Unitary transformations of optical beam arrays

By

Aldo Camilo Martinez Becerril

A thesis completed under the supervision of J. S. Lundeen
submitted in partial fulfillment of the requirements for the

Doctorate in Philosophy degree

in **Physics**

uOttawa

Department of Physics
University of Ottawa

Abstract

The ability to control the spatial mode of light is extended across photonics from fundamental tests of quantum mechanics to telecommunications. Recently, a device known as 'Multi-plane light converter' (MPLC) has been introduced as a platform for the implementation of general spatial transformations of light. The main result of this thesis is the demonstration of an MPLC as a dynamically reconfigurable device. We do this using a system of parallel optical beams (the free space counterpart of a set of wave guides in integrated optics). We call a 'two-beam array' to the system composed of two of such parallel beams. We test the full space of unitary transformations for a two-beam array obtaining an average transformation fidelity of 0.85 ± 0.03. This high fidelity suggests MPLCs are a useful tool for quantum and classical information processing.

We also report two different results where the underlying degree of freedom is the spatial mode of light. The laser beam quality, also known as M2 parameter, is the figure of merit that compares the beam size in the near and far fields to the ones of a diffraction limited beam. We develop a method to determine M2 using the complex electric field at a single plane. This method avoids the standard procedure to get M2 which uses multiple intensity measurements, fitting and an inherent pre-characterization. Our method is particularly useful in optical design and simulations where the complex electric field is known.

In the context of quantum measurement and state characterization, we theoretically propose and experimentally demonstrate a method to perform a 'joint weak-measurement' optimizing the measurement resources. Typical joint weak-measurements use one read-out system per measured observable, while our method allows to measure correlations between pairs of observables using a single internal degree of freedom of a particle as read-out. We experimentally implement our scheme to directly measure the density matrix of photon polarization states using the photon's transverse spatial mode as read-out system.

Acknowledgements

I would like to thank my supervisor Jeff Lundeen for the continuous supervision of the projects presented in this thesis. I also thank my supervisor for creating an engaging and supportive work environment, which became invaluable for me during my doctoral studies.

I would like to thank each of the examiners in my committee, Robert Boyd, Ebrahim Karimi, Joel Carpenter, and Daniel Stolarski. Thanks for reading my thesis, and providing feedback on it.

Thanks to my co-authors whose hard work has been important for the completion of the projects in this thesis. I had the opportunity to work with two posdoctoral fellows, Lambert Giner and Raphael Abrahao, whose time, knowledge, teachings and discussions have been very helpful. I was fortunate to also work with Gabriel Bussiéres, Jordán Pagé, Li Liu, Siwei Luo, Martin Griessmann, and Sudhir S. All of them great summer students from whom I learned a lot.

Thanks to the lab mates I have met in the Lundeen lab. We all have helped each other even to find some missing optics/optomechanics needed for our experiments.

I thank my family. They have always been with me even at the distance from Mexico. Gracias ma, nano, nana, Tián y pa.

I am very grateful of working in the field of photonics. It has allowed me to work on a dynamic field of physics with a very rich ecosystem. I have had the opportunity to learn new advancements in the field attending conferences, as well as presenting our results.

I acknowledge the financial support from Mitacs, the university of Ottawa and my supervisor's research grants.

List of publications

The following articles were published during the completion of my PhD. Not all articles pertain to the work presented in this thesis.

- Aldo C. Martínez, Aldo Solis, Rafael Díaz Hernández Rojas, Alfred B. U'Ren, Jorge G. Hirsch, and Isaac Pérez Castillo. "Advanced statistical testing of quantum random number generators." Entropy 20 (2018): 886.

- Aldo C. Martinez-Becerril, Gabriel Bussières, Davor Curic, Lambert Giner, Raphael A. Abrahao, and Jeff S. Lundeen. "Theory and experiment for resource-efficient joint weak-measurement." Quantum 5 (2021): 599.

- Martin H. Griessmann, Aldo C. Martinez-Becerril, and Jeff S. Lundeen, "A method to determine the M2 beam quality from the electric field in a single plane." Opt. Continuum 2 (2023): 1833.

- Aldo C. Martinez-Becerril, et al., "Reconfigurable unitary transformations of optical beam arrays." In preparation.

Author contributions

The work presented in this thesis is a collaborative work. Common to each project is the supervision and feedback from Jeff Lundeen. A detailed list of contributions is now given.

The project reported in Chapter 3 was conceived by Jeff Lundeen. Co-authors and myself worked on early stages of the project designing and implementing the experimental setup. The numerical design was based on the code provided in Ref. [1], such code was adapted to our system by myself. I contributed with the final experimental setup, characterization, experimental debugging, and performed the final data collection. The structure of the draft of the article was conceived by Jeff Lundeen, and myself. I wrote the first version of the article and implemented a large part of the following edits. Jeff Lundeen and Raphael Abrahao have provided feedback, edits and suggestions that have been incorporated in the article draft (which is in preparation at the time of writing this thesis).

Chapter 4 describes a method to calculate the beam quality of a laser beam using the electric field at one plane. The project was conceived by Jeff Lundeen upon the public request of Orad Reschef on such a method. The idea of the method was provided by Jeff Lundeen, and details were worked out by Martin Griessmann with feedback from Jeff Lundeen and myself. Martin Griessmann performed the numerical implementation of the method. The first draft of the article was performed by Martin Griessmann and me. Later versions of the draft mostly involved Jeff Lundeen and myself.

Chapter 5 provides a method to perform joint weak-measurements optimizing the resources for such a measurement. The idea was provided by Davor Curic and Jeff Lundeen upon my arrival to the group. My contribution was developing the method, and implementing the experimental setup with help from co-authors. Final data acquisition was performed by myself as well as data analysis. I wrote the first version of the article, which was later edited by Jeff Lundeen and myself. Co-authors provided feedback on the draft, some of which was incorporated into the article.

Contents

Chapter 1

Fundamentals

Introduction

The fields of quantum information and computing have attracted attention in science and technology. The problem of factorization of an integer number is particularly relevant in this context. Modern cryptography relies on the difficulty of such a problem; as solving it is prohibitively impractical with a computer based on classical physics. This is not the case when quantum resources are used for computing, as theoretically proved by Peter Shor [2] in 1994. Since those years the scientific community has worked towards realizing a quantum computer and quantum algorithms for information processing. Different platforms have been used such as trapped ions, superconductors, and photons.

Photonics has been used as a platform to demonstrate proof of principle experiments, for the implementation of quantum gates, creation of entangled states, and the realization of quantum information protocols. Photonics is also a technologically active area, with multiple research groups in academia and private companies developing photonic hardware for quantum computing.

A universal quantum computer based on photonics needs the ability to implement transformations on the state of two photons ('two-photon gates'), which typically rely on non-linear optics. Such transformations are challenging to implement at the single photon level. A milestone came with the demonstration that linear optics, with single photon sources, and single photon detectors, is universal for quantum computing. Since then, multiple groups have worked towards a linear optics based computer/information processing.

This thesis is framed within such an effort with recent developments in the field. Particularly, we work with Multi-Plane Light Converters (MPLCs), and demonstrate its dynamical reconfigurability to implement arbitrary unitary transformations on a given system. The degree of freedom we are working with is the transverse spatial mode of a photon.

This thesis is organized as follows. In Chapter 1, we provide the fundamentals of spatial modes of light. Chapter 2 uses the quantum nature of light, defines a unitary transformation, and discusses a few applications of linear optics. We also review the concept of a 'weak measurement' used later in the thesis. In Chapter 3, we describe the experimental demonstration of the reconfigurable feature of an MPLC system. The rest of the thesis covers two different

projects whose underlying degree of freedom is the spatial mode of a photon. In Chapter 4, we theoretically propose and numerically demonstrate a method to obtain the laser beam quality or M^2 using the electric field of the light at a single plane. In Chapter 5, we report a method to perform a joint weak-measurement using a single read-out system. This optimizes the resources in order to make a weak measurement, and facilitate the use of joint weak measurement. We close this thesis with conclusions in Chapter 6.

1.1 Light: an electromagnetic wave

Light is an electromagnetic wave with its electric and magnetic vector fields described by Maxwell's equations. In many applications, the vectorial nature of light is not essential and light can be described by a single component of the electric field E, i.e., the light polarization can be decoupled in two independent transverse directions. This case will suffice for most of the work presented in this thesis. The description we now provide is based on [3, 4], where derivations and further details can be found.

The electric field $E(x, y, z, t)$ of a light wave at position (x, y, z) and time t satisfies the wave equation (a derivation of which can be found in Appendix A),

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) E(x, y, z, t) = 0, \tag{1.1}$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplacian operator and c is the speed of the light wave, which we assume to be travelling in a homogeneous medium of refractive index n so that $c = c_0/n$ with c_0 being the speed of light in vacuum. For a monochromatic wave, the time-dependence of $E(x, y, z, t)$ is described by a sinusoidal oscillation at angular frequency ω, which leads to a factorization of spatial and temporal components for $E(x, y, z, t)$,

$$E(x, y, z, t) = E(x, y, z)\, e^{i\omega t}. \tag{1.2}$$

Substituting this monochromatic wave into Eq. 1.1 we obtain a time-independent equation for $E(x, y, z)$ known as the Helmholtz equation.

$$\left(\nabla^2 + k^2\right) E(x, y, z) = 0, \tag{1.3}$$

with k being the wavevector, which is given by $k = \omega/c = 2\pi/\lambda$ with λ being the light wavelength.

Monochromatic and paraxial (defined below) optical beams is an assumption we will use through this thesis. Let's consider a monochromatic light beam propagating along the z direction forming divergence angles θ respect to z. In the paraxial wave approximation, such angles satisfy the following condition $\theta \ll 1$. The time-independent electric field $E(x, y, z)$ can be expressed as follows

$$E(x, y, z) = A(x, y, z)e^{-ikz}, \tag{1.4}$$

i.e., the electric field propagates along z as a plane wave and its transverse profile is modulated by the function $A(x, y, z)$. This function and its derivative along z, $\partial A/\partial z$, vary slowly within a distance of λ, this approximation is called the 'slowly varying approximation'. More formally, the first mentioned condition is stated as follows: $\Delta A = A(x, y, z + \lambda) - A(x, y, z) \ll A$. The difference ΔA can be expressed as $\Delta A \approx \partial A/\partial z \lambda \ll A$, which leads to the usual expression of this approximation

$$\frac{\partial A}{\partial z} \ll kA. \tag{1.5}$$

We now apply the same reasoning to the condition that $\partial A/\partial z$ should vary slowly: $\Delta \partial A/\partial z = \partial A/\partial z(x, y, z + \lambda) - \partial A/\partial z(x, y, z) \ll \partial A/\partial z$. We approximate $\Delta \partial A/\partial z$ as $\partial^2 A/\partial z^2 \lambda$. Combining the last two equations, we obtain the following condition

$$\frac{\partial^2 A}{\partial z^2} \ll k^2 A. \tag{1.6}$$

The paraxial approximation applied to Eq. 1.3 leads to an equation for the complex amplitude $A(x, y, z)$ from Eq. (1.4). Such an equation is known as the paraxial Helmholtz equation:

$$\left(\Delta_\perp - 2ik\frac{\partial}{\partial z}\right) A(x, y, z) = 0, \tag{1.7}$$

where $\Delta_\perp = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the transverse Laplacian.

1.2 Transverse spatial modes

The solution of Eq. (1.7) in different coordinate systems correspond to different sets of spatial modes. Each of those sets has an infinite number of modes, which can be labelled by the combination of two integer indices. Examples of such modes are the Hermite-Gauss (HG) and Laguerre-Gauss (LG) modes. We now discuss a Gaussian beam, then describe the HG and LG modes.

The Gaussian beam

The electric field of a Gaussian beam is given by the following equation

$$E(x, y, z) = A_0 \frac{w_0}{w(z)} e^{-\frac{x^2+y^2}{w^2(z)}} e^{-ikz - ik\frac{x^2+y^2}{2R(z)} + i\zeta(z)}, \tag{1.8}$$

we have used the standard beam parameters that we now define. The Rayleigh range z_0 is the distance from the center at which the beam doubles its area, or the beam size increase to $\sqrt{2}w_0$.

The beam size $w(z)$ is the width at which the beam intensity drops to $1/e^2$. It is a function of the propagation distance z

$$w(z) = w_0 \sqrt{1 + \left(\frac{z}{z_0}\right)^2}. \tag{1.9}$$

The minimum beam width is called the beam waist, and we denote it by w_0. The beam waist is determined by the Rayleigh range and the wavelength as stated in the following equation

$$w_0 = \sqrt{\frac{\lambda z_0}{\pi}}. \tag{1.10}$$

The radius of curvature $R(z)$ of the wave-fronts is given by

$$R\left(z\right) = z\left(1 + \left(\frac{z_0}{z}\right)^2\right). \tag{1.11}$$

The phase of the Gaussian beam at the axis is given by $\phi(z) = kz - \zeta(z)$. The term kz is same as the phase from a plane wave upon propagation. The other term $\zeta(z)$ is known as the Gouy phase and it is given by the following equation

$$\zeta(z) = \tan^{-1}\left(\frac{z}{z_0}\right). \tag{1.12}$$

The Gouy phase is an excess delay compared to the plane wave caused by the spatial confinement of the beam.

The intensity of a Gaussian beam is given next

$$I\left(x, y, z\right) = I_0\left(\frac{w_0}{w\left(z\right)}\right)^2 e^{-2\left(x^2+y^2\right)/w^2(z)}. \tag{1.13}$$

In the transverse plane, this is a Gaussian distribution centered at $x = y = 0$, with a standard deviation $\sigma(z)$ equal to $\sigma(z) = 2w(z)$. Along the z direction, the Gaussian profile is maintained. The minimum beam size occurs at $z = 0$.

The power of a Gaussian beam is obtained by integrating Eq. (1.13) along the transverse plane. A circle of radius $w(z)$ contains 86% of the total power, and 99% of the beam is inside a circle of radius $1.5w(z)$.

To fully characterize a Gaussian beam, the wavelength λ and two independent parameters are needed. Such parameters can be the the Rayleigh range z_0 and the location of the waist (typically chosen to be the origin of propagation coordinate $z = 0$). We now summarize the properties of HG and LG spatial modes.

Hermite-Gauss modes

A Hermite-Gauss (HG) mode is obtained by solving Eq. (1.7) in rectangular coordinates x, y, z. An HG mode is a product of modes in x and y, and can be labeled by two non-negative integers m and n:

$$\mathrm{E}_{m,n}\left(x, y, z\right) = A_{m,n}\frac{\omega_0}{\omega\left(z\right)}\mathrm{HG}_m\left(\frac{\sqrt{2}x}{\omega\left(z\right)}\right)\mathrm{HG}_n\left(\frac{\sqrt{2}y}{\omega\left(z\right)}\right)e^{-ikz-ik\frac{x^2+y^2}{2R(z)}+i(m+n+1)\zeta(z)}, \tag{1.14}$$

where the beam size $\omega(z)$, wave-front curvature $R(z)$ and Gouy phase $\zeta(z)$ are defined as above in Eqs. (1.11)-(1.12). The Hermite-Gauss function $\mathrm{HG}_m(u)$ is defined as follows

$$\mathrm{HG}_m(u) = \mathrm{H}_m(u)\, e^{-iu^2/2}, \tag{1.15}$$

where H_m is the Hermite polynomial of order m. Fig. 1.1 shows a plot of the first HG modes.

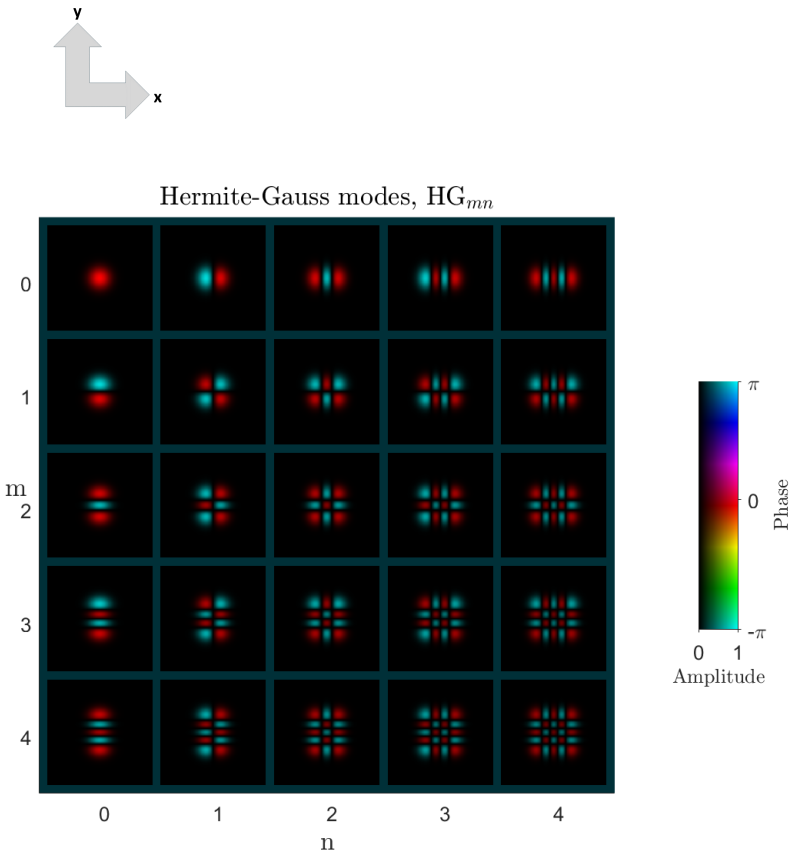


Figure 1.1: Hermite-Gauss modes ($\mathrm{HG}_{mn}(x,y)$) for m,n from zero to five. Each mode is plotted in an amplitude-phase plot. The field amplitude is encoded in the saturation and the field phase in the hue of the plot. Notice the π phase jumps surrounded by zero-intensity regions.

Laguerre-Gauss modes

Laguerre-Gauss (LG) modes are obtained by solving Eq. (1.7) in cylindrical coordinates ρ, ϕ, z. Here ρ is the radial component with values in $[0, \infty)$, ϕ is an azimuthal angle with values in $[-\pi, \pi)$, and z takes values in $(-\infty, \infty)$.

The electric field of each LG mode can be labelled by two non-negative integers p and l as $LG_{l,p}$. It is given by

$$LG_{l,p}(\rho,\phi,z) = A_{l,p}\frac{w_0}{w(z)}\left(\frac{\rho}{w(z)}\right)^l L_p^l\left(\frac{2\rho^2}{w^2(z)}\right) e^{-\frac{\rho^2}{w^2(z)}} e^{-ikz-ik\frac{\rho^2}{2R(z)}\pm il\phi+i(l+2p+1)\zeta(z)}, \quad (1.16)$$

where L_p^l represents the generalized Laguerre polynomial of orders l and p. From Eq. (1.16) we can calculate the intensity of an LG mode by taking the absolute square of the field. Such intensity remains invariant upon propagation along z, except by a magnification factor of $w(z)$. Fig. 1.2 shows a plot of the first LG modes.

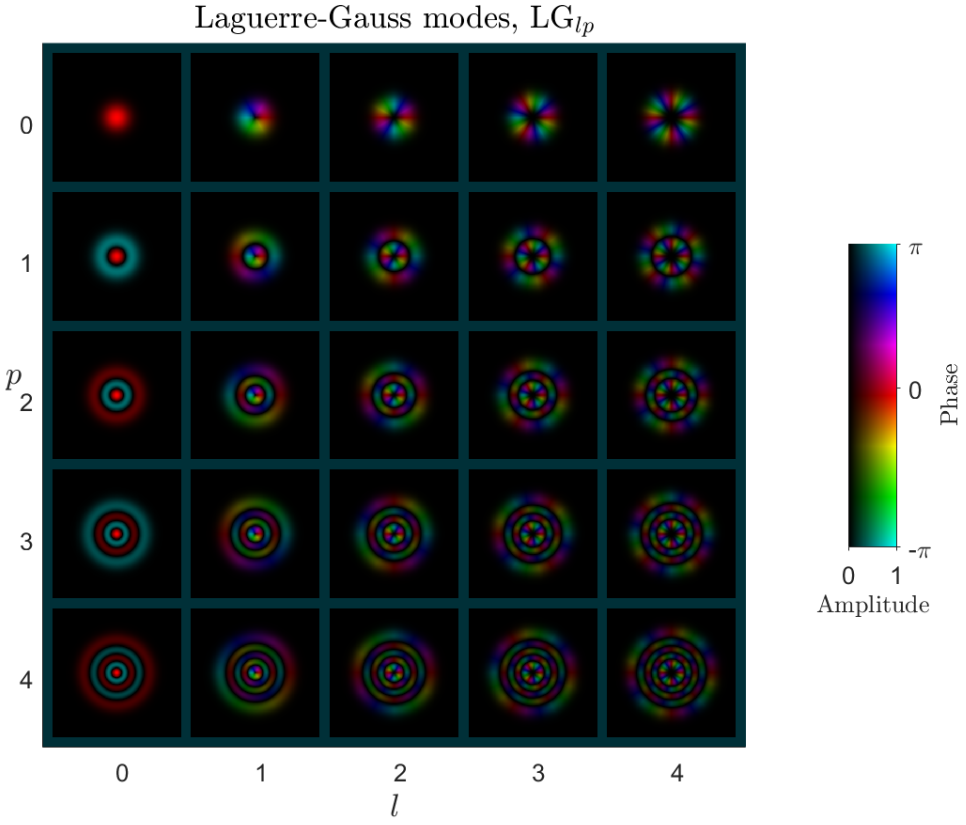


Figure 1.2: Laguerre-Gauss modes ($\mathrm{LG}_{l\rho}(\rho,\phi)$) for l,p from zero to five. Each mode is plotted in an amplitude-phase plot where the amplitude is encoded in the saturation and the phase in the hue of the plot. Notice the number p indicates there are $p+1$ regions of zero-intensity. The l number indicates the azimuthal phase that is wrapping l times within a ring of the LG mode.

Field fidelity

We define the field fidelity F of two spatial fields $E_{\psi_1}(x,y)$ and $E_{\psi_2}(x,y)$ as the following integral:

$$F\left(E_{\psi_1}, E_{\psi_2}\right) = \left| \int\int E_{\psi_1}.(x,y) E^*_{\psi_2}(x,y) dxdy \right|^2 . \tag{1.17}$$

A field E_ψ is normalized if $F\left(E_\psi, E_\psi\right) = 1$. This condition can be ensured by multiplying the field according to the following equation

$$E_\psi / \sqrt{F\left(E_\psi, E_\psi\right)}. \tag{1.18}$$

For normalized fields, the fidelity takes values between 0 and 1. Two fields are orthogonal if their fidelity equals zero.

The HG modes form an orthonormal set of modes. This means that they are normalized, and any two different HG modes $HG_{m,n}(x,y)$ and $HG_{m',n'}(x,y)$ are orthogonal between each other. The LG modes also form an orthonormal set of modes.

Basis for the spatial mode of light

HG beams form basis for the spatial mode of light. This means that any other beam $E_\psi(x,y)$ can be written as a superposition of HG modes as follows

$$E_\psi(x,y) = \sum_{n,m} c_{n.m} \mathrm{HG}_{n,m}(x,y), \tag{1.19}$$

where $c_{n,m} = \int\int E_\psi(x,y) \mathrm{HG}^*_{n,m}(x,y) dxdy$ are the coefficients of the superposition. LG modes also form an orthonormal basis for spatial modes.

An optical beam array

In integrated optics it is common to encode information in optical paths in which a photon can exist. Such optical paths are often different waveguides. The free space analogue is an optical beam array.

