International Journal of Quantum Information
Vol. 12, No. 4 (2014) 1430004 (38 pages)
© World Scientific Publishing Company
DOI: 10.1142/S0219749914300046

World Scientific
www.worldscientific.com

# Quantum information in communication and imaging*

David S. Simon†

*Department of Physics and Astronomy,*
*Stonehill College, 320 Washington Street,*
*Easton, MA 02357 USA*

*Department of Electrical and Computer Engineering,*
*Boston University, 8 Saint Marys Street,*
*Boston, MA 02215, USA*
*simond@bu.edu*

Gregg Jaeger

*Department of Electrical and Computer Engineering,*
*Boston University, 8 Saint Marys St.,*
*Boston, MA 02215, USA*

*Department of Physics, Boston University,*
*590 Commonwealth Avenue, Boston, MA 02215, USA*

*Division of Natural Sciences and Mathematics,*
*Boston University, Boston, MA 02215, USA*
*jaeger@math.bu.edu*

Alexander V. Sergienko

*Department of Electrical and Computer Engineering,*
*Boston University, 8 Saint Marys St., Boston,*
*MA 02215, USA*

*Department of Physics, Boston University,*
*590 Commonwealth Avenue, Boston, MA 02215, USA*
*alexserg@bu.edu*

Received 14 April 2014
Accepted 3 June 2014
Published 24 June 2014

A brief introduction to quantum information theory in the context of quantum optics is presented. After presenting the fundamental theoretical basis of the subject, experimental

*Lecture given at the "Advanced School on Quantum Foundations and Open Quantum Systems" held in João Pessoa, Brazil, from 16–28 July 2012.
†Corresponding author.

evaluation of entanglement measures are discussed, followed by applications to communication and imaging.

*Keywords*: Quantum cryptography; quantum imaging; quantum information theory.

# 1. From Bit to Qubit

## 1.1. *Introduction*

Quantum information science is a rapidly developing area of interdisciplinary investigation, which plays a significant role in a number of sub-disciplines of physics and engineering. Quantum communication (including quantum key distribution for cryptography) and quantum imaging are currently two of the most exciting applications of quantum information science. For this reason, we focus here on quantum optical systems, a natural choice because communication and imaging are typically optical in the current era. Further, interferometry is central to quantum information processing and interferometry has primarily progressed through optical physics. Quantum theory was developed by Einstein, Bohr, Schrödinger, Heisenberg, Dirac and others, and given a unified formalization first by Dirac[a] and later by von Neumann[b] in the first third of the 20th century. It serves as a basis for understanding quantum field theory, wherein Dirac again played a key role. By the end of the 20th century, quantum information science, which was developed entirely within this formalism, became a subject in its own right. In practice, it can be best understood as a range of interferometric systems acting as realizations of specifically quantum mechanical physical communications layers, protocols, and algorithms. These are primarily based on the use of the quantum information unit, the "qubit." The term "qubit" originated with Benjamin Schumacher, who replaced "the classical idea of a binary digit with a quantum two-state system…These quantum bits, or "qubits," are the fundamental units of quantum information."[1]

The quantum difference from classical information arises from the *superposition principle* of quantum mechanics. This means that, despite its being two-valued in the chosen computational basis, a qubit system can be in one of an infinite number of physically significant states: while a bit is capable of being in only one of two significant states at a given moment, a qubit system in general can be considered as potentially being in one measurable state and the other opposite state at the same time. In further contrast to classical states, a single unknown state of a qubit system cannot generally be found by a single measurement, but rather requires an ensemble of them to be determined. It is precisely the superposition of individual qubits that provides the possibility of secure quantum key distribution, for example.

One striking consequence of superposition in quantum mechanics, is the possibility of *entanglement*, in which the state of a composite system can not be factored into a product of states describing each of its subsystems separately. To be more

[a] In his *The Principles of Quantum Mechanics.*

[b] In his *Mathematische Grundlagen der Quantenmechanik.*

specific, consider a bipartite composite system, formed from the pair of subsystems $A$ and $B$; for example, $A$ and $B$ may be labels for two photons, two atoms, or any other pair of quantum systems. These two systems may be separated by arbitrarily large distances. We may then form the composite system $A \otimes B$, whose Hilbert space is the product of the two individual Hilbert spaces: $\mathcal{H}_{AB} = \mathcal{H}_A \times \mathcal{H}_B$. Consider a pure state $|\psi_{AB}\rangle$ of the composite system. Then the state is said to be *separable* if it can be written in some basis in the form $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, where $|\psi_A\rangle \in \mathcal{H}_A$ nd $|\psi_B\rangle \in \mathcal{H}_B$. The state is then defined to be entangled if it is not separable. We will describe the consequences of entanglement and how it may be quantified in more detail below. As we are primarily concerned here with optical systems, we will also describe in detail the process of spontaneous parametric downconversion (SPDC), which provides a convenient and versatile means of producing entangled pairs of photons. As we introduce measures of information and of entanglement, we will apply them to the downconversion process and its applications.

Our main focus is on quantum communication. The most thoroughly studied application of quantum communications is quantum cryptography, also known as quantum key distribution. After describing the basic ideas in this area, we move on to a topic which has been less well studied from a quantum information theoretical viewpoint, namely quantum imaging. We will take a broad view of the word communication in order to include the reconstruction of images over a distance. In order to quantify our ability to communicate, it will be necessary to investigate the amount of information extracted per photon and the amount of entanglement per pair of photons, as well as the amount of mutual information carried per pair. We will exclusively discuss *bipartite* systems, i.e. those with two subsystems.

One related topic we will not discuss in detail is quantum computing, which applies the quantum superposition principle to a collection of *stored* qubits, which can be thought of as forming a compound quantum system. The space of possible quantum states available to such multiple-qubit systems grows more rapidly than does the space of states available to multiple-bit systems. The size of the parameter space describing a quantum system for information encoding and computing grows exponentially in the number of qubits — in a classical system it grows only linearly in the number of bits. This also provides a unique sort of computational parallelism, which can be harnessed to make tractable some important computational tasks that are thought intractable using classical means only. This improvement in efficiency is known as "quantum speedup." Multiple-qubit states however are also very fragile, being susceptible to decoherence effects.[2,3] After a short period of time, the initially pure quantum states described are inevitably altered by interactions with their environments and must then be described instead by a mixed quantum states.

In the remainder of this section, we introduce basic notions of quantum communication theory in the context of quantum optics, including a detailed discussion of spontaneous parametric down conversion, which is the principle source of entangled optical states for experiments. Section 2 moves on to applications with a discussion of quantum cryptography, followed by a discussion of how the main ideas generalize from the context of qubits to so-called qudits. Section 3 follows up by discussing a specific realization of qudits in the form of orbital angular momentum (OAM) states. These same states are then applied in the context of imaging, leading to the idea that the mutual information shared by entangled states may serve as probes of geometric symmetries.

### 1.2. *From bits to qubits*

The properties of a qubit system are two-valued and can be probabilistically predicted like a classical system that randomly takes one of two computationally relevant values. But unlike the classical system, which can only be in one of the two states at any time irrespective of how it may be measured, a qubit can be in both states simultaneously. The unit of classical information is sometimes referred in quantum information science as *c-bits*.[4] A putative inherently probabilistic bit can be called a *probabit*.[5] The probabilities of the outcomes of measurements of any classical system are due only to *ignorance* of the actual state of the system. In the quantum case, it arises also from a *fundamental* indeterminacy of properties, entirely so in the case of the pure quantum states, defined below. The quantum bit is, therefore, not reducible to the probabilistic bit.[c]

Let us begin by considering various representations of qubit states. Recall that quantum states are associated with a complex Hilbert vector space, $\mathcal{H}$, via a special class of linear operators acting in it, the *statistical operators*, $\hat{\rho}$, constituting the quantum state-space. For pure qubit states, the statistical operators are projectors onto one-dimensional subspaces. The projective operators $P(|\psi\rangle)$ can be uniquely associated with points on the boundary of the Bloch ball, known as the Poincaré–Bloch sphere. These states can also be, and typically are, represented by the state-vectors

$$|\psi\rangle \in \mathcal{H}, \tag{1}$$

spanning them. The remaining states of the Bloch ball are essentially statistical or mixed states, defined as those which are not pure but still satisfy the definition of a density operator. The mixed states can be formed from these projectors by appropriate linear combinations and lie in the interior of the Bloch ball. However, the mixed states *cannot* be written as linear combinations of state vectors.

The set of statistical states available to a qubit system is representable by the $2 \times 2$ complex Hermitian trace-one matrices

$$[\hat{\rho}_{ij}] \in H(2). \tag{2}$$

[c] Note that we will here use "qubit" and "quantum bit" to refer to both physical systems on which quantum information can be encoded, as well as to the quantum bit of information in the sense of information theory, depending on context to make clear which is intended in any given instance.

By contrast, for the full physical state description of a quantum system *in spacetime*, an infinite-dimensional spatial representation is required in which the state-vectors are called *wavefunctions*. Quantum information theory is based on the behavior of qubits and has thus far overwhelmingly dealt with quantities with discrete eigenvalue spectra in the non-relativistic regime, the state-vectors considered here are usually taken to lie within finite-dimensional Hilbert spaces constructed by taking the tensor product of multiple copies of *two-dimensional* complex Hilbert space, or other finite-dimensional spaces. The Hilbert spaces considered here are only finite-dimensional *subspaces* of the larger full physical state-spaces of particles, the other subspaces of which are rarely taken into account in the study of quantum information processing. In many cases, we consider the polarization states or OAM states of photons, without considering the corresponding full photon wavefunctions.

The *purity*, $\mathcal{P}$, of a quantum state specified by the statistical operator $\hat{\rho}$ is the trace of its square,

$$\mathcal{P}(\hat{\rho}) = \mathrm{tr}\hat{\rho}^2, \tag{3}$$

where $\frac{1}{d} \leq \mathcal{P}(\hat{\rho}) \leq 1$ and $d$ is the dimension of the Hilbert space, $\mathcal{H}$, attributed to the system it describes. The quantum state is *pure* if $\mathcal{P}(\hat{\rho}) = 1$, i.e. if it spans a one-dimensional subspace of $\mathcal{H}$, one can then naturally define state *mixedness* as the complement of purity, $\mathcal{M}(\hat{\rho}) \equiv 1 - \mathcal{P}(\hat{\rho})$. The *Unitary* linear operators, $U$, are those for which $U^\dagger U = UU^\dagger = \mathbb{I}$, where "$\dagger$" indicates Hermitian conjugation. Here, the time-evolution is prescribed by the Schrödinger equation, assuming a time-independent Hamiltonian. (In general, temporal evolution in quantum mechanics is not always so simple; cf. Sec. 2.1 of Sakurai.[6]) The purity and mixedness of a quantum state are invariant under transformations of the form $\hat{\rho} \to U\hat{\rho}\ U^\dagger$, where $U$ is unitary, most importantly under the *dynamical mapping* $U(t,t_0) = e^{-\frac{i}{\hbar}H(t-t_0)}$, where $H$ is the Hamiltonian operator, which can readily be seen upon recalling that the trace operation $\mathrm{tr}(\cdot)$ is cyclic. Pure states are those which are *maximally specified* within quantum mechanics. A quantum state is pure if and only if the statistical operator $\hat{\rho}$ is *idempotent*, i.e.

$$\hat{\rho}^2 = \hat{\rho}, \tag{4}$$

providing a convenient test for maximal state purity. It is also a *projector*, $P(|\psi_i\rangle)$, where $|\psi_i\rangle$ is the normalized vector representative of the corresponding one-dimensional subspace of its Hilbert space. Rays cannot be added, whereas vectors $|\psi_i\rangle$ can be, making the latter better for use in calculations involving pure states, where superpositions are formed by addition. A Hermitian operator $P$ acting in a Hilbert space $\mathcal{H}$ is a *projector* if and only if $P^2 = P$. It follows from this definition that $P^\perp \equiv \mathbb{I} - P$, where $\mathbb{I}$ is the identity operator, is also a projector. The projectors $P$ and $P^\perp$ project onto orthogonal subspaces within $\mathcal{H}$, $\mathcal{H}_s$, and $\mathcal{H}_s^\perp$, respectively, thereby providing a decomposition of $\mathcal{H}$ as $\mathcal{H}_s \oplus \mathcal{H}_s^\perp$; two subspaces are said to be *orthogonal* if every vector in one is orthogonal to every vector in the other. In the case of a general state of a single qubit, one may write $\hat{\rho} = p_1 P(|\psi\rangle) + p_2 P(|\psi^{\perp}\rangle)$, where the weights $p_i$ are the eigenvalues of the statistical operator $\hat{\rho}$.

A quantum state is thus mixed if it is *not* a pure state, that is, if $\mathcal{P}(\hat{\rho}) < 1$. In the Dirac notation, projectors are written

$$P(|\psi_i\rangle) \equiv |\psi_i\rangle\langle\psi_i|. \tag{5}$$

Consider a finite set, $\{P(|\psi_i\rangle)\}$, of projectors corresponding to distinct, orthogonal pure states $|\psi_i\rangle$. Any state $\hat{\rho}'$ that can be written

$$\hat{\rho}' = \sum_i p_i P(|\psi_i\rangle), \tag{6}$$

with $0 < p_i < 1$ and $\sum_i p_i = 1$, is then a normalized mixed state. The superposition principle implies that any (complex) linear combination of qubit basis states, such as $|0\rangle$ and $|1\rangle$, i.e.

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle \tag{7}$$

with $a_i \in \mathbb{C}$ and $|a_0|^2 + |a_1|^2 = 1$, is *also* a physical state of the qubit and is, as we have seen, also a pure state. The scalar coefficients $a_0$ and $a_1$ are called quantum probability amplitudes, because their square magnitudes, $|a_0|^2$ and $|a_1|^2$, are the probabilities $p_0$ and $p_1$, respectively, of the qubit described by state $|\psi\rangle$ being found in these basis states $|0\rangle$ and $|1\rangle$, respectively, upon measurement.

The superposition principle is ultimately the source of many of the quantum phenomena that we will use in the forthcoming Sections. In particular it underlies entanglement, interference phenomena, and the inability to distinguish non-orthogonal states, all of which will be used for applications in Secs. 2 and 3. Consider the normalized sums

$$|\nearrow\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad \text{and} \quad |\searrow\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{8}$$

of two orthogonal pure state-vectors $|0\rangle \doteq (1\ 0)^{\mathrm{T}}$ and $|1\rangle \doteq (0\ 1)^{\mathrm{T}}$ of a qubit, the r.h.s.'s being given in the matrix representation and $(\cdots)^{\mathrm{T}}$ indicating matrix transposition. The superpositions in Eq. (8) are pure states, as can be immediately verified by taking their square moduli. The corresponding projectors are $P(|\nearrow\rangle) = |\nearrow\rangle\langle\nearrow|$, $P(|\searrow\rangle) = |\searrow\rangle\langle\searrow|$. However, the normalized sum of a pair of *projectors*, for example, $P(|0\rangle)$ and $P(|1\rangle)$ corresponding to pure states $|0\rangle$ and $|1\rangle$, namely,

$$\hat{\rho}_+ = \frac{1}{2}(P(|0\rangle) + P(|1\rangle)), \tag{9}$$

is a *mixed* state that can also be written

$$\hat{\rho}_+ = \frac{1}{2}(P(|\nearrow\rangle) + P(|\searrow\rangle)). \tag{10}$$

Finally, note that the statistical operator corresponding to the normalized sum of $|\nearrow\rangle$ and $|\searrow\rangle$ is $P(|0\rangle) \neq \hat{\rho}_+$.

