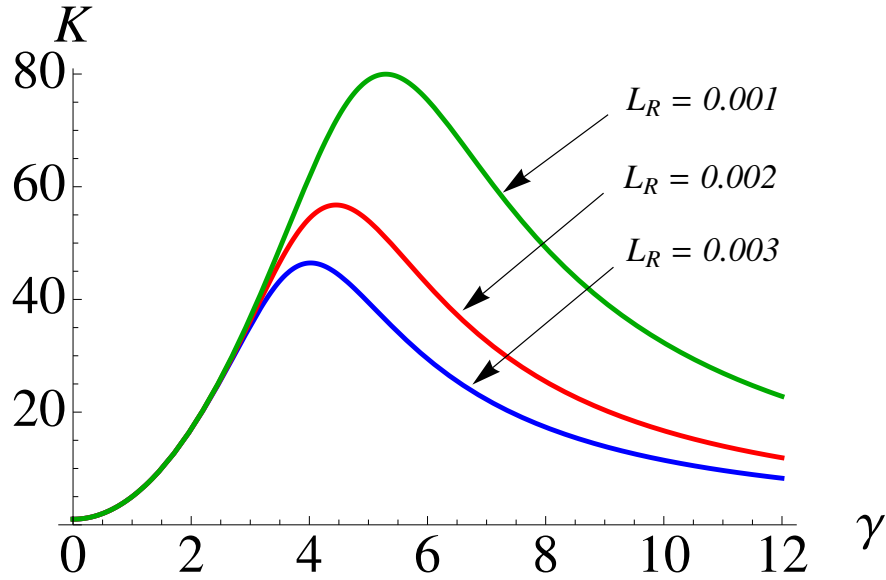
Figure 7.4: An example of a measurement bandwidth as a function of γ for three different values of L_R: 0.001, 0.002 and 0.003.

K_{gen} is defined in (7.12). Therefore in the optimal case we have $K = K_{\mathrm{gen}}/\sqrt{2}$.

We define short crystal lengths as $L_R \ll \pi/4\gamma^4$, for which the generation bandwidth is large, meaning that the measurement bandwidth is dominated by the image plane overlap of the detection modes with the pump. This gives a measurement bandwidth of

$$K \approx K_{\mathrm{ip}} = 1 + 4\gamma^2. \tag{7.22}$$

Note that this short crystal limit is characterised by an independence of K on the crystal length. In fact, it can be seen in figure 7.3 that the leftmost part of the measurement bandwidth curves is flat (for the $\gamma = 7$ curve this is not visible in this plot, but the slope of eq. (7.20) near the origin is zero for any γ), and that the range of values of L_R over which they stay flat is inversely proportional to γ^4. For much longer crystals, $L_R \gg \pi/4\gamma^4$, the measurement bandwidth, as modified by the limiting overlap in the far field, becomes dominant, giving

$$K \approx K_{\mathrm{FF}} = 1 + \frac{\pi}{L_R\gamma^2}. \tag{7.23}$$

In figure 7.4 we plot three different curves, that describe the value of the measurement bandwidth as a function of γ, for three different values of L_R. Note that for each choice of L_R there is always an optimal value of γ which maximises K, and it corresponds to the optimal value given in (7.21).

110

It is not an easy matter to determine the requisite parameters for existing experiments. Most of our own experiments, however, correspond to values of γ in the range 1.5 up to about 4. In order to achieve higher degrees of entanglement in OAM, corresponding to larger Schmidt number, our analyses suggest that it would be desirable to press towards higher values of γ.

7.6 Conclusions

We have shown that two parameters determine the OAM bandwidth for entangled states produced by parametric down-conversion. These parameters are the ratio of the widths of pump and detection modes $\gamma = w_p/w_{s,i}$, and the crystal thickness normalised to the Rayleigh range of the pump $L_R = L/z_r$.

A simple geometrical argument approximates the analytical results extremely well and allows us to suggest what needs to be adjusted in order to enhance the dimensionality of the entanglement. We have restricted our analysis to a detection system that is sensitive to the LG $p = 0$ modes only. It is for this reason that the measurement bandwidth can never reach that of the generation bandwidth for any combination of parameters. It is possible, however, to identify an optimum value of γ to maximise the measurement bandwidth for any normalised crystal length L_R. We would like to stress that we didn't employ the thin crystal approximation nor the gaussian approximation of the phase matching function.

Chapter 8

Sector phase masks approach to spiral bandwidth analysis

And now for something completely different

Monty Python, *Monty Python's Flying Circus*, 1969

8.1 Introduction

This chapter is based on the paper "Determining the dimensionality of bipartite orbital-angular-momentum entanglement using multi-sector phase masks", New J. Phys. **14**, 073046 (2012). This work was a joint effort of all the authors, and I do not take credit for the experimental activities.

The dimensionality of orbital angular momentum (OAM) entanglement produced in spontaneous parametric down-conversion can be probed by using multi-sector phase masks. In this chapter we show how it's possible to use a spatial light modulator to implement these analysers [86], and use them to measure a Schmidt number of about 50. As we saw in the introduction, the OAM of light is a property that is associated with phase structures of the form $e^{i\ell\phi}$, where $\ell\hbar$ is the OAM carried by each photon [6]. As ℓ is an integer and is theoretically unbounded, OAM offers a natural discrete space for exploring high-dimensional entanglement [65]. And as we saw in the previous chapters, in SPDC with a Gaussian pump, the spectrum of OAM correlations between the signal and idler photons is peaked at $\ell = 0$, with tails towards high $|\ell|$ values. The width and shape of the spectrum, and therefore the number of joint OAM modes that con-

stitute the two-photon entangled state, can be engineered directly by manipulating the structure of the beam pumping the crystal [20, 65, 77] or by tuning the phase-matching conditions in SPDC [87], which will be the solution adopted in the experimental procedure described here.

Although it was the central concern of chapter 7, it is important to stress once more the distinction between between the entanglement that is generated in SPDC, and the entanglement that is detected by the measurement stage of the system [24]. The effective number of entangled OAM states that are ultimately measured is in fact dependent upon both the generated down-converted OAM spectrum and the detection capabilities of our analysers. The Schmidt number K is introduced as a measure of the generated entanglement [21], while the Shannon dimensionality D represents a measure of the detection capabilities of the analysers [25]. The measured dimensionality M of the system generally depends on both K and D, via a relationship $M(K,D)$ that will be specified below. It may however be more difficult to measure K directly than it is to determine M and D. In those cases, we can first determine M (from a direct measurement) and D (which is fixed by the design of the phase masks), and invert the relation for M, from which the Schmidt number $K(M,D)$ can then be deduced.

Determining the measured dimensionality M of tailored high-dimensional entangled states can be carried out by performing appropriate selective projective measurements [57], i.e. different joint eigenstates can be singled out and detected, the rates of coincidence counts will then give the spiral bandwidth. This method, however is wasting many of the photons, as single eigenstates are being measured for each setting. An alternative method is based on pairs of multi-sector phase masks, placed in the two arms of a down-conversion system. Each mask has N azimuthal angular sectors, each of which introduces in an alternating way a π phase shift or no phase shift at all. The number and angular width of the sectors of the phase analysers placed in each of the signal and idler arms define the superposition of OAM eigenmodes in which the two-photon state produced by down-conversion is projected. It is clear that any design which is strictly azimuthal will not affect the radial components of a field. By optimising the binary phase profile of the phase analysers, it is possible to maximise the Shannon dimensionality D of the measurement apparatus, for any number of sectors N of the two masks [88]. The optimisation can be performed by writing D as a function of the set of $2N-1$ free angles that determine the position and size of each sector, and then by running a simulated annealing algorithm that maximises D by moving the $2N-1$ angular variables around the circle. We have

no certainty about the optimality of the results, but they are good enough to be successfully implemented. We will spend some words on that below.

By using angular phase analysers we infer the Schmidt number K, characterising the effective number of azimuthal entangled OAM modes. In contrast to previous works, which used quartz phase masks prepared with optical lithography [88], we can implement N-sector angular phase analysers using spatial light modulators (SLMs) [89]. Computer-controlled SLMs provide a fast, convenient and reliable way of producing holographic phase masks with arbitrary orientations and numbers of sectors, to be used in the measurement of the dimensionality of OAM entanglement. The use of multi-sector phase masks to probe high-dimensional states, as opposed to narrow single-sector analysers [27, 90], allows the measurements of tight angular correlations whilst maintaining high optical throughput.

8.2 Theory

8.2.1 Amplitude and phase masks

The angular measurement of a light field can be achieved by employing an angular slit. The idea behind this approach is the angular analog of a linear slit: a linear slit measures a field at one linear coordinate, with an uncertainty that depends on the width of the slit. An angular slit, on the other hand, measures a field at one angular coordinate with an uncertainty that depends on the angular width of the slit. As the angular position and orbital angular momentum form a pair of conjugate observables, the tighter the angular correlations, the larger the spread in the OAM observable [91, 92]. If one seeks high measurement accuracy, the angular slit has to be very narrow and, in turn, this means that much of the light is blocked, yielding the problem of lower number of counts as the measurement uncertainty is decreased. The solution to this problem is to employ phase masks, instead of amplitude masks.