An optical beam array is a set of co-propagating Gaussian beams $E(x,y) = G(x)G(y)$, where each G is given by Eq. (1.8). In Chapter 3, we focus on the case where all modes are located at a line parallel to the y axis, and two adjacent beams are separated in y by a distance Δy. If $\Delta y > 3w_0$, two adjacent beams are nearly orthogonal to each other. Thus, if there are M beams with a separation larger than $3w_0$, we can label each beam state by an index m, and write it in Dirac notation as $|m\rangle$. We call this an M-beam array. A general state $|\psi\rangle$ in the beam array can be written as a superposition of each of the $|m\rangle$ beams as

$$|\psi\rangle = \sum_{m=1}^{M} c_m |m\rangle , \tag{1.20}$$

where $c_m = \langle m|\psi\rangle = \int\int \psi(x,y)E_m^*(x,y)dxdy$. Besides the transverse spatial profile, another degree of freedom of a photon is the polarization that we now describe.

1.3 Polarization

Light's polarization is a degree of freedom that refers to the oscillation direction of the electric field. It is a vectorial property of light and it can have variations across the three dimensional space. In the paraxial approximation, the electric field can be expressed vectorially as $\boldsymbol{E}(x,y,z,t) = E_x(x,z,t)\hat{\boldsymbol{x}} + E_y(y,z,t)\hat{\boldsymbol{y}}$, where $\hat{\boldsymbol{x}}$, and $\hat{\boldsymbol{y}}$ are unit vectors along x and y respectively. We have selected rectangular coordinates, but a similar expression can be written for other coordinates. For a plane wave, such components are given by the following equations:

$$\begin{aligned} E_x(x,z,t) &= u_x Re[e^{i\omega(t-z/c)+\delta_x}], \\ E_y(y,z,t) &= u_y Re[e^{i\omega(t-z/c)+\delta_y}], \end{aligned} \tag{1.21}$$

with u_x and u_y real field amplitudes. The fields E_x and E_y can be visualized as the components of an ellipse in the transverse plane:

$$\frac{E_x^2}{u_x^2} + \frac{E_y^2}{u_y^2} - 2\cos(\delta_x - \delta_y)\frac{E_x E_y}{u_x u_y} = \sin^2(\delta_x - \delta_y). \tag{1.22}$$

At each propagation plane z the electric field traces this ellipse. The polarization state of the light is determined by the field amplitudes u_x, u_y and the phase difference $\delta = \delta_x - \delta_y$. There are important polarization states that we now define.

Polarization states

Linear polarization. The first case is the one in which the ellipse collapses in a straight line. This case occurs if $\delta = 0$ or π. The direction of oscillation of the electric field is given by the ratio $\pm u_y/u_x$. Thus, we have horizontal, vertical, diagonal and anti-diagonal polarization if $u_y = 0, u_x = 0$, $u_y/u_x = 1$ or $u_y/u_x = -1$ respectively.

Circular polarization. If the phase difference δ is $\pm\pi/2$ and both field amplitudes are equal $u_x = u_y$, the light is circularly polarized as the electric field traces a circumference. It is a convention to call right circular polarization the state with $\delta = +\pi/2$, and left circular polarization when $\delta = -\pi/2$.

Poincaré sphere.

Light polarization is a two-level system. Any polarization state can be expressed with any two mutually exclusive polarization states, this is a basis for polarization. Such a basis can be the one of horizontal and vertical polarization states $\{|H\rangle, |V\rangle\}$, or the one formed by diagonal and anti-diagonal polarizations $\{|D\rangle = (|H\rangle + |V\rangle)/\sqrt{2}, |A\rangle = (|H\rangle - |V\rangle)/\sqrt{2}\}$. A third

one is composed by right and left circular polarization states $\{|R\rangle = (|H\rangle + i\,|V\rangle)/\sqrt{2}, |L\rangle = (|H\rangle - i\,|V\rangle)/\sqrt{2}\}$.

Taking the $|H\rangle, |V\rangle$ basis, an arbitrary polarization state $|\psi\rangle$ can be written as the following superposition:

$$|\psi\rangle = c_H\,|H\rangle + c_V\,|V\rangle\,, \tag{1.23}$$

with $|c_H|^2 + |c_v|^2 = 1$. Thus, we can parametrize the coefficients as $c_H = e^{i\phi_H}\cos(\theta/2)$ and $c_V = e^{i\phi_V}\sin(\theta/2)$. Eq. (1.23) is rewritten as :

$$|\psi\rangle = \cos(\theta/2)\,|H\rangle + e^{\phi}\sin(\theta/2)\,|V\rangle\,, \tag{1.24}$$

where $\phi = \phi_V - \phi_H$. Notice we omitted the phase $e^{i\phi_H}$, which does not make a difference in the physical predictions in standard quantum mechanics. Thus, any quantum state is parametrized by $\theta \in [0, 2\pi)$ and $\phi \in [-\pi, \pi)$. These are interpreted as a polar and an azimuthal angles in a three-dimensional space. These angles define a point on the surface of a sphere of radius one. This is known as the Poincaré sphere, and it is shown in Fig. 1.3. In this representation, orthogonal states are diametrically opposite with respect to the center of the sphere.

So far we have described pure polarization states. The most general polarization state is described by a density matrix ρ that incorporates information of classical statistical mixtures of polarization states:

$$\rho = \sum_j p_j\,|\psi_j\rangle\,\langle\psi_j|\,, \tag{1.25}$$

where p_j is the probability amplitude of measuring the polarization state $|\psi_j\rangle$. These mixed states can also be represented as points in the interior of the Poincaré sphere.

1.4 Propagation of light in free space

In this section we describe the propagation of light in free space, and discuss its numerical implementation. This problem refers to solving Eq. (1.7) which is a linear equation. Also free space is a shift invariant system. Thus, free space propagation is characterized by its impulse function $h(x, y)$, or equivalently by its transfer function $H(k_x, k_y)$, which is now determined.

Consider a plane wave $U(x, y, z')$ at $z = z'$. We want to propagate it to the output plane $z = z' + d$. Using a plane wave $U(x, y, z) = e^{-i(k_x x + k_y y + k_z z)}$ such a transfer function is easy to obtain. The planes wave at the input and output planes are related by $H(k_x, k_y)$ as follows:

$$\begin{aligned} U(x, y, z' + d) &= H(k_x, k_y)\,U(x, y, z')\,, \\ e^{-i(k_x x + k_y y + k_z (z' + d))} &= H(k_x, k_y)\,e^{-i\left(k_x x + k_y y + k_z z'\right)}. \end{aligned} \tag{1.26}$$

Thus the transfer function is given by the following expression:

$$H(k_x, k_y, d) = e^{-ik_z d} \tag{1.27}$$

$$= e^{-id\sqrt{k^2 - k_x^2 - k_y^2}}. \tag{1.28}$$

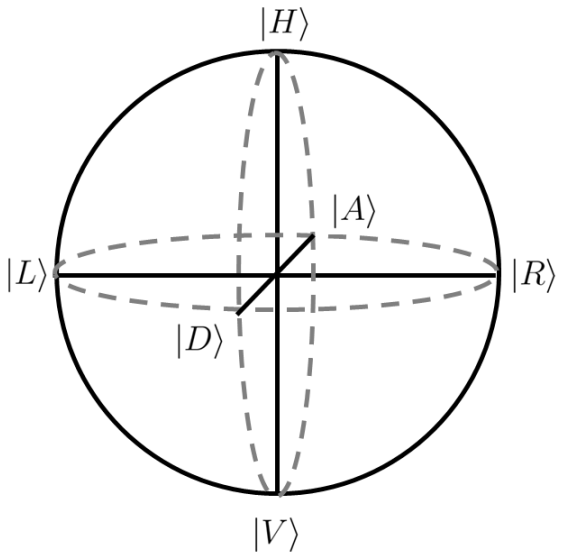


Figure 1.3: Poincaré sphere for representing polarization states. Three great circles can be defined, each of them passing through the standard basis states: horizontal and vertical states $|H\rangle, |V\rangle$, diagonal and anti-diagonal polarizations $|D\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}, |A\rangle = \frac{|H\rangle-|V\rangle}{\sqrt{2}}$, and right and left circilar polarizations $|R\rangle = \frac{|H\rangle-i|V\rangle}{\sqrt{2}}$, $|L\rangle = \frac{|H\rangle+i|V\rangle}{\sqrt{2}}$. Any pure polarization state can be represented on the surface of the Poincaré sphere. Mixed states lie in the interior of the sphere.

Notice that when the spatial frequencies satisfy $k_x^2 + k_y^2 > k^2$, the exponent in Eq. (1.28) becomes negative and the wave decays with distance z. This is an evanescent wave.

Fresnel approximation

The transfer function in Eq. (1.28) can be simplified when the spatial frequencies are much smaller that the cutoff frequency, $k_x^2+k_y^2 \ll k^2$. In this case, we use the following approximation

$$\begin{aligned} \sqrt{k^2 - k_x^2 - k_y^2} &= k\sqrt{1 - (k_x^2 + k_y^2)/k^2} \\ &\approx k\left(1 - \left(k_x^2 + k_y^2\right) / \left(2k^2\right)\right). \end{aligned} \tag{1.29}$$

Substituting this approximation in the transfer function from Eq. (1.28), we obtain the transfer function in the Fresnel approximation

$$H(k_x, k_y) = H_0 e^{i\frac{d}{2k}(k_x^2+k_y^2)}, \tag{1.30}$$

where $H_0 = e^{-ikd}$, d being the propagated distance. The Fresnel approximation is valid if the following condition is satisfied

$$N_F \theta_{max}^2 / 4 \ll 1, \tag{1.31}$$

where $N_F = \frac{a^2}{\lambda d}$ is the Fresnel number, $\theta_{max}^2 = \left(k_{xmax}^2 + k_{y,max}^2\right)/k^2$ is the sum of the squares of the maximum spatial frequencies in x and y (which are denoted by k_{xmax} and k_{ymax} respectively), and $a \approx \theta_{max} d$ is the maximum transverse shift between propagation planes.

The transfer function in Eq. (1.28) (or Eq. (1.30) in the Fresnel approximation) is used to describe the free space propagation of a beam. Such a method is described in Section 1.4. There is an equivalent method that only uses position coordinates (i.e., x and y) rather than spatial frequencies (k_x and k_y). Such a method uses the impulse response function of free space $h(x, y)$, which is the inverse Fourier transform of Eq. (1.30). For a derivation, we refer the reader to Ref. [3]. The impulse response function is

$$h(x, y) = h_0 e^{-i\frac{k}{2z}\left(x^2+y^2\right)}, \tag{1.32}$$

with $h_0 = \frac{i}{\lambda z} e^{-ikz}$. The field $U(x, y, z) = g(x, y)$ at a plane z is given as a convolution of the field at the input plane $U(x, y, z = 0) = f(x, y)$, and the impulse response function

$$g(x, y) = \int\int f\left(x', y'\right) h\left(x - x', y - y'\right) dx' dy'. \tag{1.33}$$

The idea behind this equation is that each point at plane $z = 0$ produces a new paraboloidal wave, which is transmitted to the plane z by multiplying it by the factor $h\left(x - x', y - y'\right)$. Eq. (1.33) integrates over all such point sources to obtain the resulting wave profile.

Fraunhofer approximation

The far field is a common setup in optics that exploits diffraction over large distances, which naturally separates the spatial Fourier components of an input wave. If the observation plane is a distance z sufficiently large from the input plane, then the output field gives the Fourier transform of the input wave $U(x, y, z = 0)$ just scaled by a factor $1/\lambda z$. Explicitly

$$U(x, y, z) = \frac{i}{\lambda z} e^{-ikz} U\left(\frac{k_x}{\lambda z}, \frac{k_y}{\lambda z}, z = 0\right). \tag{1.34}$$

This is known as Fraunhofer or far field approximation and it is valid if

$$N_F \ll 1, \text{ and } N_F' \ll 1, \tag{1.35}$$

where $N_F' = \frac{b^2}{\lambda d}$, and b is the minimum radius that contains the spatial profile at the input plane.

Recipe for propagating light in free space

The transfer function in Eq. (1.28) can be used to propagate an input wave $U(x, y, z = 0)$ a distance z after free space propagation $U(x, y, z)$. We now describe such a procedure.

The input wave can be expressed in momentum space by its angular spectrum $\mathcal{U}\left(\frac{k_x}{2\pi}, \frac{k_y}{2\pi}\right)$. The angular spectrum is a superposition of plane waves as follows

$$\mathcal{U}\left(\frac{k_x}{2\pi}, \frac{k_y}{2\pi}\right) = \int\int U(x, y, z = 0)\, e^{i(xk_x + yk_y)} dxdy. \tag{1.36}$$

Now that we have the plane wave components of the input wave, we can transmit each component by a propagation distance z according to the transfer function Eq. (1.28) by taking the product $H(k_x, k_y, z)\mathcal{U}(k_x, k_y)$. This is the expression in plane waves of the output wave $U(x, y, z)$. To obtain $U(x, y, z)$ we need to apply the inverse Fourier transform of $H(k_x, k_y, z)\mathcal{U}(k_x, k_y)$. Mathematically, this is written as follows

$$\begin{aligned} U(x, y, z) &= \frac{1}{4\pi^2}\int\int H(k_x, k_y, z)\mathcal{U}(k_x, k_y)\, e^{-i(xk_x + yk_y)} dk_x dk_y \\ &= \frac{1}{4\pi^2}\int\int e^{-i2\pi z\sqrt{\lambda^{-2} - k_x^2 - k_y^2}}\mathcal{U}(k_x, k_y)\, e^{-i(xk_x + yk_y)} dk_x dk_y. \end{aligned} \tag{1.37}$$

The free space propagation of a wave given by Eq. (1.37) can be easily implemented numerically. It helps to rewrite Eq. (1.37) in terms of Fourier transforms FT as follows

$$E(x, y, z) = FT^{-1}\left[e^{ik_z z} FT\left[E(x, y, 0)\right]\right], \tag{1.38}$$

where $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$. This method is accurate for all angles for a scalar field. That is, it is valid beyond the paraxial approximation from Eq. (1.6). For the implementation, one can use the native Fast Fourier Transform (FFT) common in numerical software. This general transfer function method is used through this thesis.

Fourier transform property of a lens

We now discuss a key transformation in optics, namely the realization of a Fourier transform by a lens and free space propagation. The schematic is shown in Fig. 1.4. An input wave is incident at plane $z = -d$ with an spatial profile $U(x, y, z = -d)$. A thin lens[1] of focal length f is located at $z = 0$, such a lens imparts a quadratic phase according to the following transmittance function $t(x, y)$ (the center of the transverse coordinates being at the center of the lens):

$$t_{lens}(x, y) = e^{i\frac{k}{2f}(x^2 + y^2)}. \tag{1.39}$$

Light is then transmitted to a distance f after the lens.

[1] A lens is thin if its thickness approximates zero. In practice, the thin lens approximation applies when the radii of curvature of a lens is large (thus the focal length) and the lens diameter is small [5].

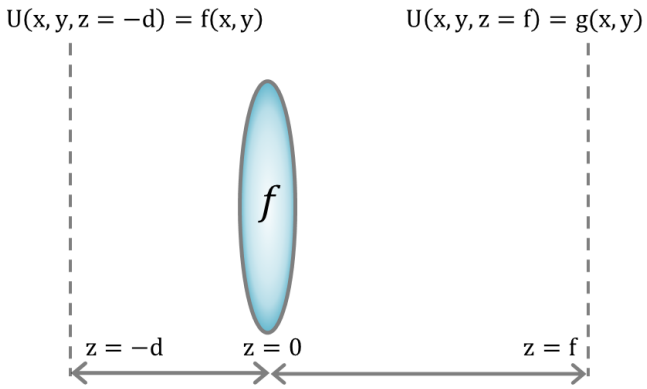


Figure 1.4: Schematic used in the derivation of the Fourier transform implementation by a lens.

In order to obtain the spatial distribution at $z = f$, we need to propagate the input field by the optical system in Fig. 1.4. The propagation from $z = -d$ to $z = 0$ can be done using the transfer function Eq. (1.30). The spatial profile before the lens is $g_1(x, y) = U(x, y, z = 0)$. Then, we apply the lens phase mask from Eq. (1.39), and propagate the field to $z = f$. We start by propagating the field before the lens, apply the lens phase and propagate to the plane $z = f$. According to the convolution approach given by Eq. (1.33), we obtain the following

$$\begin{aligned} g(x,y) &= h_0 e^{-i\frac{k}{2f}\left(x^2+y^2\right)} \int\int t_{lens}(x,y)\, g_1(x',y')\, e^{i\frac{k}{2f}\left(x'^2+y'^2\right)} e^{i\frac{k}{f}\left(xx'+yy'\right)} dx'dy' \\ &= h_0 e^{-i\frac{k}{2f}\left(x^2+y^2\right)} \int\int U(x',y',z=0)\, e^{i\frac{k}{f}\left(xx'+yy'\right)} dx'dy' \\ &= h_0 e^{-i\frac{k}{2f}\left(x^2+y^2\right)} \mathcal{U}\left(\frac{kx}{2\pi f}, \frac{ky}{2\pi f}, z=0\right) \\ &= h_0 e^{-i\frac{k}{2f}\left(x^2+y^2\right)} \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=0\right), \end{aligned} \tag{1.40}$$

where $\mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=0\right)$ is the angular spectrum of the field at $z = 0$ evaluated at the frequencies $x/(\lambda f)$ and $y/(\lambda f)$. This angular spectrum can be obtained from the angular spectrum at $z = -d$ as follows:

$$\begin{aligned} \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=0\right) &= H\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right) \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=-d\right) \\ &= H_0 e^{i\frac{d}{2k}\left(k^2x^2+k^2y^2\right)/f^2} \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=-d\right) \\ &= H_0 e^{i\frac{kd}{2f^2}\left(x^2+y^2\right)} \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=-d\right). \end{aligned} \tag{1.41}$$

Substituting this equation into Eq. (1.40),

$$\begin{aligned} g(x,y) &= h_0 H_0 e^{-i\frac{k}{2f}\left(x^2+y^2\right)} e^{i\frac{kd}{2f^2}\left(x^2+y^2\right)} \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=-d\right) \\ &= h_0 H_0 e^{i\frac{k}{2f}\left(x^2+y^2\right)(d/f-1)} \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=-d\right). \end{aligned} \tag{1.42}$$

Thus,

$$U(x,y,z=f) = h_0 H_0 e^{i\frac{k}{2f}\left(x^2+y^2\right)(d/f-1)} \mathcal{U}\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}, z=-d\right). \tag{1.43}$$

The Fourier transform of the initial wave's spatial distribution is obtained at the plane $z = f$, up to a quadratic phase factor. Such a phase can be suppressed if the initial distance is chosen to be $d = f$. This configuration would be symmetric in the input and observation planes respect to the lens. What happens in such symmetric configuration if the distance is different from $z = f$? Then a Fractional Fourier Transform is achieved. This is further explained in the following section and will be used in Chapter 5.

Fractional Fourier Transform

In this section we give an overview of the Fractional Fourier Transform (FrFT). Since its inception, the FrFT has been framed in different contexts: Condon described the FrFT mathematically in [6], while Namias framed it in quantum mechanical terms [7]. This versatility has allowed a broad range of applications ranging from differential equations to quantum optics [8], passing by signal processing, [9,10] and telecommunications [11]. Further details of the FrFT, and code for its numerical implementation can be found in Ref. [9].

The FrFT of order R of a function $f(x)$ is denoted by $f_R(u) \equiv \mathrm{FrFT}(f,R)$, and it is defined as the following transform:

$$f_R(u) = \int_{-\infty}^{\infty} K_R(u,u')f(u')du', \tag{1.44}$$

where

$$K_R(u,u') = \sqrt{1-i\cot R\pi/2}\, e^{i\pi\left(u^2\cot R\pi/2 - 2uu'\csc R\pi/2 + u'^2\cot R\pi/2\right)}, \tag{1.45}$$

for all real values of R, except for even integers. For $R = 4j$, with j an integer, $K_R(u,u') = \delta(u-u')$, and for $R = 4j+2$, $K_R(u,u') = \delta(u+u')$.

Particularly important are the $R = 0$ and $R = 1$ cases of the FrFT. These correspond, respectively, to the identity and the standard Fourier Transform (FT) operators: $f_0 = f(x)$, and $f_1 = \mathrm{FT}(f)(p)$. A useful property of the FrFT is that it is additive in R, i.e.,

$$\mathrm{FrFT}\left(\mathrm{FrFT}\left(f,R_2\right),R_1\right) = \mathrm{FrFT}\left(f,R_1+R_2\right). \tag{1.46}$$

For example, $\mathrm{FrFT}\left(\mathrm{FrFT}\left(f,1\right),1\right) = \mathrm{FrFT}\left(f,2\right) = f\left(-x\right)$ is the parity operator. Therefore, the FrFT is a transformation that interpolates the identity and a normal FT operators.

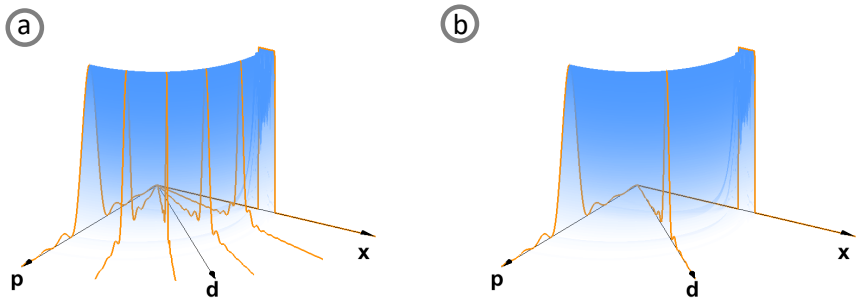


Figure 1.5: The Fractional Fourier Transform (FrFT) allows a continuous transformation of a spatial function to its Fourier transform, passing through other FrFT domains. The spatial function $f(x)$ in this figure is a rectangular function. In frame (a), we are plotting $|\mathrm{FrFT}(f, R)|^2$, for different FrFT orders R. Some of them are highlighted in orange, each plot is renormalized for better visualization. In frame (b), we highlight the cases of $R = 0$, 0.5 and 1, which correspond to a transformation to the $\boldsymbol{x}$, $\boldsymbol{d}$ and $\boldsymbol{p}$ domains respectively. In the **x**-axis, we have a square function, while on the **p**-axis we observe a sinc function as we expect to be the standard Fourier Transform of the square function, in our case FrFT for $R = 1$ corresponding to a $\pi/2$ rotation. The $\boldsymbol{d}$ domain is characterized by an equal weight of $\boldsymbol{x}$ and $\boldsymbol{p}$, and it shows an intensity profile between a square and a sinc functions.

The FrFT is strongly rooted in position-momentum phase-space. Associated with every FrFT order R, there is a different quadrature $\boldsymbol{u}$ that corresponds to a different superposition of position $\boldsymbol{x}$ and momentum $\boldsymbol{p}$. For dimensionless variables, such domain can be written as $\boldsymbol{u} = \boldsymbol{x}\cos(R\pi/2) + \boldsymbol{p}\sin(R\pi/2)$. The connection to phase-space is made through the Wigner distribution of a function. The action of a FrFT of order R is a clockwise rotation in phase-space of the Wigner distribution by an angle $R\pi/2$. A relevant result states that a projection of the Wigner distribution (a marginal of the Wigner distribution) on a domain at an angle $R\pi/2$, corresponds to the absolute squared of the FrFT of order R. In other words, performing a FrFT on a function allows one to obtain the function representation in a different FrFT domain. Fig. 1.5 illustrates an example. In the **x**-axis we have a square function, while on the **p**-axis we observe a sinc function as we expect to be the standard Fourier Transform of the square function, in our case FrFT for $R = 1$ corresponding to a $\pi/2$ rotation.

It is widely known that a FT can be performed optically by allowing free space propagation to the far field. The FrFT completes this picture of wave propagation. Diffraction is a continuous process of FrFTs as demonstrated in [12]. Optical implementations of the FrFT have been proposed utilizing lenses, graded-index media, and waveguide arrays [13–16]. Now we describe the setup to perform a FrFT utilized in our implementation.