The pure states of the qubit can be represented by vectors in the two-dimensional complex Hilbert space, $\mathcal{H} = \mathbb{C}^2$. Any orthonormal basis for this space can be put in correspondence with two bit values, 0 and 1, in order to act as the single-qubit *computational basis*, sometimes also called the *rectilinear basis*, and written $\{|0\rangle, |1\rangle\}$. The vectors of the computational basis can be represented in matrix form as:

$$|0\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{11}$$

Other commonly used bases are the *diagonal basis*, $\{|\nearrow\rangle, |\searrow\rangle\}$, sometimes also written $\{|+\rangle, |-\rangle\}$, and the *circular basis* $\{|\mathrm{r}\rangle, |\mathrm{l}\rangle\}$:

$$|\mathrm{r}\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \quad |\mathrm{l}\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle), \tag{12}$$

sometimes also written $\{|\circlearrowright\rangle, |\circlearrowleft\rangle\}$, is also useful for quantum cryptography, being conjugate to both the computational and diagonal bases.

All three of the above bases are mutually conjugate and are used in protocols for quantum key distribution (Sec. 2.2); the probabilities of qubits in the states $|\mathrm{r}\rangle$ and $|\mathrm{l}\rangle$ being found in the states $|0\rangle$, $|1\rangle$, $|\nearrow\rangle$, and $|\searrow\rangle$ are all $\frac{1}{2}$, and vice-versa. The generic mixed state, $\hat{\rho}$, lies in the interior of the Bloch ball, can be written as a convex combination of basis-element projectors corresponding to the pure-state bases described above. The effect of a general operation on a qubit can be viewed as a (possibly stochastic) transformation within this ball; for illustrations of this in practical context, see Ref. 7. The parametrization required to adequately describe mixed states is now discussed in detail.

The density matrix and the *Stokes four-vector*, $S_\mu$, are related by

$$\hat{\rho} = \frac{1}{2}\sum_{\mu=0}^{3} S_\mu \sigma_\mu, \tag{13}$$

where $\sigma_\mu$ ($\mu = 1, 2, 3$) are the *Pauli operators* which, together with the identity $\sigma_0 = \mathbb{I}_2$, are represented in the matrix space $H(2)$ by the Pauli matrices. The Pauli matrices form a basis for $H(2)$, which contains the qubit density matrices. The qubit density matrices themselves are the positive-definite, trace-class elements of the set of $2 \times 2$ complex Hermitian matrices $H(2)$ of unit trace, i.e. for which the total probability $S_0$ is one, as prescribed by the Born rule for quantum probabilities and the well-definedness of quantum probabilities as such. Density matrices are similarly defined for systems of countable dimension. The non-trivial products of the four Pauli matrices — those between the $\sigma_i$ for $i = 1, 2, 3$ — are given by

$$\sigma_i \sigma_j = \delta_{ij}\sigma_0 + i\epsilon_{ijk}\sigma_k, \tag{14}$$

which defines their algebra. Appropriately exponentiating the Pauli matrices provides the rotation operators, $\mathrm{R}_i(\xi) = e^{-i\xi\sigma_i/2}$, for Stokes vectors about the corresponding directions $i$[6]; these rotations realize the group $SO(3)$.

The Stokes parameters $S_\mu$ $(\mu = 0, 1, 2, 3)$ also allow one to directly visualize the qubit state geometrically in the Bloch ball via $S_1, S_2, S_3$. The Euclidean length of this three-vector (also known as the *Stokes vector*, or *Bloch vector*) is the radius $r = (S_1^2 + S_2^2 + S_3^2)^{1/2}$ of the sphere produced by rotations of this vector. With the matrix vector $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ and the three-vector $\vec{S} = (S_1, S_2, S_3)$, one has

$$\hat{\rho} = \frac{1}{2}(S_0\mathbb{I}_2 + S_1\sigma_1 + S_2\sigma_2 + S_3\sigma_3), \tag{15}$$

known as the *Bloch-vector representation* of the statistical operator, in accord with Eq. (15). In optical situations, where $\vec{S}$ describes a polarization state of a photon, the degree of polarization is given by $P = r/S_0$, where $S_0$ is positive. For the qubit, when the state is normalized so that $S_0 = 1$, $S_0$ corresponds to total quantum probability. The density matrix of a single qubit is then of the form

$$\hat{\rho} \doteq \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}, \tag{16}$$

where $\rho_{00} + \rho_{11} = 1$, $\rho_{ii} = \rho_{ii}^*$ with $(i = 0, 1)$, and $\rho_{10} = \rho_{10}^*$, where $^*$ indicates complex conjugation. One can write the Pauli matrices for $\mu = 1, 2, 3$ in terms of outer products of computational basis vectors, as follows. The Stokes parameters are expressed in terms of the density matrix as

$$S_\mu = \mathrm{tr}(\hat{\rho}\sigma_\mu), \tag{17}$$

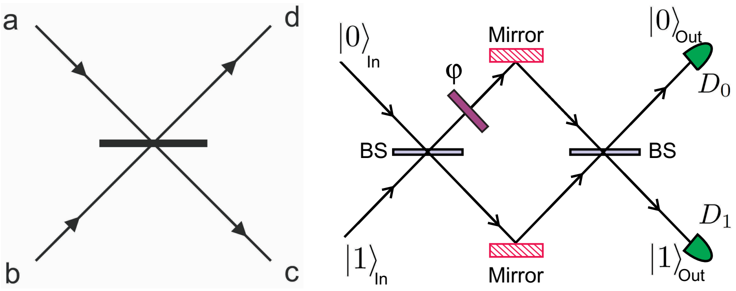
which are probabilities corresponding to ideal normalized counting rates of measurements in the standard eigenbases.

### 1.3. *Optical qubits*

For specificity, let us now take the system in question to be a photon. Light is easy to produce and to detect, and has properties that are both well understood and easily controlled. As a result, most experiments in quantum information and communication are carried out on optical systems. Consequently, we will focus henceforth exclusively on quantum optical systems. We begin by describing how optical qubits can be created.

Consider the beam splitter (BS) shown in Fig. 1(a). A BS is a device for splitting a single optical beam into two: a portion of the beam is transmitted through the BS, while a portion is reflected. Throughout, we assume that all BSs used are 50–50, i.e. that equal amounts of light are reflected and transmitted. We also consider only nonpolarizing BSs. A BS is a linear, passive four-port device, with two input ports ($a$ and $b$) and two output ports ($c$ and $d$). To describe its action, we form the operator-valued column vectors

$$\begin{pmatrix} \hat{a}^\dagger \\ \hat{b}^\dagger \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \hat{c}^\dagger \\ \hat{d}^\dagger \end{pmatrix}, \tag{18}$$



(a) (b)

Fig. 1. (a) A 50/50 beamsplitter. A photon entering either input port, **a** or **b**, has equal probability of being transmitted or reflected out either output port, **c** or **d**. (b) The Mach–Zehnder interferometer providing a range of qubit states as the input qubit amplitudes $\mathrm{a}_i$ and phases $\phi_i$ are changed. The detectors provide count rates proportional to the probability of lying in the output computational-basis states described by state-projectors $P(|0\rangle)$ and $P(|1\rangle)$, for input amplitudes $a_0 = 0$, $a_1 = 1$, namely, $p(0) = \sin^2[(\phi_0 - \phi_1)/2]$ and $p(1) = \cos^2[(\phi_0 - \phi_1)/2]$.

where $\hat{a}^\dagger, \hat{b}^\dagger, \hat{c}^\dagger, \hat{d}^\dagger$ are the creation operators for photon states at the corresponding ports. Then we may denote the action of the BS by a matrix $T$ relating ingoing and outgoing operators,

$$\begin{pmatrix} \hat{c}^\dagger \\ \hat{d}^\dagger \end{pmatrix} = T \begin{pmatrix} \hat{a}^\dagger \\ \hat{b}^\dagger \end{pmatrix}.$$

The form of this matrix is easy to determine: the photon is unchanged when it is transmitted and picks up a phase of $\frac{\pi}{2}$ when reflected, so the BS matrix is

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}. \tag{19}$$

We may now think of photons entering or leaving from above the BS (i.e. ports $a$ and $d$) as representing state $|0\rangle$, while those entering or leaving below the BS (i.e. $b$ and $c$) represent $|1\rangle$ states. This provides a representation of physical qubits as spatial modes, and then allows us to think of the BS matrix $T$ as taking combinations of input bits to combinations of output bits. In particular, if a *bit* 0 is input, the resulting output is the *qubit* $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Thus, we have a simple way of producing spatial qubits from classical bits.

We may form more general spatial qubits with the Mach–Zehnder interferometer (Fig. 1(b)). This is equivalent to a double-slit-like arrangement where only two directions are available to the self-interfering system, so that the exit ports of a BS act as "slits". In this interferometer, a photon enters from the left into a BS, with two exit paths on the right. It provides a spatial qubit, consisting of occupation of one and/or the other interior beam path. Each path then encounters a mirror, a phase shifter, a second BS, and finally a particle detector. Since only the relative phase between arms matters, the phase shift in one path can be set to zero without loss of generality. One can also use this interferometer to prepare a *phase qubit* by selecting only those systems entering a single initial input port and exiting a single final output port.

The action of the interferometer may be described by the matrix $B = T\Phi T$, where $T$ is the BS matrix above and the phase shift is described by the matrix

$$\Phi = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix}. \tag{20}$$

Multiplying out the matrices, we find that the action on an incoming bit $|0\rangle$ is:

$$|0\rangle \to \frac{1}{2}[(e^{i\phi} - 1)|0\rangle + i(e^{i\phi} + 1)|1\rangle], \tag{21}$$

allowing construction of a family of phase qubits.

Rather than spatial or phase qubits, we may consider superposition states of some other degree of polarization. A common choice is the polarization qubit, such as

$$|\psi\rangle = a_0|\uparrow\rangle + a_1|\rightarrow\rangle, \tag{22}$$

or in the diagonal basis:

$$|\psi\rangle = a_0'|\nearrow\rangle + a_1'|\searrow\rangle. \tag{23}$$

The next subsection shows one means of creating entangled polarization qubits. A further type of optical qubit, formed from superpositions of OAM states is considered in Sec. 3.

### 1.4. *Spontaneous parametric down conversion*

The most reliable and versatile means of producing entangled photon pairs is via *spontaneous parametric down conversion* (SPDC) inside a nonlinear crystal, such as $\beta$-barium borate (BBO) or potassium titanyl phosphate (KTP). In this process, a high frequency incoming photon (the *pump*) is converted into a pair of lower frequency outgoing photons (known for historical reasons as the *signal* and *idler* photons). Although the signal and idler beams are individually spatially and temporally incoherent, the signal and idler are mutually coherent, in the sense that the two photons in a given pair always leave the interaction point with a stable relation between their phases. The resulting photons are entangled in a number of different variables: position, momentum, frequency, time, polarization, and OAM. In fact, the eigenstates of these multiple variables for the two photons are intertwined through entanglement; for example, the joint signal polarization-idler momentum states are entangled, a phenomenon which is known as *hyperentanglement.*[8–10] Note that the output is entangled in both continuous *and* discrete degrees of freedom. In later sections, we will use the entangled state produced in down conversion for several communication and cryptography applications; this state has found a number of other uses in diverse areas such as dispersion and aberration cancelation, quantum optical coherence tomography, and precision measurement of polarization mode dispersion.[11–19]

When an electric field is applied to a material with a nonlinear response, the polarization may be expanded in powers of the field. Here we concentrate on the second-order term, $\hat{P}_i^{(2)} = \chi_{ijk}^{(2)} \hat{E}_j \hat{E}_k$, where the indices label spatial components and repeated indices are summed over. The corresponding interaction Hamiltonian is

$$\hat{H}_{\mathrm{int}}(t) = \epsilon_0 \int d^3 r \hat{P}^{(2)} \cdot \hat{\boldsymbol{E}} = \epsilon_0 \int d^3 r \chi_{jkl}^{(2)} \hat{E}_{\mathrm{p}j} \hat{E}_{\mathrm{s}k} \hat{E}_{\mathrm{i}l}. \tag{24}$$

The labels $p$, $s$, $i$ have been added to distinguish the pump, signal and idler fields. We may expand each field in terms of plane wave components,

$$\hat{E}_j(\boldsymbol{r}, t) = \int d^3 k [\hat{E}^-(\boldsymbol{k}) e^{-i(\omega t - \boldsymbol{k} \cdot \boldsymbol{r})} + \hat{E}^+(\boldsymbol{k}) e^{i(\omega t - \boldsymbol{k} \cdot \boldsymbol{r})}], \tag{25}$$

where, for quantization volume $V$, the positive and negative frequency parts are given by

$$\hat{E}_j^{(-)}(\boldsymbol{k}) = i\sqrt{\frac{2\pi\hbar\omega}{V}} \hat{a}_j^\dagger(\boldsymbol{k}), \quad \hat{E}_j^{(+)}(\boldsymbol{k}) = -i\sqrt{\frac{2\pi\hbar\omega}{V}} \hat{a}_j(\boldsymbol{k}). \tag{26}$$

Substituting Eqs. (25) and (26) into Eq. (24) and keeping only the terms that give a non-zero result when wedged between a one-photon incoming state and two-photon outgoing state, the result is:

$$\begin{aligned} \hat{H}_{\mathrm{int}}(t) = C \int d^3 k_{\mathrm{s}} d^3 k_{\mathrm{i}} e^{i(\omega_{\mathrm{p}} - \omega_{\mathrm{s}} - \omega_{\mathrm{i}})t} \int_0^L dz e^{i(k_{\mathrm{p}z} - k_{\mathrm{s}z} - k_{\mathrm{i}z})z} \\ \times \int_A d^2 r_\perp e^{i(k_{\mathrm{s}\perp} + k_{\mathrm{i}\perp}) r_\perp} \hat{a}^\dagger(\boldsymbol{k}_{\mathrm{s}}) \hat{a}^\dagger(\boldsymbol{k}_{\mathrm{i}}) + h.c. \end{aligned} \tag{27}$$

Here we have assumed that the incoming intensity is high enough to treat the pump as a classical field, and we have swept all of the overall constants into a single constant, $C$. It has also been assumed that the pump is a plane wave aligned along the $z$-axis, with no transverse momentum. In addition, the $\sqrt{\omega}$ terms coming from Eq. (26) are very slowly varying compared to the exponentials, and so were treated as constants. $L$ is the length of the crystal in the $z$ direction, and $A$ is the area of the interaction region, i.e. the region of the crystal where the pump is intense enough for significant downconversion to take place. Since the interaction area $A$ is normally much larger than the wavelength, we may approximate by taking $A \to \infty$, making the transverse integral trivial:

$$\int_A d^2 r_\perp e^{i(\boldsymbol{k}_{s\perp} + \boldsymbol{k}_{i\perp}) \cdot \boldsymbol{r}_\perp} = 2\pi \delta^{(2)}(\boldsymbol{k}_{s\perp} + \boldsymbol{k}_{i\perp}). \tag{28}$$

Defining the longitudinal momentum mismatch, $\Delta k = k_{\mathrm{p}z} - \boldsymbol{k}_{\mathrm{s}z} - \boldsymbol{k}_{\mathrm{i}z}$, the longitudinal integration may also be carried out:

$$\Phi(\Delta k L) \equiv \int_0^L dz e^{i\Delta k z} \equiv 2 e^{i\Delta k L} \frac{\sin \frac{\Delta k L}{2}}{\Delta k L} = e^{i\Delta k L} \mathrm{sinc}\left(\frac{\Delta k L}{2}\right), \tag{29}$$

where the sinc function is defined by $\mathrm{sinc}(x) = \frac{\sin x}{x}$. In the limit of a long crystal, $L \to \infty$, this phase-matching function becomes a delta function for the longitudinal momenta: $\lim_{L\to\infty}\Phi(\Delta kL) = \pi\delta(\Delta k)$.