The design of the phase mask consists of an angular step mask, which is characterised by a number of sectors N and by a set of $2N-1$ angles that describe the position and width of each sector [88]. Each alternate sector applies a π phase shift. The overall action of a phase mask of this kind is therefore to flip the phase of a light field in each sector, about the centre of the mask, and to leave the phase unchanged everywhere else. Note that the action of a phase mask does not affect the radial degree of freedom, as the design is radially invariant. This means that there is no coupling between different eigenmodes of the radial degree of

freedom, which allows us to restrict ourselves to the azimuthal content of the measured state. This effect can be described in terms of OAM. A plane wave is turned into a superposition of different OAM eigenstates. The range of OAM eigenstates of which the superposition consists depends on the number of sectors and on their relative positions and widths [88]. Such effect is analogous to the effect of an amplitude mask, without the drawback of letting less and less light through as the angular uncertainty is decreased.

8.2.2   Measurement of Hilbert space dimensionality

In order to optimise the azimuthal profile of a sector mask the most useful fact to keep in mind is precisely that we are trying to place the phase shifts so that the beams are localised within a narrow angle. If we indicate the azimuthal phase shift with a 2π periodic function $A(\theta, \alpha)$, where θ is the azimuthal angle and α is the angle of rotation of the whole phase mask with respect to a fixed position, the projection of a flat field that passed through a phase mask onto a second phase mask is given by

$$\gamma(\theta_1, \theta_2) = \int_0^{2\pi} A(\theta_1, \alpha) A^*(\theta_2, \alpha) d\alpha \tag{8.1}$$

The symmetry of the problem gives $\gamma(\theta_1, \theta_2) = \gamma(\theta_1 - \theta_2, 0)$. We have therefore a function that indicates how much of the light that passed through the first phase mask passes through a second phase mask, oriented at an angle with respect to the first one. The optimisation procedure involves precisely minimising the "area under the curve", i.e. by making the localisation more peaked at one specific location.

This is the idea of how we proceed: with a pair of optimised identical phase masks we can probe the OAM content of the SPDC state by putting one on each "arm" of the system and then detecting the pairs in coincidence while rotating one of the phase masks. The area under the curve will be a consequence of both the design of the mask and of the OAM content. What's fundamental here is that the design (i.e. the function A) can be represented in Fourier space, and each coefficient simply multiplies the relative OAM coefficient of the SPDC state to deliver the coincidence curve. Hence, knowing one of the two "widths", given by the design, D, the other (the Schmidt number K) can be immediately retrieved.

In other words, we start from an SPDC state

$$|\Psi\rangle = \sum_{\ell=-\infty}^{\infty} c_\ell |\ell\rangle_s \otimes |-\ell\rangle_i \tag{8.2}$$

described exclusively in the OAM basis, where $|\ell\rangle_s$ and $|\ell\rangle_i$ correspond to the states of the signal and idler photons respectively. By expressing the projection state associated with a phase-mask analyser oriented at an angle α as the superposition

$$|A(\alpha)\rangle = \sum_\ell \lambda_\ell |\ell\rangle \, e^{i\ell\alpha}, \tag{8.3}$$

the coincidence probability for a pair of identical analysers, oriented at α and β respectively, is given by

$$P(\alpha, \beta) = \left| \langle A(\alpha), A^*(\beta) | \Psi \rangle \right|^2, \tag{8.4}$$

where $\langle A(\alpha), A^*(\beta) | = \langle A(\alpha)| \otimes \langle A^*(\beta)|$. The coefficients $\gamma_\ell = |\langle \ell | A(0)\rangle|^2 = |\lambda_\ell|^2$ (with $\sum_\ell \gamma_\ell = 1$), defined by the profile of the N-sector phase masks, determine the respective OAM spectrum. It is possible to design the arc sectors of each N-sector phase mask in a way that maximises the dimensionality D of the analyser, which is given by

$$D = \frac{1}{\sum_\ell \gamma_\ell^2} \tag{8.5}$$

much alike the Schmidt number. For each N we used the optimal arrangement of sectors, based on results from [88]. The maximum number of modes D that can in principle be measured by each of such optimal N-sector masks can then be inferred from the theoretical distribution of eigenmodes γ_ℓ, as shown in [88].

The maximal dimensionality D of each N-sector mask used here was found through a numerical model. It was found that D increases linearly with N, in particular it was found that adding a pair of $\pi/0$ sectors can lead to a different optimal design that adds about 6 extra modes to the dimensionality D. The numerical simulation considered the distribution of the overlap between two identical N-sector masks within a two-dimensional region with a Gaussian profile, as the orientation β of one of the masks was rotated with respect to the other, α. Given the coincidence probability distribution $P(\alpha - \beta)$ obtained from the numerical model, the Shannon dimensionality can also be directly calculated as

$$D = \frac{2\pi}{\int_0^{2\pi} P(\alpha - \beta) \, d(\alpha - \beta)}, \tag{8.6}$$

where angle β is measured with respect to α. We implemented multi-sector masks with a number N of sectors between 1 and 16. The dimensionalities D obtained from the numerical and theoretical models, the latter obtained from the
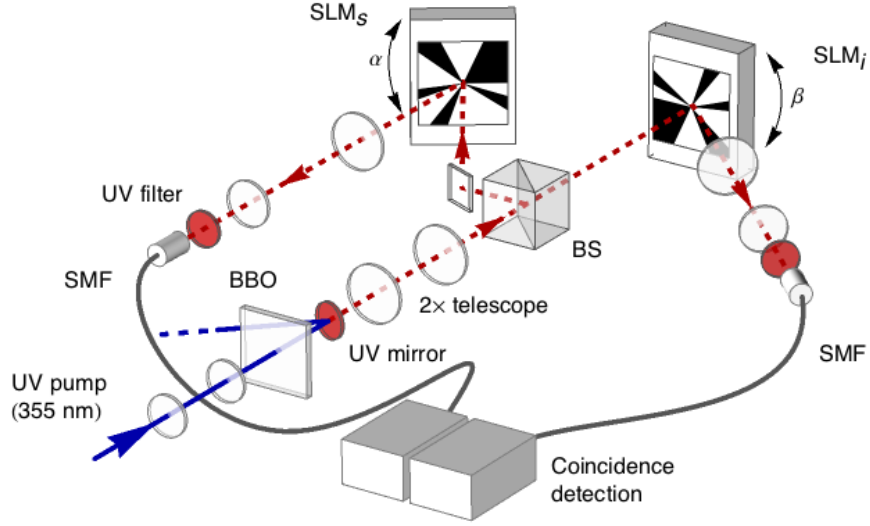
Figure 8.1: Experimental setup implementing SLM-based coincidence detection with angular-sector phase masks (see text for details).

decomposition of each N-sector mask into eigenmodes γ_ℓ, were found to differ by less than 2.5%, meaning that SLMs do indeed deliver a reliable representation of a phase mask.

8.3 Experimental results

8.3.1 Experimental setup

Implementing phase-mask analysers using computer-controlled spatial light modulators allows for quick and effective measurement of the Schmidt number K of the entangled state produced by SPDC. No optical elements need to be fabricated, physically rotated or replaced when using a different multi-sector mask, as the measurement process simply involves displaying one of a set of different rotated N-sector holograms on the SLMs and performing coincidence detection.

A 5 mm-thick β-barium borate (BBO) non-linear crystal cut for type-I collinear SPDC acts as our source of entangled photon pairs. The crystal is pumped by a 1 W UV laser to produce frequency-degenerate entangled photon pairs at 710 nm. The co-propagating signal and idler photons are separated by a non-polarizing beam splitter, and redirected to SLMs. The SLMs, onto which the crystal output face is imaged by a 2× telescope, are encoded with N-sector phase holograms. The SLMs are then imaged onto single-mode fibers (SMFs), which couple the SLM output to single-photon photo-diode detectors (figure 8.1), whose output is
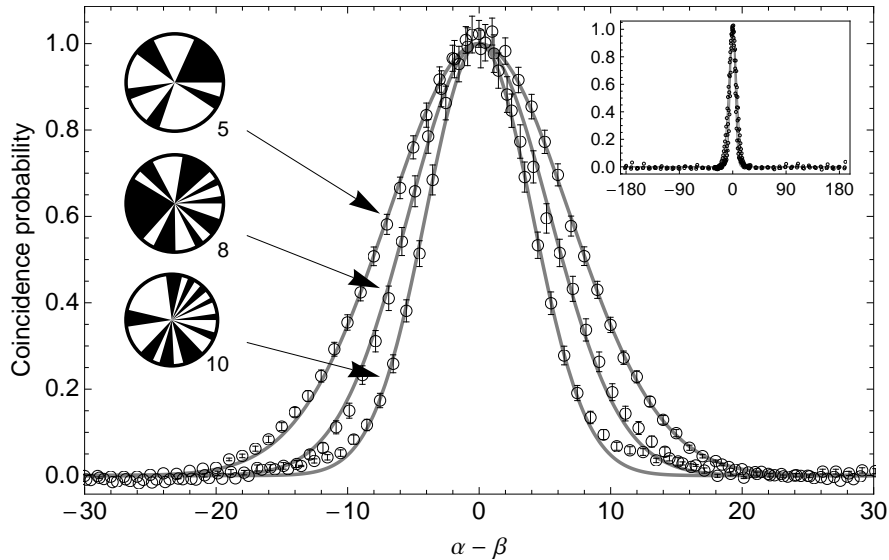
Figure 8.2: Typical best Gaussian fits of coincidence probability distributions $P(\alpha - \beta)$ are shown for $N = 5, 8, 10$ and $\phi = 0$. The inset shows the full $-180° < \alpha - \beta < 180°$ range used in the measurements, for the aforementioned values of N. Background subtraction was performed by assessing the experimental accidental coincidences.

routed to coincidence-counting electronics. The coincidence counting has a timing window of 10 ns. Narrow-band, 2 nm interference filters are placed in front of the detectors to ensure that the frequency spread of the detected down-converted fields is small compared to the central frequencies. SLMs introduce great flexibility in our measurements, but this comes at the price of an overall lower detection efficiency, as the diffraction efficiency of SLMs is around 50 %.