Lohman [13] proposed a setup based on a lens that performs a FrFT of order R. Fig. 1.6 reproduces such setup. The transverse initial position $\boldsymbol{x}$ of an optical mode is fractional Fourier transformed by a lens of focal length f, and free space propagation by a distance $z = f\tan\left(\frac{R\pi}{4}\right)\sin\left(\frac{R\pi}{2}\right)$ before and after the lens f.

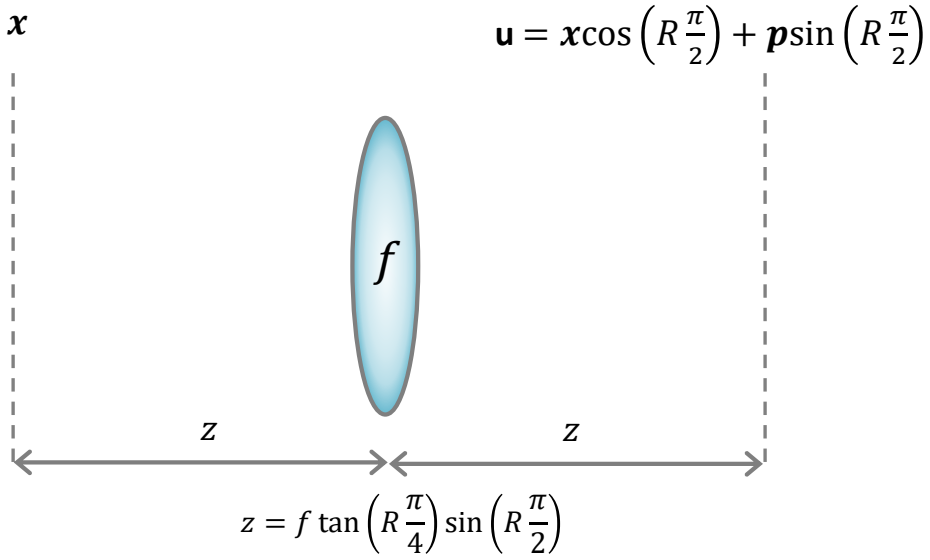


Figure 1.6: The transverse position distribution of a light mode is mapped to a different FrFT domain by means of a lens and free space propagation.

Chapter 2

Unitary transformations

In this chapter we discuss the concept of a unitary transformation in the context of spatial modes of light. At the end of the chapter, we mention a couple of example applications where it is desired to use a device to implement such transformations.

Photon states

In Section 1.2 we described the transverse spatial mode of light and the Hermite-Gauss and Laguerre-Gauss modes as basis for any state. This statement holds at the fundamental level of a single photon.

The quantization of the electromagnetic field (EM) leads us to the conclusion that a photon is an excitation of the electromagnetic field in a certain mode (e.g., a plane wave mode characterized by its wave vector k) [17]. The quantized EM field is described as a set of independent quantum harmonic oscillators. States of such field can be described by creation $\boldsymbol{a}^\dagger$ and annihilation $\boldsymbol{a}$ operators of a mode. The action of such operators on a state $|n\rangle$ of n photons is given as follows:

$$\boldsymbol{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, \tag{2.1}$$

$$\boldsymbol{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{2.2}$$

Such operators obey the bosonic commutation relations $[\boldsymbol{a},\boldsymbol{a}] = [\boldsymbol{a}^\dagger,\boldsymbol{a}^\dagger] = 0$ and $[\boldsymbol{a},\boldsymbol{a}^\dagger] = \mathbf{1}$.

The state with zero photons in it ($n = 0$) is called the vacuum state $|0\rangle$. Number states are orthogonal to each other $\langle n|m\rangle = \delta_{n,m}$, and they form a basis for the Hilbert space of n photons. Thus, any n photon state $|\psi_n\rangle$ can be written as a superposition of $|n\rangle$ states as follows:

$$|\psi_n\rangle = \sum_n c_n\,|n\rangle\,, \tag{2.3}$$

where $c_n = \langle\psi|n\rangle$ are complex coefficients. They satisfy the normalization condition $\sum_n |c_n|^2 = 1$.

2.1 Unitary transformations in quantum mechanics

In this section we give a description of a unitary transformation in quantum mechanics. Mathematically, a unitary transformation $\boldsymbol{U}$ is defined by the following equation

$$\boldsymbol{U}^\dagger \boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^\dagger = 1, \tag{2.4}$$

where $\boldsymbol{U}^\dagger$ is the Hermitian conjugate[1] of $\boldsymbol{U}$, and 1 is the identity operator. A unitary transformation preserves the inner product between any two states ψ_1 and ψ_2:

$$\begin{aligned} \langle \boldsymbol{U}\psi_1 | \boldsymbol{U}\psi_2 \rangle &= \left\langle \left(\boldsymbol{U}^\dagger\right)^\dagger \psi_1 \middle| \boldsymbol{U}\psi_2 \right\rangle \\ &= \langle \psi_1 | \boldsymbol{U}^\dagger \boldsymbol{U}\psi_2 \rangle \\ &= \langle \psi_1 | \psi_2 \rangle . \end{aligned} \tag{2.5}$$

In the second step we used the definition of a Hermitian conjugate operator given in footnote 1. In the last step we used the unitary property of $\boldsymbol{U}$. We can apply this result to sets of states. In particular, an orthonormal basis of states is mapped into a different basis of states, which is also orthonormal.

A general unitary of dimension N has N^2 free parameters as we show now. Now we count the number of dependent parameters for a unitary matrix, we look at the entries of the matrix product $\boldsymbol{U}\boldsymbol{U}^\dagger = 1$, which gives us an algebraic N dimensional system. There are N parameters from the diagonal entries, plus $2\sum_{n=0}^{N-1} n$ from the upper elements (the lower elements give identical equations up to complex conjugation), the factor of two corresponds to the fact that the system of equations is complex. Thus, the total number of free parameters in a unitary is $2N^2 - (N + 2\sum_{n=0}^{N-1} n) = 2N^2 - (N + (N-1)N) = N^2$.

We can understand the action of a unitary in two different ways [18, 19]. It can represent a change of basis where only the representation of the states is changed, i.e., using different basis to represent the same state. The other use of a unitary is the description of a dynamical evolution of a quantum state. Such an evolution has to be unitary to preserve the probability distribution associated with a quantum state. The action of a beam splitter, which is described in Section 2.3, is an example of the later use of an unitary. We now describe a change of basis on a photon's spatial mode.

2.2 Unitary transformations of a photon's spatial mode

We have described a photon's state in a certain spatial mode. In general, there is a set of different modes, each of them with a creation operator $\boldsymbol{a}_m^\dagger$ for $m = 0, ..., M$, where M is the maximum number of modes of the system. Examples of such mode sets are plane waves, HG or LG modes. Here it is shown how a unitary transformation can be used to express a state in a different modal basis.

[1] Given an operator $\boldsymbol{T}$, its Hermitian conjugate $\boldsymbol{T}^\dagger$ is defined as the operator which fulfills $\langle \boldsymbol{T}^\dagger \psi_1 | \psi_2 \rangle = \langle \psi_1 | \boldsymbol{T}\psi_2 \rangle$ for any states ψ_1 and ψ_2.

A change of basis is done by a unitary transformation $\boldsymbol{U}$ between the modes $\boldsymbol{a}_m^\dagger$ and the new modes $\boldsymbol{b}_m^\dagger$. The new modes are given using the matrix form of $\boldsymbol{U}$:

$$\boldsymbol{b}_n^\dagger = \sum_{m=0}^{M} U_{n,m} \boldsymbol{a}_m^\dagger, \tag{2.6}$$

for every mode $m = 0, ..., M$. As expected, the election of the mode basis is arbitrary.

There are some states that are invariant under any change of basis. One of such states is the vacuum state $|0\rangle$. This is useful as any state of n photons in mode m, which is denoted as $|n\rangle_m$, can be generated applying $\boldsymbol{b}_m^\dagger$ to the vacuum state:

$$\begin{aligned} |n\rangle_m &= \frac{1}{\sqrt{n!}} \left(\boldsymbol{b}_m^\dagger\right)^n |0\rangle , \\ &= \frac{1}{\sqrt{n!}} \left(\sum_{l=0}^{M} U_{m,l} \boldsymbol{a}_l^\dagger \right)^n |0\rangle . \end{aligned} \tag{2.7}$$

A change of basis is useful for simplifying the description of a state. For example, a single photon in the HG_{00} mode would be expressed by a single creation operator of such a mode. The same state can also be expressed in terms of plane waves, which would require a superposition of many of them to describe the same mode. This is a specific case of the fact that any single photon state $|\psi\rangle$ can always be expressed as a single mode. Suppose $|\psi\rangle$ is written as follows

$$|\psi\rangle = \sum_n c_n |1\rangle_n . \tag{2.8}$$

We need to find the unitary that transforms the $\boldsymbol{a}^\dagger$ into new $\boldsymbol{b}^\dagger$ modes according to Eq. (2.6). This has to be done in such a way that $|\psi\rangle$ can be expressed as $\boldsymbol{b}_1^\dagger |0\rangle$. We can achieve this if we choose the entries of the unitary $U_{1,n}$ as the coefficients c_n in Eq. (2.8). The rest of the rows $U_{m,n}$, for $m > 1$, can be chosen as the coefficients of other normalized states, which are orthogonal to $|\psi\rangle$ and to each other. This ensures the transformation is unitary. Let's now calculate $\boldsymbol{b}_1^\dagger |0\rangle$,

$$\begin{aligned} \boldsymbol{b}_1^\dagger |0\rangle &= \sum_n U_{1,n} \boldsymbol{a}_n^\dagger |0\rangle \\ &= \sum_n c_n \boldsymbol{a}_n^\dagger |0\rangle \\ &= |\psi\rangle . \end{aligned} \tag{2.9}$$

This shows a single photon state is single mode, and shows how to theoretically find such a mode using a unitary transformation. We now discuss how a unitary transformation can be experimentally realized.

2.3 Implementation of an N dimensional unitary

The pioneering work by Reck [20] showed a constructive algorithm to implement any unitary of dimension N. This work was framed in the context of optical paths. The dimension N of the unitary is the number of optical paths in which a photon can exist. Mathematically, such a photon state $|\psi\rangle$ is written as follows

$$|\psi\rangle = \sum_m c_m |m\rangle , \tag{2.10}$$

where $|m\rangle$ is the state of the m^{th} optical path and $c_m = \langle\psi|m\rangle$ are complex amplitudes satisfying $\sum_m |c_m|^2 = 1$. Such a single photon state is transformed by the unitary $\boldsymbol{U}$, which gives a different state $|\psi'\rangle$. The spatial distribution of $|\psi'\rangle$ can be measured by placing detectors at the output of each optical path. We now describe the elements appearing in Reck's scheme: phase shifters and beam splitters.

Phase shifter

A phase shifter is a device that adds a phase ϕ to the input photon. This can be propagation by a distance d through a slab of a material of refractive index n, in which case the imparted phase is $\phi = n\frac{2\pi}{\lambda}d$. In terms of creation operators, a phase shifter performs the following operation: $\boldsymbol{a}^\dagger_{out} = e^{i\phi}\boldsymbol{a}^\dagger_{in}$.

Beam splitter

A beam splitter (BS) is depicted in Fig. 2.1. A BS has two input ports with creation operators $\boldsymbol{a}^\dagger, \boldsymbol{b}^\dagger$ and two output ports with creation operators $\boldsymbol{c}^\dagger, \boldsymbol{d}^\dagger$. A BS is characterized by a coupling parameter θ and the phase ϕ it imparts between the output ports. A photon entering a single input port will exit as a superposition of the two output ports. In terms of the creation operators the BS transformation is written as:

$$\boldsymbol{d}^\dagger = \cos\theta\boldsymbol{a}^\dagger + i\sin\theta e^{-i\phi}\boldsymbol{b}^\dagger, \tag{2.11}$$

$$\boldsymbol{c}^\dagger = i\sin\theta e^{i\phi}\boldsymbol{a}^\dagger + \cos\theta\boldsymbol{b}^\dagger. \tag{2.12}$$

Two-dimensional unitary transformations

The building block in Reck's scheme is a two-level transformation that we now describe. The general unitary $\boldsymbol{U}$ of dimension two is given by

$$\boldsymbol{U}(\theta,\phi) = \begin{pmatrix} \cos\theta & ie^{-i\phi}\sin\theta \\ ie^{i\phi}\sin\theta & \cos\theta \end{pmatrix}, \tag{2.13}$$

where $\theta \in [0, \pi/2]$ is the coupling parameter, and $\phi \in [-\pi, \pi]$ is a phase. Eq. (2.13) for $\boldsymbol{U}(\theta,\phi)$ is a two-mode beam splitter transformation. In this context, θ is related to the reflectivity R of

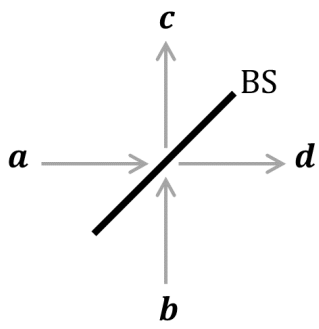


Figure 2.1: Schematic of a beam splitter (BS) with input ports a and b, and output ports c and d.

the beam splitter as $|R| = \sin^2\theta$. The case $\theta = \pi/4$ corresponds to a 50/50 beam splitter. The case $\theta = \pi/4$ and $\phi = -\pi/2$ corresponds to the Hadamard transformation (with the modes swapped) widely used in quantum information [21]. Pauli gates can also be obtained for some values of θ and ϕ.

Reck's scheme used such two-dimensional transformations from Eq. (2.13) as buiding blocks to implement any unitary of dimension N. Fig. 2.2 shows the general concept of such a decomposition. A detailed description of the algorithm can be found in Refs. [20, 21]. Reck's decomposition is not unique. There exist at least another design reported in [22]. The experimental realization of Reck's scheme requires stabilizing a large interferometer, which is challenging in bulk optics. The work from Ref. [23] implemented the Reck scheme in integrated optics (basen on silicon photonics) with six optical modes and the ability to input different photon number states. This device uses thermal phase shifters to achieve different unitary transformations in time, i.e., it is reconfigurable. Recently, a free-space approach has emerged as a Multi-Plane Light Converter (MPLC). This is the system we report in this thesis, and it will be further explained in Chapter 3.

2.4 Example applications of unitary transformations

The realization of arbitrary unitary transformations on spatial modes has a wide range of applications in classical and quantum information processing. In this section we mention some of them.

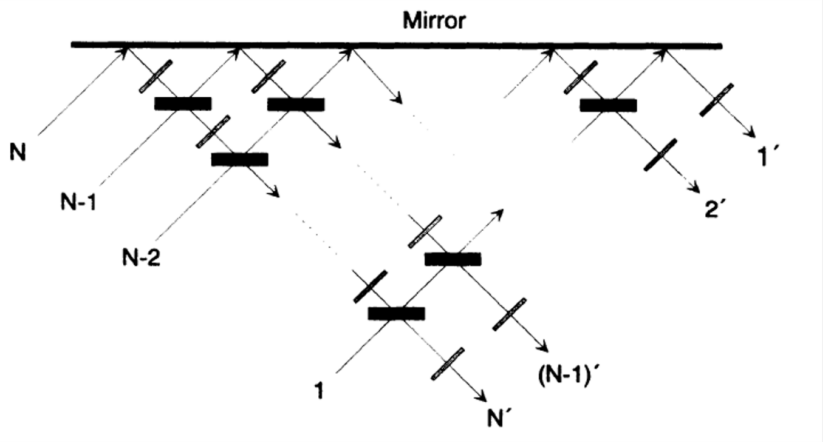


Figure 2.2: Schematic of Reck's scheme for realizing an N dimensional unitary using beam splitters and phase shifters. Image credits Ref. [20].

Quantum computing with photons

Quantum computing promises to speed up computations that are practically prohibited in classical computing. There exist multiple hardware platforms for quantum computing such as trapped ions, superconducting qubits, and linear optics. Independently of the platform, a device capable of implementing arbitrary unitary transformations in the spatial mode of light would be an asset towards quantum computing.

One of the requirements for a quantum computing platform is the ability to implement universal gates [24]. This is related to Reck's scheme and to the present work as well. Therefore, we will dedicate this subsection to describing what we mean by a universal gate set for quantum computing.

Quantum information is typically encoded in two-level ($|0\rangle$ and $|1\rangle$) quantum systems called qubits. Unitary transformations act on a qubit or a group of qubits. In this context, a transformation is called a gate. A universal gate set is a set of transformations that can implement any other gate on the space of N qubits.

Gates acting on a single qubit can be decomposed in terms of the Pauli gates σ_x, σ_y and σ_z. The gate σ_x maps $|0\rangle$ to $|1\rangle$, and the other way around. In matrix form,

$$\sigma_X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.14}$$

The gate σ_z applies a π phase shift to $|1\rangle$. In matrix form,

$$\sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.15}$$

The gate σ_y can be obtained as the product of the other two Pauli gates as $\sigma_y = i\sigma_X\sigma_Z$. In

matrix form,

$$\sigma_Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{2.16}$$

Another gate frequently used is the Hadamard gate H, which maps each state $|0\rangle$ or $|1\rangle$ into an equal superposition of both states. In matrix form

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{2.17}$$

Any single qubit transformation U can be expressed in terms of the Pauli gates.

Two-qubit gates act on two different qubits, taking one qubit as a control. These gates act conditionally on such a control qubit, e.g. if control qubit is in state $|1\rangle$, then apply a π phase shift to the other qubit. Examples of two-qubit gates are the CNOT and CSIGN gates.

The CNOT gate flips the state of the second qubit (from $|0\rangle$ to $|1\rangle$, and the other way around) if the control qubit is in state $|1\rangle$. Otherwise, it acts as the identity. This can be described by the following definition:

$$|p\rangle\,|q\rangle \xrightarrow{CNOT} |p\rangle\,|p \oplus q\rangle\,, \tag{2.18}$$

where $\oplus$ means addition modulo 2.

Another two-qubit gate is the CSIGN gate which applies a π phase shift only if the control qubit is in state $|1\rangle$. Mathematically,

$$|p\rangle\,|q\rangle \xrightarrow{CSIGN} (-1)^{pq}\,|p\rangle\,|q\rangle\,. \tag{2.19}$$

The case of an N qubit gate is based on above gates. It was demonstrated that an N qubit gate can be implemented in a circuit containing single qubit gates and one two-qubit gate, the CNOT gate. This construction is presented in [21]. Furthermore, Ref. [25] showed that almost any two-qubit gate with single qubit gates is sufficient for a universal gate set. The election of such a two-qubit gate is arbitrary. For photons, these two-qubit gates pose the challenge of making two photons interact. This can be done using nonlinear optics. We briefly mention such an approach as well as the KLM scheme, which is entirely based on linear optics.

The Kerr effect is the change of refractive index of a medium due to the intensity of a light beam:

$$n = n_0 + \chi^{(3)} I, \tag{2.20}$$

where n_0 is the linear refractive index, $\chi^{(3)}$ is the third-order susceptibility of the medium and I the intensity of the light beam. Such an intensity could be due to the beam itself or from a second 'probe' beam. The last one is often referred to as the cross-Kerr effect which is the one we will refer to.

Kerr effect and optical non-linearities in general are weak. In a highly non-linear medium, $\chi^{(3)}$ is in the order of 10^{-22} $\mathrm{m^2V^{-2}}$. Assuming a single photon focused in a non-linear medium with $\chi^{(3)} = 2 \times 10^{-22}$ $\mathrm{m^2V^{-2}}$, and a volume of $0.1 \times 1\,\mathrm{cm^3}$, the phase shift is in the order of 10^{-18} [26]. The quest of larger non-linearities is an ongoing research topic.

A milestone work for optical quantum computing was the one by KLM [27], where it was shown that linear optics with single photon sources, photon detection and feed forward of classical information are universal. Photon detection is an important element in KLM's protocol. The result of such a measurement indicates the rest of the N qubits are in some known state. The preparation of the N qubits can be chosen such that a target state can be induced by measurement and feed forward of classical information. This leads to the implementation of probabilistic and heralded gates.

The use of probabilistic gates for quantum information and computing leads to a reduction in the probability of implementing a protocol. Let's assume all gates can be implemented with probability p, a protocol with n gates will succeed with probability p^n. KLM addresses this issue by using quantum teleportation to implement the gate 'offline'. The gate is performed independently using ancillary qubits, and teleported to the main quantum circuit when it is successfully achieved. However, teleportation requires the realization of measurements in the Bell (entangled) basis, which is probabilistic in linear optics [28, 29]. The main breakthrough of KLM is to construct a near-deterministic linear optics scheme that uses n ancillary qubits to increase the probability of teleportation to $n^2/(n+1)^2$ which gives a probability of 1 in the limiting case of $n \to \infty$ [30].

A crucial gate in the KLM scheme is the non-linear sign gate (NS) defined by the following transformation

$$\alpha\,|0\rangle + \beta\,|1\rangle + \gamma\,|2\rangle \xrightarrow{NS} \alpha\,|0\rangle + \beta\,|1\rangle - \gamma\,|2\rangle\,. \tag{2.21}$$

This gate can be achieved by heralding, i.e., post-selecting upon the result of a measurement. With NS gates it is possible to implement a heralded CNOT gate. This was implemented in [23].

KLM scheme is challenging to implement in a practical scenario because of the overhead in the number of ancillary qubits needed. Improvements to KLM's protocol have been proposed. Another approach to quantum computing is the cluster state one, which also uses linear optics as well as large entangled states as resources. We refer the interested reader to Ref. [30] for further details.

Ultimately, the quest of a quantum computer promises quantum advantage, i.e., an exponential speed up for the solution of certain problems gained by using quantum resources. We now discuss another approach for quantum advantage. This one does not contain non-linear elements, neither classical feed forward of information.

Boson sampling

Boson sampling consists on n bosonic particles passing through a network of m modes, with $m \gg n$ i.e., the number of modes is much larger than the number of photons injected into the network. Calculating the output probability distribution of the network is a hard problem, in the sense that it is practically prohibited using classical resources only. However, such a system can be physically simulated e.g., with photons in a linear optical network where the output photon distribution can be sampled by placing detectors at the output of the network.

The root of the complexity of boson sampling is the calculation of a large bosonic state. It involves the calculation of a matrix permanent, which ensures the obtained state is symmetric under the permutation of any two particles of the system. A matrix permanent is calculated similarly to the matrix determinant, with the only difference of having positive signs in every term. The permanent and the determinant contain $n!$ terms which need to be calculated. However, there are efficient algorithms for the calculation of a determinant using a polynomial number of operations [31, 32]. There is no such an algorithm for a permanent.

Early experimental implementations of boson sampling were performed with a few photons e.g., up to four in Refs. [33–36]. Subsequent experiments [23, 37, 38] increased the number of photons and modes, and implemented verification tests to certify the experimental samples indeed corresponded to multiphoton interference. Ref. [39] implemented boson sampling experiments using $n = 20$ photons in $m = 60$ modes. boson sampling is a very active research field. For further details on the subject we refer the reader to the review articles [40, 41].

We have mentioned a couple of applications of an optical network to highlight its relevance in science and technology. Chapter 5 describes a method to perform a 'joint weak-measurement' and an experimental implementation of such method. We now provide the fundamentals of a weak measurement.

2.5 Weak measurement

In this section, we introduce a theoretical model for quantum measurement, von Neumann's model, that is typically used to describe weak measurement. The model involves a measured quantum system $\mathcal{S}$ and a pointer system $\mathcal{P}$ [42]. The latter indicates the measured value, the read-out, on a meter. A key aspect of the model is that the pointer is also quantum mechanical. Before the measurement, $\mathcal{S}$ and $\mathcal{P}$ are in an initial product state, $|I\rangle_{\mathcal{S}} \otimes |\phi\rangle_{\mathcal{P}}$, here $\otimes$ indicates a tensor product between different Hilbert spaces and the subscript is a label of the system. Both of these symbols will be omitted in the rest of the chapter. As usual, we assume that the pointer's initial spatial wave function $\phi(x)$ is a Gaussian centered at zero [42]:

$$\begin{aligned} \phi(x) &\equiv \langle x|\phi\rangle \\ &= \frac{1}{(2\pi\sigma_x^2)^{\frac{1}{4}}} e^{-\frac{x^2}{4\sigma_x^2}}, \end{aligned} \tag{2.22}$$

where σ_x is the standard deviation of the position probability-distribution.