The result, finally, is that the relevant part of the interaction Hamiltonian may be written as:

$$H_{\mathrm{int}}(t) = C' \int d^3k_{\mathrm{s}} d^3k_{\mathrm{i}} \Phi(\Delta kL) e^{i(\omega_{\mathrm{p}}-\omega_{\mathrm{s}}-\omega_{\mathrm{i}})t} \delta^{(2)}(\boldsymbol{k}_{\mathrm{s}\perp} + \boldsymbol{k}_{\mathrm{i}\perp}) \hat{a}^\dagger(\boldsymbol{k}_{\mathrm{s}}) \hat{a}^\dagger(\boldsymbol{k}_{\mathrm{i}}) + h.c. \quad (30)$$

This, of course, must be supplemented by the appropriate dispersion relations connecting the frequencies to the wavevectors in the birefringent crystal. The resulting *phase matching conditions* (equivalent to energy–momentum conservation) that must be satisfied by the outgoing fields are thus dependent on the polarizations of the photons. The down conversion is called *Type I* if the signal and idler have the same polarization (opposite to the pump), and *Type II* if the signal and idler have opposite polarizations to each other. Henceforth, we assume Type II parametric down conversion, $\mathrm{e} \to \{\mathrm{e}, \mathrm{o}\}$, with o being the idler and the pump and signal both being e-polarized. (o and e denote ordinary and extraordinary polarizations.)

For a weak interaction Hamiltonian $\hat{H}_{\mathrm{int}}$ which is only non-zero for times in the interval $-T < t < T$, perturbation theory tells us that $\hat{H}_{\mathrm{int}}$ will transform an initial vacuum state (before the interaction) $|\mathrm{vac}\rangle$ into a new state $|\Psi\rangle$ afterwards:

$$|\Psi\rangle = \int_{-T}^{T} dt e^{-\frac{i}{\hbar}\hat{H}_{\mathrm{int}}t} |\mathrm{vac}\rangle = (1 - \frac{i}{\hbar}\int_{-T}^{T} dt H_{\mathrm{int}} + \cdots)|\mathrm{vac}\rangle. \quad (31)$$

Taking $T \to \infty$, the time integration becomes $\int_{-\infty}^{\infty} dt e^{i(\omega_{\mathrm{p}}-\omega_{\mathrm{s}}-\omega_{\mathrm{i}})t} = 2\pi\delta(\omega_{\mathrm{p}} - \omega_{\mathrm{s}} -\omega_{\mathrm{i}})$. Using the Hamiltonian of Eq. 30, we have the biphoton state:

$$\begin{aligned} |\Psi\rangle = -\frac{iC'}{\hbar} \int & d^3k_{\mathrm{e}} d^3k_{\mathrm{o}} \delta(\omega_{\mathrm{e}} + \omega_{\mathrm{o}} - \omega_{\mathrm{p}}) \\ & \times \delta^{(3)}(\boldsymbol{k}_{s\perp} + \boldsymbol{k}_{i\perp}) \Phi(\Delta k\ L) \hat{a}_H^\dagger(\boldsymbol{k}_{\mathrm{e}}) \hat{a}_V^\dagger(\boldsymbol{k}_{\mathrm{o}})|vac\rangle. \end{aligned} \quad (32)$$

Using the dispersion relations, the $k$ integrations may be rewritten as frequency integrations, so:

$$|\Psi\rangle = \int d\omega_{\mathrm{e}} d\omega_{\mathrm{o}} e^{-i\Delta kL/2} E(\omega_{\mathrm{e}} + \omega_{\mathrm{o}}) \Phi(\omega_{\mathrm{e}}, \omega_{\mathrm{o}}) |\omega_{\mathrm{e}}\rangle_H |\omega_{\mathrm{o}}\rangle_V, \quad (33)$$

where we have generalized the situation to include a non-plane-wave pump with envelope $E(\omega_{\mathrm{p}}) = E(\omega_{\mathrm{e}} + \omega_{\mathrm{o}})$. The momentum mismatch is now written in terms of frequency:

$$\Delta k = \frac{1}{c}\left[n(\omega_{\mathrm{p}})\omega_{\mathrm{p}} - n(\omega_{\mathrm{e}})\omega_{\mathrm{e}} - n(\omega_{\mathrm{o}})\omega_{\mathrm{o}}\right]. \quad (34)$$

Due to the non-factorability of $\Phi(\omega_{\mathrm{e}}, \omega_{\mathrm{o}})$ into a product of terms each involving only one of the frequencies, the state of Eq. (33) is clearly entangled in terms of the various frequency states. It is also entangled in polarization; in particular, if the frequencies are held fixed (by means of filters, for example), we have $|\Psi\rangle = \frac{1}{2}[|H\rangle_s|V\rangle_i + |V\rangle_s|H\rangle_i]$. The latter is a realization of the well-known Bell state $|\psi^+\rangle$, i.e. a maximally entangled bipartite state.

We now look at several ways of quantifying the entanglement of the biphoton state.

### 1.5. *Concurrence in down conversion*

If the state of a system is known, then one readily computable measure of entanglement is the concurrence. Given a two-qubit pure state $|\Psi\rangle$ on Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, define the *spin-flipped state* $|\tilde{\Psi}\rangle = \sigma_2^{(A)} \otimes \sigma_2^{(B)}|\Psi\rangle$. More generally, the spin-flipped state corresponding to two-qubit state $\hat{\rho}$ is $\tilde{\rho} = \sigma_2^{(A)} \otimes \sigma_2^{(B)}\hat{\rho}\sigma_2^{(A)} \otimes \sigma_2^{(B)}$. Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ denote the eigenvalues of the density operator $\sqrt{\hat{\rho}\tilde{\rho}}$, in descending order. Then the *concurrence* of the bipartite system is defined to be

$$C = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}. \tag{35}$$

For a pure state, this reduces to an inner product: $C = \langle\Psi|\tilde{\Psi}\rangle$. (A more general definition applying to bipartite systems of arbitrary dimension can be given.[5])

The concurrence in the frequency spectrum of Type II SPDC has been calculated by Grice and Walmsley[20] in approximate form and more exactly by Erenso.[21] Here we follow the latter.

If the spectral bandwidth of downconversion is relatively small, the phase matching function $\Phi(\omega_1, \omega_2)$ is approximately symmetric in the frequencies. However, due to the birefringence of the crystal, the function necessarily shows noticeable asymmetry at larger bandwidths. As a result, we write the Type II two-photon down conversion state as:

$$|\Psi\rangle = C\int d\omega_{\mathrm{e}}d\omega_{\mathrm{o}}E(\omega_{\mathrm{e}} + \omega_{\mathrm{o}})(\Phi(\omega_{\mathrm{e}}, \omega_{\mathrm{o}})|\omega_{\mathrm{e}}\rangle_H|\omega_{\mathrm{o}}\rangle_V + \Phi(\omega_{\mathrm{o}}, \omega_{\mathrm{e}})|\omega_{\mathrm{o}}\rangle_H|\omega_{\mathrm{e}}\rangle_V), \tag{36}$$

where

$$\Phi(\omega_{\mathrm{e}}, \omega_{\mathrm{o}}) = \frac{\sin([k_{\mathrm{o}}(\omega_{\mathrm{o}}) + k_{\mathrm{e}}(\omega_{\mathrm{e}}) - k_{\mathrm{p}}(\omega_{\mathrm{o}} + \omega_{\mathrm{e}})]L)}{[k_{\mathrm{o}}(\omega_{\mathrm{o}}) + k_{\mathrm{e}}(\omega_{\mathrm{e}}) - k_{\mathrm{p}}(\omega_{\mathrm{o}} + \omega_{\mathrm{e}})]L} \tag{37}$$

and we take the spectral envelope function of the pump to be

$$E(\omega_{\mathrm{e}} + \omega_{\mathrm{o}}) = \exp\left\{-\frac{[2\omega_0 - (\omega_{\mathrm{o}} + \omega_{\mathrm{e}})]^2}{s\sigma_{\mathrm{p}}^2}\right\}. \tag{38}$$

Here, $2\omega_0$ and $\sigma_{\mathrm{p}}$ are the central frequency and bandwidth of the pump. Expanding $k_{\mathrm{e}}$ and $k_{\mathrm{o}}$ about $\omega_0$, and $k_{\mathrm{p}}$ about $2\omega_0$, we find that

$$\Phi(\omega_1, \omega_2) = \frac{\sin(\tau_1\nu_1 + \tau_2\nu_2)}{[\tau_1\nu_1 + \tau_2\nu_2]L}. \tag{39}$$

In the latter expression, $\nu_j = \omega_j - \omega_0$ (for $j = 1, 2$) are the frequency detunings of the two photons, while $\tau_j = (\frac{\partial k_{\mathrm{p}}}{\partial\omega}|_{\omega=2\omega_0} - \frac{\partial k_j}{\partial\omega}|_{\omega=\omega_0})\, L$ are the differences in time delay of the pump photon relative to photon $j$ during transit through the crystal. So the state is

$$|\Psi\rangle = \mathcal{N}\int d\nu_{\mathrm{e}} d\nu_{\mathrm{o}} \{f(\nu_{\mathrm{e}},\nu_{\mathrm{o}})|\omega_{\mathrm{e}}\rangle_H|\omega_{\mathrm{o}}\rangle_V + f(\nu_{\mathrm{o}},\nu_{\mathrm{e}})|\omega_{\mathrm{o}}\rangle_H|\omega_{\mathrm{e}}\rangle_V\}, \tag{40}$$

where

$$f(\nu_i,\nu_j) = \mathcal{N} e^{-\frac{(\nu_i+\nu_j)2}{2\sigma_{\mathrm{p}}^2}} \frac{\sin([\tau_1\nu_1+\tau_2\nu_2])}{[\tau_1\nu_1+\tau_2\nu_2]L}. \tag{41}$$

Computing the density operator and its eigenvalues, the concurrence is then given by

$$C = \mathcal{N}^2 \int d\nu_{\mathrm{e}} d\nu_{\mathrm{o}} e^{-\frac{(\nu_{\mathrm{e}}+\nu_{\mathrm{o}})2}{2\sigma_{\mathrm{p}}^2}} \frac{\sin([\tau_1\nu_{\mathrm{e}}+\tau_2\nu_{\mathrm{o}}])}{[\tau_1\nu_{\mathrm{e}}+\tau_2\nu_{\mathrm{o}}]L}\, \frac{\sin([\tau_1\nu_{\mathrm{o}}+\tau_2\nu_{\mathrm{e}}])}{[\tau_1\nu_{\mathrm{o}}+\tau_2\nu_{\mathrm{e}}]L}. \tag{42}$$

This expression may be readily plotted as a function of transit times for a given pump beam and crystal. Examples of such plots were constructed by Erenso,[21] in which it can be seen that when the transit time of the pump beam through the crystal is small compared to that of the signal and idler, the concurrence is close to one. As the pump transit time decreases relative to the others, the concurrence decays. Thus, one means to control the degree of spatial entanglement is to alter the frequency and polarization dependence of the index of refraction, thus altering the transit times.

### 1.6. *Schmidt number and von Neumann entropy*

One of the most useful tools in quantum information theory is the *Schmidt decomposition.*[22] In Schmidt form, a bipartite state vector is “diagonal”, in the sense that the basis vectors of the first and second Hilbert spaces are matched up in one-to-one fashion,

$$|\Psi\rangle = \sum_{i=1}^{d_{\min}} \sqrt{\lambda_i}|u_i\rangle|v_i\rangle, \tag{43}$$

where $d_{\min}$ is the dimension of the smaller of the two Hilbert spaces. $\lambda_i$ is the $i$th eigenvalue of the density matrix, and so gives the probability of measuring the $i$th term in the expansion, $p_i = \lambda_i$. The quantum correlations present in entangled systems are now manifest, in this form: whenever the first system is measured to be in state $|u_i\rangle$, the second system is guaranteed to be in state $|v_i\rangle$. The number of non-zero terms in the expansion is known as the *Schmidt number*, $K$, and serves as a simple measure of entanglement: $K = 1$ for an unentangled product state, and increasing with increasing number of entangled states in the sum. Interpreting $\lambda_k$ as the probability of the $k$th state, the *average* probability per state in the sum is

$\sum_k p(k)\lambda_k = \sum_k \lambda_k^2$, so the average effective number of non-zero components in the decomposition is $1/\sum \lambda_i^2$. Thus, if the Schmidt decomposition is known, then the Schmidt number can be computed from the coefficients[23]:

$$K = 1 \Big/ \sum \lambda_i^2 \tag{44}$$

Being essentially a count of available states, the Schmidt number is bounded above by the number of states that can fit into the phase space volume accessible to the system. So $K$ is finite, even for continuous degrees of freedom, as long as the available phase space volume is finite.

Once the system is put into Schmidt form, the *von Neumann entropy* can then be computed:

$$S(\hat{\rho}) = -\mathrm{tr}\hat{\rho}\log_2\hat{\rho} = -\sum_i \lambda_i \log_2 \lambda_i. \tag{45}$$

The von Neumann entropy is a measure of the mixedness of a state: $S(\hat{\rho}) = 0$ for a pure state $\hat{\rho} = |\psi\rangle\langle\psi|$ and attains a maximum of $\log_2 d$ for the maximally mixed state $\hat{\rho} = \frac{1}{d}\hat{I}$. The von Neumann entropy is the quantum analog of the Shannon entropy to be discussed in the next section, and is essentially a measure of the information gained by measurement of the state. An explicit recipe can be constructed for putting a state into Schmidt form. Consider some pure state $|\Psi\rangle = \sum_{ij} C_{ij}|u_i'\rangle|v_j'\rangle$, so that the density operator is of the form

$$\hat{\rho} = \sum_{ijkl} C_{ij}C_{kl}^{*}|u_i'\rangle\langle u_k'| \otimes |v_j'\rangle\langle v_l'|. \tag{46}$$

We first rotate from the $|u_i'\rangle$ basis to the basis $|u_i\rangle$ in which $\hat{\rho}_u = \mathrm{tr}_v\hat{\rho}$ is diagonal. $\sum_{ij} C_{ij}C_{kl}^{*} = \delta_{ik}|g_i|^2$ in this basis, for some constants $g_i$, and $\hat{\rho}_u = \sum_i |g_i|^2|u_i\rangle\langle u_i|$. For each non-zero $g_i$, we also define a new basis for the second Hilbert space, $|v_i\rangle = \sum_j \frac{C_{ij}}{|g_i|}|v_j'\rangle$. We then find that

$$\hat{\rho} = \sum_{ik} |g_i||g_k||u_i\rangle\langle u_k| \otimes |v_i\rangle\langle v_k| = |\Psi\rangle\langle\Psi|, \tag{47}$$

with corresponding state vector $|\Psi\rangle = |g_i||u_i\rangle|v_i\rangle$, which is of Schmidt form. Therefore, $\sqrt{\lambda_i} = |g_i|$.