This detection configuration is insensitive to any overall phase factors. Therefore, while the conservation of OAM in the SPDC process would require placing mutually phase-conjugate N-sector phase masks in the detection arms (i.e., the 0 and π phase-shifted sectors are inverted between the two masks), two identical phase masks can be used instead. These phase masks are self-conjugate in case of a π phase shift. The finite pixel size of the SLMs places a restriction on the width of the sectors that can be displayed on the holograms. We show that we can implement optimal multi-sector phase masks with $N = 1, \ldots 16$. Suppressing the centres of the holograms, where the N angular sectors meet in a very limited spatial region of the SLM displays, did not turn out to be necessary.

The phase-matching conditions of the down-conversion process for the BBO crystal were adjusted by slightly changing the orientation of the crystal with respect to the propagation direction of the pump beam [26]. This allowed to in-
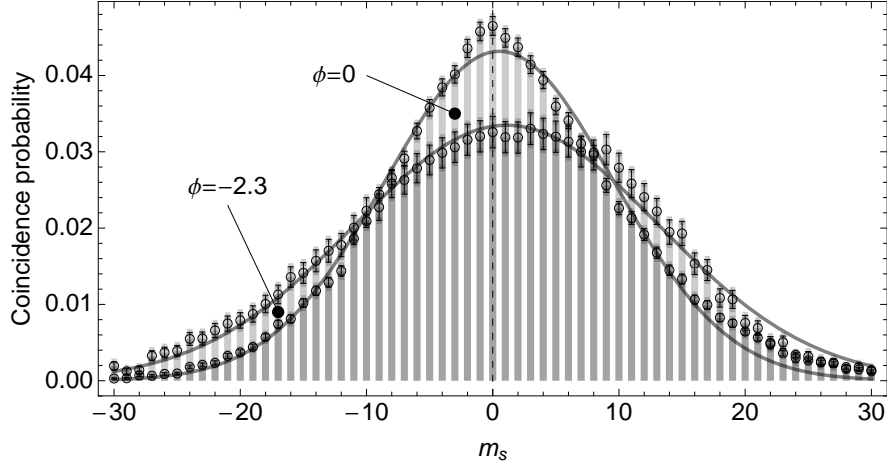
Figure 8.3: Detected orbital angular momentum spectrum with projective measurements. Shown is the coincidence probability $P(\ell_i, \ell_i = -\ell_s)$ for $\ell_s \in [-30, 30]$, for collinear ($\phi = 0$) and near-collinear ($\phi = -2.3$) phase-matching conditions.

crease the width of the orbital angular momentum spectrum, and thus decrease the width of the angular correlations distribution. The intensity profile of the down-conversion emission can be expressed as: $I \propto |E|^2 \propto \text{sinc}(\phi + c\xi)^2$, where ξ is the external emission angle in air, $c = (|\mathbf{k}|_s + |\mathbf{k}_i|)L/(2n)^2$ is a constant depending on the experimental parameters, and $\phi = (|\mathbf{k}_p| - |\mathbf{k}_s| - |\mathbf{k}_i|)L/2$ (with L length of the crystal) determines the degree of non-collinearity of the process [88]. Measurements were performed for collinear ($\phi = 0$) and near-collinear ($\phi = -2.3$) phase-matching conditions.

The coincidence probability distribution $P(\alpha - \beta)$ was obtained by changing the orientation β of the second phase mask over the range $\alpha \pm 180°$, where α is the orientation of the first. We found that, for the purposes of the experiment, a Gaussian distribution is an excellent empirical fit for the coincidence probability distributions (figure 8.2). The detected number of modes M, dependent on both the source and the detectors' properties, was obtained by substituting the Gaussian fit of the measured coincidences to $P(\alpha - \beta)$ in eq. 8.6.

8.3.2   Results and discussion

The phase masks used have no radial structure. Any sensitivity of the detection apparatus to the radial quantum number p is therefore due to the spatial selectivity of the SMF coupling. Unlike full projective measurements of the down-conversion entangled state over a range of OAM eigenstates $|\ell\rangle$, a phase-mask determination of the dimensionality K of the source does not provide any direct

information on the shape of the OAM spectrum.

We consider the OAM spectrum generated by the source and the coincidence probability obtained from the numerical model, both fitted with Gaussian distributions, from which the dimensionalities K and D can respectively be obtained. We performed projective measurements of idler and signal over $|\ell\rangle$ and $|-\ell\rangle$ respectively to verify the validity of the assumption concerning the OAM spectrum (figure 8.3). The numerical model, from which the values of D for different N are obtained, calculates the overlap in eq. 8.4. Although the theoretical shape of the distribution of the coefficients γ_ℓ, which characterise the action of a phase mask on a light field, has own distinctive features, its implementation on an SLM makes it possible to approximate it with a Gaussian distribution. In fact, as the hologram representing the phase mask is rotated on the surface of the SLM, any imperfections (finite pixel size, surface roughness, slight unevenness of the phase shift, electrical fluctuations) influence the effect of the phase mask in a stochastic fashion. As this a very small overall effect, it does not change the dimensionality of the phase mask but rather smooths out the distribution of the γ_ℓ coefficients.

The measured dimensionality M is given by the normalised convolution of the two other distributions (namely, the coincidence distribution associated with the effective Schmidt number of the system and that obtained from the numerical model for the sector phase masks), as the coincidence probability depends on the properties of both the source and the detectors. M can then be approximated by the following formula, valid in the case of two gaussian distributions:

$$M \simeq \frac{DK}{\sqrt{D^2 + K^2}}, \tag{8.7}$$

from which the Schmidt number K can be derived:

$$K \simeq \frac{DM}{\sqrt{D^2 - M^2}}. \tag{8.8}$$

Consequently, the source dimensionality K can be inferred from the theoretical coincidence probabilities and measured coincidence distribution. The calculated dimensionality for each N-sector phase mask is shown in figure 8.4 (collinear phase-matching) and 8.5 (near-collinear phase-matching). We measured Schmidt numbers of 35 ± 2 for collinear down-conversion, and 49 ± 2 in the near-collinear case. The results for source dimensionality obtained from the phase-mask analysis are compatible with the Schmidt number derived from projective spiral bandwidth measurement (figure 8.3), with the assumption of per-
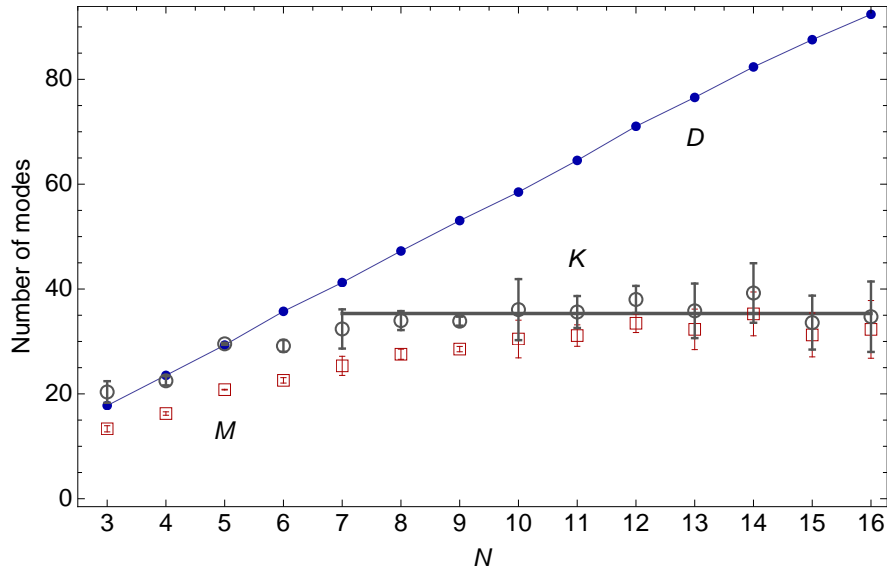
Figure 8.4: *Collinear phase-matching conditions*. Phase mask dimensionality D from numerical model (blue points), measured M (red squares), and calculated dimensionality K (gray circles). We observe that $D < K$ for $N < 7$; therefore, the calculated M saturates to D for any given $N < 7$. The solid gray line shows the best estimation for the number of modes K of the source, 35 ± 2.

fect single-mode detection.

The mean visibility achieved in the experiment, defined here as the ratio between the mean baseline of the measured coincidence probability and the peak at $\alpha - \beta = 0$, without background subtraction, is 90 % for collinear phase-matching, and 92 % for near-collinear. Systematic errors due to misalignment are found to be much larger than photon statistics uncertainties.