The pointer's initial state happens to be same as the ground state of a harmonic oscillator. Thus, following Ref. [43], we define a lowering operator $\boldsymbol{a}$ as the operator that annihilates this pointer state, $\boldsymbol{a}\,|\phi\rangle = 0$. By this logic, from here on we label the pointer's initial state as $|0\rangle = |\phi\rangle$. As a standard lowering operator, $\boldsymbol{a}$ can be written in terms of the position $\boldsymbol{x}$ and momentum $\boldsymbol{p}$ of the pointer as follows $\boldsymbol{a} = \boldsymbol{x}/(2\sigma_x) + i\boldsymbol{p}\sigma_x/\hbar$. (Note, we use the natural length-scale σ_x of the system in place of the mass m and angular frequency ω that usually appear in the harmonic oscillator: $\sigma_x = \sqrt{\hbar/(2m\omega)}$.) Associated with $\boldsymbol{a}$, there is a raising operator $\boldsymbol{a}^\dagger$ that fulfills $[\boldsymbol{a}, \boldsymbol{a}^\dagger] = \mathbf{1}$. Similarly we can define number states $|n\rangle = \frac{(\boldsymbol{a}^\dagger)^n}{\sqrt{n!}}\,|0\rangle$.

Formulating the model in terms of lowering, raising operators and number states has proven fruitful in the past [43–45] and will be important for what follows.

Suppose we want to measure observable $\boldsymbol{A}$ of $\mathcal{S}$. Then, in the von Neumann model, one couples $\mathcal{S}$ to $\mathcal{P}$ by the following Hamiltonian,

$$\begin{aligned} \boldsymbol{H} &\equiv g\boldsymbol{A}\boldsymbol{p} \\ &= i\frac{g\hbar}{2\sigma_x}\boldsymbol{A}(\boldsymbol{a}^\dagger - \boldsymbol{a}), \end{aligned} \tag{2.23}$$

here g is a real parameter that indicates the interaction strength and we have used the usual decomposition of $\boldsymbol{p}$ in terms of $\boldsymbol{a}$ and $\boldsymbol{a}^\dagger$. We stress that there is no physical harmonic potential in the system and thus no quantum harmonic oscillator. We are following Ref. [43] and simply using the formalism of raising and lowering operators to analyze the effect of the interaction on the pointer state.

We now consider the state of the total system after the unitary evolution induced by $\boldsymbol{H}$: $\boldsymbol{U}_A \left|I\right\rangle \left|0\right\rangle = e^{-i\frac{t\boldsymbol{H}}{\hbar}} \left|I\right\rangle \left|0\right\rangle = \sum_{n=0}^{\infty} \frac{\gamma^n}{n!} \boldsymbol{A}^n (\boldsymbol{a}^\dagger - \boldsymbol{a})^n \left|I\right\rangle \left|0\right\rangle$. Here, $\gamma \equiv \frac{gt}{2\sigma_x}$ is a unitless parameter that quantifies the measurement strength. In general, the evolved system is in an entangled state between $\mathcal{S}$ and $\mathcal{P}$. For a strong interaction ($\gamma \gg 1$), in each trial, a measurement of the position of the pointer will unambiguously indicate the value of $\boldsymbol{A}$ (though it is not particularly obvious in this harmonic oscillator formulation).

So far the model is general and independent of the measurement strength. Now we consider the weak measurement regime. A weak measurement is characterized by $\gamma \ll 1$, which allows one to approximate the evolved state as $\boldsymbol{U}_A \left|I\right\rangle \left|0\right\rangle = \left|I\right\rangle \left|0\right\rangle + \gamma \boldsymbol{A} \left|I\right\rangle \left|1\right\rangle$ to first order in γ. In the weak regime, the entanglement between the pointer and measured system is reduced, and the initial state $\left|I\right\rangle$ of the particle is largely preserved. Following the work in [46], a post-selection on a final system state $\left|F\right\rangle$ is performed. Mathematically, this amounts to projecting onto $\left\langle F\right|$ and renormalizing, after which the pointer's final state is $\left|\phi'\right\rangle = \left|0\right\rangle + \gamma \frac{\left\langle F\right|\boldsymbol{A}\left|I\right\rangle}{\left\langle F|I\right\rangle} \left|1\right\rangle$. Thus the pointer's final state is largely left unchanged. That is, it is mostly left in $\left|0\right\rangle$, but a small component proportional to γ, is transferred to $\left|1\right\rangle$ due to the interaction with $\mathcal{S}$.

Our goal is to identify in what manner the pointer is shifted by the interaction. To this end, we find the expectations of the position and momentum of the final pointer. These respectively appear as the real and imaginary parts of $\left\langle \boldsymbol{a}\right\rangle \equiv \left\langle \phi'|\boldsymbol{a}|\phi'\right\rangle = \frac{1}{2\sigma_x} \left\langle \boldsymbol{x}\right\rangle + i\frac{\sigma_x}{\hbar} \left\langle \boldsymbol{p}\right\rangle$. Thus, using $\left|\phi'\right\rangle$ from just above, one finds

$$\begin{aligned} \left\langle \boldsymbol{a}\right\rangle &= \gamma \frac{\left\langle F\right|\boldsymbol{A}\left|I\right\rangle}{\left\langle F|I\right\rangle} \\ &\equiv \gamma \left\langle A\right\rangle_w . \end{aligned} \tag{2.24}$$

Consequently, the pointer is shifted from having $\left\langle \boldsymbol{x}\right\rangle = \left\langle \boldsymbol{p}\right\rangle = 0$ to indicating an average outcome $\left\langle A\right\rangle_w = \frac{1}{2\sigma_x\gamma} \left\langle \boldsymbol{x}\right\rangle + i\frac{\sigma_x}{\hbar\gamma} \left\langle \boldsymbol{p}\right\rangle$. This average pointer shift was introduced by Aharonov, Albert, and Vaidman in Ref. [46] and is called the 'weak value'. Unlike in the standard expectation value,

$|F\rangle \neq |I\rangle$ and, thus, the weak value is a potentially complex quantity. In summary, the real and imaginary parts of the weak value are the average shifts of the position and momentum of the pointer, which, in turn, are given by the expectation value of the lowering operator.

Now that we have framed the context of this thesis, we move on to describe the three research projects that compose this thesis.

Chapter 3

Reconfigurable unitary transformations of optical beam arrays

3.1 Introduction

The transformation of optical spatial modes in a reconfigurable and reliable way would have a wide range of applications, ranging from fundamental studies in optics to applications in telecommunications, classical computing, and quantum information processing. Most passive optical information processing tasks are unitary transformations $\boldsymbol{U}$ (i.e., $\boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{U}^\dagger\boldsymbol{U} = 1$) since they preserve information in the field. A device that creates a reconfigurable unitary could implement matrix multiplication [47], optical quantum gates [48, 49] and quantum random walks [50–52]. Furthermore, adding a non-linear optical element to such a device would allow one to obtain an optical neural network [53–57]. In this work, we experimentally implement arbitrary unitary transformations in the space of two parallel beams (a 'two-beam array') using a platform known as multi-plane light converter (MPLC).

In Section 2.2, we showed an N-dimensional unitary $\boldsymbol{U}$ possess N^2 independent real parameters that need to be specified to fully determine $\boldsymbol{U}$. To physically implement a unitary on spatial modes such as optical beams, one typically needs two elements: a mode coupler and the ability to impart phases to each mode. Two approaches can be distinguished, one that uses two-mode beam splitters as mode couplers, and another that uses an optical Fourier Transform or diffraction to couple modes. The first approach was discussed in Section 2.3.

We now discuss the diffraction approach. An MPLC is composed of layers alternating between a phase-mask plane and a diffraction layer. The latter is either created by free space propagation or an optical Fourier Transform. The MPLC was conceived and demonstrated in Refs. [58–60]. Fig. 3.1 contains the concept of an MPLC. Initially, Refs. [58, 59] focused on the use of MPLCs to perform mode multiplexing, and the phase-mask profiles were obtained by numerical optimization. In Ref. [1] a particular numerical optimization, wavefront matching, was adapted to MPLCs. They converted each beam in a two-dimensional grid to a different Laguerre-Gauss optical mode; achieving the impressive result of converting 210 optical modes

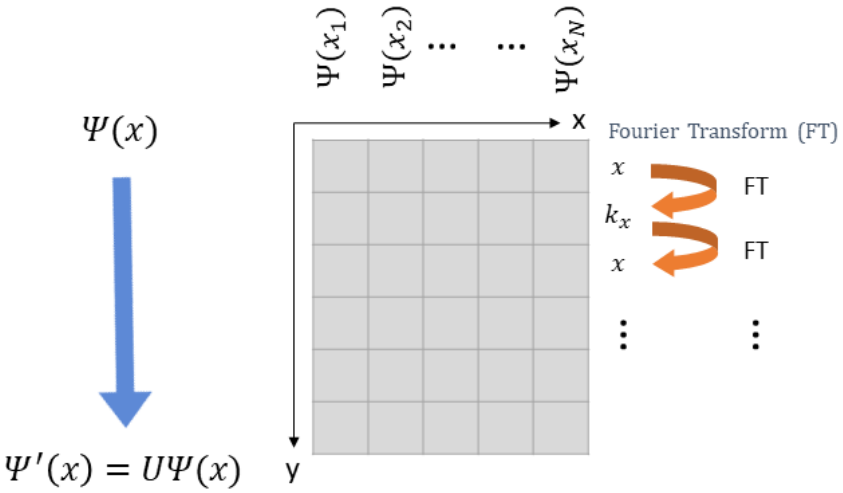


Figure 3.1: Initial concept of a multi-plane light converter (MPLC) as introduced in Refs. [58–60]. The photon's spatial state $\psi(x)$ is transformed by a sequence of alternating phases in x and k_x spaces. The transformation between x and k_x is achieved by a Fourier Transform (FT). Such a system composed of a number of layers of these phases, implements a unitary transformation $\boldsymbol{U}$. The final photon state is given by this unitary as $\psi'(x) = \boldsymbol{U}\psi(x)$.

using seven phase planes. In the recent years, there has been an increasing number of works that use MPLCs to perform different tasks such as diffractive networks [61] and high dimensional quantum gates using orbital angular momentum [48]. Other works are looking at different platforms like dielectric slabs [62]. The MPLC has drawn much interest for both applications and fundamental science.

An MPLC presents a series of advantages over the beam splitter approach. To mention a few: (i) in an MPLC one has more phase control channels, as there are of the order of 10^6 pixels on a single SLM. In the beam splitter approach, the number of phase controls is, in practice, limited to the order of magnitude 10^2 [63]; (ii) An MPLC can be fully implemented in free space, thus it does not suffer from insertion losses to chip and wave guides; (iii) An MPLC is composed of common optical elements, thus it is an easier system to be widely implemented and generalized to higher dimensions.

One of the main drawbacks of the MPLC, compared to the beam splitter mesh, is the lack of a deterministic analytical algorithm for finding the phase masks. Consequently, in an MPLC system, the minimum number of phase planes needed to create a given unitary of dimension N is unknown. The said, if there are N pixels in each phase mask, there must be at least N planes so that one can set all N^2 free parameters of the unitary. This sets a bound for the minimum for number of required phase planes. The seminal work in Ref. [64] provided an existence proof that an MPLC could implement a general unitary transformation. More recent work [65] gave an analytical method to implement any unitary but it requires $6N$ phase planes, far from the minimum, albeit all but N were fixed phase distributions. Compounding this

drawback is the fact that, unlike the discrete space theory mentioned above, a real MPLC acts in a continuous position space to perform transformations. Moreover, in practice, diffraction replaces the Fourier transform used in the above mentioned theory. In short, our analytical understanding of the MPLC is far-removed from how it is used in practice. The lack of an analytical algorithm means that the optimal transformation fidelity and its dependence on the MPLC system parameters are unknown. With so much unknown about MPLC design and performance it is important to characterize it for a wide range of tasks.

Most previous implementations of MPLC systems have created a single transformation. Moreover, usually that transformation was special in some way, e.g., Laguerre Gauss modes are propagation eigenstates. It was unclear how well an MPLC would perform for an arbitrary target unitary. In this work, we experimentally demonstrate an MPLC's versatility and reconfigurability by implementing a set of transformations that densely samples and fully covers the space of possible unitary transformations. In this way, we show that an MPLC can be a universal unitary.

Like the original Reck scheme, we aim to create unitary transformations of a linear array of parallel beams. We consider a beam array because it is a common information encoding in photonics for classical and quantum information processing. Moreover, a linear array matches the square grid of pixels in an SLM. Specifically, a linear beam array can be aligned with the top row of the grid so that subsequent phase planes use lower rows, as shown in Fig. 3.1. Or, more abstractly, if the SLM has N^2 pixels, that would be sufficient, in principle, to implement an N-mode unitary. In short, using a beam array on a line nominally makes an effective use of the active area of the SLM, which could be helpful to scaling an MPLC to higher dimension unitary transformations.

We continue in the next section by introducing the MPLC in more detail, and describing the wavefront matching algorithm that we use to find our phase planes for each unitary in our full set. We then proceed to describe our experimental setup and to characterize its performance using the designed phase planes, thereby creating a map of the transformation fidelity over all unitary transformations. Finally, we summarize the work of this Chapter and point future research directions.

3.2 Multi-plane light converter

A phase mask can be applied with a spatial light modulator (SLM). An SLM is composed of an array of liquid crystal pixels, each square with side-length s, that can be individually addressed to apply a chosen phase, see Fig. 3.2. Thus the imparted phase $P(x, y)$ is a sum of rectangular functions times the phase ϕ at each pixel $P(x, y) = \sum_{k,l} \phi(x_k, y_l)\text{rect}(\frac{x-x_k}{s})\text{rect}(\frac{y-y_l}{s})$. Our SLM size is 792×600 pixels, see Section 3.3 for further details of our SLM. However a more common SLM size is 1920×1080 pixels, which gives of the order of 10^6 addressable channels, three orders of magnitude higher than current beam splitter meshes.

A single phase mask, e.g., a single phase applied on an SLM, is insufficient to create a general spatial unitary. To see this, consider a general two-mode unitary transformation. Any thin phase-grating will couple more than just these two modes [66], see Fig. 3.3. For example,

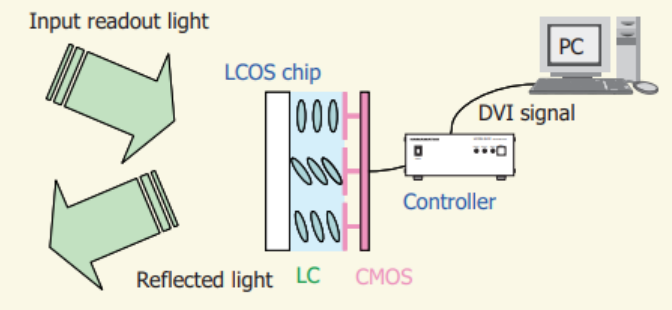


Figure 3.2: Schematic of a spatial light modulator (SLM). The SLM is composed of liquid crystals on a substrate e.g., silicon. A CMOS interface drives the liquid crystals cells to impart a spatially varying phase to light reflected off the SLM. An SLM is typically controlled from a computer, in this figure via a DVI signal. Image credits of Hamamatsu.

a sinusoidal or blazed diffraction grating will diffract light into new orders upon successive applications. This is Raman-Nath diffraction. On the other hand, thick gratings are able to limit coupling to two and only two modes through e.g., Bragg diffraction [66], as required by a two-mode unitary. A thick grating can be pictured as a series of different phase distributions (each of them contained in a plane) interleaved with spatial propagation, i.e., diffraction. This arrangement, the MPLC, is exactly what is needed to create a general spatial unitary.

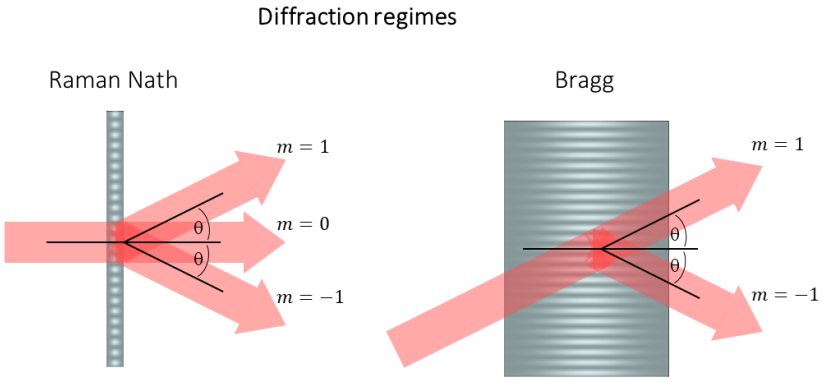


Figure 3.3: Diffraction regimes. A medium with a sinusoidal refractive index acts as a diffraction grating. Two regimes can be distinguished depending on the thickness of the medium: Raman-Nath and Bragg diffraction. In the Raman-Nath regime, the grating is thin. An incoming beam is transformed into multiple diffraction orders. In the Bragg regime, the grating is thick and the input beam is transformed into two diffraction orders only. Varying the thickness of the grating changes the splitting ratio between such orders.

Applying a phase distribution and diffraction are unitary transformations that we denote by $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{D}}$, respectively. An MPLC effectively performs a unitary transformation $\boldsymbol{U}$ by concatenating $\boldsymbol{\mathcal{DP}}$ blocks (assuming the numerical aperture of the system is large enough):

$$\boldsymbol{U} = \boldsymbol{\mathcal{D}}_m \boldsymbol{\mathcal{P}}_m \ldots \boldsymbol{\mathcal{D}}_1 \boldsymbol{\mathcal{P}}_1.$$

A schematic of an MPLC is shown in Fig. 3.4. A flat mirror and an SLM parallel to each other and L distance apart form the MPLC system. That is, instead of using a separate SLM for each phase, we reflect the light back to the same SLM. The light is inserted to the MPLC at angle τ so that each incidence on the SLM is at a different y position, utilizing a different area on the SLM. In this figure, we also show a linear beam array along x (red circles) as the set of input modes on which the transformation is being performed.

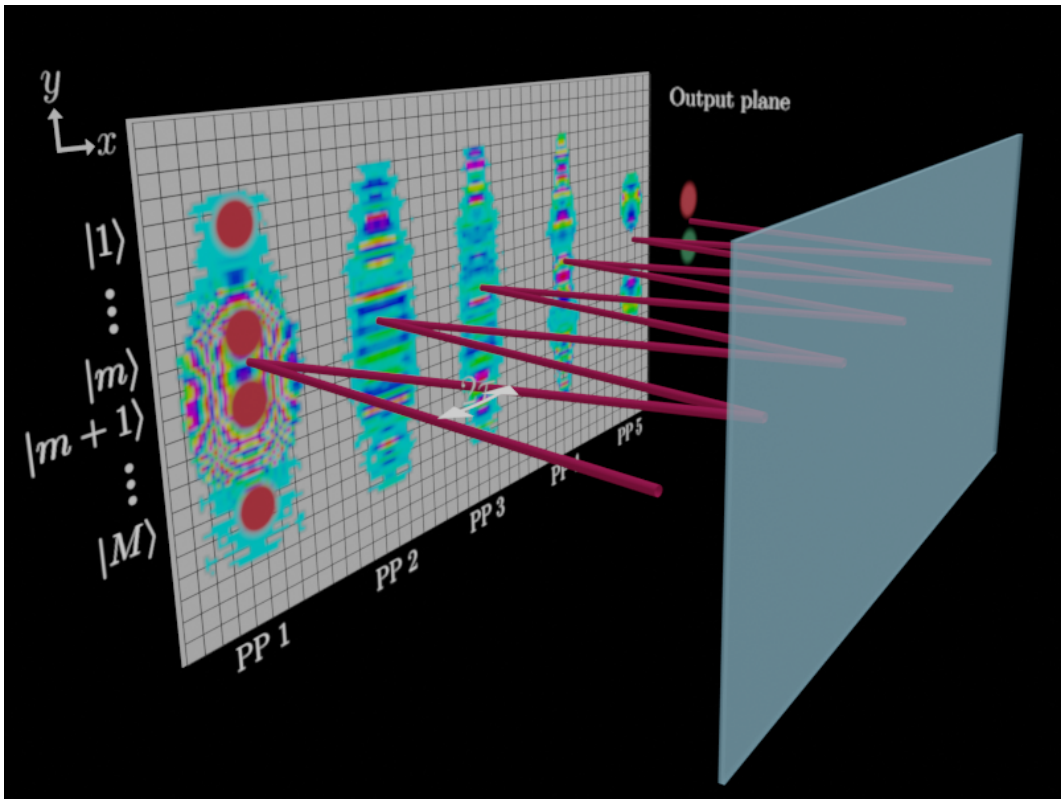


Figure 3.4: Schematic of a multi-plane light converter (MPLC), composed by a spatial light modulator (SLM) and a flat mirror. These optical components are separated by a distance L. In an MPLC, a set of P phase planes (PP) (color indicates phase, red $= 0$ and green $= \pi$) and diffraction implement a unitary transformation of dimension N. A beam array (circles at PP1) is at the input of the MPLC with insertion angle τ. At the detection plane after the SLM, we show $\boldsymbol{U}\,|m\rangle$, i.e., the output state after the MPLC when the input is beam $|m\rangle$ of the beam array.

Unitary transformations in a beam array

A beam array was formally defined in Section 1.2. The case of a general unitary $\boldsymbol{U}$ transformation on a two-beam array is given by eq. (2.13), and it is repeated here for convenience

$$\boldsymbol{U}(\theta,\phi)=\begin{pmatrix}\cos\theta & ie^{-i\phi}\sin\theta\\ ie^{i\phi}\sin\theta & \cos\theta\end{pmatrix}, \tag{3.1}$$

where $\theta\in[0,\pi/2]$ is the coupling parameter, and $\phi\in[-\pi,\pi]$ is a phase. The case $\theta=\pi/4$ and $\phi=-\pi/2$ transforms a single beam into an equal superposition of the two beams in the array, i.e., a Hadamard transformation (with the beams swapped), which is widely used in quantum information [21].

Optimization procedure

Now we describe the general idea behind the optimization procedure to find the required phase planes for achieving a given unitary. We follow Ref. [1] and use a wavefront matching algorithm [67, 68] to obtain the phase distribution at each phase plane in our system. In a general MPLC (as the one shown in Fig. 3.4), there are P phase planes (plus one more additional output plane) equally separated by a distance $\Delta z=2\times L$. The phase distribution at plane p is $\Phi_p(x,y)$. At the first plane, we have the N input beam modes $|m\rangle$, where m inside the ket is used as an index to label each beam $m=1,\ldots,N$. The target beam array states $|\psi_{t_m}\rangle$ at the last phase plane are determined by the target unitary $\boldsymbol{U}_t$ i.e., $|\psi_{t_m}\rangle=\boldsymbol{U}_t|m\rangle$. The design states $|\psi_{d_m}\rangle$ are the theoretical output states of a given MPLC design (i.e., a specific set of designed phase planes), thus $|\psi_{d_m}\rangle=\boldsymbol{U}_d|m\rangle$, where $\boldsymbol{U}_d=\boldsymbol{\mathcal{D}}_P\boldsymbol{\mathcal{P}}_P\ldots\boldsymbol{\mathcal{D}}_1\boldsymbol{\mathcal{P}}_1$.