The two-photon state in Type II SPDC can be written

$$|\psi\rangle = \int \Phi(\omega_1,\omega_2)a_H^{\dagger}(\omega_1)a_V^{\dagger}(\omega_2)|0\rangle_H|0\rangle_V d\omega_1 d\omega_2, \tag{48}$$

and the spectral amplitude $\mathcal{A}(\omega_1,\omega_2)$ then decomposed into Schmidt form:

$$\mathcal{A}(\omega_1,\omega_2) = \sum_n \sqrt{\lambda_n}\psi_n(\omega_1)\phi_n(\omega_2), \tag{49}$$

where the eigenvalues and eigenfunctions $\lambda_n$, $\psi_n$, and $\phi_n$ are solutions to the integral equations

$$\int K_1(\omega,\omega')\psi_n(\omega')d\omega' = \lambda_n\psi_n(\omega), \tag{50}$$

$$\int K_2(\omega,\omega')\phi_n(\omega')d\omega' = \lambda_n\psi_n(\omega). \tag{51}$$

The integral kernels in these equations are given by

$$K_1(\omega,\omega') = \int \mathcal{A}(\omega,\omega_2)\mathcal{A}^*(\omega',\omega_2)d\omega_2, \tag{52}$$

$$K_2(\omega,\omega') = \int \mathcal{A}(\omega_1,\omega)\mathcal{A}^*(\omega_1,\omega')d\omega_1. \tag{53}$$

The eigenfunctions $\psi_n$ and $\phi_n$ can be used to define a new set of effective creation operators for horizontally and vertically polarized photons,

$$\hat{b}_n^\dagger = \int \psi_n(\omega_1)\hat{a}_H^\dagger(\omega_1)d\omega_1, \quad \hat{c}_n^\dagger = \int \phi_n(\omega_2)\hat{a}_V^\dagger(\omega_2)d\omega_2. \tag{54}$$

In terms of these, we can rewrite the Schmidt decomposition of the biphoton state as

$$|\psi\rangle = \sum_n \sqrt{\lambda_n}\hat{b}_n^\dagger\hat{c}_n^\dagger|\text{vac}\rangle_H|\text{vac}\rangle_V. \tag{55}$$

For SPDC, we may split the amplitude into a pump envelope and a phase-matching function $\Phi$: $\mathcal{A}(\omega_1,\omega_2) = \tilde{E}(\omega_1+\omega_2)\Phi(\omega_1,\omega_2)$. Law, Walmsley, and Eberly[24] have calculated the eigenvalues for this case and found that the sizes of the terms in the sums of Eqs. (49) and (55) drop rapidly, leaving only a small number of eigenvalues of non-negligible size. As a result, the *effective Schmidt number K* of the spectrally-entangled system is in fact relatively small. In fact, for the parameter values they used, the authors found that 96% of the state could be accounted for by the first six eigenvalues. The von Neumann entropy computed from these first six eigenvalues gives a value $S = 1.4$, compared to the large $K$ limit of 1.8. By narrowing the bandwidth, correlations between the spectral components increases. As a result, the von Neumann entropy and the effective Schmidt number both increase. Bandwidth therefore determines the level of entanglement present in the current situation.

Rather than frequency entanglement, we can take a similar approach to quantify the entanglement in some other degree of freedom, for example the spatial entanglement carried by the momentum vectors. This was investigated by Law and Eberley,[25] as follows. Let $\boldsymbol{k}$ and $\boldsymbol{q}$ be the transverse spatial momenta of the two photons. (Transverse here means perpendicular to the direction of the pump beam, taken to be along the $z$-axis.) As a simple model of downconversion that allows analytic calculation of the Schmidt number, take the biphoton amplitude in transverse momentum space to be of Gaussian form,

$$\mathcal{A}(\boldsymbol{k},\boldsymbol{q}) = E(\boldsymbol{k}+\boldsymbol{q})\Phi(\boldsymbol{k}-\boldsymbol{q}) = C_g e^{-\frac{|\boldsymbol{k}+\boldsymbol{q}|^2}{\sigma^2}} e^{-b^2|\boldsymbol{k}-\boldsymbol{q}|^2}, \tag{56}$$

where the two terms represent the pump envelope and the phase matching function in momentum space. For this form, the Schmidt number can be found exactly[25]: $K = \frac{1}{4}(b\sigma + \frac{1}{b\sigma})^2$. The degree of entanglement thus depends only on $b\sigma$, the ratio of widths of the two exponentials. $K$ increases whenever $b\sigma \gg 1$ or $b\sigma \ll 1$, with a minimum at $b\sigma = 1$.

The Gaussian form given above is unrealistic. A more realistic approximation for the amplitude is given by replacing the second exponential (the phase-matching term) by a sinc function, as we have seen in Sec. 1.4:

$$\mathcal{A}(\boldsymbol{k}, \boldsymbol{q}) = E_{\text{p}}(\boldsymbol{k} + \boldsymbol{q})\Phi(\boldsymbol{k} - \boldsymbol{q}) = C_g e^{-\frac{|\boldsymbol{k}+\boldsymbol{q}|^2}{\sigma^2}} \text{sinc}(b^2|\boldsymbol{k} - \boldsymbol{q}|^2), \tag{57}$$

where $b^2 = L/4k_{\text{pump}}$. The Schmidt number now has to be calculated numerically, but the result is qualitatively similar to the Gaussian model, with $K$ becoming large whenever $b\sigma$ is either much larger or much smaller than 1.[25] Thus, spatial entanglement can be increased by, for example, increasing the transverse momentum spread. For some parameter ranges, the effective number of states $K$ can be in the hundreds, but not all of these states are necessarily accessible. We will return to this issue in Sec. 3.

The analysis of Law and Eberly[25] has been generalized by van Exter *et al.*[26] Among other things, these authors showed that a one-dimensional Schmidt number $K_{\text{1d}}$ can be calculated for photon pairs confined to propagate in a single plane, and that the full two-dimensional Schmidt number is simply $K_{\text{2d}} = K_{\text{1d}}^2$. They also added in the effect of a finite-sized detection aperture (diameter $a$), showing that in this case

$$K_{\text{2d}} = \frac{(1/\sigma^2 + b^2 + a^2)^2}{(1/\sigma^2 + b^2 + a^2)^2 - (\frac{1}{\sigma^2} - b^2)^2}. \tag{58}$$

This decreases asymptotically to $K_{\text{2d}} = 1$ as $a \to \infty$, demonstrating the role of spatial filtering by the detector and reminding us that the degree of entanglement, as well as the information content, will be dependent on our measuring devices and is not entirely intrinsic to the system being measured.

How is the Schmidt number measured experimentally? It can be shown[27] that for transverse spatial modes in the quasi-homogeneous approximation, the Schmidt number can be written in a form analogous the étendue[28] of an optical system:

$$K = \frac{1}{\lambda^2} \frac{[\int I_S(\boldsymbol{x}) d\boldsymbol{x}]^2}{\int I_S^2(\boldsymbol{x}) d\boldsymbol{x}} \times \frac{[\int I_{\text{FF}}(\theta) d\theta]^2}{\int I_{\text{FF}}^2(\theta) d\theta}, \tag{59}$$

where $I_S$ and $I_{\text{FF}}$ are the near-field (source) and far-field intensities. Thus intensity measurements in two planes suffice to determine the Schmidt number.

The Schmidt number for the output of SPDC depends strongly on the properties of both the pump beam and the crystal. For some parameter ranges, it can be extremely large; for example, in the experiment of Dixon *et al.*,[29] the number of product states superposed in the outgoing spatially-entangled biphoton state was $K \sim 1400$! In contrast, we have seen that for the parameter values considered by

Law, Walmsley, and Eberly,[24] the effective number of polarization-entangled terms was very small, on the order of $K \approx 2$. This is one of the reasons that down conversion is such an important source for optical experiments: by appropriately tuning the input parameters or measuring different variables, we can exert a great deal of control over the output state and can vary its properties over a very wide range.

### 1.7. *Other measures of entanglement*

Many other measures of entanglement have been defined (see Plenio and Virmani[30] for comprehensive reviews). Here, we briefly mention a couple of these.

For a bipartite system on a Hilbert space $\mathcal{H}_A \times \mathcal{H}_B$, the *partial transpose* operations $T_A$ or $T_B$ consist of taking the transpose of the part of an operator's action only on one of the two subsystems. Thus, for example, the partial transpose of a density operator relative to the $A$ subsystem is defined by $\langle i_A j_B|\hat{\rho}^{T_A}|k_A l_B\rangle = \langle k_A j_B|\hat{\rho}|i_A l_B\rangle$. According to the *Peres-Horodečki criterion*,[31,32] a system is entangled if (either) partial transpose of the density matrix is negative. This can be associated with a numerical measure by defining the *negativity*:

$$\mathcal{N}(\hat{\rho}) = \frac{1}{2}(||\hat{\rho}^{T_A}||_1 - 1), \tag{60}$$

where $||\hat{\mathcal{O}}||_1 \equiv \mathrm{tr}\sqrt{\mathcal{O}^\dagger \mathcal{O}}$ is the trace-norm of Hermitian operator $\mathcal{O}$. This may also be expressed as $\mathcal{N}(\hat{\rho}) = |\sum_i \lambda_i|$, where $\lambda_i$ represent the *negative* eigenvalues of $\hat{\rho}^{T_A}$. The negativity is bounded by the concurrence, $\mathcal{N}(\hat{\rho}) \leq c(\hat{\rho})$.

A further fundamental entanglement measure that can be related to the concurrence is the entanglement of formation: $E_f(\hat{\rho}) = h(c(\hat{\rho}))$, where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$; see Ref. 5 for more information.

## 2. Communication and Cryptography

### 2.1. *Information and channel capacity*

In the 1940's, the major problems in telecommunications included the questions of how to quantify the amount of information being carried on a communication channel, how to determine what the maximum information a given channel could carry, and how to understand the effect of noise on information capacity. These questions were largely answered by Shannon and his contemporaries. Here we briefly discuss these questions and their generalizations to quantum theory. Then in the following subsections, we look at some communication phenomena that exist only in the quantum case.

The most basic quantity in classical information theory is the Shannon entropy. Given random variable $X$, we will denote the possible values that it can take by $x_1, x_2, x_2, \ldots$; $x$ will be used to denote a generic value. These values occur according to some probability distribution $p(X)$. We will restrict ourselves here to discrete distributions for simplicity. Then the *Shannon entropy* associated with variable $X$ is

$$H(X) = -\sum_i p(x_i)\mathrm{log}_2 p(x_i) = -E[\log_2(X)], \tag{61}$$

where $E$ denotes expectation value or mean.

The significance of $H(X)$ is that it tells you the "surprise value" or average amount of new knowledge you gain from a measurement of $X$. For example, consider a variable $X$ which can take on two values $x_1$ and $x_2$. If $p(x_1) = 1$ and $p(x_2) = 0$ (a state of maximal *a priori* knowledge), then the Shannon information vanishes; this is as expected from the fact that we know $X$ will always take the value $x_1$, so a measurement tells us nothing new. In contrast, if $p(x_1) = p(x_2) = \frac{1}{2}$ (the state of maximal *a priori* uncertainty) the entropy reaches its maximum value ($H(X) = \log_2 2 = 1$), since in this case we learn the most from each measurement.

The entropy depends only on the probability distribution associated with the random variable, $H(X) = H(p(X))$, is concave, and is non-negative: $H(X) \geq 0$ for all $X$, with equality if and only if only a single value of $X$ has non-zero probability. Conceptually, the Shannon entropy is a measure of how much redundancy is occurring in a message, or equivalently how much the message can be compressed. This is the content of the Shannon noiseless coding theorem: a message of length $n$ can be coded by a string of only $nH$ bits, as $n \to \infty$.

Similarly, the *Shannon noisy coding theorem* tells how much additional redundancy must be encoded into a message transmitted over a noisy channel in order to allow for error correction. For the simplest case, a binary symmetric channel with error probability $q$ per bit, the theorem says that each binary digit may carry no more than $1 - h(q)$ bits of information, where $h(q) = q\mathrm{log}_2(q) + (1-q)\mathrm{log}_2(1-q)$ is the entropy of the error probability distribution. $h(q)$ serves as a measure of the amount of redundancy that must be built into a message to enable error correction.

The *von Neumann entropy* was introduced in the last section, and can be viewed as the quantum analog of the Shannon entropy for a quantum state $\hat{\rho}$:

$$S(\hat{\rho}) = -\mathrm{tr}\hat{\rho}\mathrm{log}_2\hat{\rho} = -\sum_i \lambda_i \log_2 \lambda_i, \tag{62}$$

where the $\lambda_i$ are the Schmidt coefficients. $S(\hat{\rho})$ is a measure of the mixedness of the state: for a system of dimension $n$, the von Neumann entropy is bounded by $0 \leq S(\hat{\rho}) \leq \mathrm{log}_2 n$, with the lower limit reached by pure states and the upper limit achieved for the maximally mixed states $\hat{\rho} = \frac{1}{n}\hat{I}$.

For a statistical mixture of states $\hat{\rho} = \sum_i p_i \hat{\rho}_i$ it can be shown that

$$\sum_i p_i S(\hat{\rho}_i) \leq S(\hat{\rho}) \leq H(p_1, \ldots, p_n) + \sum_i p_i S(\hat{\rho}_i). \tag{63}$$

The left-hand inequality simply expresses the concavity of the von Neumann entropy; as for the right hand inequality, the first term on the right describes the classical uncertainty due to the statistical mixture of the states, while the second term describes the uncertainty inherent in the quantum states themselves. If the $\hat{\rho}_i$ are pure states, the latter terms vanish, so that $S(\hat{\rho}) \leq H(p_1, \ldots, p_n)$; thus the quantum uncertainty is less than the uncertainty of the corresponding classical system. This is a reflection of the fact that quantum systems can contain correlations stronger than the ones that are possible classically, as reflected by the well-known Bell inequalities.[33,34]

The von Neumann entropy to a large extent plays a role in quantum systems similar to that of the Shannon entropy in classical systems. For example, there is a theorem (the Schumacher theorem[1]) for quantum systems analogous to that of the Shannon noiseless coding theorem, with $S$ replacing $H$.

Rather than investigating the formal properties of entropy and information in detail, we move on in the next Section to discuss attempts to communicate secretly by means of encryption keys shared between two parties. We will see that the laws of quantum mechanics will prevents an eavesdropper from gaining information about the key without causing disturbances that can be detected by the communicating parties.

### 2.2. *Quantum key distribution*

The goal is to generate a secret key for encrypting and decrypting messages that is shared between two legitimate users, usually known as Alice and Bob, and which cannot be broken by an eavesdropper, usually called Eve. The only truly unbreakable code is the *one-time pad* or *Vernam cipher* in which the secret key $k$ is a random string of binary digits which is used only once, and then discarded. If the text to be encoded is given as a binary string $m$, then the encoded message is given by $m \oplus k$, where $\oplus$ is base-two addition. To decode the message, Bob simply adds the same key to the encoded message: since $k \oplus k = 0$, it follows that $m \oplus k \oplus k = m$. The randomness of the key means that there are no patterns that can be used to break the code: the key has the maximum possible entropy and carries no information. However, if the same key is used multiple times, detectable patterns in the messages themselves will cause correlations in the sum $m \oplus k$, which in principle can leak information about the messages. Therefore, it is essential that each key not be reused. Although the key itself is unbreakable, there is still the problem of distribution: Alice and Bob must use the *same* key, so Eve may be able to intercept the passing of the key from one to the other, destroying the security of the message.