8.4 Conclusions

In conclusion, we have shown how multi-sector phase-mask analysers can be implemented using spatial light modulators, and used them to probe the effective number of modes in the high-dimensional bi-photon entangled state produced by parametric down-conversion. We used a set of several multi-sector analysers to infer the Schmidt number for different phase-matching conditions, and therefore, different widths of the OAM spectrum of the source.
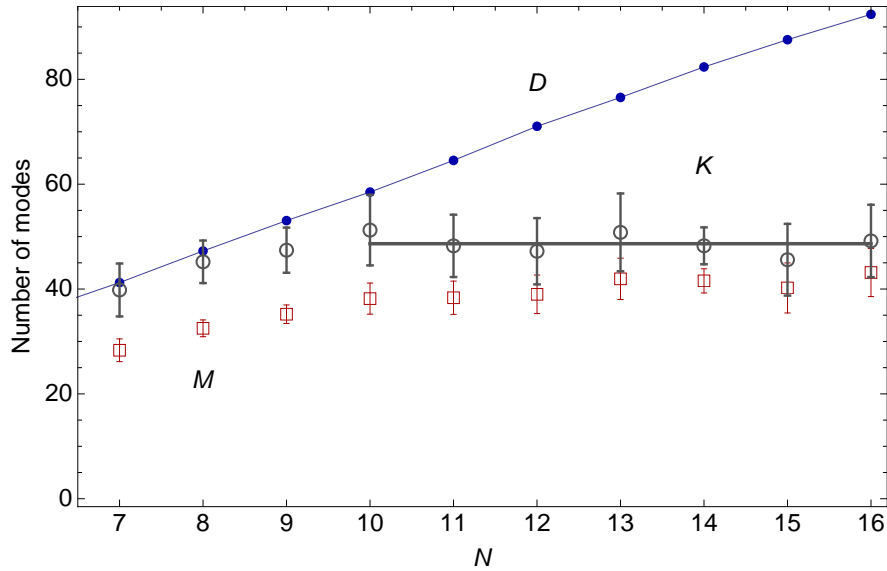
Figure 8.5: *Near-collinear phase-matching conditions.* See caption of figure 8.4. Given the wider spiral bandwidth in the near-collinear regime, K saturates to D for $N < 9$. The solid gray line shows the best estimation for the number of modes K of the source, 49 ± 2.

8.5 Appendix: optimisation of phase masks

For completeness we add this appendix on the optimisation of sector phase plates. The material in this appendix is taken and adapted from my master thesis "High dimensional photon entanglement and the design of sector phase plates" [1], completed in Leiden University in 2009.

8.5.1 Description of a phase mask

A phase mask is an object which purpose is to shift the phase of an incoming wavefront. In a quartz phase mask the shift will happen due to different path lengths in the medium, that will slow down the propagation of the wave, in an SLM-reproduced phase mask, the orientation of the liquid crystals are responsible for the phase shift. The design on the surface will be radially invariant, as we want to produce purely azimuthal phase shifts.

The plane wave that approaches the phase mask will have a different shift according to the point where it will enter, either 0 or π.
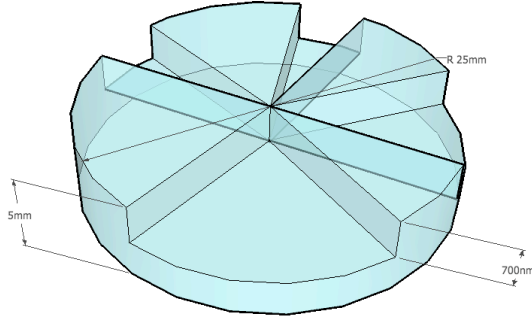
Figure 8.6: Quartz phase mask with greatly exaggerated slice thickness. The real thickness is around $800nm$, while the total thickness is around $2mm$. The diameter is $2.5cm$.

8.5.2 Optimisation procedure

The problem of optimising a phase mask has been addressed before [25], but the results were not optimal. The process can be split into an analytical part and a numerical part: although an analytical expression for the Shannon number D for any number of sectors is found, its full expression is too complicated to work with, therefore a numerical approach is the fastest way to retrieve a solution.

Here we describe the details of a single sector mask, the generalisation will be straightforward. We want to show a way to optimise a phase mask of N sectors. By optimisation we mean the maximisation of the measure that corresponds to the number of modes of OAM that a phase mask can reach, namely the Shannon number D. There is therefore an optimal way of choosing the angles between the sectors, and their size too that will maximise the Süssmann measure of the distribution of the γ_m factors, defined below. We start by describing a general sector phase mask.

Given the angle β of the sector, we define the normalised profile function

$$A_\beta(\theta, \alpha) = \frac{1}{\sqrt{2\pi}} t_\beta(\theta, \alpha) \tag{8.9}$$

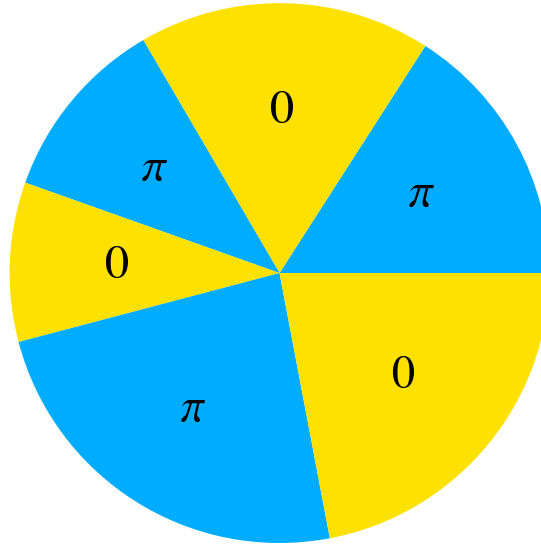where $t_\beta(\theta, \alpha)$ is a sum of Heaviside theta functions that jumps between +1 to

123

Figure 8.7: Example of 3-sectors phase mask. In this representation we show the phase shift that different parts of the mask will imprint on the wave front. Note that there is no radial dependence: the phase shifts are a purely angular feature of the design of the mask.

-1 when the variable θ becomes equal to the parameter α and jumps back to +1 when $\theta = \alpha + \beta$. The function t is periodic in 2π: we have

$$t_\beta(\theta, \alpha) = H[(\theta - \alpha)\mathrm{mod}\ 2\pi] - H[(\theta - \alpha)\mathrm{mod}\ 2\pi - \beta] \tag{8.10}$$

For the rotational nature of the problem, all the angles can be shifted together by the same amount without giving any change in the final result of our computation. This fact allows us to set the parameter α at will, and in particular we have that $A(\theta, \alpha) = A(\theta - \alpha, 0)$. Such a sum of distributions describes how the phase of a wavefront gets shifted according to the position and size of the sectors in a sector phase mask. Define the function $\gamma(\theta_1, \theta_2)$:

$$\gamma(\theta_1, \theta_2) = \int_0^{2\pi} A(\theta_1, \alpha)A^*(\theta_2, \alpha)\,d\alpha \tag{8.11}$$

This function represents the angular correlation of two superposed phase masks, each one rotated by an angle θ_i, with $i = 1, 2$. This function can be used to study the angular localisation of a phasemask with respect to its twin. For the problem that we need to solve, the function γ is a hermitian Hilbert-Schmidt kernel (it is continuous in a compact interval), positive semidefinite. Therefore it is possible
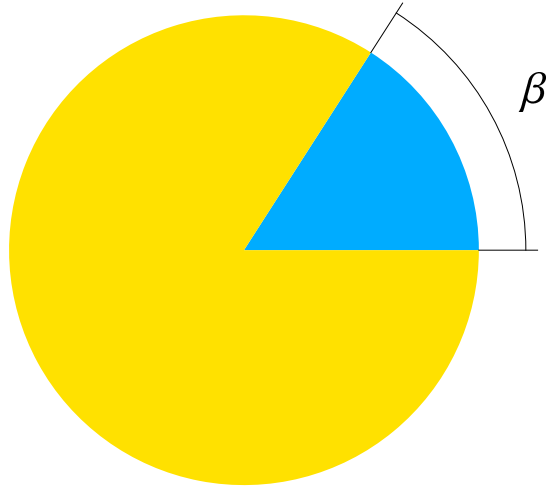
124

Figure 8.8: The angle β that appears in the description below is represented in this picture. Notice that the first edge of the first sector corresponds to an angle $\phi_0 = 0$.
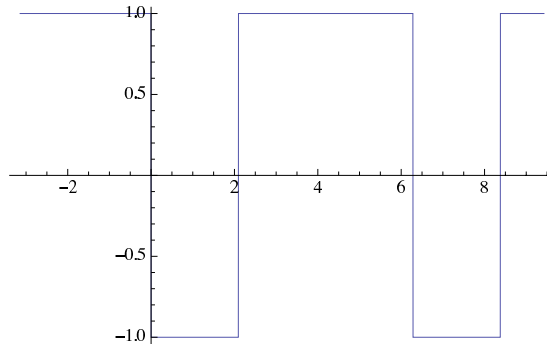

Figure 8.9: The function $t_\beta(\theta, 0)$ varying with respect to θ. The value of α determines the overall position, here is set to 0, the value of β determines for how long the function stays in -1.

to choose a complete othonormal basis $\{u_m(\theta)\}$ and expand $\gamma(\theta_1, \theta_2)$ in this basis:

$$\gamma(\theta_1, \theta_2) = \sum_\ell \gamma_\ell u_\ell(\theta_1) u_\ell^*(\theta_2) \tag{8.12}$$