The phase distributions are obtained through an inverse design optimization. The goal of which is to minimize the phase difference between input and target states at every propagation plane by using a few phase distributions Φ_p located at different propagation planes. The phase mismatch of corresponding input-target states at plane p is given by

$$\int\int \arg\left(m\left(x,y,z=\text{plane p}\right)\psi^*_{t_m}\left(x,y,z=\text{plane p}\right)\right)dxdy. \tag{3.2}$$

A superposition of these phase errors is formed each of them being weighted by the overlap of the corresponding input-target states of the MPLC system. The updated phase at plane p is given as the phase of such a superposition, this is mathematically written in Eq. 3.3. We now give an example of how to update the first phase Φ_1. We need to back propagate the target states $|m'\rangle$ to the first phase plane. The phase at plane one is updated as

$$\Delta\Phi_1(x,y)=-\arg\left(\sum_m m\left(x,y,\Delta z\right)\psi^*_{t_m}\left(x,y,\Delta z\right)\right.$$
$$\left.e^{-i\int\int\arg\left(m(x,y,\Delta z)\psi^*_{t_m}(x,y,\Delta z)\right)dxdy}\right). \tag{3.3}$$

A similar procedure is used to update the other phase planes by propagating the input and target states to the proper phase plane. For example, to update the p^{th} phase plane, one needs to propagate the input beams to this plane as $\boldsymbol{\mathcal{D}}_p\boldsymbol{\mathcal{P}}_p \ldots \boldsymbol{\mathcal{D}}_1\boldsymbol{\mathcal{P}}_1 |m\rangle$. A single iteration of the algorithm is conducted by repeating this phase update for all P planes. This iteration is then repeated until the figure of merit is reduced below a threshold value or a number of iterations is achieved. As mentioned in Subsection 3.2, in some cases the optimal value of the optimization is not the last iteration. Thus, when reaching the maximum number of iterations we selected the optimal value and the corresponding design. The figure of merit is the gate fidelity F_G between the target unitary $\boldsymbol{U}_t$ and the one implemented by the designed MPLC $\boldsymbol{U}_d$,

$$F_G(\boldsymbol{U}_t, \boldsymbol{U}_d) = \frac{1}{M} \sum_{m=1}^{M} \left| \langle m | \boldsymbol{U}_t^{\dagger} \boldsymbol{U}_d | m \rangle \right|^2. \tag{3.4}$$

The gate fidelity can be obtained from the electric field of the target and output design states as

$$\begin{aligned} F_G(\boldsymbol{U}_t, \boldsymbol{U}_d) &= \frac{1}{M} \sum_{m=1}^{M} \left| \langle \psi_{t_m} | \psi_{d_m} \rangle \right|^2 \\ &= \frac{1}{M} \sum_{m=1}^{M} \left| \int\int \psi_{d_m}(x,y) \psi^{*}_{t_m}(x,y) dx dy \right|^2. \end{aligned} \tag{3.5}$$

For further details of the algorithm we refer the reader to Ref. [68].

We point out that we are not using the SLM in its most common mode of operation, which we now describe. Typically, when creating or manipulating optical modes with SLMs, the transformation is encoded on top of a modulation of a regular saw-tooth phase pattern, a blazed grating [69, 70]. This is a unitary process with each diffraction order carrying different information. Particularly, the light diffracted into the first-order contains the desired mode. Selecting a single diffraction order makes the process inherently non-unitary. This mode of operation of the SLM is used to achieve amplitude and phase modulation with a single hologram [69,70]. It is also used as a selection mechanism due to the fact that SLM's diffraction efficiencies are lower than 100%. In contrast, in the mode of operation we utilize, each phase the SLM imparts is directly given by the wavefront matching algorithm. There is no underlying grating. While there may be technical sources of loss (e.g., fill-factor), our mode of operation can be fundamentally unitary.

Choice of parameters

In this subsection we describe how we selected the parameters of the beam array and the MPLC system. As mentioned in the Introduction, the realization of an N-dimensional unitary requires at least N phase planes. We estimated the number of possible reflections on an MPLC system, details are given below in Subsection 3.2. For our SLM's dimensions, we found there are up to ten reflections that can be achieved with realistic parameters. In general, the insertion angle τ

and plane-to-plane distance L limit the number of reflections one can achieve. Larger values of L achieve less number of reflections as there is overlap between reflections leading to unwanted interference. As L decreases more reflections can be obtained, but it also decreases the amount of diffraction between phase planes. Thus, in order to achieve general unitary transformations, it is not possible to decrease L arbitrarily. From the experimental point of view, decreasing L is better in the implementation of an MPLC as a small value of L avoids the need of angles τ smaller than $1°$. These numbers give an idea of the order of magnitude of the τ and L parameters. We now describe the details of estimating the number of reflections in an MPLC system.

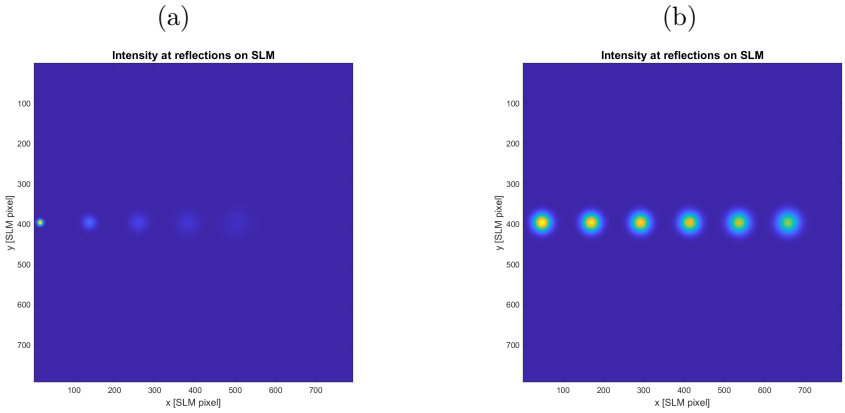


Figure 3.5: Intensity of a beam on the SLM of an MPLC system with a plane-to-plane spacing of $L = 10\,\text{cm}$. The beam has an insertion angle $\tau = 0.7\,\text{deg}$ and the beam waist equals $200\,\mu\text{m}$ and $600\,\mu\text{m}$ in a) and b) respectively. A value of ω_0 as small as $200\,\mu\text{m}$ leads to overlap between two reflections at the fifth and sixth phase-mask plane. Instead, a value of $\omega_0 = 600\,\mu\text{m}$ achieves six reflection on the SLM without overlap. Increasing the value of ω_0, for example $\omega_0 = 800\,\mu\text{m}$, gives overlap between reflections one and two.

Number of reflections in an MPLC system

The geometry of a general MPLC is shown in Fig. 3.4. The distance between SLM and mirror is L and the reflection angle of the input beam at the SLM is τ. The distance between two consecutive reflections on the SLM is $2L\tan\tau$. The optical path length between two reflections is $2L\sec\tau$. We assume a Gaussian beam with beam waist w_0 at the first plane. The beam waist is $w(z) = w_0(1 + (z/z_0)^2)^{1/2}$ where z is the propagation distance. Thus, the beam waist at reflection N on the SLM is given by $w_0\left(1 + (2LN\sec\tau/z_0)^2\right)^{1/2}$. Now we count how many reflections on the SLM are possible. One limitation is the SLM physical dimensions. The other condition to count the number of reflections is avoiding overlap of reflections. Given that 99% of the beam power lies within a circle of radius $1.5w(z)$, we use such radius as the size of one reflection. To avoid overlap we require two beams to be separated by at least a couple of pixels on the SLM. We search the number of reflections under these conditions

in MATLAB. An example of this task is shown in Fig. 3.5. This task is repeated for other values of the parameters τ, ω_0 and L. The results are shown in Fig. 3.6, where we plot the number of reflections as a function of w_0 and τ for different distances L. Fig. 3.7 shows an early implementation of an MPLC system with a flat mirror instead of the SLM, this image is displayed for showing multiple reflections between two reflecting surfaces as in an MPLC system. Such a figure shows ten reflections on the mirror and it corresponds to $L = 10\,\text{cm}$.

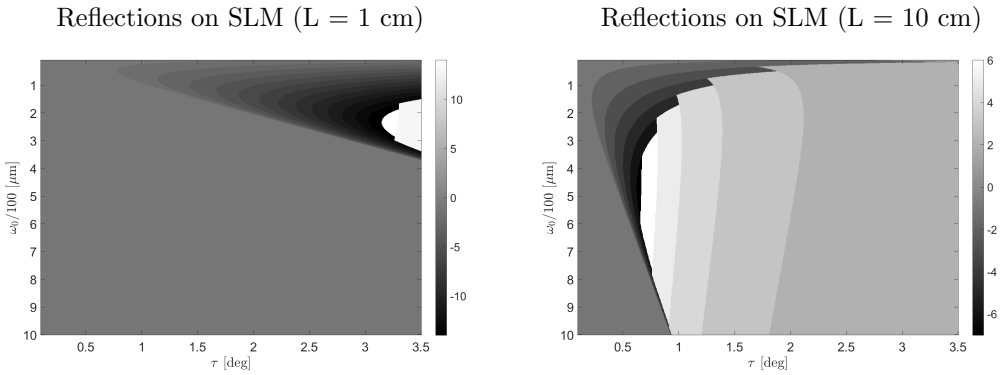


Figure 3.6: Number of reflections on an SLM in an MPLC system as a function of the insertion angle τ and the beam waist ω_0 at the first reflection plane. We show two cases of SLM-mirror distance L, the values are 10 cm and 1 cm. The regions with negative values indicate overlap between reflections at that particular reflection. For $L = 10\,\text{cm}$ the angle τ is restricted to be larger than 0.7° to achieve multiple reflections. As L decreases, the beam size is constrained to a few hundreds of microns.

Figure 3.7: Image showing multiple reflections between two reflecting surfaces as in an MPLC system. This image was captured in the summer of 2016 when I did a research internship in the Lundeen group. The mirror-mirror separation is 10 cm and ten reflections are achieved on a mirror of diameter 2.54 cm. I used a webcam for taking this image, the webcam introduced visual distortion probably due to an ultra wide-angle lens. Mirror appears ellipsoidal rather than circular, and spots appear circular when they were produced to be elliptical by using cylindrical lenses. Reflections do not lie on a straight line due to the use of a cylindrical mirror for implementing this MPLC. Notice that subsequent results used a new implementation of the MPLC as described in Section 3.3.

We proceed to perform an optimization of the system using the wavefront matching algo-

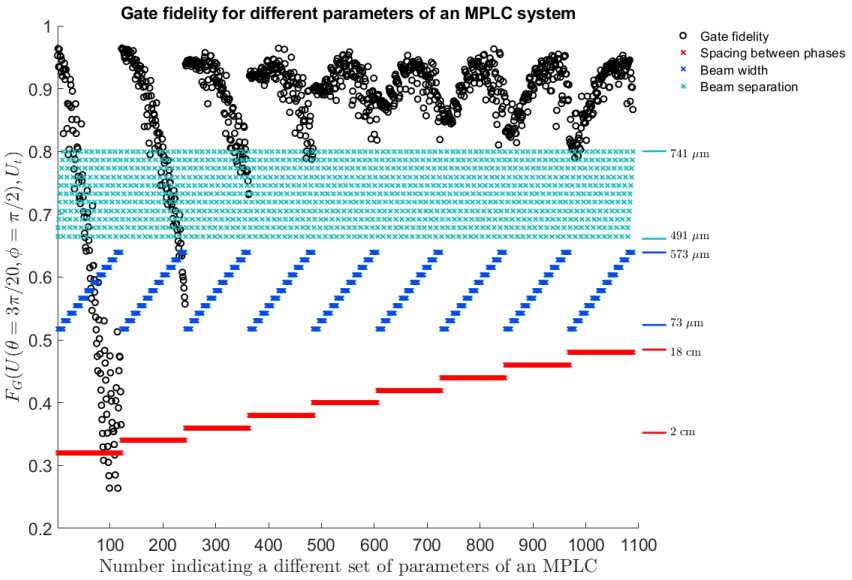
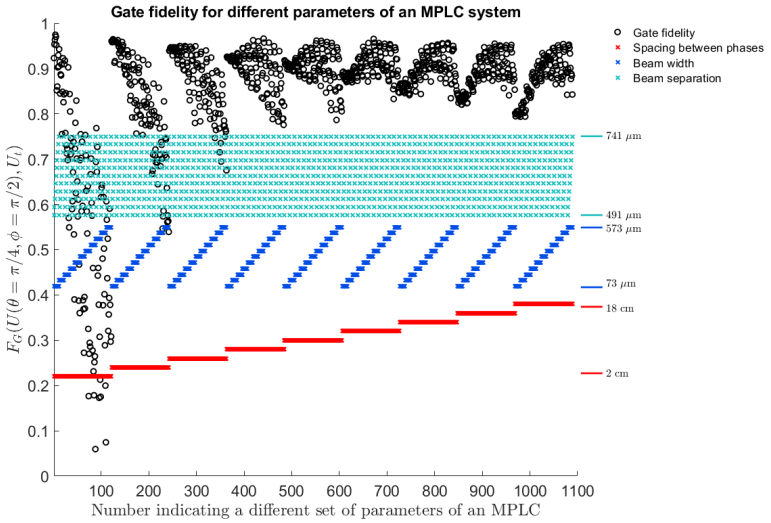


Figure 3.8: Gate fidelity for the designed unitary $\boldsymbol{U}(\theta = 3\pi/20, \phi = \pi/2)$ using different MPLC systems. The difference is given by the following parameters: distance between phase distributions, beam width, and beam spacing. Each set of parameters is labeled by an integer index displayed in the x axis. For each set of parameters, the wave front matching algorithm was used to find the phase distributions to achieve the target unitary. Gate fidelity is shown in black markers. Each of the parameters of the MPLC was scaled to be shown in the same plot with its maximum and minimum sampled values.

rithm as we explain next. We used the wavefront matching algorithm to sample the space of the design parameters. We implemented the unitary transformation given by $\boldsymbol{U}(\theta = 3\pi/20, \phi = \pi/2)$ for different design parameters. We found the MPLC achieves a gate fidelity $> 90\%$ for a wide range of parameters, as shown in Fig. 3.8. An spacing between phases smaller than 5 cm leads to small amount of diffraction and a gate fidelity smaller than 60%. For a spacing of 10 cm, the MPLC achieves a gate fidelity of 0.9 with variations of 5%, which suggests the system is robust to experimental imperfections. We selected a spacing between phase distributions equal to 10 cm, and experimentally achievable values of beam spacing and beam width for our experimental implementation. With the chosen parameters, we verified the MPLC system achieved a similar performance when changing the target unitary. This is shown for a couple of other transformation in Fig. 3.9.

We then proceed to experimentally implement the MPLC system with the chosen parameters and characterized the beams and MPLC (see details in the section below). This characterization was used to update the parameters of the design in order to match the experimental ones. We now list the final design parameters. The wavelength is 637 nm and the SLM's pixel size is 20 µm. We are using $P = 5$ phase planes and a plane spacing of $2L = 10$ cm. Adjacent input beams have a center-to-center separation of $\Delta y = 704$ µm. Each beam has a waist (half

(a)



(b)

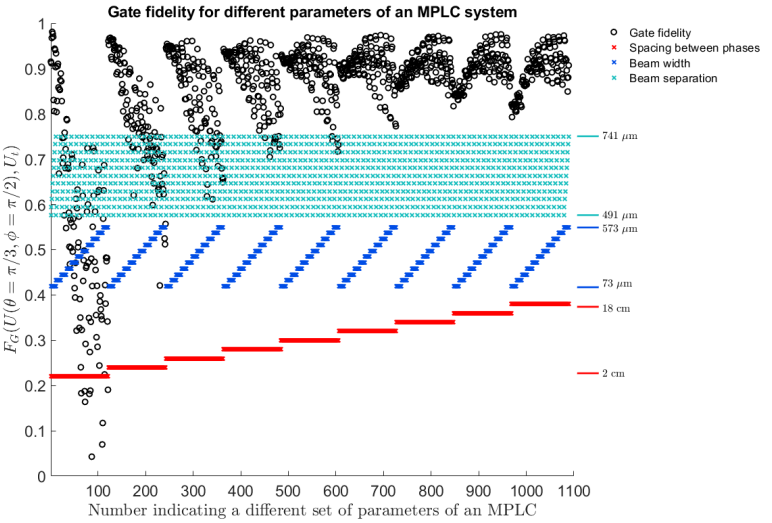


Figure 3.9: Gate fidelity for a couple of different designed unitary transformations using different MPLC systems. This figure is similar to Fig. 3.8. The difference is the target transformations, which are $\boldsymbol{U}(\theta = \pi/4, \phi = \pi/2)$ for a) and $\boldsymbol{U}(\theta = \pi/3, \phi = \pi/2)$ for b).

width at $1/e^2$ intensity) of 161.5 µm located 2.08 cm before the first phase plane. We apply a linear phase $e^{iy\pi/(18*\mathrm{Pixel})}$ on one of the beams to account for the beam tilt experimentally observed in our setup. The output beam array is the same as the input one, with the beam size and beam separation demagnified by a factor of 4. This demagnification helps the optimization algorithm to compensate for the natural diffraction and reduces losses due to diffraction out of

the light path.

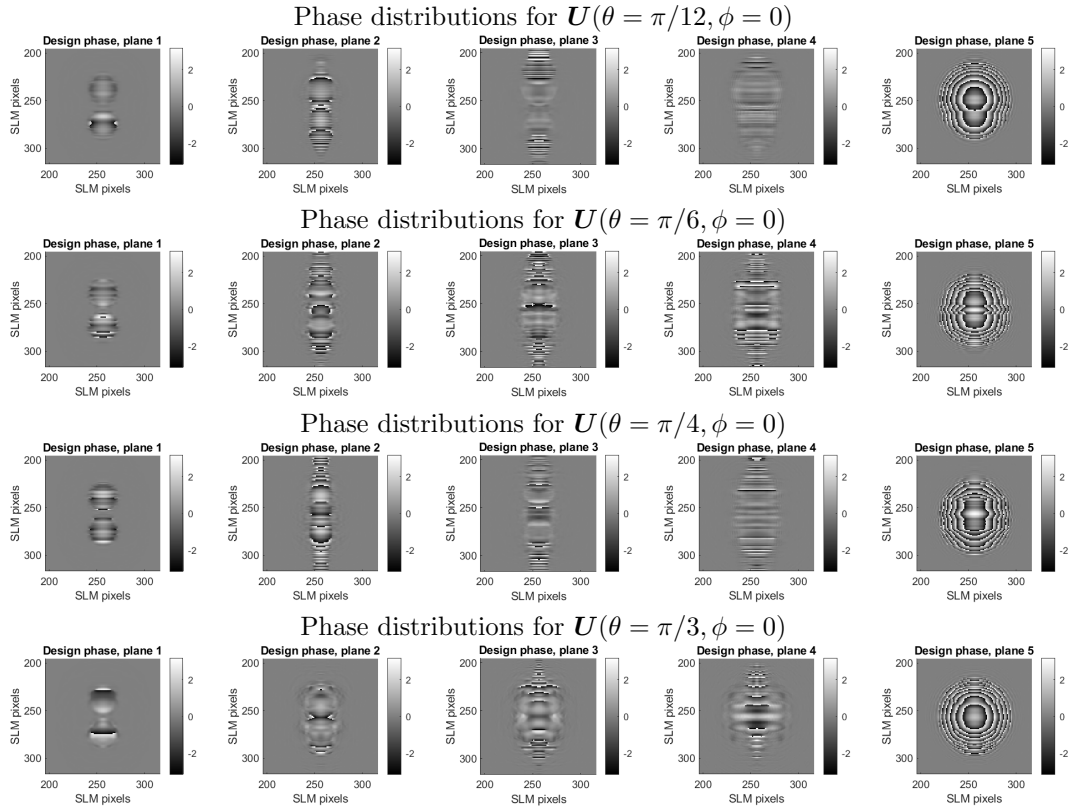


Figure 3.10: Phase distributions with the wave front matching optimization algorithm for different unitary transformations $\boldsymbol{U}(\theta, \phi = 0)$ using the optimization parameters listed in Section 3.2. The MPLC system uses five phase-mask planes (which experimentally correspond to five reflections on the SLM) and the output of the MPLC is recorded at phase-plane number six (the plane to be imaged experimentally).

Optimization results

The goal of the optimization is to design the MPLC that will be experimentally implemented and reach the target fidelity performance for the given target unitary, $\boldsymbol{U}_t$. We matched the parameters of the designed and the experimental beam array. This was done after the beam array characterization (see details in the section below). We now list the design parameters. The wavelength is 637 nm and the SLM's pixel size is 20 µm. We are using $P = 5$ phase planes and a plane spacing of $2L = 10$ cm. Adjacent input beams have a center-to-center separation of $\Delta y = 352$ µm. Each beam has a waist (half width at $1/e^2$ intensity) of 161.5 µm located 2.08 cm before the first phase plane. We apply a linear phase $e^{iy\pi/(18*\mathrm{Pixel})}$ on one of the beams to account for the beam tilt experimentally observed in our setup. The output beam array is

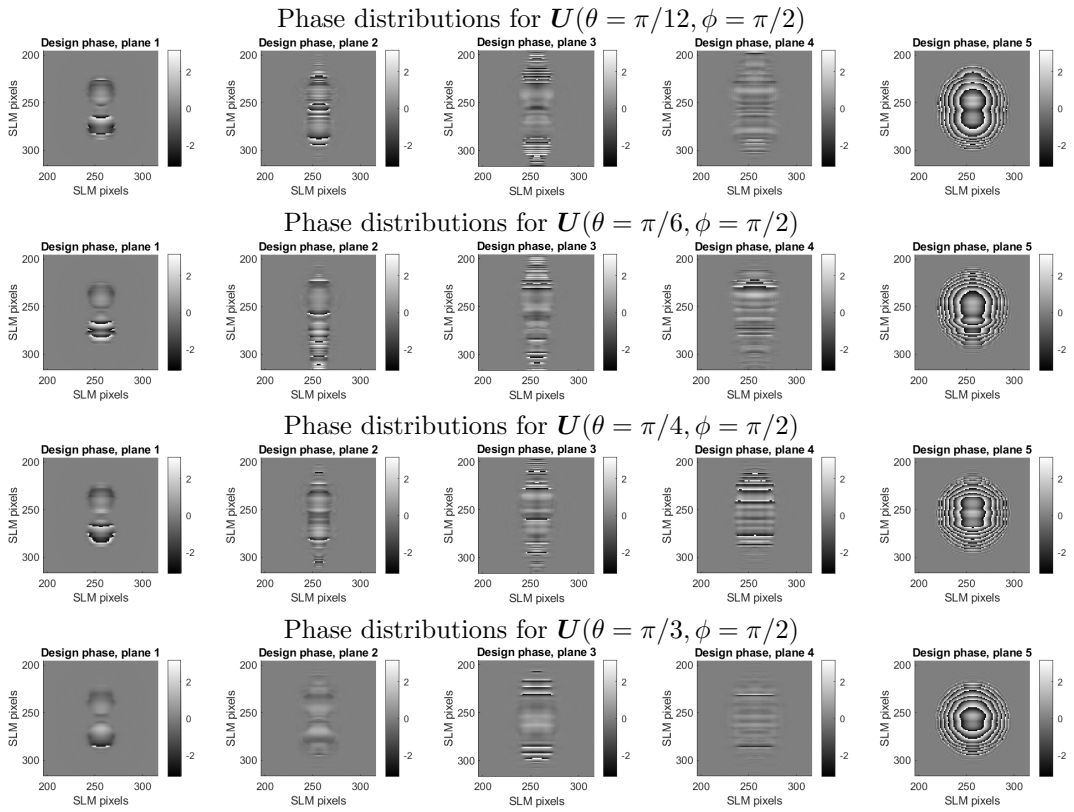


Figure 3.11: Phase distributions with the wave front matching optimization algorithm for different unitary transformations $\boldsymbol{U}(\theta, \phi = \pi/2)$ using the optimization parameters listed in Section 3.2

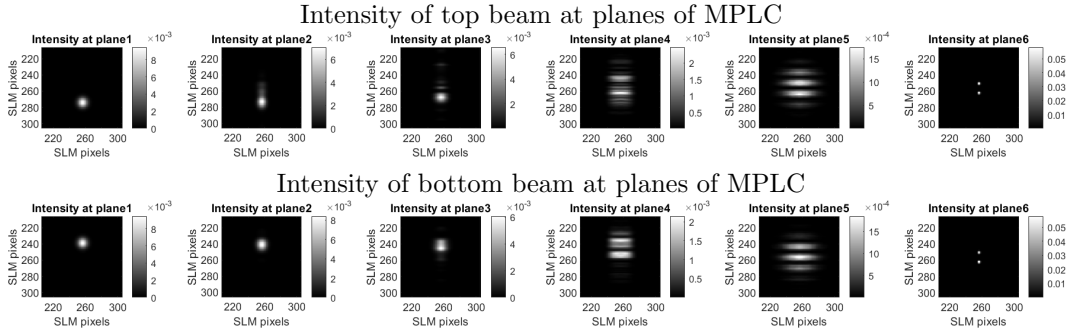


Figure 3.12: Intensity of input beams when passing by the MPLC system with a target unitary $\boldsymbol{U}(\theta = \pi/4, \phi = \pi/2)$ using the optimization parameters listed in Section 3.2.

the same as the input one with the beam size and beam separation demagnified by a factor of 4. This demagnification helps the optimization algorithm to compensate for the natural

diffraction and reduces losses due to diffraction out of the light path.