The Vernam cipher solves the problem of encrypting a message in an unbreakable manner, once the participants share a random key. However, this does not solve the problem of *distributing* the key among the legitimate users without it being intercepted. Classically there is no foolproof means for completely secure key distribution; this is where quantum mechanics becomes essential. In *quantum cryptography* or *quantum key distribution* (QKD), the goal is to generate a one-time encryption key and to share it between the two legitimate users, Alice and Bob, while using the laws of quantum mechanics to prevent illegitimate eavesdroppers from obtaining the key undetected. We will see that, although the eavesdropping itself is not preventable, it will always be possible to detect it if it is occurring, so that it will be ineffective. If Eve is detected, Alice and Bob know their communication line has been compromised, so they must stop using it and seek another communication channel.

What makes QKD possible is the existence of non-commuting operators in quantum mechanics. Suppose we have two Hermitian operators $\hat{\mathcal{O}}$ and $\hat{\mathcal{O}}'$ which fail to commute: $[\hat{\mathcal{O}}, \hat{\mathcal{O}}'] \neq 0$. We assume that either (i) Alice prepares a state, makes a measurement on it and sends it to Bob, or else (ii) a third party sends an entangled pair of states (half of the pair to Alice, the other to Bob), in which the values of the relevant operators are either correlated or anticorrelated between the two states. Alice chooses randomly to measure the value of either $\hat{\mathcal{O}}$ or $\hat{\mathcal{O}}'$ on the state, obtaining some value $o_A$ which is an eigenvalue of whichever operator was used. This determines the value $o_B$ Bob will measure *if* he measures the same operator. However, if he measures the other operator, the value he obtains is random (indeterminate), due to the fact that $\hat{\mathcal{O}}$ and $\hat{\mathcal{O}}'$ are incompatible observables. The operators should be chosen so that application of one operator makes all possible eigenvalues of the other equally likely; such operators are called *mutually unbiased*, *conjugate*, or *incompatible*. The communication procedure then consists schematically of the following steps: (i) Alice generates a sequence of states. (ii) For each state, she randomly chooses $\hat{\mathcal{O}}$ or $\hat{\mathcal{O}}'$ and makes a measurement. (iii) Bob then randomly chooses $\hat{\mathcal{O}}$ or $\hat{\mathcal{O}}'$ for each state and also makes a measurement. (iv) Alice and Bob then communicate over a classical communication channel. This channel can be completely public. They tell each other which measurement operator they chose for each state, but *not* the result of the measurement. (v) They keep only those values for which they made the *same* choice, discarding the rest. This process is called *sifting.* (vi) They randomly select a subset of the sifted trials to subject to a security check. They compare the values obtained on these trials, and check to see if the these values have the correlation (or anticorrelation) expected. Unexpected drops in correlation signal the activity of an eavesdropper. (vii) If the security trials have the expected level of correlation, then they can be certain that no eavesdropping occurred. They can therefore use the values they measured on the remaining trials (after sifting and security trials) as the digits of the one-time key. Although they have not told each other their values, the fact that they measured the same operator on these correlated or anticorrelated states guarantees that each can deduce the other's value from their own.

This procedure is safe, because if Eve is intercepting the states and making her own measurements, she has no way of knowing whether Alice chose to measure $\hat{\mathcal{O}}$ or $\hat{\mathcal{O}}'$ on each trial. She has to guess, and has only a 50% probability per trial of guessing correctly. Suppose on a given trial Alice measures $\hat{\mathcal{O}}$. Then if Eve also measures $\hat{\mathcal{O}}$, she will measure the correct value $o_A$ and can generate a copy of the state to send to Bob. Bob (if he also measures $\hat{\mathcal{O}}$) will also determine the value $o_A$, and so the tampering will not be detected. But when Eve chooses to measure the wrong operator

$\hat{\mathcal{O}}'$, she will sometimes (let us say with probability $p$, where $p < 1$) measure the correct value $o_A$, but will also sometimes measure an incorrect value $o'_A$ with probability $1 - p$. When this happens, it will show up during the security check: instead of Alice and Bob agreeing 100% of the time when they used the same operator, they will find that they now only agree on a fraction $\frac{p+1}{2}$ of the trials. This drop in correlation between their values immediately signals the presence of an eavesdropper. (The strategy used here by Eve is called the *intercept-resend strategy*. It is the simplest type of eavesdropping attack. More sophisticated attacks are also possible, which may require additional safeguards.[35–39])

To make the protocol as safe as possible, we want the overlap between each eigenstate of one operator with all of the eigenstates of the other to be as uniform as possible, i.e. we want mutually unbiased operators. In the most common case, the different operators represent projections onto polarization states measured along the axes of different bases. For example, $\hat{\mathcal{O}}_j = |\psi_j\rangle\langle\psi_j|$ and $\hat{\mathcal{O}}'_k = |\psi'_k\rangle\langle\psi'_k|$. In that case, we specify the different operators by specifying the bases, and the eigenvectors represent the basis vectors. For the case of two incompatible bases, the best possible choice is $\langle\psi_i|\psi'_j\rangle = \frac{1}{\sqrt{2}}$ for all $i \in \{1, 2\}$ and $j \in \{1, 2\}$ (so that $p = 1 - p = \frac{1}{2}$). Here, $|\psi_i\rangle$ and $|\psi'_j\rangle$ are respectively the eigenvectors of $\hat{\mathcal{O}}$ and $\hat{\mathcal{O}}'$. More generally, we may use $m$ incompatible operators $\hat{\mathcal{O}}_1, \ldots, \hat{\mathcal{O}}_m$, such that

$$\langle\psi_i^{(\mu)}|\psi_j^{(\nu)\prime}\rangle = \frac{1}{\sqrt{m}}, \tag{64}$$

for all $i, j \in \{1, 2\}$ and all $\mu, \nu \in \{1, 2, \ldots m\}$. (The superscripts $\mu, \nu$ label the operator, while subscripts $i, j$ label the states within the set of eigenstates of each operator.) Bases satisfying the conditions Eq. (64) are called *mutually unbiased* or *conjugate* bases. Mutually unbiased bases are such that a measurement in one basis gives no information about the value in the other basis: a measurement in one basis completely randomizes values in the other, with a uniform probability distribution (for a review of mutually unbiased bases and their construction, Durt *et al.*[40]).

To make this more concrete, let us consider the first successful QKD method, invented by Bennet and Brassard[41] which is known as the BB84 protocol. Here, we take the states to be polarization states of a photon, and the operator $\hat{\mathcal{O}}$ to be the polarization operator in a coordinate system defined by a pair of perpendicular axis, the horizontal ($H$) and vertical ($V$) axes. We take $o_A = 0$ if the polarization is horizontal and $o_A = 1$ if it is vertical, with corresponding eigenvectors $|0\rangle = |H\rangle$ and $|1\rangle = |V\rangle$. The second operator $\hat{\mathcal{O}}'$, incompatible with $\hat{\mathcal{O}}$, is the polarization operator in a system defined by two axes ($|\nearrow\rangle$ and $|\searrow\rangle$) at $\pm 45^\circ$ to the horizontal. We will denote the eigenvectors $|0'\rangle = |\nearrow\rangle$ and $|1'\rangle = |\searrow\rangle$, and the eigenvalues $o'_A = \{0, 1\}$. These two bases are clearly mutually unbiased.

To generate a secure key, Alice randomly selects one of the two bases for each photon and makes a measurement of the polarization in that basis. She then sends the photon on to Bob, who similarly makes a random choice among the two bases and measures polarization. If they both chose the same basis, then they should always measure the same value for polarization. However, if they make different choices, then (due to the incompatibility of the bases), the result of Bob's measurement should be completely random and independent of the basis. This is the key to the security. Alice and Bob select a random subset $\mathcal{S}$ of photons to use for a security check, and tell each other (over a classical and potentially public channel) both their basis choices and the results of their measurements. The trials on which they used different bases are discarded. For the rest, they compare their measurements. Assuming ideal conditions (negligible noise, perfect detectors, etc.), their measurements should match 100% of the time if there is no eavesdropping, but only $(\frac{100-\eta}{2})\%$ of the time if Eve has intercepted and resent a fraction $\eta$ of the photons. The presence of eavesdropping is therefore immediately detectable, unless the eavesdropping rate $\eta$ is so small that Eve cannot obtain significant information anyway. If no eavesdropping has been detected, then for the remaining photons (those not in $\mathcal{S}$) classical information is again exchanged between Alice and Bob, but only concerning the choice of bases, *not* the actual polarization values in those bases. The photons for which the choices disagreed are again discarded. For the remainder, the polarizations are guaranteed to match. These polarizations then form a random sequence which is then used as the key.

A common variation on the BB84 idea is the E91 protocol. Here, rather than Alice sending a photon to Bob, a third party sends out a pair of *entangled* photons, one to Alice and one to Bob. Usually these photons were produced in Type II SPDC, so their polarizations are perfectly anticorrelated. Now Alice and Bob proceed as before, choosing bases, discarding trials on which the choices differ, checking for security by comparing measurement results on $\mathcal{S}$, and using the random sequence determined by the remaining trials as a key. (In this variation, one possible means of verifying the absence of eavesdropping is to verify that there is no decrease in Bell inequality violations.[33,34])

Other protocols are possible as well, including one that only requires the use of two non-orthogonal states.[42] A slightly different approach is to use the visibility of interference patterns instead of correlations between polarizations.[43] (For an interference pattern that oscillates between intensity values $I_{\text{min}}$ and $I_{\text{max}}$, the visibility is defined by $V = (I_{\text{max}} - I_{\text{min}})/(I_{\text{max}} + I_{\text{min}}.)$ The interference used here is not the familiar interference between *amplitudes* that occurs, for example, in the Young two-slit experiment, rather it is interference between *intensities*, involving the fourth order correlation function between the two fields or second order correlation between intensities. Two independent detectors measure intensities at different output ports of the interferometer, then each detector feeds its signal into a computer which measures the correlation function between the signals. At low intensities, when only a single photon at a time is likely to be striking the detectors, this becomes a coincidence counting setup, in which an event is registered only when both detectors see a photon simultaneously (i.e. within a very short coincidence time window). Such coincidence counting or intensity correlation experiments are common for investigating entanglement effects in quantum optics. Quantum correlations can lead to very high visibility in these experiments, close to 100%, whereas the presence of background terms reduce the maximum visibility to 70.7% in the classical case. This classical visibility limit is directly analogous to (and stems from the same source as) the Bell inequality for correlations. When an eavesdropper interferes with the photon traveling to Bob, it is detectible by a sudden drop in the visibility of the interference pattern. For more details on this approach, see Sergienko *et al.*[43]

### 2.3. *Quantum ghost imaging and secure image distribution*

In order to prepare for applications in the next section, we now discuss that idea of forming images through spatial correlations between pairs of photons. This is two-photon imaging process is known as *ghost imaging* or *correlated imaging*. The spatial correlations involved may be either classical correlations or quantum mechanical correlations due to entanglement. Although ghost imaging has been achieved using classically-correlated light sources,[44–50] we focus here on the original version of ghost imaging (*quantum ghost imaging*),[51,52] which relies on pairs of entangled photons produced by SPDC. The essential idea, shown schematically in Fig. 2(a) is that one photon encounters the object to be imaged, then passes on to a single-pixel detector, known as a *bucket detector*, $D_A$, which has no spatial resolution. This detector only registers the presence or absence of a photon, recording no information about the spatial location or momentum. The other photon does not encounter the object at all, but proceeds directly to a second, spatially-resolving detector, $D_B$. Clearly, neither detector by itself is capable of imaging the object: one detector gains no spatial information, the other detector only sees photons that never interact with the object. But when the detectors are connected to a coincidence circuit, the image reappears in the coincidence rate between the detectors (i.e in the intensity correlations). The

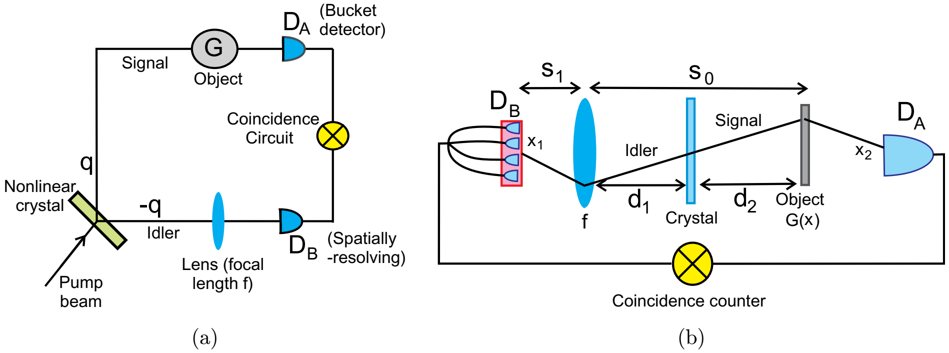


Fig. 2. (a) Quantum ghost imaging with entangled photons. (b) Klyshko backward-wave picture, in which the signal and idler are treated as a single ray passing through the crystal from one detector to the other.

photon interrogating the object essentially acts as a gate, which opens the detection window for the second photon only when the first photon has not been blocked by the object. The second photon then provides the spatial information needed to reconstruct the image.

The signal is transmitted from the object, then detected by bucket detector $D_A$. $D_A$ should be large enough to collect all of the signal photons arriving at the right end of the apparatus. It only registers whether the photons passed through the object or were blocked. Detector $D_B$ on the other hand has high spatial resolution: it can be a CCD camera, an array of avalanche photodiodes, or a single small detector scanned over the imaging region. The lens in branch 2 has focal length $f$. Let $d_1$ and $d_2$ be the distances from the source to object and source to lens, and let $s_0 = d_1 + d_2$. $s_1$ is the distance from the lens to detector $D_B$. The distances $s_0$ and $s_1$ satisfy the imaging condition $\frac{1}{s_0} + \frac{1}{s_1} = \frac{1}{f}$. When the information from the two detectors is combined via coincidence counting, the image reappears if the coincidence rate is plotted versus position in $D_B$. The imaging process is therefore highly non-local; in fact, the original motivation of this line of inquiry was to investigate the non-local causal structure of quantum mechanics and the EPR "paradox". Although it is now clear that only classical correlation between the spatial degrees of freedom is required for the information between the two photons to be correctly integrated, we will look below at a variation in which true entanglement is needed.

The imaging property of the apparatus is more clearly shown by displaying a schematic version drawn in the Klyshko"backward wave" picture[53,54] (Fig. 2(b)). Here, we view the signal and idler as a single photon passing through the crystal. The signal is viewed as traveling backward from the object, into the crystal, where it converts into the forward-moving idler, then travels onward to the detector $D_B$. The detector $D_A$ acts like the source in this view. Alternatively, we could fold the picture over, so the signal appears to reflect off the crystal in order to form the idler. In this latter version, the crystal acts as a mirror, and the pump determines the properties of that mirror. We will assume that the pump beam is approximately a plane wave, so the crystal acts as a planar mirror.

The ghost imaging apparatus has improved resolution compared to the image formed in ordinary imaging with a comparable single lens, and in fact beats the usual diffraction limit by a factor of 2. Effectively, the resolution is determined by the shorter pump frequency, rather than the longer signal or idler frequencies of the detected photons. This fact has formed the basis for the process of *quantum lithography*,[55,56] in which two-photon imaging is used to write subdiffraction-sized structures onto a semiconductor surface. The idea has been extended to $N$-photon imaging with $N > 2$, although the prospects for this to become practical seem limited, due to the difficulty of producing sufficiently entangled states of more than two photons.