As we are interested in the orbital angular momentum decomposition we will choose the OAM eigenfunctions as a basis for the expansion, $u_\ell(\theta) = \frac{1}{\sqrt{2\pi}} e^{i\ell\theta}$. The eigenvalues γ_ℓ are found by solving the Fredholm (because γ defines a compact operator) integral equation [24]:

$$\int_0^{2\pi} \gamma(\theta, \theta') u_\ell(\theta') d\theta' - \gamma_\ell u_\ell(\theta) = 0 \tag{8.13}$$

125

We solve it by substituting the definition of $\gamma(\theta, \theta')$ into the equation, changing the variable from θ' to $x = \theta' - \alpha$ as well as $y = \theta - \alpha$:

$$\gamma_\ell = u_\ell^{-1}(\theta) \frac{1}{2\pi} \int_0^{2\pi} A(\theta, \alpha) e^{i\ell\alpha} \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} A^*(x, 0) e^{i\ell\theta'} d\theta' \, d\alpha \tag{8.14}$$

$$= \sqrt{2\pi} e^{-i\ell\theta} \mathscr{F}^*[A(x, 0)](\ell) \frac{1}{2\pi} \int_0^{2\pi} A(\theta, \alpha) e^{i\ell\alpha} d\alpha \tag{8.15}$$

$$= e^{-i\ell\theta} \mathscr{F}^*[A(x, 0)](\ell) \frac{1}{\sqrt{2\pi}} e^{i\ell\theta} \int_0^{2\pi} A(y, 0) e^{-i\ell y} dy \tag{8.16}$$

$$= \mathscr{F}^*[A(x, 0)](\ell) \mathscr{F}[A(y, 0)](\ell) = |\mathscr{F}[A](\ell)|^2 \tag{8.17}$$

Where the $\mathscr{F}$ denotes the Fourier Transform. Here we showed that the expansion coefficients are indeed the Fourier transforms of the normalised profile function A, evaluated in the discrete modes ℓ. We give a new name to the coefficients, for convenience of notation: $\mathscr{F}[A](\ell) = c_\ell$.

The eigenvalues γ_ℓ represent the coupling strength of the analyser (the phase mask) to the field mode $u_\ell(\theta)$. The set includes the natural normalisation condition $\sum_{\ell=-\infty}^{\infty} \gamma_\ell = 1$.

We have shown that $\gamma_\ell = |\mathscr{F}[A](\ell)|^2 = |c_\ell|^2$, we now calculate the coefficients c_ℓ. For $\ell \neq 0$, using 8.9 and 8.10

$$c_\ell = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} A(\theta, 0) e^{-i\ell\theta} d\theta \tag{8.18}$$

$$= \frac{1}{2\pi} \left[-\int_{\phi_0=0}^{\phi_1} e^{-i\ell\theta} d\theta + \int_{\phi_1}^{\phi_2} e^{-i\ell\theta} d\theta - \int_{\phi_2}^{\phi_3} e^{-i\ell\theta} d\theta + \cdots + \int_{\phi_{2N-1}}^{2\pi} e^{-i\ell\theta} d\theta \right] \tag{8.19}$$

$$= \frac{1}{2\pi} \left[-\frac{e^{i\ell\phi_1} - e^{i\ell\phi_0(=0)}}{-i\ell} + \frac{e^{i\ell\phi_2} - e^{i\ell\phi_1}}{-i\ell} - \frac{e^{i\ell\phi_3} - e^{i\ell\phi_2}}{-i\ell} \cdots + \frac{e^{i\ell2\pi} - e^{i\ell\phi_{2N-1}}}{-i\ell} \right] \tag{8.20}$$

$$= \frac{i}{2\pi\ell} \left[1 + e^{i\ell2\pi} - 2\left(e^{-i\ell\phi_1} - e^{-i\ell\phi_2} + e^{-i\ell\phi_3} - \cdots + e^{-i\ell\phi_{2N-1}} \right) \right] \tag{8.21}$$

$$= \frac{i}{\pi\ell} \left[1 + \sum_{n=1}^{2N-1} (-1)^n e^{-i\ell\phi_n} \right] \tag{8.22}$$

so we can multiply the Fourier transform of the profile by its complex conjugate (for the passage to the second line remember that we set $\phi_0 = 0$. We now add the

subscript N to stress the dependence to the total number of sectors):

$$\gamma_{\ell,N} = \frac{1}{\pi^2 \ell^2} \left[1 + \sum_{n=1}^{2N-1} (-1)^n e^{-i\ell\phi_n}\right] \left[1 + \sum_{n=1}^{2N-1} (-1)^n e^{i\ell\phi_n}\right] \tag{8.23}$$

$$= \frac{1}{\pi^2 \ell^2} \left[\sum_{n=0}^{2N-1} (-1)^n e^{-i\ell\phi_n}\right] \left[\sum_{n=0}^{2N-1} (-1)^n e^{i\ell\phi_n}\right] \tag{8.24}$$

$$= \frac{1}{\pi^2 \ell^2} \sum_{n,k=0}^{2N-1} (-1)^{n+k} e^{-i\ell(\phi_n - \phi_k)} \tag{8.25}$$

$$= \frac{1}{\pi^2 \ell^2} \sum_{n,k=0}^{2N-1} (-1)^{n+k} \cos[\ell(\phi_n - \phi_k)] \tag{8.26}$$

This last passage is justified by the fact that we are summing over all possible pairs of $\{n,k\}$, so we get a factor 1/2 because we need only half of the terms and a factor 2 from the cosine representation of the real part, that cancel out. If one wants to plot the distribution of the γ_ℓ coefficients (that are equal to $|c_\ell|^2$), this expression is not computationally convenient, because given N sectors, for each γ_ℓ factor one has to evaluate and sum $4N^2$ terms. There is a way to express the same coefficient evaluating and summing $4N$ factors, as we will show now. If we write the cosine of the difference that appears in eq 8.26, we obtain:

$$\gamma_\ell = \frac{1}{\pi^2 \ell^2} \sum_{n,k=0}^{2N-1} (-1)^{n+k} \cos[\ell(\phi_n - \phi_k)] \tag{8.27}$$

$$= \frac{1}{\pi^2 \ell^2} \sum_{n,k=0}^{2N-1} (-1)^{n+k} [\cos(\ell\phi_n)\cos(\ell\phi_k) + \sin(\ell\phi_n)\sin(\ell\phi_k)] \tag{8.28}$$

$$= \left(\sum_{n=0}^{2N-1} \frac{(-1)^n \cos(\ell\phi_n)}{\ell\pi}\right)^2 + \left(\sum_{n=0}^{2N-1} \frac{(-1)^n \sin(\ell\phi_n)}{\ell\pi}\right)^2 \tag{8.29}$$

For the last passage notice that the double sum is on two different indices. Now that we have introduced two expressions for the γ_ℓ factors, we introduce the Süssmann measure as a measure of width of the distribution of modes generated by a phase mask. It has a profound similarity with the Schmidt number.

8.5.3   Süssmann measure

As we mentioned above, the dimensionality value associated to a given sector phase mask is the Shannon number of distinct eigenstates that such a mask

could address and is defined as [24]:

$$D_N = \frac{1}{\displaystyle\sum_{\ell=-\infty}^{\infty} \gamma_{\ell,N}^2} \tag{8.30}$$

so we need to calculate the square of the γ_ℓ factors. To do it we keep the first form 8.26, as it contains a single occurrence of the cosine, although the sum is performed on two indices.

The square of a sum can be written by doubling the indices and squaring the argument:

$$\gamma_{\ell,N}^2 = \frac{1}{\pi^4} \sum_{n,k,n',k'=1}^{2N-1} (-1)^{n+k+n'+k'} \frac{\cos[\ell(\phi_n - \phi_k)]\cos[\ell(\phi_{n'} - \phi_{k'})]}{\ell^4} \tag{8.31}$$

We can perform the sum over ℓ first, as there is a convenient closed form for it. We start by cutting the sum in half and keeping only the terms with positive ℓ, as they are symmetric about the value $\ell = 0$:

$$\sum_{\substack{\ell=-\infty \\ m \neq 0}}^{\infty} \frac{\cos[\ell(\phi_n - \phi_k)]\cos[\ell(\phi_{n'} - \phi_{k'})]}{\ell^4} \tag{8.32}$$

$$= 2\sum_{\ell=1}^{\infty} \frac{\cos[\ell(\phi_n - \phi_k)]\cos[\ell(\phi_{n'} - \phi_{k'})]}{\ell^4} \tag{8.33}$$

Let's take a simple version of the sum we are about to perform:

$$\sum_{\ell=1}^{\infty} \frac{\cos \ell x}{\ell^4} \tag{8.34}$$

This sum converges if $-2\pi \leq x \leq 2\pi$ to a fourth degree polynomial in x:

$$\sum_{m=1}^{\infty} \frac{\cos(\ell x)}{\ell^4} = \frac{\pi^4}{90} - \frac{\pi^2 x^2}{12} + \frac{\pi|x|^3}{12} - \frac{x^4}{48} \tag{8.35}$$

We cannot apply it as straight away to the sum 8.31, because of the double cosine. But by the trigonometric identities we can write

$$\frac{\cos[\ell(\phi_n - \phi_k)]\cos[\ell(\phi_{n'} - \phi_{k'})]}{\ell^4} = \frac{\cos[\ell(\phi_n - \phi_k + \phi_{n'} - \phi_{k'})] + \cos[\ell(\phi_n - \phi_k - \phi_{n'} + \phi_{k'})]}{2\ell^4}$$

$$\tag{8.36}$$

So now we have two sums, each one is written in the form 8.34. The domain of the sums, though, is different: now the angles can range between -4π and 4π, so

we need to make the argument periodic by hand.