We used the optimization algorithm to sample the whole $\boldsymbol{U}(\theta, \phi)$ space with the following grid: $\theta \in [0, \pi/2]$ in steps of $\pi/120$, and $\phi \in [-\pi, \pi]$ in steps of $\pi/60$. Figs. 3.10-3.11 show some phase distributions obtained after the optimization procedure. Each design unitary the result of an optimization with 100 iterations. Each optimization takes 53.806 s to run in a desktop computer with an 8 core Intel i7-10700 processor at 2.9 GHz with 16 GB of RAM (this same computer was used in the other data processing tasks). Fig. 3.12 shows the intensity of the input beams at each plane of the MPLC when the system performs the unitary $\boldsymbol{U}(\theta = \pi/4, \phi = \pi/2)$.

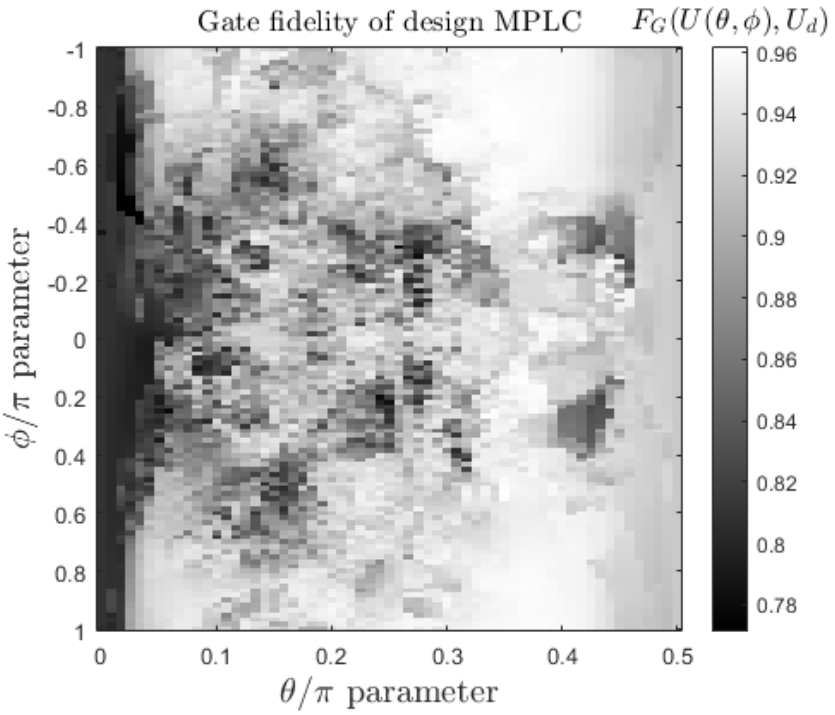


Figure 3.13: Predicted performance of the designed MPLC system. We plot the gate fidelity $F_G\left(\boldsymbol{U}\left(\theta,\phi\right),\boldsymbol{U}_d\right)$ between the target unitary transformation $\boldsymbol{U}(\theta,\phi)$, and the design one $\boldsymbol{U}_d$ obtained with the wavefront matching algorithm. The plot samples the full space of two dimensional unitary transformations.

The optimization results are shown in Fig. 3.13. Averaging F_G over the $\boldsymbol{U}(\theta, \phi)$ space, we found the average gate fidelity to be $\bar{F_G} = 0.90 \pm 0.04$, where the uncertainty is the standard deviation. However, regions with significantly higher and lower fidelity values than average can be distinguished. In particular, there is a strip region in $\theta \approx 0$ which gives the lowest fidelity, which is surprising since those unitary transformations are close to the identity. The low fidelity regions are mostly due to phase mismatch in the output states, which the wave front matching algorithm was not able to correct even after increasing the number of iterations. In this regard we now dedicate a subsection to some observations of the performance of the wavefront matching algorithm.

Observations of performance of wave front matching algorithm

The drop in the gate fidelity shown in Fig. 3.13 is due to phase mismatch mostly. The field amplitude of the design output states is same as the target ones. This can be seen by plotting an intensity gate fidelity F_{G_I} defined similarly to Eq. 3.5 with the difference being the use of intensities rather than fields:

$$F_{G_I}(\boldsymbol{U}_t, \boldsymbol{U}_d) = \frac{1}{M}\sum_{m=1}^{M}\int\int I_{d_m}(x,y)I_{t_m}(x,y)dxdy, \tag{3.6}$$

where I_{d_m} and I_{t_m} are the intensities (renormalized so that F_{G_I} takes values between zero and one) of the design and target output states when the input state is the beam $|m\rangle$. Such intensities are renormalized according to the following equation

$$I(x,y) = |\psi(x,y)|^2 / \left(\int\int |\psi(x,y)|^2 dxdy\right)^{1/2}. \tag{3.7}$$

The intensity gate fidelity for our design MPLC system is shown in Fig. 3.14. Averaging over the unitary space we obtain $F_{G_I} = 0.983 \pm 0.005$.

The wavefront matching algorithm is an optimization which can suffer from not reaching the optimal phase masks to achieve a target unitary (a global minima for the problem) but finding a set of phase distributions and not being able to further optimize them (a local minima). This behaviour is also present in a few cases of our implementation and it is shown in Fig. 3.15. We also note that the final design (the one reported in this thesis) was taken with the phase masks that optimized the performance of the system.

In Section 3.4, we show how the phase-mismatch in Fig. 3.13 can be corrected by a single phase plane at the detection plane. We now describe the experimental implementation of unitary transformations in a two-beam array. Here we exploit the reconfigurability of the MPLC.

3.3 Experimental unitary transformation in a two-beam array

The experimental setup is shown in Fig. 3.16. Our light source is a laser diode with a wavelength of 637.7 nm with a FWHM bandwidth of 1.3 nm coupled to a single-mode fibre. We can describe the setup in three stages: beam array preparation, MPLC and output state reconstruction. The two-beam array is created with a birefringent beam displacer and polarization optics (see caption of Fig. 3.16 for details). We characterized the input beam array by measuring $\omega(z)$ at several propagation planes z. The measured beam characteristics were given in Subsection 3.2. In the beam array preparation, we have the ability to choose the input mode ($|1\rangle$ or $|2\rangle$), to the MPLC.

Our MPLC uses five reflections on the SLM as the $P = 5$ phase planes, see Fig. 3.17 for an image of the experimental MPLC system. The output of the MPLC is located at a sixth

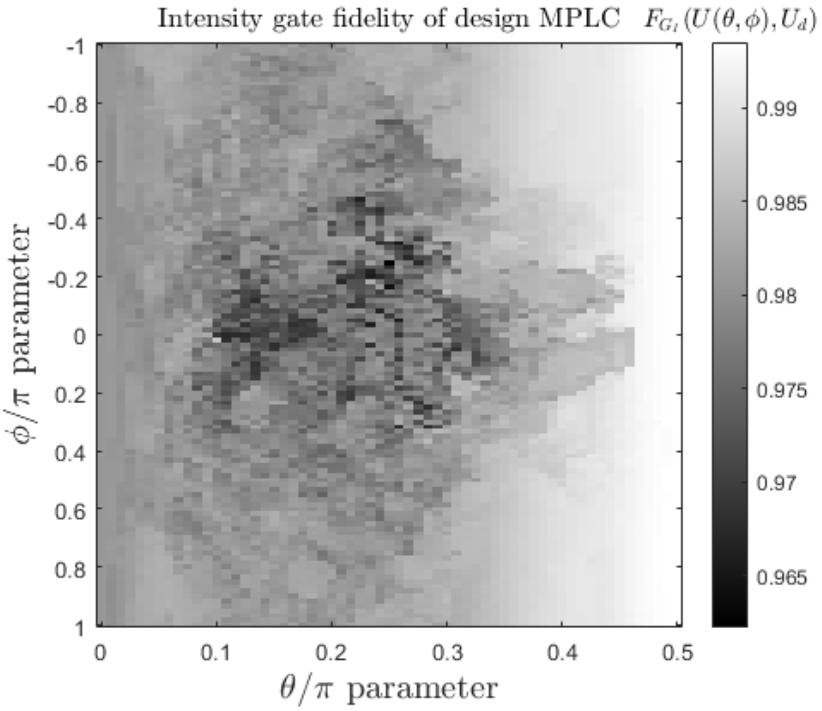


Figure 3.14: Predicted performance of the design MPLC system. We plot the intensity gate fidelity $F_{G_I}\left(\boldsymbol{U}\left(\theta,\phi\right),\boldsymbol{U}_d\right)$ (Eq. 3.6) between the target unitary transformation $\boldsymbol{U}(\theta,\phi)$, and the design one $\boldsymbol{U}_d$ obtained with the wavefront matching algorithm. The plot samples the full space of two dimensional unitary transformations.

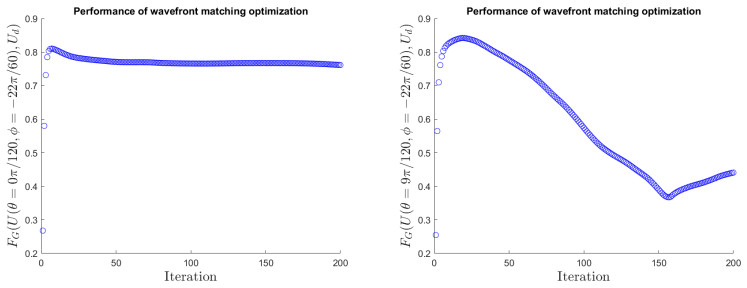


Figure 3.15: Gate fidelity $F_{G_I}\left(\boldsymbol{U}\left(\theta,\phi\right),\boldsymbol{U}_d\right)$ (Eq. (3.6)) between the target unitary transformation $\boldsymbol{U}(\theta,\phi)$, and the design one for a couple of transformations showing the last iteration of the optimization does not necessarily lead to the optimal performance of the MPLC.

phase-mask plane after the SLM, this is the plane to be imaged for measuring the performance of the MPLC. The SLM is a phase-only liquid crystal SLM (Hamamatsu X10468-07) with an effective area of $15.8 \times 12\,\text{mm}$ and a fill factor of 98%. The SLM size is 792×600 pixels, with a pixel pitch of $20\,\mu\text{m}$. The SLM has high throughput efficiency from 620 to 1100 nm. Each pixel can be set to a value ranging from 0 to 256. We use the calibration provided by the

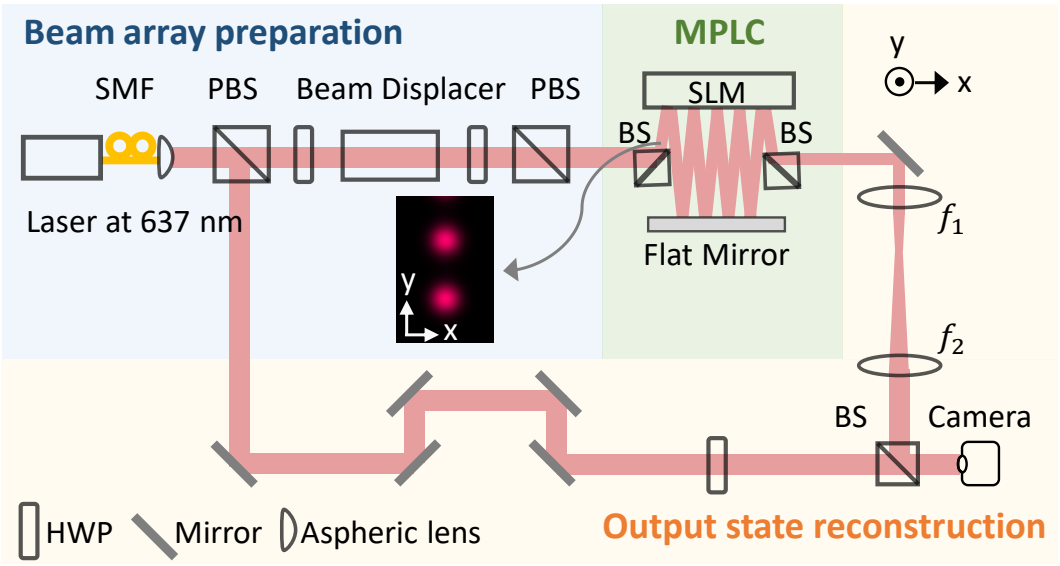


Figure 3.16: Experimental setup for the realization of an arbitrary two-dimensional unitary transformation using a multi-plane light converter (MPLC). **Beam array preparation**. A polarizing beam splitter (PBS) and a half wave plate (HWP) produce light linearly polarized at 45°. A beam displacer shifts the spatial mode of photons with vertical polarization, and transmits spatially unshifted photons with horizontal polarization. The beam displacer creates the two-beam array. After the beam displacer, a HWP and a PBS select the input state. The transmitted light at the PBS is horizontally polarized, which is the working polarization of the spatial light modulator (SLM). **MPLC**. The MPLC is created with an SLM and a flat mirror. The beam array is inserted to the MPLC by a non-polarizing 50:50 beam splitter. Five reflections are achieved in the SLM. After which, a second 50:50 beam splitter is used to extract the output states from the MPLC. A sixth phase-mask plane after the beam splitter corresponds to the output plane of the MPLC. The efficiency of the five reflections on the MPLC (without considering the two beam splitters) is 10%. **Output state reconstruction**. A reference beam is created at the first PBS. The reference beam propagates through a delay line which compensates the optical path difference of the MPLC. A HWP at 45° sets the reference beams's polarization to be horizontal. The output plane of the MPLC is imaged on a CMOS image sensor (camera) by a 4f-lens pair (magnification 1.25). The imaged output states and the reference beam interfere after a beam splitter and the interference pattern is recorded the camera. Output state reconstruction is performed computationally from the interference pattern.

manufacturer, which gives a pixel value-phase linear response and a pixel value of 118 for a 2π modulation at 637 nm. We do not use the phase to correct aberrations provided by the manufacturer. The phase Φ_{SLM} on the SLM is obtained as follows

$$\Phi_{SLM} = \frac{118}{2\pi} \mod (\Phi, 2\pi), \tag{3.8}$$

where Φ is the phase at the five phase planes, directly obtained from the wavefront matching algorithm.

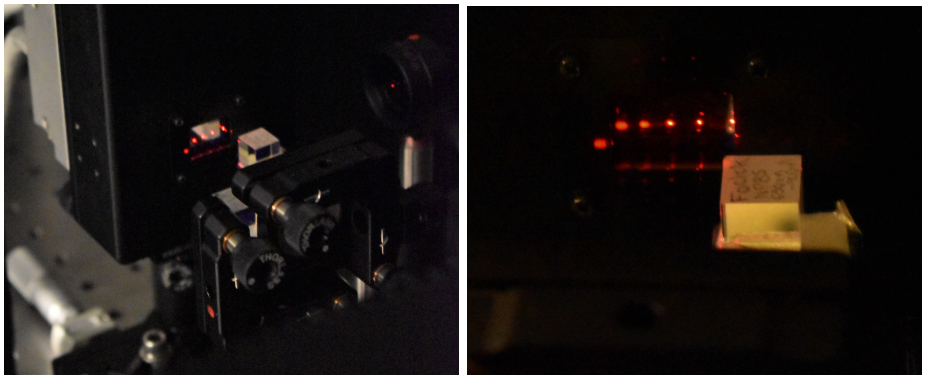


Figure 3.17: Images of the experimental MPLC. The SLM can be appreciated with the five reflections on it as well as the beam splitter used to insert the beams into the MPLC.

We now describe the steps for characterizing the beam array. The beam size of every mode of the beam array is measured at several z planes to characterize it. Each of the modes has a beam waist of $w(z = 0) = 161.5\,\mu\text{m}$ and it is located 2 cm before the first plane. At this stage one can detect experimental imperfections as we found with the tilt between the two beams. From this characterization, we know the beam size of the beam array at every phase plane. This information is used for setting the parameters of the optimization algorithm and as a diagnosis tool of the MPLC alignment.

Centring each of the phase distributions

Locating the centers of the reflections on the SLM is performed through random phase distributions as described next. The method is based on the razor method for Gaussian beam characterization [71]. A razor is moved along the y direction across the beam. The transmitted intensity is used to obtain the center of the beam, and the beam waist [71] at the plane of the knife edge. The location of the second beam in the beam array can be obtained from the location of the first beam and the knowledge of the beam separation. This technique can be emulated with an SLM. A random phase distribution diffracts light in all directions, thus, in the far field, it creates loss. A random phase that gradually moves in the y direction effectively acts as a razor. A power meter at the far field can detect the power as a function of the random phase distribution position in y. This data is used for obtaining the center and beam waist of the beams at each reflection. The drawback of this method is that it requires post-processing.

A more practical modification of the razor technique can be used instead. Create a random phase mask, and move it through the SLM. Place a power meter after the SLM. Initially, the total power of a beam is measured. As the random phase is moved on the SLM, the beam disappears when the random mask is on one of the beam reflections. The random phase mask can be cropped to a rectangle to match the size of the beam. This provides an estimate of the beam center and beam size. This technique is less accurate than the full razor one, but it is simpler and faster to implement.

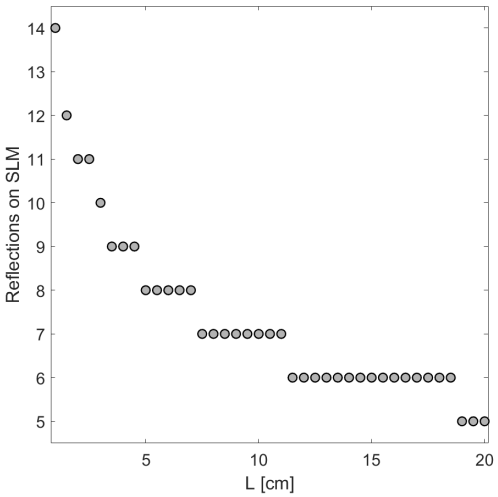


Figure 3.18: Maximum number of reflections on an MPLC system as a function of the mirror-SLM separation L. Small values of L provide the most reflections in an MPLC system. As L decreases, the diffraction also decreases possibly affecting the coupling between the transformed modes. Thus, L can not be arbitrarily small. The maximum number of reflections is around ten (depending on the beam array parameters).

Fine tuning

The centers of the phase distribution need to be further tuned. We now describe such a method. A $\boldsymbol{U}(\theta, \phi)$ transformation is the target unitary. We start by the first plane, the position of the first phase distribution (x_1, y_1) is moved in the y direction in steps of one pixel. The experimental intensity $I_e(x, y)$ is measured on the camera. We compare $I_e(x, y)$ with the design intensity $I_d(x, y)$, which corresponds to the same experimental configuration i.e., with only the first phase applied. The figure of merit in this comparison is the 'Intensity fidelity' $F_I(I_d, I_e)$ between I_d and I_e, $F_I(I_d, I_e)$, defined as follows

$$F_I\left(I_d, I_e\right) = \int I_d\left(y\right) \times I_e\left(y\right) dy, \tag{3.9}$$

where $I_d\left(y\right) = \int I_d\left(x, y\right) dx / \left(\int \left(\int I_d(x, y) dx\right)^2 dy\right)^{1/2}$ is the design intensity along y. The normalization of such intensity ensures that $F_I\left(I_d\left(y\right), I_d\left(y\right)\right) = 1$. The experimental intensity $I_e\left(y\right)$ is calculated in an similar way. We move the phase distribution, pixel by pixel along y, maximizing $F_I\left(I_d, I_e\right)$. We repeat the same procedure for the x direction. We noticed that the y direction is more sensitive than the x direction.

Having optimized the first phase, we apply the first two phase distributions and repeat the fine tuning procedure for the second phase. We repeat this procedure with the other phase distributions. Ideally, these positions are fixed and should work for any $\boldsymbol{U}(\theta, \phi)$. Our data

$\boldsymbol{U}(\theta, \phi = \pi/2)$	F_G five PPs	F_G two PPs
$\theta = \pi/2$	0.91 ± 0.01	0.914 ± 0.009
$\theta = \pi/3$	0.859 ± 0.008	0.85 ± 0.05
$\theta = \pi/4$	0.69 ± 0.06	0.86 ± 0.06
$\theta = \pi/5$	0.83 ± 0.02	0.85 ± 0.06
$\theta = \pi/6$	0.83 ± 0.02	0.82 ± 0.01

Table 3.1: Gate fidelity $F_G(\boldsymbol{U}(\theta, \phi = \pi/2), \boldsymbol{U}_e)$, calculated using eq. (3.5), between target unitary transformations $\boldsymbol{U}(\theta, \phi = \pi/2)$ and the experimentally obtained ones $\boldsymbol{U}_e$ using five and two phase planes (PP). The positions of the PPs were not changed when reducing the number of planes. We found using two PPs give same or better performance as five PPs.

acquisition was performed with fixed positions for the phases, even when changing the target unitary. However, we noted that some transformations experimentally provided higher values of the intensity fidelity after small changes (shifting one or two phase masks a few pixels) in the phase's centers.

Once the centring of the five phases is done, the designed phases for different unitary transformations can be applied. The field of each output state is reconstructed by off-axis holography (for a more detailed description of the method, see for example [72, 73]). Then the gate fidelity can be calculated using eq. (3.5). We exploit the reconfigurability of the SLM by experimentally sampling the two-dimensional unitary space $\boldsymbol{U}(\theta, \phi)$. The sampling used the following grid: $\theta \in [0, \pi/2]$ in steps of $\pi/120$ and $\phi \in [-\pi, \pi)$ in steps of $\pi/60$.

The data acquisition used solely the last two, of the five design phase distributions, for realizing each unitary. We now explain the reason of this. While doing the experiment, we noticed the last two design phase planes are the most important ones to realize the unitary transformations $\boldsymbol{U}(\theta, \phi)$. To that point, Table 3.1 shows the gate fidelity between target transformations $\boldsymbol{U}(\theta, \phi = \pi/2)$ and the experimentally obtained ones. The performance is same or better using only the last two phases, while setting the first three to zero. When using all five design phases for different unitary transformations, the MPLC produced a complicated scattered intensity distribution. Such an intensity was qualitatively different than the expected one. The physical mechanism for this behavior is unknown and a subject for future investigation. All results presented in this Chapter were taken with the last two phase distributions for each unitary.

As for data acquisition, we set the phase planes on the SLM to perform a unitary $\boldsymbol{U}(\theta, \phi)$. For every input state $|\psi\rangle = |m\rangle$ $(m = 1, 2)$, we record five images of the interference pattern. Each image is used to numerically reconstruct the optical field of the output state by off-axis holography. Acquiring this set of ten images takes 25 s. The off-axis holography routine takes 44.12 s using MATLAB (most of the time is taken by geometric transformations such as magnifying and flipping that account for the last 4f system in our experimental setup). The alignment procedure, and data acquisition are performed in LABVIEW.