One question that could be asked at this point is whether we can use quantum mechanics to securely transmit images from Alice to Bob. Of course, the answer is obviously yes, since we can encode an image digitally and then encode with a quantum key, as in Sec. 2.2. But can we use some variation on ghost imaging to accomplish secure quantum image distribution in an analog manner, without digitizing? Suppose that the object whose image is to be transmitted is in Alice's lab, along with the bucket detector, $D_A$. The spatially-resolving detector $D_B$ is in Bob's lab, and Alice wishes to send the object's image to him, while keeping it safe from eavesdroppers. Note that if Alice and Bob have detectors with sufficient time resolution and which have been well-synchronized (taking the transit time from $A$ to $B$ into account), then in place of using a coincidence counter they can simply compare the times at which they have detected photons and discard those times at which they did not make simultaneous detections. So Alice may send a list of her detection times (via a classical channel) to Bob, who then compares it with his list of detection times, thus determining the coincidence times. Since Bob also knows the spatial locations (the specific pixels) of each detection event, he can now reconstruct the image. To anyone eavesdropping on the classical channel, the list of random detection times is meaningless, unless they have also intercepted the quantum channel (Bob's photon), which contains the spatial information. To prevent this, Alice and Bob can use the same means as in the E91 protocol: they place polarizers, randomly switching between two bases, in front of the detectors. Alice then sends the choice of polarization basis along with the detection times. They keep only events on which they chose the same basis. By comparing the polarizations measurements on a random subset, eavesdropping may be detected, exactly as before.

The procedure is essentially a three-dimensional version of the E91 protocol, where the two transverse spatial dimensions plus time replace the two-dimensional qubit space of the conventional E91 case. This indicates that it might be advantageous to investigate more generally what happens when we replace our two-dimensional qubits with quantum degrees of freedom belonging to higher-dimensional Hilbert spaces. This will be the topic of the next section.

## 3. Qudits and Imaging

The generalization from qubits built from a two-dimensional effective Hilbert space spanned by states $|0\rangle$ and $|1\rangle$ to a $d$-dimensional *qudit* on a space spanned by states $|0\rangle, \ldots, |d-1\rangle$ is obvious:

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle \cdots + a_{d-1}|d-1\rangle, \tag{65}$$

with $\sum_{i=0}^{d-1} |a_i|^2 = 1$. These are known as *qutrits* for $d = 3$ and *ququats* for $d = 4$.

Since we will again be looking at QKD, we need to find sets of mutually unbiased (or mutually complementary) bases for these states. Let $m$ be the number of bases we seek, so we seek sets of basis vectors $|\psi_i^{(\mu)}\rangle$, where $\mu = 1, \ldots, m$ labels the basis, while $i = 0, 1, \ldots, d-1$ labels the vector in that basis. We then require orthonormality,

$$|\langle\psi_i^{(\mu)}|\psi_j^{(\mu)}\rangle| = \delta_{ij}, \quad \text{for all } \mu, i, j \tag{66}$$

and mutual complementarity (mutually unbiasedness),

$$|\langle\psi_i^{(\mu)}|\psi_j^{(\nu)}\rangle| = \frac{1}{\sqrt{d}}, \quad \text{for all } i, j, \mu \neq \nu. \tag{67}$$

Wootters and Fields[57] showed that there exist $m = d + 1$ mutually unbiased bases whenever $d = p^k$, with $p$ prime and $k$ non-negative integer.

For $m = 2$ bases and $d$ dimensions, we can take one basis arbitrarily, $|\psi_0^{(1)}\rangle, |\psi_1^{(1)}\rangle, \ldots, |\psi_{d-1}^{(1)}\rangle$, then construct the second basis according to

$$|\psi_k^{(2)}\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} e^{2\pi ikn/d} |\psi_n^{(1)}\rangle. \tag{68}$$

It has been shown that higher values of $d$ and $m$ can lead to both higher capacity and improvements in security against eavesdropping.[58–62] In addition, they maintain their security in the face of greater amounts of noise. Consider a pure state $|\psi\rangle$, and add some some admixture of noise $F$ $(0 \leq F \leq 1)$ by defining the density operator $\hat{\rho} = (1 - F)|\psi\rangle\langle\psi| + F\hat{\rho}_{\text{noise}}$, where $\hat{\rho}_{\text{noise}} = \frac{1}{9}I$ is the density matrix for a completely chaotic system. Einstein's conditions on a physical theory, represented in the EPR assumptions, have come to be known in the physics literature as *local realism*. These preconditions have turned out to be too strong but do not preclude either locality or realism.[5,63,64] For $F$ too large, apparent Bell inequality violations can be due to noise-induced errors, and so the security of quantum cryptography breaks down. The value $F$ for which this occurs is $\frac{2-\sqrt{2}}{2} = 0.293$ for maximally entangled qubits $(d = 2)$; in contrast, this value increases to $\frac{11-6\sqrt{3}}{2} = 0.304$[65] for maximally entangled qutrits $(d = 3)$, and to 0.309 for $d = 4$.[66] Thus, QKD can be carried out in the presence of larger amounts of noise as the dimension of the Hilbert space increases.

A number of realizations of qudits have been carried out experimentally, including polarization entangled four-photon states,[67] time-energy entangled qutrits using single photon in a three-arm interferometer,[68] and time-bin-entangled photons is produced by a train of laser pulses.[69] Here, however, we will concentrate one specific realization, optical OAM, which we introduce in the next subsection.

### 3.1. *Orbital angular momentum*

In addition to the intrinsic or spin angular momentum that leads to the existence of polarization states, it is somewhat less well known that photons can also carry OAM. This OAM is due to the possibility of the photon state having non-trivial spatial structure. It wasn't until the 1990's that a thorough investigation of optical OAM began to be carried out and that a simple way was found to produce it controllably. After the seminal paper of Allen, *et al.*,[70] a flood of papers began which continues to grow today. A number of excellent reviews of the subject exist.[71–73]

The key observation is that if an approximate plane wave is given, an azimuthally-dependent phase shift of the form $e^{il\phi}$, where $\phi$ is the angle about the propagation axis, $z$, the resulting wave has angular momentum about the $z$-axis given by $L_z = l\hbar$. (Note that single-valuedness of the field forces the *topological charge* $l$ to be quantized to integer values.) This phase factor has the effect of tilting the wavefronts by an increasing amount as the axis is circumnavigated, so that the wavefronts have a corkscrew shape. The Poynting vector $\boldsymbol{S}$ must be perpendicular to the wavefront, so it is at an angle to the propagation axis. $\boldsymbol{S}$ therefore rotates about the axis as the wave propagates, leading to the existence of non-zero OAM.

A number of different beam modes can carry OAM, including higher-order Bessel or Hermite–Gauss modes. Here, we focus on Laguerre–Gauss (LG) modes. The LG wavefunction with OAM $l\hbar$ and with $p$ radial nodes is[74]

$$u_{lp}(r,z,\phi) = \frac{C_{\mathrm{p}}^{|l|}}{w(z)}\left(\frac{\sqrt{2}r}{w(z)}\right)^{|l|} e^{-r^2/w^2(r)} L_{\mathrm{p}}^{|l|}\left(\frac{2r^2}{w^2(r)}\right) \times e^{-ikr^2z/\left(2(z^2+z_R^2)\right)} e^{-i\phi l + i(2p+|l|+1)\arctan(z/z_R)}, \tag{69}$$

with normalization $C_{\mathrm{p}}^{|l|} = \sqrt{\frac{2p!}{\pi(p+|l|)!}}$ and beam radius $w(z) = w_0\sqrt{1+\frac{z}{z_R}}$ at $z$. $z_r = \frac{\pi w_0^2}{\lambda}$ is the Rayleigh range and the arctangent term is the Gouy phase.

There are a number of ways to generate optical OAM states, the most common being the use of spiral phase plates (plates whose optical thickness varies azimuthally according to $\frac{l\phi}{k(n-1)}$[75]), computer generated holograms of forked diffraction gratings,[76] which convert Guassian modes into OAM modes in first-order diffraction, or spatial light modulators (SLM).

### 3.2. *Entangled OAM pairs*

The SPDC-generated biphoton state is most often written as an expansion in the space of transverse linear momenta of the outgoing signal and idler, as was done in Sec. 1. Now, though, we instead wish to expand in the space of OAM. Consider a pump beam of spatial profile $E(\boldsymbol{r}) = u_{l_0p_0}(\boldsymbol{r})$ encountering a $\chi^2$ nonlinear crystal, producing two outgoing beams via SPDC. For fixed beam waist, the range of OAM values produced by the crystal is roughly inversely proportional to the square root of the crystal thickness $L$.[77] We wish a broad OAM bandwidth, so we assume a thin crystal located at the beam waist ($z=0$). The output is an entangled state,[78] with a superposition of terms of form $u_{l_1',p_1'}u_{l_2',p_2'}$, angular momentum conservation requiring $l_0 = l_1' + l_2'$. We will take the pump to have $l_0 = 0$, so that the OAM values just after the crystal are equal and opposite: $l_1' = -l_2' \equiv l$. The $p_1', p_2'$ values are unconstrained, although the amplitudes drop rapidly with increasing $p'$ values (see Eq. (82) below). The output of the crystal may be expanded as a superposition of signal and idler LG states:

$$|\Psi\rangle = \sum_{l_1',l_2'=-\infty}^{\infty}\sum_{p_1',p_2'=0}^{\infty} C_{p_1'p_2'}^{l_1',l_2'}|l_1',p_1';l_2',p_2'\rangle\delta(l_0-l_1'-l_2'), \tag{70}$$

where the coupling coefficients are given by

$$C^{l'_1,l'_2}_{p'_1p'_2} = \int d^2r\Phi(\boldsymbol{r})\Big[u_{l'_1p'_1}(\boldsymbol{r})u_{l'_2p'_2}(\boldsymbol{r})\Big]^*. \tag{71}$$

Explicit expressions for the $C^{l'_1,l'_2}_{p'_1p'_2}$ coefficients have been calculated Torres, Alexandrescu, and Torner.[77]

How entangled are the angular momenta of the beams? One way to answer this is to again compute the Schmidt number. Taking $p_1 = p_2 = 0$ for simplicity, Eq. (70) reduces to

$$|\Psi\rangle = \sum_{l=0}^{\infty} C^{l,-l}_{00}|l,0\rangle|-l,0\rangle \equiv \sum_{l=0}^{\infty}\sqrt{\lambda_l}|l\rangle|-l\rangle, \tag{72}$$

from this, $\lambda_l = (C^{l,-l}_{00})^2$ can be calculated explicitly.[77] The state is already in Schmidt form. From the $\lambda_l$, the Schmidt number and von Neumann entropy can then be found.

Salakhutdinov *et al.*[79] examine the Schmidt number for parametric down conversion in detail. Looking at the case of a Gaussian pump ($l = p = 0$) and vanishing radial quantum numbers for both signal and idler ($p_{\rm s} = p_{\rm i} = 0$), they found that a pump beam of waist $w = 325\,\mu$m and wavelength $\lambda_{\rm p} = 413\,$nm on a crystal of thickness $L = 2\,$mm, the total Schmidt number was $K \approx 350$. However, those associated with entangled azimuthal degrees of freedom (OAM) only accounted for roughly $K_{\rm az} \approx 2\sqrt{K} \approx 37$ of them. A more detailed analysis found that pure radially entangled modes ($p_s$ and $p_i$ values entangled) account for a further $K_r \approx \sqrt{K} \approx 18$ modes. The remainder are modes of radial-azimuthal cross-correlation, with $p$ and $l$ values jointly entangled.

### 3.3. *Quantum cryptography with OAM*

The first successful demonstration of QKD with OAM was achieved by Gröblacher *et al.*,[62] using qutrits formed by superpositions of $l = 0, \pm 1$ states. A Gaussian beam ($l = p = 0$) was used to pump a nonlinear crystal. Parametric downconversion then produced photon pairs with opposite momenta $\pm\, l$. Only pairs with $l = 0, 1$ were used. A transmission hologram was placed in each outgoing beam. When the beam strikes on-axis, the hologram changes $l$ by one unit, so for example $l_{\rm initial} = 0$ changes to $l_{\rm final} = 1$. However, when the beam strikes the hologram off-center, the result is a superposition of $l_{\rm initial}$ and $l_{\rm final}$. By changing the displacement of the hologram within the beam, we may control the relative weights in this superposition. By this means, it is possible to produce a maximally-entangled state $|\psi\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |1,-1\rangle + |-1,1\rangle)$ where the two numbers in each ket represent the OAM in arms A and B, respectively. In this manner, the successful construction of a quantum key shared between two experimenters was demonstrated.[62] 150 qutrits were sent in, and security was maintained by checking the parity of 3-qutrit blocks and discarding those with parity mismatch. The result was a final key of 72 qutrits, which was used to code and decode a 72-bit message without error. The Bell parameter (a measure of supposed violation of local realism) for this experiment was $S_3 = 2.688$, well above the classical upper limit of $S_3 = 2$.

### 3.4. *Digital spiral imaging*

*Digital spiral imaging.*[80,81] is a form of angular momentum spectroscopy in which properties of an object are reconstructed based on how it alters the OAM spectrum of light used to illuminate it (Fig. 3). The input and output light may be expanded in LG functions, with the object acting by transforming the coefficients of the ingoing expansion into those of the outgoing expansion. Information about the transmission profiles of both phase and amplitude objects may be retrieved[80,77]

The idea naturally arises of trying to use the measured OAM spectrum to reconstruct an image of the object. But, since only intensities are measured, the lack of phase information prevents this. One way to extract the necessary phase is to use pairs of beams, in some sort of interferometric arrangement. This leads naturally to the idea of using pairs of photons with entangled OAM states, so we next investigate OAM entanglement in SPDC.

### 3.5. *Joint OAM spectra*

We now investigate the use of two beams, rather than one, in combination with spiral imaging. The full benefits of doing this will emerge in Sec. 3.7. In this section, we focus on examination of the OAM correlations. We begin with an entangled version, where the light source is parametric downconversion in a nonlinear crystal such as $\beta$-barium borate (BBO). Imagine an object in the signal beam (Fig. 4). Since OAM conservation holds exactly only in the paraxial case, we assume the signal and idler are produced in *collinear* downconversion, then directed into separate branches by a BS.

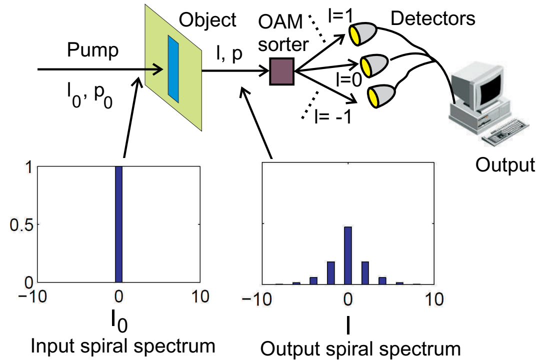


Fig. 3. Digital spiral imaging: the presence of an object in the light beam alters the distribution of angular momentum values in the outgoing light.