That is accomplished by using modulus functions ($x \to x \mod 2\pi$) so that the sum 8.34 can be made valid for every x:

$$\sum_{\ell=1}^{\infty} \frac{\cos(\ell x)}{\ell^4} = \frac{\pi^4}{90} - \frac{\pi^2 (x \mod 2\pi)^2}{12} + \frac{\pi |x \mod 2\pi|^3}{12} - \frac{(x \mod 2\pi)^4}{48} \qquad (8.37)$$

so that the original sum can be written in the following form:

$$2 \sum_{\ell=1}^{\infty} \frac{\cos[\ell(\phi_n - \phi_k)]\cos[\ell(\phi_{n'} - \phi_{k'})]}{\ell^4} \qquad (8.38)$$

$$= \left(\frac{\pi^4}{90} - \frac{\pi^2 \chi_+^2}{12} + \frac{\pi \chi_+^3}{12} - \frac{\chi_+^4}{48} \right) + \left(\frac{\pi^4}{90} - \frac{\pi^2 \chi_-^2}{12} + \frac{\pi \chi_-^3}{12} - \frac{\chi_-^4}{48} \right) \qquad (8.39)$$

where the χ_+ and χ_- are the periodicised sums of the angles:

$$\chi_\pm = (\pm\phi_n \pm \phi_k \mp \phi_{n'} \mp \phi_{k'}) \mod 2\pi \qquad (8.40)$$

Since we have to perform a sum over all possible values of the indices of the angles, every permutation of every set of choices is present in the sum, so it is possible to take only one of the two parts of eq (18) and multiply by 2:

$$D_N^{-1} = \frac{1}{\pi^4} \sum_{n,k,n',k'=0}^{2N-1} (-1)^{n+k+n'+k'} \left(\frac{\pi^4}{45} - \frac{\pi^2 \chi_+^2}{6} + \frac{\pi \chi_+^3}{6} - \frac{\chi_+^4}{24} \right) + \gamma_{0,N}^2 \qquad (8.41)$$

$\gamma_{0,N}^2$ is the $\ell = 0$ term, which is trivial to calculate: we just need to set ℓ to zero in equation 8.18:

$$\gamma_{0,N}^2 = \left(1 + \frac{1}{\pi} \sum_{n=1}^{2N-1} (-1)^n \phi_n \right)^4 \qquad (8.42)$$

Equation 8.41 is usable by numerical optimisation software until a maximum number of variables around 15, with more variables the software is not efficient anymore and the result is not optimal.

It is possible to simplify it though, allowing a faster optimisation and the handling of a higher number of variables.

8.5.4 How to simplify the computational optimisation

By separating the sum of the polynomial in four sums of the four parts it's easy to notice that it's possible to eliminate the factor $\frac{\pi^4}{45}$ because the sum of $(-1)^{n+k+n'+k'}$ is equal to zero.

Then let's take what we're left with and let's concentrate on the set of the four indices. Since all possible combinations are present, and since some of them produce a null result, or the same result, or pairs of them give opposite results that cancel out, we can reduce the set.

We need to look at the symmetries of the function

$$\left(\frac{\pi^2 \chi_+^2}{6} - \frac{\pi \chi_+^3}{6} + \frac{\chi_+^4}{24} \right) = \tag{8.43}$$

$$= \frac{1}{24} \chi_+^2 (2\pi - \chi_+)^2 = \frac{\chi_+^2 \chi_-^2}{24} \tag{8.44}$$

So that equation 8.41 becomes

$$D_N^{-1} = \frac{1}{24\pi^4} \sum_{n,k,n',k'=0}^{2N-1} (-1)^{n+k+n'+k'} \chi_+^2 \chi_-^2 + \gamma_{0,N}^2 \tag{8.45}$$

It is much easier now to reduce the set $\{\{n,k,n',k'\}|n,k,n',k' \in 1,2,\ldots 2N-1\}$ of all the possible combinations of the indices. Below we list a series of rules that we can use to eliminate single occurrences:

1. the elements $\{a,a,a,a\}$ give 0

2. the elements $\{a,b,a,b\}$ give 0

3. if we swap the first and second index, or the third and the fourth or both, the result is the same. So the elements $\{a,b,c,d\}$, $\{b,a,c,d\}$, $\{a,b,d,c\}$ and $\{b,a,d,c\}$ give the same result

4. if we swap the first *and* the second with the third and fourth, the result is the same. So the elements $\{a,b,c,d\}$ and $\{c,d,a,b\}$ give the same result

5. if two indices cancel out, the result is independent on their specific value, so the elements $\{a,b,c,b\}$, $\{b,a,b,d\}$, $\{a,b,b,d\}$ and $\{b,a,d,b\}$ (for each of them we can also apply rules 3 and 4) not only give the same result, but also are independent on the value of b.

By using these rules we get from a number of 256 possible combinations for a given set of four different values to a set of 27 (per set of four different indices, even less if the set contains less different values). Defining a function $Z(a,b,c,d)$ that counts the multiplicity of the same occurrences (accordingly with rules 3, 4 and 5), then we can multiply the argument of the sum by the function Z and we can sum only over *different* indices ($\sum_{n,k,n',k'} \rightarrow \sum_{n<k<n'<k'}$). This can computationally be accomplished, and the huge speedup that we obtain comes from the

fact that now, instead of $16N^4$ terms in the sum for an N sectors mask, we have 216N terms. When $N \geq 3$ the number of sets of indices is larger than 1 and it's possible to apply the rules above to eliminate terms that belong to different sets. This process reduces the size of the function D_N^{-1} and reduces the computational complexity of the optimisation process. The optimisation procedure that the software applies (a customised simulated annealing algorithm) becomes much faster than before, especially for a large number of sectors. The simulated annealing algorithm used starts 50 evaluation points at the same time in the variables space and moves them like they are particles in a potential, according to the value of a parameter that simulates the temperature during the cooling of a piece of material. Suppose that a point is stuck in a local minimum, then if the temperature is high enough it could (statistically) hop out and proceed for a better minimum. This process eventually ends when the temperature is low enough to consider sufficiently unlikely that a jump could occur. The results of simulated annealing are not guaranteed to be optimal, as the search tree is not completely probed, but the speed and overall quality of the results are satisfactory enough. Without the simplification we were able to find the optimal mask for $N = 9$ in almost three days, while with the simplification we could find the optimal mask for $N = 10$ in less than three minutes.

By means of equation 8.45 we were able to calculate the optimal set of angles for phase masks with up to 16 sectors.

Chapter 9

Conclusions and Outlook

9.1 Summary of the results

In the course of this thesis several results were obtained. The structure of the thesis reflects the way such results can be grouped together, namely we have two main results about the SPDC state and three main results on the detection of entanglement in the SPDC state. The two main results on the SPDC state are the analytical Schmidt decomposition, which is described in chapter 4, and the numerical Schmidt decomposition, which is described in chapter 5. The difference between these two forms of the Schmidt decomposition, a part from their analytical or numerical nature, is that the former employs the gaussian approximation of the phase matching function (a consequence of which is the impossibility of including a phase mismatch), while the latter makes use of the most general form of the phase matching function.

The main results on the detection of entanglement in SPDC are two sets of projections, either on a complete LG family (chapter 6), or on $p = 0$ LG modes (chapter 7), and a different approach of measurement, namely with sector phase masks (chapter 8). The difference between the two chapters on projection, a part from the type of LG modes, is that the former employs the thin crystal approximation, while the latter uses the full sinc form (albeit with no phase mismatch). Chapter 7, in particular, is very interesting because it shows how a complex task like the projection onto LG modes of a rather complicated wave function (the sinc is a function which turned out to be mathematically hard to handle) can be reached by simple heuristic reasoning. Chapter 8 is a realisation of my master thesis, which was centred on the optimisation of phase masks, and it was an extraordinary personal gratification to see those results in action.

9.2 Outlook and Acknowledgements

As we anticipated in the historical section in chapter 1, there are still unanswered questions on SPDC. In the future I would like to concentrate my efforts on hyperentanglement, where I hope I will be able to use my expertise on spatial entanglement. Quantities like crystal size, beam width and wavelength, can affect the temporal part of the entanglement too, and it is likely that optimality of spatial entanglement does not mean optimality of temporal entanglement. The compromise will lay somewhere in between, and it will be a major opportunity to be able to address this problem as a whole. However, there are still unanswered question exclusively on the spatial part, which I am going to summarise below.

It is certain that some of the results that we have found in this couple of years will have an important future, i.e. those results that can be immediately used, results that are useful for further enquiry into the SPDC state. in particular I would mention the Schmidt decompositions. Both the analytical and the numerical versions will hopefully give useful insight on *how* to perform ideal measurements. Once the settings for an experiment are ready, the recipes for Schmidt modes are in chapters 4 and 5. If the phase mismatch is zero the solution is even analytical.