3.4 Performance assessment of the implemented unitary

We first present the results of the method as described so far. We will then describe how the use of one additional phase plane improved the average fidelity considerably.

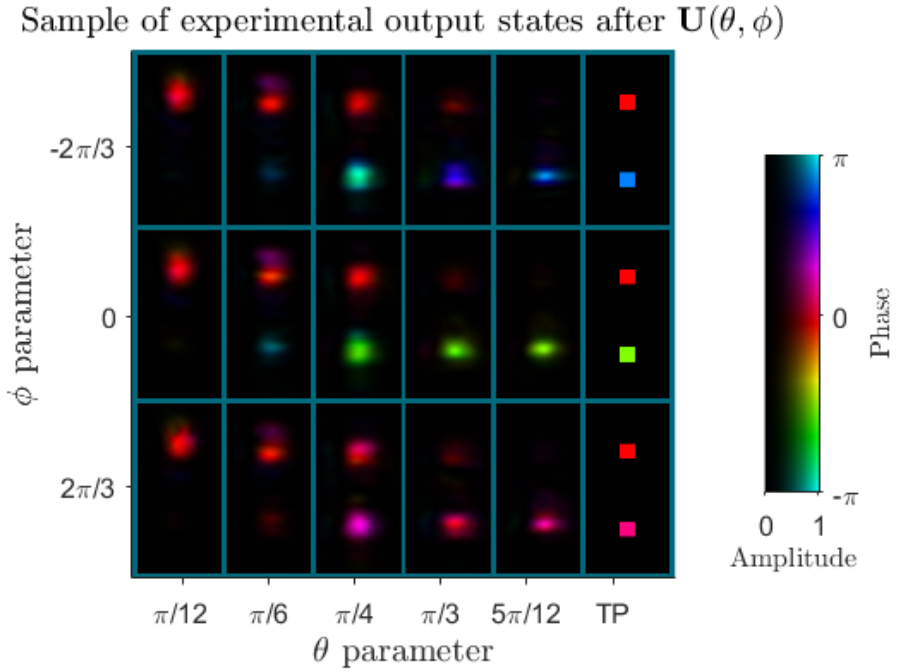


Figure 3.19: Field of experimentally obtained output states for different transformations $\boldsymbol{U}(\theta, \phi)$. The blue grid separates different states corresponding to a different $\boldsymbol{U}(\theta, \phi)$. The values of θ and ϕ are indicated on the axes. In this sample the input state is the beam $|1\rangle$ of the beam array. The last column shows the target phase for the fields on the same row. The field amplitude of each state was normalized to have the same maximum amplitude, this was done for image visibility.

Results - Part I: No correcting mask

Fig. 3.19 shows a sample of experimentally obtained output states for different unitary transformations $\boldsymbol{U}(\theta, \phi)$ (values of θ and ϕ are indicated in the axes creating a grid shown in blue color) when the input state is the beam $|1\rangle$. The amplitude and phase of the field's states are shown in the brightness and hue of the plot respectively. This figure explicitly demonstrates the ability of an MPLC to implement the full gamut of unitary transformations. The obtained output states cover the full range of amplitude and phase a beam array state can have. As a result, we are also able to generate any state in the two-beam array. Different degrees of amplitude and phase mismatch can be appreciated in Fig. 3.19.

The results for Part I–No correcting mask are shown in Fig. 3.20. We show the gate fidelity

$F_G(\boldsymbol{U}_t, \boldsymbol{U}_e)$ (obtained by using eq. (3.5)), between target unitary $\boldsymbol{U}_t$ and the experimentally implemented by the MPLC $\boldsymbol{U}_e$. Averaging over all the unitary space, we got an averaged gate fidelity $\bar{F}_G = 0.67 \pm 0.17$. This plot shows our experiment is able to implement a large set of transformations with a large gate fidelity (larger than 80%). However, there are regions with a gate fidelity smaller than 0.5 in the center of the plot. We next investigated how to improve the performance of the MPLC over the unitary space.

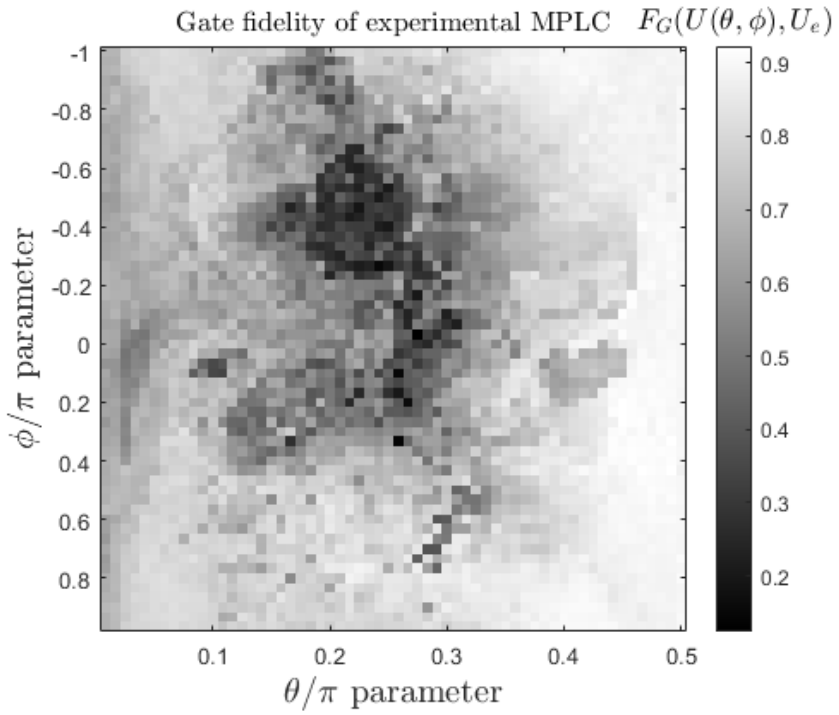


Figure 3.20: Experimental gate fidelity $F_G\left(\boldsymbol{U}(\theta,\phi), \boldsymbol{U}_e\right)$ (calculated using eq. (3.5)) between target unitary $\boldsymbol{U}\left(\theta,\phi\right)$ and the unitary $\boldsymbol{U}_e$ experimentally implemented by the MPLC system. These results correspond to Part I: No correcting mask. We sample the full space of the unitary transformations in a two-beam array given by eq. (2.13). This plot is the experimental analogue of Fig. 3.13.

A possible method to improve the gate fidelity is to move the location of each phase distribution for every transformation. The field distribution in the actual MPLC might be slightly offset from the nominal field used in the optimization. To account for that, one might consider 're-centring', shifting each phase distribution on the SLM by a few pixels with the goal of improving the output states. To be clear, we did not do this for the data in Fig. 3.20. The problem of performing such a re-centring optimization is that it is not theoretically needed as the designed MPLC has fixed input beams. Thus, re-centring does not guarantee improving the gate fidelity. Additionally, the centring procedure is time-consuming which challenges the reconfigurability feature of our experiment.

Instead, we followed a different approach for improving the averaged gate fidelity of our

experiment. In Fig. 3.20 there is a region around $\phi = \pi/2$ where the fidelity is higher than 0.8 independent of θ. Such a strip of transformations accounts for the value of θ of the transformation $\boldsymbol{U}(\theta, \phi)$. We use an additional phase to adjust the phase ϕ to create any target unitary $\boldsymbol{U}(\theta, \phi)$. This is the method experimentally used and reported in the next subsection.

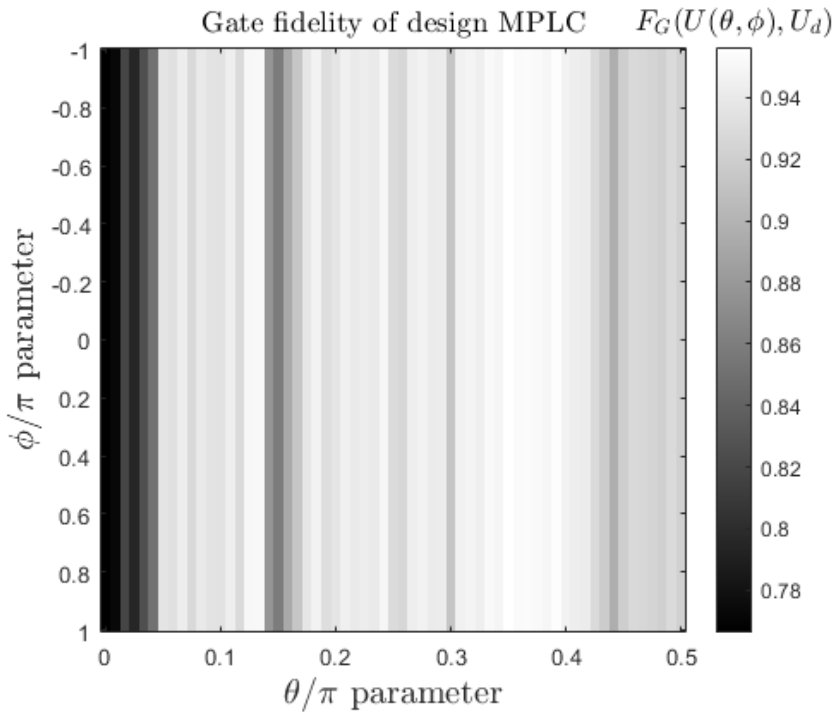


Figure 3.21: Performance of the design MPLC system plus a correcting phase plane, as explained in Results - Part II: With correcting mask. We plot the gate fidelity $F(\boldsymbol{U}(\theta,\phi), \boldsymbol{U}_d)$ (eq. (3.5)) between the target unitary $\boldsymbol{U}(\theta,\phi)$, and the one implemented by the MPLC $\boldsymbol{U}_d$. Gate fidelity is calculated for every unitary of dimension two. This figure uses the set of transformations $\boldsymbol{U}(\theta, \phi = \pi/2)$ from Fig. 3.13 and an additional phase plane to generate the unitary space.

Results - Part II: With correcting mask

Our experimental results presented in Fig. 3.20 show a non-ideal performance of our MPLC over the two dimensional unitary space. In other words, there are regions with contrasting fidelities. To improve the performance of the MPLC, we use the line of designs along $\boldsymbol{U}(\theta, \phi = \pi/2)$ and an additional phase located at the detector plane to achieve any other transformation $\boldsymbol{U}(\theta, \phi)$. Such a phase plane applies a phase difference $\Delta\phi = (\phi - \pi/2)/2$ to the top beam and $-\Delta\phi$ to the bottom one. The design gate fidelity with this approach is shown in Fig. 3.21. It has an averaged gate fidelity $\bar{F}_G = 0.92 \pm 0.04$, which has a percentage difference of 2.2% higher respect to the one of Fig. 3.13. Thus, our correcting mask approach preserves the performance of the

original design MPLC $\boldsymbol{U}_d$. The lower performance of transformations with $\theta = 0 - 6\pi/120$ is due to a phase gradient of π across each beam, which is not fixed by the wave front matching algorithm.

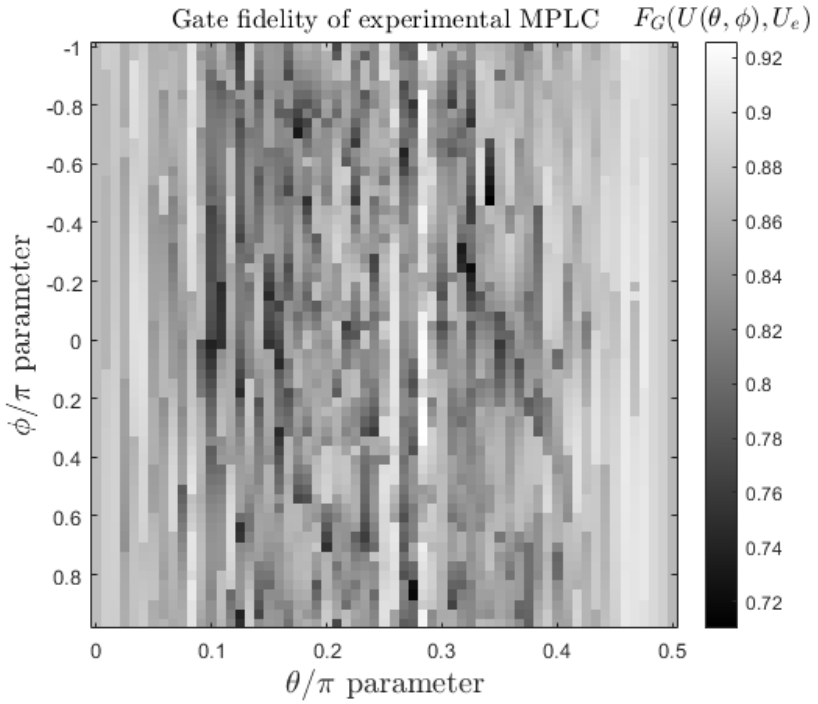


Figure 3.22: Experimental gate fidelity $F_G(\boldsymbol{U}(\theta,\phi),\boldsymbol{U}_e)$ (Eq. (3.5)) between target unitary $\boldsymbol{U}(\theta,\phi)$ and the unitary $\boldsymbol{U}_e$ experimentally implemented. These results correspond to Part II: With correcting mask..

For the experimental implementation of the correcting mask approach, we used an additional SLM located at the output plane of the MPLC system from Fig. 3.16. This SLM is a Pluto-2 phase only SLM with a dielectric mirror at its back plane to enhance reflectivity in the range of $730 - 950\,\text{nm}$. It achieves a maximum modulation of 2.4π with a pixel value of 255 at 637 nm. Implementing the correcting mask approach yielded the results shown in Fig. 3.22. we obtained an experimental averaged gate fidelity of $\bar{F}_G = 0.85 \pm 0.03$. This is a much improved result when compared to Fig. 3.20 (note the change in scale).

3.5 Experimental vs designed performance

We found a mismatch in the performance of the designed MPLC (Fig. 3.13) and the experimentally obtained one (Fig. 3.20). The reason of this mismatch may be due to experimental imperfections that effectively produce additional phases not taken into account in the design. Examples of such imperfections include light scattering, misalignment of the system, mismatch between design and experimentally available beams, positioning of phase planes, SLM's surface

curvature, inherent grid in pixel structure of an SLM, etc. Modelling these imperfections is challenging and it gets harder as more phase distributions are used. By using a non unrealistic phase distortion, we calculated the effect of a phase perturbation in an MPLC. Such a phase perturbation has an average (averaged over the five locations of the reflections on the SLM) gradient equal to 0.058 ± 0.02 rad(SLM pixel)$^{-1}$. This perturbation causes a 20% drop in gate fidelity. The details of this calculation are now given.

The procedure is as follows. Use the designed phase planes from the wave front matching algorithm to implement a unitary $\boldsymbol{U}(\theta, \phi)$, add a perturbation phase $\alpha\Phi_{pert}$, with α being an amplification factor, and Φ_{pert} the additional phase. Thus the final phase imparted by the perturbed MPLC system is the sum of the phase obtained by the wave front matching algorithm and $\alpha\Phi_{pert}$. We propagated the input beams by such a perturbed MPLC system and calculated the gate fidelity $F_G(\boldsymbol{U}(\theta, \phi), \boldsymbol{U}_{pert})$ between the target unitary and the perturbed MPLC. We took Φ_{pert} to be the one provided by the manufacturer to correct for non-flatness of the SLM, i.e., the SLM applies an spatially varying phase even when no phase is being displayed on it, the correction pattern Φ_{pert} is supposed to cancel such an effect. Fig. 3.23 shows the gate fidelity of a perturbed MPLC with a target unitary $\boldsymbol{U}(\theta = \pi/4, \phi = \pi/2)$. The perturbation phase $\alpha\Phi_{pert}$ was applied with different values of α. The trend of Fig. 3.23 also appears for other values of θ. We found the gate fidelity decreases even for values of $\alpha < 1$. The phase masks obtained from this example are shown in Fig. 3.24. For $\alpha = 1$, such a phase perturbation has an average gradient equal to 0.058 ± 0.02rad(SLM pixel)$^{-1}$. The average was taken over the five reflections on the SLM.

This example shows that simple phase patterns can alter the performance of the MPLC. The effects of misalignment, parameter mismatch due to experimental imperfections, positioning of phase planes, spatial variations across the SLM, etc. overall contribute to a phase pattern (in general different from the correction pattern Φ_{pert}) that causes the mismatch between the design MPLC and the experimentally obtained one. This phase perturbation is harder to correct as more planes are used. We did not use the correction pattern Φ_{pert} experimentally, because it strongly distorted the beams even when no other phase was applied. The perturbation phase is given by $\alpha\Phi_{pert}$, it is added to the design phase planes of the design MPLC system. Figs. 3.24 and 3.25 show the design phase planes, the correction pattern with different values of α and the sum of both phases. This final phase is the one used by the perturbed MPLC system.

Potential future directions

An MPLC could be used as an auxiliary device in a quantum computer e.g., for state preparation or in the implementation of gates. For such an application an SLM with higher efficiency is needed. Nowadays, there are SLMs with $> 97\%$ light utilization efficiency, which is achieved by using a dielectric mirror at the end of the SLM.

The realization of reconfigurable unitary transformations in a high dimensional beam array is the ultimate goal of an MPLC system. Such a device would allow one to use linear optics in free space for implementing and studying boson sampling experiments, the implementation of quantum gates and random walks which have many applications, see for example [74]. We

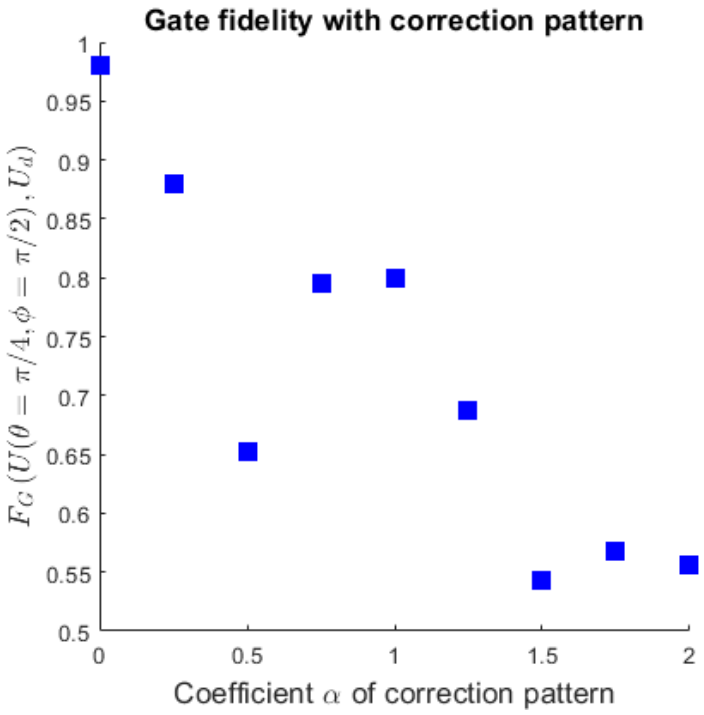


Figure 3.23: Gate fidelity $F_G\left(\boldsymbol{U}_t, \boldsymbol{U}_d = \boldsymbol{U}_{pert}\right)$ between target unitary $\boldsymbol{U}_t(\theta = \pi/4, \phi = \pi/2)$ and the unitary $\boldsymbol{U}_{pert}$ implemented by a perturbed MPLC system. The perturbation phase is given by $\alpha\Phi_{pert}$ with Φ_{pert} being the correction phase pattern provided by the manufacturer of the SLM. This perturbation phase was used as an example of a realistic phase perturbation that can appear in an MPLC system.

think it would be interesting to use an MPLC for specific 'meta-optics' transformations. For example the conception of a 'spaceplate' [75] has similarities with an MPLC system. Thus combining both systems could lead to an improvement of the performance of these devices.

3.6 Conclusions

As suggested by the experimentally obtained results, an MPLC system is a challenging system to characterize. It is sensitive to experimental errors which are hard to eliminate. On the other hand, the reconfigurability feature of an MPLC allows experimentalists to adjust input states and phase distributions of the MPLC system, and to measure the output states after such changes. Such experimental information can potentially be used to optimize the phase distributions of the MPLC, and account for experimental imperfections that were not taken into account in the design of the MPLC. There are different approaches to this experimentally optimized version of an MPLC system. For example, the phase distributions could be optimized by machine learning techniques, genetic algorithms, stochastic methods or by the optimization of an analytical model for the phases. This is undergoing work in our lab and we leave details

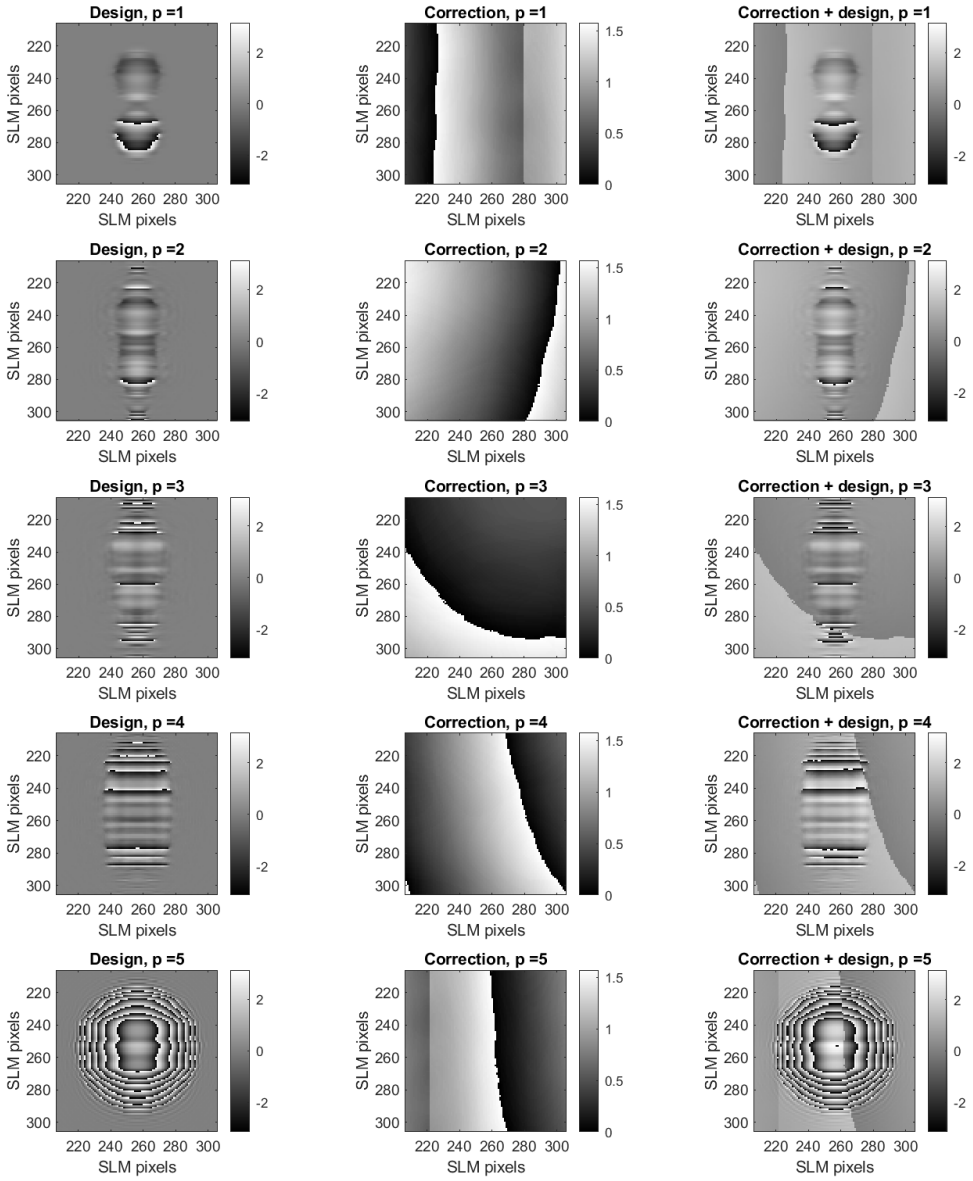


Figure 3.24: Phase planes for the design and perturbed MPLC systems. First column shows the design phase planes at each of the $p = 1, ..., 5$ phase planes for the unitary $\boldsymbol{U}(\theta = \pi/4, \phi = \pi/2)$. Second column shows the perturbation phase $\alpha\Phi_{pert}$ with $\alpha = 0.25$. The third column shows the phase addition of the columns one and two.

for future work.