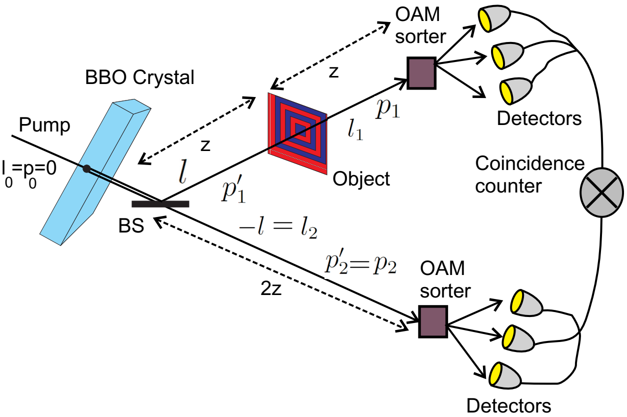


Fig. 4. Setup for analyzing object via OAM of entangled photon pairs.

(We assume all BSs are 50–50.) Assume perfect detectors for simplicity (imperfect detectors can be accounted for by the method in Ref. 82).

Let $P(l_1, p_1; l_2, p_2)$ be the joint probability for detecting the signal with quantum numbers $l_1, p_1$ and the idler with values $l_2, p_2$. The marginal probabilities at the two detectors (probabilities for detection of a single photon, rather than for coincidence detection) are

$$P_s(l_1, p_1) = \sum_{l_2, p_2} P(l_1, p_1; l_2, p_2), \quad P_i(l_2, p_2) = \sum_{l_1, p_1} P(l_1, p_1; l_2, p_2). \tag{73}$$

Then the mutual information for the pair is

$$I(s, i) = \sum_{l_1, l_2 = l_{\min}}^{l_{\max}} \sum_{p_1, p_2 = 0}^{p_{\max}} P(l_1, p_1; l_2, p_2) \log_2 \left( \frac{P(l_1, p_1; l_2, p_2)}{P_s(l_1, p_1) P_i(l_2, p_2)} \right). \tag{74}$$

The most common experimental cases are when (i) $p_{\max} = \infty$ ($p_1$ and $p_2$ are not measured, so all possible values must be summed), or (ii) $p_{\max} = 0$. Except when stated otherwise, we will use $l_{\max} = -l_{\min} = 10$ and $p_{\max} = 0$.

Suppose the transmission profile for the object is $T(\boldsymbol{x})$, where $\boldsymbol{x}$ is position in the plane transverse to the beam axis. The goal is to determine the function $T(\boldsymbol{x})$ from measurements of OAM correlations *only*. The coincidence probabilities $P(l_1, p_1; l_2, p_2) = |A^{l_1 l_2}_{p_1 p_2}|^2$ have amplitudes

$$A^{l_1 l_2}_{p_1 p_2} = C_0 \sum_{p_1'} C^{-l_2, l_2}_{p_1' p_2} a^{-l_2, l_1}_{p_1' p_1}(z), \tag{75}$$

$$a^{l_1' l_1}_{p_1' p_1}(z) = \int u_{l_1' p_1'}(\boldsymbol{x}, z) [u_{l_1 p_1}(\boldsymbol{x}, z)]^* T(\boldsymbol{x}) d^2 x, \tag{76}$$

where $C_0$ is a normalization constant. Here it is assumed that the total distance in each branch is $2z$ (see Fig. 4). We define an operator $\hat{T}$ to represent the effect of the object on the beam. We may expand this operator in the position basis,

$$\hat{T} = \int d^2r d^2r' |\boldsymbol{r}'\rangle T(\boldsymbol{r},\boldsymbol{r}')\langle \boldsymbol{r}| = \int d^2r |\boldsymbol{r}\rangle T(\boldsymbol{r})\langle \boldsymbol{r}|, \tag{77}$$

where the last line assumes that the operator is local, i.e. diagonal in the position space basis. So the function $T(\boldsymbol{r})$ is then given by

$$T(\boldsymbol{r}) = \langle \boldsymbol{r}|\hat{T}|\boldsymbol{r}\rangle. \tag{78}$$

Alternately, the object operator may be expanded in the Laguerre–Gauss basis,

$$\hat{T} = \sum_{ll'}\sum_{pp'} d^{l'l}_{p'p} |l'p'\rangle\langle lp|. \tag{79}$$

Making use of these definitions and of Eq. (76), it follows immediately that

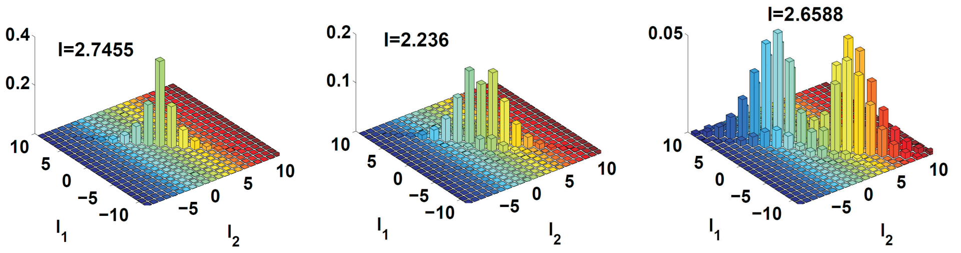
$$d^{l_1',l_1}_{p_1',p_1} = \langle l_1'p_1'|\hat{T}|l_1p_1\rangle = a^{l_1,l_1'}_{p_1,p_1'}. \tag{80}$$

Using this result in Eq. (79), then applying Eq. (78) and the fact that $u_{lp}(\boldsymbol{r}) = \langle \boldsymbol{r}|lp\rangle$, we find that determination of the $a^{l_1',l_1}_{p_1',p_1}$ coefficients is equivalent to reconstructing the object, since

$$T(\boldsymbol{r}) = \langle \boldsymbol{r}|\hat{T}|\boldsymbol{r}\rangle = \sum_{ll'}\sum_{pp'} a^{l_1',l_1}_{p_1',p_1} u_{l_1p_1}(\boldsymbol{r})[u_{l_1'p_1'}(\boldsymbol{r})]^*. \tag{81}$$

That the object's size and shape affect the coincidence rate is easy to see. For example, Fig. 5 shows the calculated spectrum when a single opaque strip of width $d$ is placed in the beam. We see a clear effect from changing an object parameter (the strip width). Similarly, if the corresponding mutual information is calculated it is found to vary with width, exihibiting a minimum at $d = w_0$.

The central peak of the spectrum (Fig. 5) broadens as $d$ increases from zero, reducing the correlation between $l_1$ and $l_2$; the mutual information between them thus declines over the range $d/w_0 < 1$. But at $d/w_0 \approx 1$, the central peak in $\{l_1, l_2\}$



(a) (b) (c)

Fig. 5. An opaque strip of width $d$ placed in the signal path. The widths are (a) $d = 0.1w_0$, (b) $d = 0.9w_0$, (c) $d = 2.5w_0$. The outgoing joint angular momentum spectra are plotted. As the width increases, the peak in the spectrum broadens, then (at $d = w_0$) splits into two peaks.

space bifurcates into two narrower peaks (right side of Fig. 5); the information thus goes back up as the peaks separate in the region $d/w_0 > 1$, after passing through the minimum at $d = w_0$. If we continue to sufficiently large $d$, the two peaks once again broaden and the mutual information decays gradually to zero. In addition, the total intensity getting past the opaque strip will continue to drop, so coincidence counts decay rapidly.

### 3.6. *Mutual information and symmetry*

Figure 6 shows the computed mutual information for several simple shapes. It can be seen that $I$ depends strongly on the size and shape of the object, so that for object identification from among a small set a comparison of the $I$ values rather than of the full probability distribution may suffice.

If the object has rotational symmetry about the pump axis, then its transmission function $T(r)$ depends only on radial distance $r$, not on azimuthal angle $\phi$. The angular integral in Eq. (76) is then $\int_0^{2\pi} e^{-i\phi(l-l')} d\phi = 2\pi\delta_{l,l'}$. So the joint probabilities reduce to the form $P(l_1, l_2) = f(l_1)\delta_{l_1,l_2}$ (assuming $p_1 = p_2 = 0$) for some function $f$. The marginal probabilities for each arm reduce to $P_1(l_1) = f(l_1)$ and $P_2(l_2) = f(l_2)$. The mutual information $I(L_1, L_2) = S_1(L_1)$ where $S_1(L_1) = -\sum_{l_1} f(l_1) \ln f(l_1)$ is the Shannon information of the object arm OAM spectrum. Thus in the case of rotational symmetry, the second arm becomes irrelevant from an information standpoint. In this sense, the quantity $\mu(L_1, L_2) \equiv |I(L_1, L_2) - S_1(L_1)|$ is an order parameter, capable of detecting breaking of rotational symmetry.

More generally, suppose that the object has a rotational symmetry group of order $N$; i.e. it is invariant under $\phi \to \phi + \frac{2\pi}{N}$. From Eqs. (69) and (76), it follows that the coefficients must then satisfy $a^{l'_1 l_1}_{p'_1 p_1} = e^{\frac{2\pi i}{N}(l'_1 - l_1)} a^{l'_1 l_1}_{p'_1 p_1}$, which implies $a^{l'_1 l_1}_{p'_1 p_1} = 0$ except when $\frac{l'_1 - l_1}{N}$ is integer. When $N$ goes up (enlarged symmetry group), the number of non-zero $a^{l'_1 l_1}_{p'_1 p_1}$ goes down; with the probability concentrated in a smaller number of configurations, correlations increase and mutual information goes up. This may be seen in the three right-most objects of Fig. 6, for example.

The first experimental use of this correlated spiral spectrum method has recently been carried out for the purpose of object identification.[83] Figure 7 shows a simple example: if the coincidence rate is plotted versus the OAM of two entangled photons,

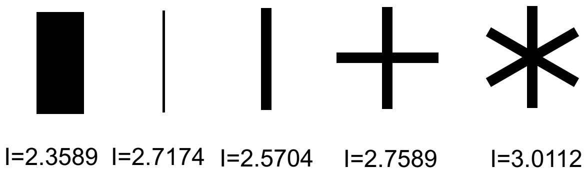
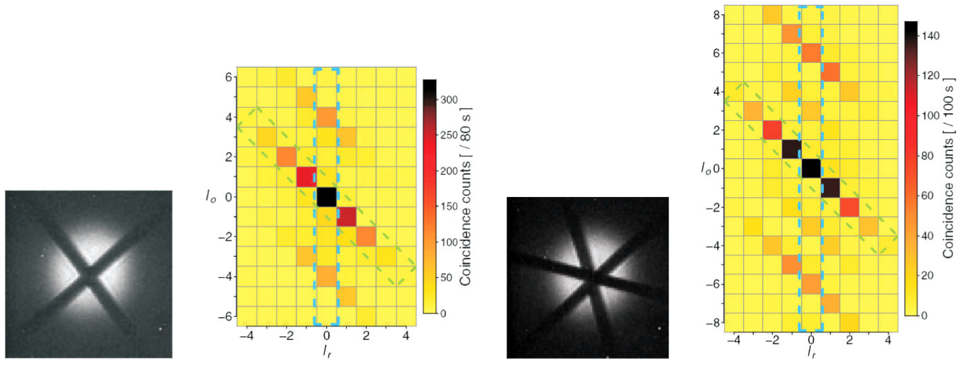


Fig. 6. The mutual information depends strongly on size and shape of the object. Here, the two objects on the left have widths $1.5w_0$ and $0.2w_0$; all other widths are $0.4w_0$.



(a) (b) (c) (d)

Fig. 7. An object with four-fold rotational symmetry (a) is placed in one output beam of a down conversion crystal. The joint OAM spectrum (b) of the two outgoing beams shows, in addition to the OAM-conserving main diagonal, a pair of secondary bands displaced from the diagonal by four units. Similarly a six-fold symmetric object (c) has joint OAM spectrum with bands displaced by six units (d).

normally angular momentum conservation forces the coincidence rate to vanish off the diagonal. However, as seen in the plots when objects with four-fold and six-fold rotational symmetry are placed in one beam off-diagonal terms appear, shifted respectively by three or four units from the diagonal. This allows the symmetry structure of the object to be easily determined, opening up possible applications such as rapid recognition of defective items on an assembly lines or irregular and diseased cells in a tissue sample.

### 3.7. *Imaging with entangled OAM*

The inability of digital spiral imaging to produce images due to loss of phase information has been pointed out. But a variation on the entangled OAM setup can be used to find the expansion coefficients *including phase.*

Assume that the beam waist for the OAM expansion (which is determined by the size and location of the detector aperture) is equal to the pump waist. Then, for the case of a Gaussian pump ($l_0 = p_0 = 0$) the expansion coefficients of Sec. 3.2 reduce to[82]:

$$C^{l,-l}_{p_1,p_2} = \sum_{m=0}^{p_1} \sum_{n=0}^{p_2} \frac{\left(\frac{2}{3}\right)^{m+n+l}(-1)^{m+n}\sqrt{p_1!p_2!(l+p_1)!(l+p_2)!}(l+m+n)!}{(p_1-m)!(p_2-n)!(l+m)!(l+n)!m!n!}. \qquad (82)$$

Using the latter expression for the coefficients, it can be shown[84] that determining the coincidence amplitudes $A^{l_1 l_2}_{p_1 p_2}$ is sufficient to determine the $a_{p_1' 0}$ coefficients, including phase. The measurement of the $A^{l_1 l_2}_{p_1 p_2}$ is accomplished by inserting a BS to mix the signal and idler beams before detection, as in Fig. 8, erasing information about which photon followed which path. We then count singles rates in the two detection stages, rather than the coincidence rate. If value $l$ is detected at a given

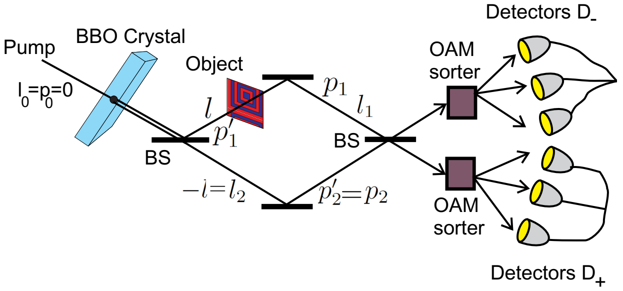


Fig. 8. A configuration allowing image reconstruction via phase-sensitive measurement of entangled OAM content.

detector it could have arrive by two different paths, so interference occurs between these two possibilities. The detection amplitudes in the two sets of detectors $D_+$ and $D_-$ involve factors $A_+ \sim (1 + ia_{00}^{l_0-l_2,l_1})$ and $A_- \sim (i + a_{00}^{l_0-l_2,l_1})$, with detection rates $R_\pm \sim 1 + |a_{00}^{l_0-l_2,l_1}|^2 \pm 2i\,\mathrm{Im}\,a_{00}^{l_0-l_2,l_1}$. From these counting rates, both the amplitudes and the relative phases of all coefficients can be found, allowing full image reconstruction, by inversion of Eq. (76) to find the object transmission function $T(\boldsymbol{r})$.