I think the outlook of the results in chapter 4 are on the experimental side, as Schmidt modes are the ideal modes to use in order to maximise the shared information. It would be extremely interesting to check the results on imperfect measurements. If it is feasible to implement initial states with strong entanglement, an imperfect measurement could in any case yield a fraction of shared bits that can be large enough for the purpose. Experimental conditions to generate very high dimensional entanglement are relatively easy to achieve, especially if one is not concerned with luminosity: a short crystal would be enough to generate enough initial modes. The choice of detection basis, however, should try to be as close to the Schmidt basis as possible.

There are several points raised in chapter 5 that are worth further investigation: the question if a simple, heuristic rule like the α-modification that comprises a phase mismatch $\Phi \neq 0$ exists, the physical intuition about the families of displaced HG-like modes that fill the radial sectors and their properties as field modes. However, perhaps the most interesting is the actual behaviour of the radial modes, especially at high p. The proposed splitting of p in two different quantum numbers might be valid up to a limited upper value of p, but this precise behaviour, could be used to explore the meaning and the validity of the radial modes. The plots in figures 5.6 and 5.8 make one wonder about the problem

of tunnelling between different sites, and its similarity with topics in quantum integrable systems, quantum chaos and the Berry-Tabor conjecture [93], in reference to the potential impossibility of telling in which site the wavefunction is going to jump as the value of p is increased.

Chapter 6 raises some interesting points that find significance in the context of thin crystals, i.e. when the source can be considered to produce an infinite number of modes. In particular, questions like the behaviour of the results in the absence of OAM, i.e. equation (6.20) for $\ell = 0$, the optimal size that maximises the overlap between pump and generated modes for a fixed $|\ell|$, and potential benefits of sudden splittings like the ones observed in figures 6.9 and 6.10, which could be used to detect a difference in beam size. Additionally, the relations that determine the branching in figure 6.10 could be potentially easy to find using the properties of Laguerre polynomials.

The geometrical result in chapter 7 might feel a bit of a "disappointment", for its incredible simplicity and ability of reproducing the results found with much harder calculations. However, it is regarding only the $p = 0$ subspace. A similar result for $p \neq 0$ might not be as easy to find, but it is certainly worth some thought. A way of checking the eventual findings is clearly at hand, by numerically evaluating the amplitudes. Even if a result for any p is found, the hardest step would probably be putting together the results in a meaningful way, and to yield similar curves for a general $\Phi = 0$, sinc phase matching generation and detection scheme.

I'm particularly proud of chapter 7, for two reasons. The first is that the results were unexpected. It is by working with the full sinc and the gaussian approximation of the phase matching *together* that we realised their true difference. The second reason is that I learned a lot by working in close contact with an experimental group. The ability of an experimentalist of approaching problems is remarkable, the method itself is what I admired most and I hope I'll be able to use the same intuition and build a model of reality in my head and use it with the same passion and effectiveness as I've seen people doing so often during those weeks of intense collaboration.

Chapter 8 is a different take on the topic of detection. Its advantage is that it might greatly speed up the process of characterisation of a source of OAM entanglement, both for the fact that it focusses on *K directly*, without inferring it from a set of projection amplitudes, and because it is not affected by the same problem of an amplitude mask, so the coincidence rate is maximised. As the technology of spatial light modulators becomes better and better, might it even be

superseded, the basic idea of employing an azimuthal phase structure can always be used, and masks that span a D-dimensional Hilbert space can in principle be prepared for any D. The bottom line of chapter 8 is that we showed a proof of principle of a technique that in the future might be perfected and used to characterise sources in a very fast and reliable way.

Bibliography

[1] F. M. Miatto. *High dimensional photon entanglement and the design of sector phase plates*. Leiden: Master thesis, 2009.

[2] R. W. Boyd. *Nonlinear Optics*. Academic Press, 2008.

[3] M. Nielsen and I. Chuang. *Quantum computation and quantum information*. Cambridge: Cambridge University Press, 2000.

[4] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics*. Cambridge Univ. Press, 1995.

[5] A. K. Jha et al. "Fourier relationship between the angle and angular momentum of entangled photons". In: *Physical Review A* **78** (2008), p. 43810.

[6] L. Allen et al. "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes". In: *Physical Review A* **45** (1992), pp. 8185–8189.

[7] M. J. Padgett and L. Allen. "The Poynting vector in Laguerre-Gaussian laser modes". In: *Optics Communications* **121** (1995), pp. 36 –40.

[8] V. D. Salakhutdinov, E. R. Eliel, and W. Löffler. "Full-Field Quantum Correlations of Spatially Entangled Photons". In: *Physical Review Letters* **108** (2012), p. 173604.

[9] G. Bacciagaluppi and A. Valentini. "Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference". In: *quant-ph/0609184v2* (2006).

[10] N. Harrigan and R. Spekkens. "Einstein, Incompleteness, and the Epistemic View of the Quantum States". In: *Foundations of Physics* **40** (2 2010), pp. 125–157.

[11] D. Howard. "Einstein on locality and separability". In: *Studies In History and Philosophy of Science Part A* **16** (1985), pp. 171 –201.

[12] J. Bell. "On the Einstein-Podolsky-Rosen paradox". In: *Physics* **1** (1964), pp. 195 –200.

[13] A. Aspect, P. Grangier, and G. Roger. "Experimental Tests of Realistic Local Theories via Bell's Theorem". In: *Physical Review Letters* **47** (1981), pp. 460–463.

[14] P. G. Kwiat et al. "New High-Intensity Source of Polarization-Entangled Photon Pairs". In: *Physical Review Letters* **75** (1995), pp. 4337–4341.

[15] J. T. Barreiro et al. "Generation of Hyperentangled Photon Pairs". In: *Physical Review Letters* **95** (2005), p. 260501.

[16] C. Hong and L. Mandel. "Theory of parametric frequency down conversion of light". In: *Physical Review A* **31** (1985), pp. 2409–2418.

[17] A. Joobeur et al. "Coherence properties of entangled light beams generated by parametric down-conversion: Theory and experiment". In: *Physical Review A* **53** (1996), pp. 4360–4371.

[18] C. H. Monken, P. H. S. Ribeiro, and S. Pádua. "Transfer of angular spectrum and image formation in spontaneous parametric down-conversion". In: *Physical Review A* **57** (1998), pp. 3123–3126.

[19] A. Mair et al. "Entanglement of the orbital angular momentum states of photons". In: *Nature* **412** (2001), pp. 313–316.

[20] J. P. Torres, A. Alexandrescu, and L. Torner. "Quantum spiral bandwidth of entangled two-photon states". In: *Physical Review A* **68** (2003), p. 050301.

[21] C. Law and J. Eberly. "Analysis and Interpretation of High Transverse Entanglement in Optical Parametric Down Conversion". In: *Physical Review Letters* **92** (Mar. 2004), p. 127903.

[22] S. S. R. Oemrawsingh et al. "How to Observe High-Dimensional Two-Photon Entanglement with Only Two Detectors". In: *Physical Review Letters* **92** (2004), p. 217901.

[23] A. Aiello et al. "Nonlocality of high-dimensional two-photon orbital angular momentum states". In: *Physical Review A* **72** (2005), p. 052114.

[24] B. Pors et al. "Angular phase-plate analyzers for measuring the dimensionality of multimode fields". In: *Physical Review A* **77** (Mar. 2008), p. 033845.

[25] J. B. Pors et al. "Shannon Dimensionality of Quantum Channels and Its Application to Photon Entanglement". In: *Physical Review Letters* **101** (2008), p. 120502.

[26] J Romero et al. "ntangled Optical Vortex Links". In: *Physical Review Letters* **106** (2011), p. 100407.

[27] A. C. Dada et al. "Experimental high-dimensional two-photon entanglement and violations of generalized Bell inequalities". In: *Nature Physics* **7** (2011), pp. 677–680.

[28] Y. M. Mikhailova, P. A. Volkov, and M. V. Fedorov. "Biphoton wave packets in parametric down-conversion: Spectral and temporal structure and degree of entanglement". In: *Physical Review A* **78** (2008), p. 062327.

[29] M. V. Fedorov et al. "Spontaneous parametric down-conversion: Anisotropical and anomalously strong narrowing of biphoton momentum correlation distributions". In: *Physical Review A* **77** (2008), p. 032336.

[30] E. Brambilla et al. "Spatiotemporal structure of biphoton entanglement in type-II parametric down-conversion". In: *Physical Review A* **82** (2010), p. 013835.

[31] A. Gatti et al. "X Entanglement: The Nonfactorable Spatiotemporal Structure of Biphoton Correlation". In: *Physical Review Letters* **102** (2009), p. 223601.

[32] F. M. Miatto, T. Brougham, and A. M. Yao. "Cartesian and polar Schmidt bases for down-converted photons". In: *European Physics Journal D* **66** (7 2012), p. 183.

[33] F. M. Miatto et al. "Schmidt modes generated in parametric down-conversion". In: *European Physics Journal D (in press)* (2012).

[34] E. M. Gauger et al. "Sustained Quantum Coherence and Entanglement in the Avian Compass". In: *Physical Review Letters* **106** (2011), p. 040503.

[35] B. J. Pors et al. "Transport of Orbital-Angular-Momentum Entanglement through a Turbulent Atmosphere". In: *Opt. Express* **19** (2009), pp. 6671–6683.