### 3.8. *Pixel entanglement*

In the previous sections, we have discussed entanglement between orbital angular momenta. Another avenue for investigating the spatial entanglement in imaging situations is via *pixel entanglement.*[85] Here, a spatially-resolving detector is used in each branch of a correlated-imaging setup, and spatial correlations are found by measuring coincidence counts between corresponding pixels in the two detectors. If the detectors are in the imaging plane, the result is the position-space correlation function, while if the detectors are in the Fourier plane of the imaging system, then the momentum space correlations are measured instead. The mutual information carried by the spatial correlations has been studied[29] using a Gaussian model similar that used in subsection 1.6: in the notation of Dixon *et al.*,[29] the biphoton state is

$$|\psi\rangle = \int d^2x_{\mathrm{a}} d^2x_{\mathrm{b}}\; f(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{x}_{\mathrm{b}})\, \hat{a}_{\mathrm{a}}^\dagger \hat{a}_{\mathrm{b}}^\dagger |vac\rangle = \int d^2k_{\mathrm{a}} d^2k_{\mathrm{b}} \tilde{f}(\boldsymbol{k}_{\mathrm{a}}, \boldsymbol{k}_{\mathrm{b}})\, \hat{a}_{\mathrm{a}}^\dagger \hat{a}_{\mathrm{b}}^\dagger |vac\rangle, \tag{83}$$

where the position- and momentum-state amplitudes are

$$f(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{x}_{\mathrm{b}}) = \mathcal{N} e^{-\frac{|\boldsymbol{x}_{\mathrm{a}}-\boldsymbol{x}_{\mathrm{b}}|^2}{4\sigma_{\mathrm{c}}^2}} e^{-\frac{|\boldsymbol{x}_{\mathrm{a}}+\boldsymbol{x}_{\mathrm{b}}|^2}{16\sigma_{\mathrm{p}}^2}}, \tag{84}$$

$$\tilde{f}(\boldsymbol{k}_{\mathrm{a}}, \boldsymbol{k}_{\mathrm{b}}) = (4\sigma_{\mathrm{p}}\sigma_{\mathrm{c}})^2 \mathcal{N} e^{-\sigma_{\mathrm{c}}^2|\boldsymbol{k}_{\mathrm{a}}-\boldsymbol{k}_{\mathrm{b}}|^2} e^{-4\sigma_{\mathrm{p}}^2|\boldsymbol{k}_{\mathrm{a}}+\boldsymbol{k}_{\mathrm{b}}|^2}, \tag{85}$$

and $\mathcal{N} = \frac{1}{2\pi\sigma_{\mathrm{c}}\sigma_{\mathrm{p}}}$. Defining joint and marginal probability densities $p(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{x}_{\mathrm{b}}) = |f(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{x}_{\mathrm{b}})|^2$, $p(\boldsymbol{x}_{\mathrm{a}}) = \int p(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{x}_{\mathrm{b}}) d^2x_{\mathrm{b}}$, and $p(\boldsymbol{x}_{\mathrm{b}}) = \int p(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{x}_{\mathrm{b}}) d^2x_{\mathrm{a}}$, the mutual information (the same for either the position or momentum basis) is

$$I(A:B) = -\int p(\boldsymbol{x}_{\rm a}, \boldsymbol{x}_{\rm b}) \log_2 \frac{p(\boldsymbol{x}_{\rm a}, \boldsymbol{x}_{\rm b})}{p(\boldsymbol{x}_{\rm a}) p(\boldsymbol{x}_{\rm b})} d^2 x_{\rm a} d^2 x_{\rm b} = \log_2 \left( \frac{4\sigma_{\rm p}^2 + \sigma_{\rm c}^2}{4\sigma_{\rm c}\sigma_{\rm p}} \right)^2 . \tag{86}$$

For the limiting case $\frac{\sigma_{\rm p}}{\sigma_{\rm c}} \gg 1$, we find the simple result $I(A:B) = \log (\frac{\sigma_{\rm p}}{\sigma_{\rm c}})^2$. For the values $\sigma_{\rm p} = 1500\,\mu$m and $\sigma_{\rm c} = 40\,\mu$m, this predicts $I(A:B) \approx 10.5$ bits per photon; experimentally, the value was found to be $7.1 \pm 0.7$ in the position basis and $7.2 \pm 0.3$ in the momentum basis.[29] For these approximate Gaussian states, the Schmidt number is approximately $K \approx (\frac{\sigma_{\rm p}}{\sigma_{\rm c}})^2 = 1400$, indicating a highly entangled state.

## 4. Conclusion

We have seen that the unique properties of quantum systems, superposition and entanglement in particular, allow for a number of interesting and useful phenomena in communication that are not possible in classical systems. Focusing on cryptography and imaging, we have seen that not only does the non-commutativity of quantum operators allow truly secret communication, but that many of the applications examined have involved extraction of multiple bits of information in a single photon. Thus, further developments along these lines offer the promise of somebody being able to securely transmit enormous amounts of information with a handful of photons. It is impossible to say what further interesting phenomena may be uncovered in the course of these investigations.

## References

1. B. W. Schumacher, *Phys. Rev. A* **51** (1995) 2738.
2. R. Omnès, *Phys. Rev. A* **56** (1997) 3383.
3. H. D. Zeh, *Found. Phys.* **1** (1970) 69.
4. C. Caves and C. A. Fuchs, Quantum information: How much information is in a state vector?, in *Sixty years of EPR*, Ann. Phys. Soc., Israel, A. Mann and M. Revzen, (eds.) (1996), pp. 226–257.
5. G. Jaeger, *Quantum Information — An Overview* (Springer, Berlin, 2007).
6. J. J. Sakurai, *Modern Quantum Mechanics*, revised edn. (Addison-Wesley, Reading, MA, 1994)
7. N. Peters, J. Altepeter, E. Jeffrey, D. Branning and P. Kwiat, *Q. Info. Comp.* **3** (2003) 503.
8. M. Atatüre, G. Di Giuseppe, M. D. Shaw, A. V. Sergienko, B. E. A. Saleh and M. C. Teich, *Phys. Rev. A* **66** (2002) 023822.
9. A. V. Sergienko, G. S. Jaeger, G. Giuseppe, B. E. A. Saleh and M. C. Teich, Quantum metrology and quantum information processing with hyper-entangled quantum states. in *Quantum Communication and Information Technologies (NATO Science Series)*, A. S. Shumovsky and V. I. Rupasov, (eds.) (Springer, Dordrecht, The Netherlands, 2003).
10. C. Bonato, D. Simon, P. Villoresi and A. V. Sergienko, *Phys. Rev. A* **79** (2009) 062304.
11. A. M. Steinberg, P. G. Kwiat and R. Y. Chiao, *Phys. Rev. Lett.* **68** (1992) 2421.

12. J. D. Franson, *Phys. Rev. A* **45** (1992) 3126.
13. A. Abouraddy, M. B. Nasr, B. E. A. Saleh, A. V. Sergienko and M. C. Teich, *Phys. Rev. A* **65** (2002) 053817.
14. D. S. Simon and A. V. Sergienko, *Phys. Rev. A* **80** (2009) 053813.
15. O. Minaeva, C. Bonato, B. E. Saleh, D. S. Simon and A. V. Sergienko, *Phys. Rev. Lett.* **102** (2009) 100504.
16. D. S. Simon and A. V. Sergienko, *Phys. Rev. A* **82** (2010) 023819.
17. D. S. Simon and A. V. Sergienko, *J. Opt. Soc. Am. B* **28**, (2011) 247.
18. A. Fraine, D. S. Simon, O. Minaeva, R. Egorov and A. V. Sergienko, *Optics Express* **19** (2011) 22820.
19. A. M. Fraine, O. M. Minaeva, D. S. Simon, R. Egorov and A. V. Sergienko, *Optics Express* **20** (2012) 2025.
20. W. P. Grice and I. A. Walmsley, *Phys. Rev. A* **56** (1997) 1627.
21. D. Erenso, *Opt. Res. Lett.* **2009** (2009) Article ID 387580.
22. A. Ekert and P. L. Knight, *Am. J. Phys.* **63** (1995) 415.
23. B. Grobe, K. Rzazewski and J. H. Eberly, *J. Phys. B: At. Mol. Opt. Phys.* **27** (1994) L503.
24. C. K. Law, I. A. Walmsley and J. H. Eberly, *Phys. Rev. Lett.* **84** (2000) 5304.
25. I. A. Law and J. H. Eberly, *Phys. Rev. Lett.* **92** (2004) 127903.
26. M. P. van Exter, A. R. Aiello, S. S. Oemrawsingh, G. Nienhuis and J. P. Woerdman, *Phys. Rev. A* **74** (2006) 012309.
27. H. Di Lorenso Pires, C. H. Monken and M. P. Van Exter, *Phys. Rev. A* **80** (2009) 022307.
28. J. Mertz, *Introduction to Optical Microscopy* (Robert and Co. Publishers, 2010).
29. P. B. Dixon, G. A. Howland, J. Schneeloch and J. C. Howell, *Phys. Rev. Lett.* **108** (2012) 143603.
30. M. B. Plenio and S. Virmani, *Quant. Inform. Comp.* **7** (2001) 001.
31. A. Peres, *Phys. Rev. Lett.* **77** (1996) 1413.
32. M. Horodečki, P. Horodečki and R. Horodečki, *Phys. Lett. A* **223** (1996) 1.
33. J. S. Bell, *Rev. Mod. Phys.* **38** (1966) 447.
34. J. S. Bell, *Physics* **1** (1964) 195.
35. G. Brassard, N. Lütkenhaus, T. Mor and B. C. Sanders, *Phys. Rev. Lett.* **85** (2000) 1330.
36. A. Vakhitov, V. Makarov and D. R. Hjelme, *J. Mod. Opt.* **48** (2001) 2023.
37. V. Makarov and D. R. Hjelme, *J. Mod. Opt.* **52** (2005) 691.
38. C.-H. F. Fung, B. Qi, K. Tamaki and H.-K. Lo, *Phys. Rev. A* **75** (2007) 032314.
39. B. Qi, C.-H. F. Fung, H.-K. Lo and X. Ma, *Quant. Info. Compu.* **7** (2007) 73.
40. T. Durt, B. G. Englert, I. Bengtsson and K. Zyczkowski, *Int. J. Quant. Inform.* **8** (2010) 535.
41. C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proc. IEEE Int. Conf. Computers, Systems, and Signal Processing*, Bangalore (1984), pp. 175.
42. C. Bennett, *Phys. Rev. Lett.* **68** (1992) 3121.
43. A. V. Sergienko, M. Atatüre, Z. Walton, G. Jaeger, B. E. A. Saleh and M. C. Teich, *Phys. Rev. A* **60** (1999) R2622.
44. R. S. Bennink, J. Bentley, R. W. Boyd and J. C. Howell, *Phys. Rev. Lett.* **92** (2004) 033601.
45. A. Gatti, E. Brambilla, M. Bache and L. A. Lugiato, *Phys. Rev. A* **70** (2004) 013802.
46. A. Gatti, E. Brambilla, M. Bache and L. A. Lugiato, *Phys. Rev. Lett.* **93** (2004) 093602.
47. Y. J. Cai and S. Y. Zhu, *Opt. Lett.* **29** (2004) 2716.
48. Y. J. Cai and S. Y. Zhu, *Phys. Rev. E* **71** (2005) 056607.
49. A. Valencia, G. Scarcelli, M. D'Angelo and Y. H. Shih, *Phys. Rev. Lett.* **94** (2005) 063601.

50. F. Ferri, D. Magatti, A. Gatti, M. Bache, E. Brambilla and L. A. Lugiato, *Phys. Rev. Lett.* **94** (2005) 183602.
51. A. V. Belinskii and D. N. Klyshko, *Sov. Phys. JETP* **78** (1994) 259.
52. T. B. Pittman, Y. H. Shih, D. V. Strekalov and A. V. Sergienko, *Phys. Rev. A* **52** (1995) R3429.
53. D. Klyshko, *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* **83** (1982) 1313 [*Soviet Journal of Experimental and Theoretical Physics* **56** (1982) 753].
54. D. Klyshko, *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* **94** (1988) 82.
55. A. Boto, P. Kok, D. Abrams, S. Braunstein, C. Williams and J. Dowling, *Phys. Rev. Lett.* **85** (2000) 2733.
56. P. Kok, A. Boto, D. Abrams, C. Williams, S. Braunstein and J. Dowling, *Phys. Rev. A* **63** (2001) 063407.
57. W. K. Wootters and B. D. Fields, *Ann. Phys. (N. Y.)* **191** (1989) 363.
58. D. Bruß, *Phys. Rev. Lett.* **81** (1998) 3018.
59. H. Bechmann-Pasquinucci and A. Peres, *Phys. Rev. Lett.* **85** (2000) 3313.
60. M. Bourennane, A. Karlsson and G. Björk, *Phys. Rev. A* **64** (2001) 012306.
61. N. J. Cerf, M. Bourennane, A. Karlsson and N. Gisin, *Phys. Rev. Lett.* **88** (2002) 127902.
62. S. Groblacher, T. Jennewein, A. Vaziri, G. Weihs and A. Zeilinger, *New J. Phys.* **8** (2006) 75.
63. Th. M. Nieuwenhuizen, *Found. Phys.* **41** (2010) 580.
64. K. Hess, H. De Raedt and K. Michielsen, *Phys. Scr.* **T151** (2012) 014002.
65. J. L. Chen, D. Kaszlikowski, L. C. Kwek, C. H. Oh and M. Zukowski, *Phys. Rev. A* **64** (2001) 052109.
66. M. Genovese and P. Traina, *Adv. Sci. Lett.* **1** (2008) 153.
67. J. C. Howell, A. Lamas-Linares and D. Bouwmeester, *Phys. Rev. Lett.* **88** (2002) 030401.
68. R. Thew, A. Acin, H. Zbinden and N. Gisin, *Quant. Inf. and Comp.* **4** (2004) 93.
69. D. Stucki, H. Zbinden and N. Gisin, *J. Mod. Opt.* **52** (2005) 2637.
70. L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw and J. P. Woerdman, *Phys. Rev. A* **45** (1992) 8185.
71. A. M. Yao and M. J. Padgett, *Adv. Opt. Phot.* **3** (2011) 161.
72. J. P. Torres and L. Torner (eds.), *Twisted Photons: Applications of Light with Orbital Angular Momentum* (Wiley, Hoboken, 2011).
73. S. Franke-Arnold, L. Allen and M. Padgett, *Laser Photon. Rev.* **2** (2008) 299.
74. L. Allen, M. Padgett and M. Babiker, *Prog. Opt.* **39** (1999) 291.
75. M. W. Beijersbergen, R. Coerwinkel, M. Kristensen and J. P. Woerdman, *Opt. Commun.* **112** (1994) 321.
76. V. Yu. Bazhenov, M. V. Vasnetsov and M. S. Soskin, *JETP Lett.* **52** (1990) 429.
77. J. P. Torres, A. Alexandrescu and L. Torner, *Phys. Rev. A* **68** (2003) 050301(R).
78. A. Mair, A. Vaziri, G. Weihs and A. Zeilinger, *Nature* **412** (2001) 313.
79. V. D. Salakhutdinov, E. R. Eliel and W. Löffler, *Phys. Rev. Lett.* **108** (2012) 173604.
80. L. Torner, J. P. Torres and S. Carrasco, *Opt. Exp.* **13** (2005) 873.
81. G. Molina-Terriza, L. Rebane, J. P. Torres, L. Torner and S. Carrasco, *J. Eur. Opt. Soc.* **2** (2007) 07014.
82. X. F. Ren, G. P. Guo, B. Yu, J. Li and G. C. Guo, *J. Opt. B: Quantum Semiclass. Opt.* **6** (2004) 243.
83. N. Uribe-Patarroyo, A. M. Fraine, D. S. Simon, O. M. Minaeva and A. V. Sergienko, *Phys. Rev. Lett.* **110** (2013) 043601.
84. D. S. Simon and A. V. Sergienko, *Phys. Rev. A* **85** (2012) 043825.
85. M. N. O'Sullivan-Hale, I. Ali Khan, R. W. Boyd and J. C. Howell, *Phys. Rev. Lett.* **94** (2005) 220501.