[36] D. M. Greenberger, M. A. Horne, and A. Zeilinger. "Going Beyond Bell's Theorem". In: *quant-ph/0712.0921v1* (2007).

[37] N. Gisin. "Quantum Nonlocality: How Does Nature Do It?" In: *Science* **326** (2009), pp. 1357–1358.

[38] A. K. Ekert and P. L. Knight. "Entangled quantum systems and the Schmidt decomposition". In: *American Journal of Physics* **63** (1995), p. 415.

[39] A. E. Siegman. *Lasers.* Sausalito: University Science Books, 1986.

[40] F. Pampaloni and J. Enderlein. "Gaussian, Hermite-Gaussian, and Laguerre-Gaussian beams: A primer". In: *arXiv:physics/0410021* (2004).

[41] S. J. Van Enk and G Nienhuis. "Commutation Rules and Eigenvalues of Spin and Orbital Angular Momentum of Radiation Fields". In: *Journal of Modern Optics* **41** (May 1994), pp. 963–977.

[42] S. J. V. Enk and G Nienhuis. "Spin and Orbital Angular Momentum of Photons". In: *EPL (Europhysics Letters) (EPL)* **25** (Mar. 1994), pp. 497–501.

[43] R. P. Cameron, S. M. Barnett, and A. M. Yao. "Optical helicity, optical spin and related quantities in electromagnetic theory". In: *New Journal of Physics* **14** (2012), p. 053050.

[44] G. P. Agrawal. *Applications of Nonlinear Fiber Optics, second edition*. Elsevier, 2008.

[45] S. M. Barnett. *Quantum information*. Oxford: Oxford University Press, 2009.

[46] C. K. Law, I. A. Walmsley, and J. H. Eberly. "Continuous frequency entanglement: effective finite Hilbert space and entropy control". In: *Physical Review Letters* **84** (June 2000), pp. 5304–7.

[47] E. Abramochkin and V. Volostnikov. "Beam transformations and nontransformed beams". In: *Opt. Commun.* **83** (1991), p. 123.

[48] S. S. Straupe et al. "Angular Schmidt modes in spontaneous parametric down-conversion". In: *Physical Review A* **83** (2011), p. 060302.

[49] G. N. Watson. In: *Journal of the London Mathematical Society* (1933), pp. 189 –192.

[50] S. J. Habraken. *Light with a twist (ray aspects in singular wave and quantum optics)*. Leiden: PhD thesis, 2011.

[51] C. E. Shannon and W. Weaver. *The Mathemtiatical Theory of Communication*. Urbana: University of Illinois Press, 1949.

[52] E. T. Jaynes. "Information Theory and Statistical Mechanics". In: *Physical Review* **106** (1957), p. 620.

[53] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons, 1991.

[54] A. Réyni. In: *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability* (1960), p. 567.

[55] S. T. Flammia et al. "Topological Entanglement Rényi Entropy and Reduced Density Matrix Structure". In: *Physical Review Letters* **103** (2009), p. 261601.

[56] S. M. Barnett and S. J. D. Phoenix. "Entropy as a measure of quantum optical correlation". In: *Physical Review A* **40** (1989), p. 2404.

[57] C. H. Bennett et al. "Concentrating partial entanglement by local operations". In: *Physical Review A* **53** (1996), p. 2046.

[58] C. H. Bennett et al. "Mixed-state entanglement and quantum error correction". In: *Physical Review A* **54** (1996), p. 3824.

[59] S. M. Barnett and S. J. D. Phoenix. "Information theory, squeezing, and quantum correlations". In: *Physical Review A* **44** (1991), p. 535.

[60] M. J. W. Hall. "Quantum information and correlation bounds". In: *Physical Review A* **55** (1997), p. 100.

[61] M. J. W. Hall, E. Andersson, and T. Brougham. "Maximum observable correlation for a bipartite quantum system". In: *Physical Review A* **74** (2006), p. 062308.

[62] T. Brougham and S. M. Barnett. "Information communicated by entangled photon pairs". In: *Physical Review A* **85** (2012), p. 032322.

[63] X. Ma, C.-H. F. Fung, and H.-K. Lo. "Quantum key distribution with entangled photon sources". In: *Physical Review A* **76** (2007), p. 012307.

[64] H. Di Lorenzo Pires, C. H. Monken, and M. P. van Exter. "Direct measurement of transverse-mode entanglement in two-photon states". In: *Physical Review A* **80** (2009), p. 22307.

[65] A. M. Yao. "Angular momentum decomposition of entangled photons with an arbitrary pump". In: *New Journal of Physics* **13** (2011), p. 053048.

[66] M. P. van Exter et al. "Effect of spatial filtering on the Schmidt decomposition of entangled photons". In: *Physical Review A* **74** (2006), p. 012309.

[67] A. B. U'Ren, K. Banaszek, and I. A. Walmsley. "Photon engineering for quantum information processing". In: *Quantum Information and Computation* **3** (2003), p. 480.

[68] L. E. Vicent et al. "Design of bright, fiber-coupled and fully factorable photon pair sources". In: *New Journal of Physics* **12** (2010), p. 093027.

[69] H. J. Korsch, A. Klumpp, and D. Witthaut. "On two dimensional Bessel functions". In: *quant-ph/0608216v1* (2006).

[70] G. Dattoli et al. "Advances on the Theory of Generalized Bessel Functions and Applications to Multiphoton Processes". In: *Journal of Scientific Computing* **8** (1993), pp. 69 –109.

[71] H. Di Lorenzo Pires and M. van Exter. "Observation of near-field correlations in spontaneous parametric down-conversion". In: *Physical Review A* **79** (2009), p. 041801.

[72] C. Hong, Z. Ou, and L. Mandel. "measurement of subpicosecond time intervals between two photons by interference". In: *Physical Review Letters* **59** (1987), pp. 2044–2046.

[73] R. Zambrini and S. M. Barnett. "Quasi-Intrinsic Angular Momentum and the Measurement of Its Spectrum". In: *Physical Review Letters* **96** (2006), p. 113901.

[74] H. D. L. Pires, J Woudenberg, and M. P. van Exter. "Measurement of the orbital angular momentum spectrum of partially coherent beams." In: *Optics letters* **35** (2010), pp. 889–91.

[75] F. M. Miatto et al. "Bounds and optimisation of orbital angular momentum bandwidths within parametric down-conversion systems". In: *European Physics Journal D* **66** (7 2012), p. 178.

[76] S. S. Straupe et al. "Spatial Correlations and Angular Schmidt Modes in Spontaneous Parametric Down-Conversion". In: *quant-ph/1112.3806v1* (2011).

[77] F. M. Miatto, A. M. Yao, and S. M. Barnett. "Full characterization of the quantum spiral bandwidth of entangled biphotons". In: *Physical Review A* **83** (2011), p. 033816.

[78] D. N. Nikogosyan. "Beta Barium Borate (BBO)". In: *Applied Physics A* **52** (1991), p. 359.

[79] B. E. A. Saleh et al. "Duality between partial coherence and partial entanglement". In: *Physical Review A* **62** (2000), p. 043816.

[80] B. Jack et al. "Entanglement of arbitrary superpositions of modes within two-dimensional orbital angular momentum state spaces". In: *Physical Review A* **81** (2010), p. 043844.

[81] S. S. R. Oemrawsingh et al. "Experimental Demonstration of Fractional Orbital Angular Momentum Entanglement of Two Photons". In: *Physical Review Letters* **95** (2005), p. 240501.

[82] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics.* New York: Wiley, 2007.

[83] R. S. Bennink. "Optimal collinear Gaussian beams for spontaneous parametric down-conversion". In: *Physical Review A* **81** (2010), p. 053805.

[84] S. Franke-Arnold et al. "Two-photon entanglement of orbital angular momentum states". In: *Physical Review A* **65** (2002), p. 033823.

[85] G. N. Watson. *A treatise on the theory of Bessel functions.* Cambridge Mathematical Library. Cambridge University Press, 1995.

[86] B. J. Pors. *Entangling light in high dimensions.* Leiden: PhD thesis, 2011.

[87] H. Di Lorenzo Pires, H. C. B. Florijn, and M. P. van Exter. "Measurement of the Spiral Spectrum of Entangled Two-Photon States". In: *Physical Review Letters* **104** (Jan. 2010), p. 020505.

[88] B.-J. Pors et al. "High-dimensional entanglement with orbital-angular-momentum states of light". In: *Journal of Optics* **3** (2011), p. 064008.

[89] J. Leach et al. "Violation of a Bell inequality in two-dimensional orbital angular momentum state-spaces". In: *Optics express* **17** (May 2009), pp. 8287–93.

[90] J. Leach et al. "Quantum Correlations in Optical Angle-Orbital Angular Momentum Variables". In: *Science* **329** (2010), pp. 662–665.

[91] S. M. Barnett and D. T. Pegg. "Quantum theory of rotation angles". In: *Physical Review A* **41** (1990), p. 3427.

[92] S. Franke-Arnold et al. "Uncertainty principle for angular position and angular momentum". In: *New Journal of Physics* **6** (2004), p. 103.

[93] M. Berry and M. Tabor. "Level clustering in the regular spectrum". In: *Proc. Roy. Soc. A* **356** (1977), pp. 375–394.