In summary, we experimentally demonstrated a reconfigurable MPLC capable of implementing arbitrary unitary transformations in a two-beam array. Using an additional phase distribution, we achieved an averaged gate fidelity of $\bar{F}_G = 0.85 \pm 0.03$ which demonstrates the benefits of our method. This work demonstrates the usefulness of MPLC systems and how

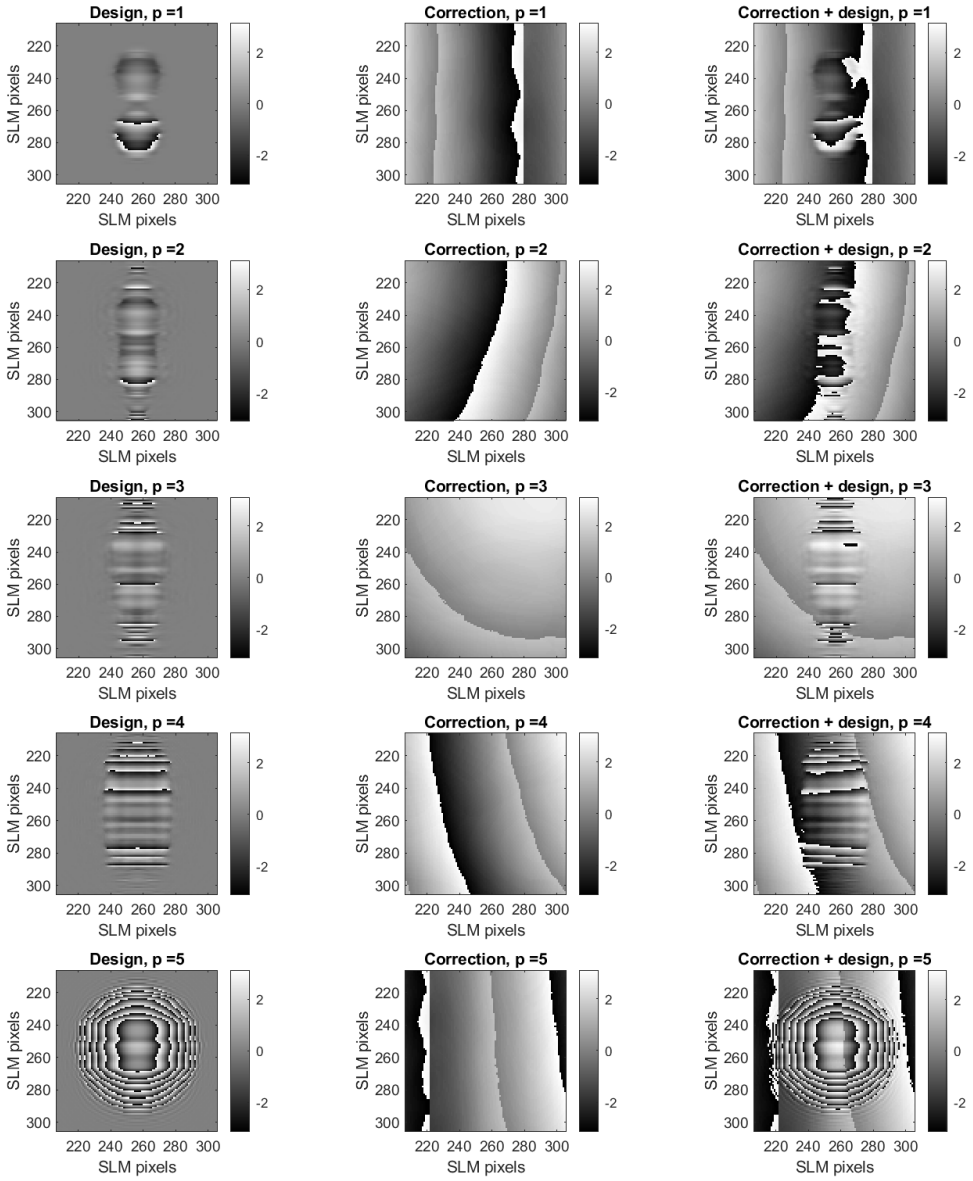


Figure 3.25: Phase planes for the design and perturbed MPLC systems. First column shows the design phase planes at each of the $p = 1, ..., 5$ phase planes for the unitary $\boldsymbol{U}(\theta = \pi/4, \phi = \pi/2)$. Second column shows the perturbation phase $\alpha\Phi_{pert}$ with $\alpha = 1.25$. The third column shows the phase addition of the columns one and two.

to use them in a reconfigurable way, leading to a novel photonics tools that can benefit both classical and quantum optics.

Chapter 4

A method to determine the M2 beam quality from the electric field in a single plane

This chapter is based on M. H. Griessmann, A. C. Martinez-Becerril, and J. S. Lundeen, 'A method to determine the M2 beam quality from the electric field in a single plane', *Opt. Continuum* 2, 1833-1848 (2023).

4.1 The problem of a laser beam's characterization

Upon its invention in 1960, the laser immediately distinguished itself from other light sources by its high quality spatial coherence, which allowed it to propagate long distances in a pencil-like beam. Though this beam quality is just one of the laser's many exemplary properties, it has proven particularly useful to science and technology, for example in distance measurements e.g., to the moon; high-resolution imaging, cutting, machining, welding, and 3D printing; the characterization of surface profiles such as the cornea; the construction of interferometric sensors e.g., for gravitational waves; and long-distance power delivery and communication. Nonetheless, it took another 30 years for the field to settle on a standard performance metric for laser beam quality. To this end, in 1990, Siegman drew attention to the M^2 parameter [76–78], the ratio of the product of the beam size in the near and far fields to the same product for a diffraction limited beam [77]. It is common to write M2 in titles instead of M^2 to make articles searchable in databases, we do this in the title of this chapter for highlighting this point. A reliable way to measure M^2 was codified in 1999 by the ISO standard 11146 [79], which called for transverse spatial intensity-profiles to be recorded at ten distances along the propagation axis. The M^2 parameter is now routinely measured and reported in scientific research and included in the advertised specifications of commercial lasers.

Since 1999, beam measurement tools and optical modelling have continued to advance beyond intensity spatial-profiles. There are now established methods and even commercial devices for experimentally measuring the electric field spatial-profile of a beam. These include newer methods such as off-axis holography [80,81] and direct measurement [82], as well as time-

tested devices such as Shack-Hartmann sensors [83–86] and shear plates [87, 88]. In addition, with advances in computation, researchers and engineers, can now perform accurate electric field modeling of a proposed device design using finite-difference time-domain, transfer-matrix, or finite element analysis methods. With such commercial or custom optical simulation software one can predict the electric field profile produced by a nano-antenna, or transmitted by a multilayer thin-film stack, or reflected by a metalens, to give some examples. The M^2 parameter is often the goal of such simulations, particularly of devices in which new optical modes are excited e.g., large area optical fibers [89–91], materials with a non-linear optical response [92], and laser cavities [93, 94]. In light of these advancements, intensity spatial-profiles may not always be the ideal choice for determining M^2.

Unlike an intensity profile, all the information about a laser beam's quality is contained in its electric field spatial-profile in any single transverse plane. This makes the numerous intensity profiles and subsequent fitting of the ISO 11146 inconveniently circuitous, particularly for optical simulations. Moreover, these numerous measurements in the ISO standard are included, in a large part, to make the procedure robust to measurement errors (e.g., background) and uncertainties, which numerical modelling does not have. On top of this, for both simulations and experiment, the ISO procedure is additionally problematic since intensity profiles must be acquired at prescribed distances from the beam waist. This requires a pre-characterization of the beam before the numerous profiles are obtained, i.e., to find the approximate location of the beam waist and the Rayleigh range. The complexity of the ISO standard has motivated work over the last decades to simplify the experimental procedure to obtain M^2; for an overview see [95, 96].

Early papers on M^2 [97–102] focused on characterizing M^2 through intensity measurements. Some papers [90, 99, 100, 103] have developed theory based on the electric field distribution to derive formulae for M^2 for specific beam shapes with analytic forms. Recently, other works have developed experimental methods to obtain M^2 from the electric field, albeit indirectly. For example, [85] demonstrated a method to determine M^2 using a Shack-Hartmann sensor along with near and far field intensity profiles, and [104] measured the modal decomposition of a beam in order to calculate M^2. Other works numerically replicated one of the experimental methods to obtain M^2. For example, [105–107] numerically propagated the electric field to different planes to replicate the ISO standard method, and [108] also performed numerical propagation of the electric field to three different planes and added numerical propagation through a cylindrical lens to simulate the experimental procedure from [101, 109]. Such a cylindrical lens has to be carefully selected every time that a different beam is used to ensure a fair sampling of the beam. In this work, using a position-angle phase-space picture, we motivate the fundamental meaning of M^2 and use that to introduce a method to find M^2 that only requires the electric field at one plane. This eliminates the need for numerical propagation or numerical projection onto a modal basis. We numerically validate such a method for two nontrivial test beams. Our method is particularly useful in optical simulations where the complex electric field profile is known.

The rest of this chapter is organized as follows. In Section 4.2, we start by defining M^2

and outline the method in ISO 11146 [79] (which we call the "ISO standard" from hereon). We then introduce an angle-position phase-space for a beam's state and relate M^2 to the state's area, which we show is conserved under paraxial propagation (i.e., the range of angles $<<$ 1 rad). We give explicit formulas to calculate M^2, the beam waist w_0 and its location z_0, the beam's angular width $\Delta\theta$, and the Rayleigh range z_R from the electric field profile at one plane. We call this the *covariance method*. In Section 4.5, we validate our covariance method by numerically comparing its predictions for M^2 to the ISO method for two non-trivial beam shapes. In Section 4.7, we discuss the advantages and drawbacks of our covariance method and the prospects for generalizing it to more complicated beams. While many of the concepts in this work have been introduced elsewhere, they have not been connected in a simple way to allow for easy use. For those solely interested in applying our covariance method, in Section 4.4 we summarize our results and provide a detailed straightforward recipe for determining M^2 from a single electric field profile of the beam anywhere.

4.2 Theoretical description of laser beams

Definitions and conventions

For simplicity we will consider beam profiles in one transverse dimension only, Section 4.6 contains a discussion towards the generalization to two dimensions. The z axis is the propagation direction while the x-axis is the transverse direction, along the beam's spatial profile. As a beam propagates along z, generally its width in position $w(z)$ will change while its angular width $\Delta\theta$ will be constant. We use the standard width conventions $w(z) = \Delta x(z) = 2\,\sigma_x(z)$ and $\Delta\theta = 2\,\sigma_\theta$, where the σ are the standard deviation of corresponding intensity distribution (i.e., in x or θ). We take $z = 0$ as the position of the beam waist w_0, the beam's minimum $w(z)$ over all z. It is important to note that, except for Gaussians, the w_0 width here differs from the definition of the beam waist width common in Gaussian optics, the $1/e^2$ intensity half-width. The Rayleigh range z_R is defined as the distance from the waist (taken as $z = 0$) at which the beam has increased in width to $w(z_R) = \sqrt{2}w_0$.

4.3 Definition of the M^2 parameter

The beam or mode quality M^2, also known as the beam propagation factor, is defined in terms of the beam waist w_0 and angular width $\Delta\theta$. The product of these two widths divided by the same product for a Gaussian beam is the definition of M^2. Since for any Gaussian this product equals λ/π [78] we have,

$$M^2 = \frac{\pi}{\lambda} w_0 \Delta\theta. \tag{4.1}$$

Whereas the product $w(z)\,\Delta\theta$ will change as the beam propagates, since w_0 and $\Delta\theta$ are independent of z their product will not change. They are characteristics of the beam as a whole, as is the beam quality M^2.

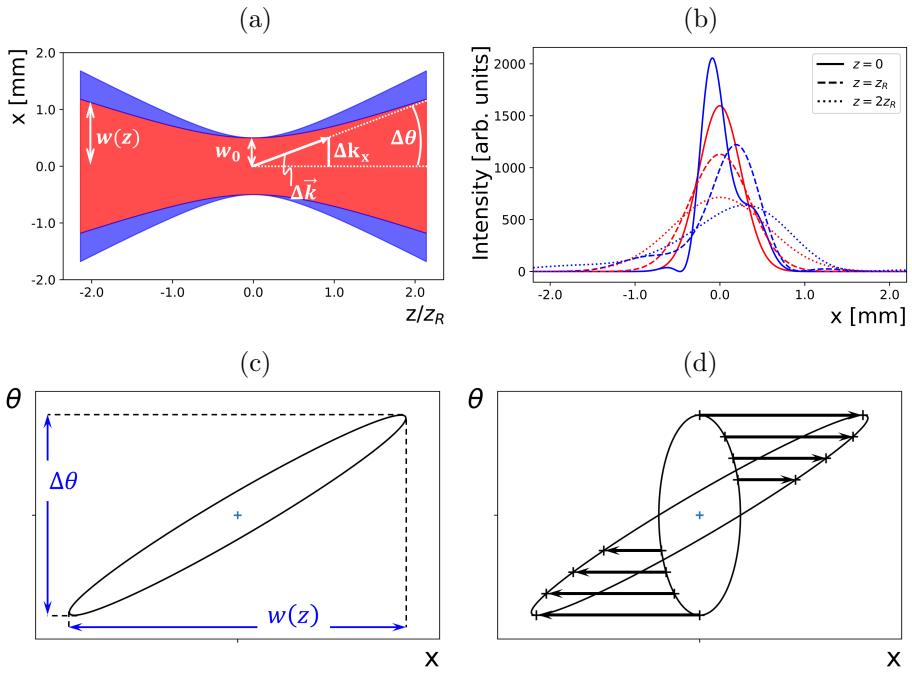


Figure 4.1: The evolution of paraxial beams under propagation. (a) Sketch showing the beam width along z for a Gaussian beam (red, $M^2 = 1$) and combination of first four Hermite-Gauss modes (blue, $M^2 = 1.5$) with $w_0 = 500\,\mu\text{m}$, $\lambda = 400\,\text{nm}$. The z-values are in units of the Gaussian's Rayleigh distance $z_R = 1.96\,\text{m}$. (b) Normalized intensity profiles of the same beams as in (a) at three z-positions. (c) The phase-space ellipse representing the beam state at a general z-position. (d) The x-shear resulting from paraxial propagation of the beam. Blue cross indicates origin.

As we will show, the value of M^2 is bounded by the Heisenberg uncertainty principle. As can be seen in Fig. 4.1(a), the relation between the widths in angle ($\Delta\theta$) and transverse wavevector (Δk_x) can be used to express the transverse momentum (Δp_x) of a photon as follows

$$\Delta p_x = \hbar \Delta k_x = \hbar k \sin \Delta\theta \approx \hbar k \Delta\theta, \tag{4.2}$$

where $k = 2\pi/\lambda$ is the total wavevector magnitude, λ is the wavelength, and the small angle approximation was applied. Using Eq. (4.2), one can rewrite M^2 as

$$M^2 = \frac{1}{2\hbar} \Delta x \Delta p_x \geq 1, \tag{4.3}$$

where the inequality comes from the Heisenberg uncertainty principle: $\sigma_x \sigma_{p_x} \geq \hbar/2$ or, equivalently, $\Delta x \Delta p_x \geq 2\hbar$. Gaussian beams are minimum uncertainty states and thus saturate inequality (4.3), $M^2 = 1$, while all other beams give strict inequality.

Although Eq. (4.1) for M^2 seems quite simple, determining its value is deceptively difficult. The angular width $\Delta\theta$ can be found from the beam width far from the waist, i.e., in the far

field. In contrast, to measure the beam waist w_0 requires that one knows its z-location. This is typically not known a priori, and it is seldomly known with good accuracy. The ISO standard avoids this accuracy issue by recommending measuring the beam width $w(z)$ in five different z-planes on each side of the beam waist, half of which should lie within one Rayleigh range z_R and the rest at least two Rayleigh ranges away from it. While one is free to choose the precise locations of the planes, to do so one still must approximately know z_R and the location of the beam waist. A rough pre-characterization can take the place of this a priori information, but this adds complexity to the procedure. The basic definition (Eq. (4.1)) is hardly applicable in either experiment or simulation.

The connection between phase-space and the electric field

We now introduce a phase-space picture of the beam in order to justify a simpler method to determine M^2. Appropriately, the 2D phase-space is made up of position x on one axis and angle θ on the perpendicular axis (though the latter could equally well be k_x or p_x, as explained in the last section). In this phase-space, the beam profile at any z-position can be used to define a tilted ellipse (see Fig. 4.1(c)) which contains information about the extent of the beam's state in phase-space. The length of the projection of the ellipse on the x-axis is the full-width intensity standard deviation, $w(z) = \Delta x$. Similarly, the projection on the θ-axis is the angular beam width, $\Delta\theta$.

When the beam has reached its waist (taken as $z = 0$), the ellipse is aligned with the phase-space axes and its full-width in x is minimized, equalling w_0. The area $\mathcal{A}$ of the ellipse is proportional to its width times its height, $\mathcal{A} = \frac{1}{4}\pi w_0 \Delta\theta$. In turn, using Eq. (4.1), we find that the beam quality is proportional to the ellipse area,

$$M^2 = \frac{4\mathcal{A}}{\lambda}. \tag{4.4}$$

As the beam propagates away from the waist location, correlations between position and angle form and the ellipse becomes tilted and its axes change in size. These correlations follow from simple geometry; waves traveling at angle θ will increase in position as $x = \theta z$ in the small angle approximation, creating a correlation between x and θ. This shears the ellipse along the x-direction in phase-space [110] as illustrated in Fig. 4.1(d). (Conversely, momentum conservation in free space ensures that each constituent wave's angle θ is constant.) Equivalently, free space propagation is given by the following transfer matrix

$$P = \begin{bmatrix} 1 & z' \\ 0 & 1 \end{bmatrix}, \tag{4.5}$$

which describes a shear transformation of a beam in phase-space along the x-direction. However, since the ellipse is now tilted, the ellipse projections (divided by two) on the x and θ axes are distinct from the width and height of the ellipse itself i.e., along its semi-minor and semi-major axes. Crucially, in geometry, a shear transformation always preserves area [111,112]

and, thus, M^2. This implies that Eq. (4.4) is true for all z, not just at the waist. In other words P has unit determinant ensures that the ellipse area is conserved under paraxial propagation, which justifies Eq. (4.9) for all z-planes. This finding has been obtained by other means in [113–115]. Thus, the problem of determining M^2 from information in a single z-plane reduces to finding the area of the corresponding tilted ellipse.

The covariance matrix

We gather the parameters of a tilted ellipse centered at the origin (i.e., the ellipse equation, $ax^2 - 2bx\theta + c\theta^2 = ac - b^2$) into a symmetric matrix [116], a method from geometry known as the *matrix of quadratic form.* This matrix sets the ellipse aspect ratio, orientation of its axis, and its size. In the context of beam widths, the matrix of quadratic form is the following covariance matrix,

$$Q(z) = \begin{bmatrix} \langle x^2 \rangle & \frac{1}{2}\langle x\theta + \theta x \rangle \\ \frac{1}{2}\langle x\theta + \theta x \rangle & \langle \theta^2 \rangle \end{bmatrix} \equiv \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \tag{4.6}$$

where $\langle\rangle$ is analogous to an average or expectation value but evaluated using the complex distribution $E(x;z)$, the transverse profile of the beam's electric field at a plane z. We give further detail on calculating $\langle\rangle$ at the end of this subsection and explicit expressions for Q will be given in Section 4.4. For now, we point out that the diagonals of Q are the variances, $\langle x^2 \rangle = \sigma_x^2 = w^2/4$ and $\langle \theta^2 \rangle = \sigma_\theta^2 = (\Delta\theta)^2/4$ for a beam with $\langle x \rangle = \langle \theta \rangle = 0$. When the beam has reached its waist $z = 0$, the ellipse is aligned with the phase-space axes and Q is diagonal. As explained in the previous subsection, at other z, correlations exist between angle and position. These are the off-diagonal covariance terms in Q. Explicitly the covariance matrix can be propagated via the propagation law

$$Q(z + z') = P\ Q(z)\ P^T, \tag{4.7}$$

where T indicates transpose. This simple formula is valid in both the Fresnel and Frauhnhofer diffraction regimes. We use Eq. (4.7) to propagate an arbitrary $Q(z)$ by a distance z',

$$Q(z + z') \equiv Q' = \begin{bmatrix} a' & b' \\ b' & c' \end{bmatrix} = \begin{bmatrix} a + bz' + (cz' + b)\, z' & b + cz' \\ b + cz' & c \end{bmatrix}. \tag{4.8}$$

The Q matrix is closely related to the beam matrices defined in terms of the Wigner function; namely, it is the Weyl transform of the matrices in [101, 117] and Part 3 of the ISO standard [118]. Unlike those matrices, Q is directly computed from the electric field distribution $E(x;z)$ in a single z-plane.

The advantage of this matrix formulation is that the determinant of the matrix Q is proportional to the ellipse area, regardless of any tilt. The ellipse area is now simple to find by $\mathcal{A} = \pi\sqrt{\det Q}$. Combining this with Eq. (4.4) we arrive at our central idea, the determinant of the covariance matrix is directly related to M^2 via

$$M^2 = \frac{4\pi}{\lambda}\sqrt{\det Q} = \frac{4\pi}{\lambda}\sqrt{ac - b^2}. \tag{4.9}$$

This relation is valid at arbitrary z-planes as long as the (paraxial) beam propagates only through first-order optical systems (e.g., spherical mirrors and lenses) or freely in space [117].

In summary, one can evaluate M^2 using the elements of the covariance matrix Eq. (4.6). These expectation values use the complex field $E(x;z)$ (in analogous way to the quantum wavefunction). That is, the expectation value of a general operator v acting on E is defined by $\langle v\rangle = \frac{1}{n}\int E^* v E dx$, where E^* is the complex conjugate of the electric field, and $n = \int |E(x;z)|^2 dx$ is a normalization factor. The angle operator is $\theta = -\frac{i}{k}\frac{\partial}{\partial x}$, while x is simply a multiplication by x. Note, the action of these two operators changes if their order changes, which may motivate why each off-diagonal in the covariance matrix incorporates both orderings. In general, the phase-space ellipse might not be centered at the origin, i.e., $\langle x\rangle \neq 0$ and/or $\langle\theta\rangle \neq 0$. We can account for this by substituting x and θ in the covariance matrix (Eq. (4.6)) with $x - \langle x\rangle$ and $\theta - \langle\theta\rangle$, respectively. We give explicit formulae for elements a, b, and c for such offset beams in Section 4.6. For example, the beam width can be evaluated as

$$\Big(w(z)\Big)^2 = \frac{4}{n}\int |E(x;z)|^2 \Big(x - \langle x\rangle\Big)^2 dx. \tag{4.10}$$

From Eq. (4.8) we can retrieve the common beam parameters. First, we note that c is unchanged so, trivially,

$$\Delta\theta = 2\sqrt{c}. \tag{4.11}$$

Next, we find the distance to the waist from z, the plane Q was determined at, by setting Q' to be diagonal, $b' = 0 = b + cz'$. Taking the position of the waist (Q') to be z_0 so that $z' = z - z_0 \equiv -\Delta z$, we then find

$$\Delta z = \frac{b}{c} = \frac{\langle x\theta + \theta x\rangle}{2\langle\theta^2\rangle}, \tag{4.12}$$

where a positive Δz means the waist is further along the beam propagation direction from the plane that $E(x;z)$ (and Q) was determined in. With $z' = -b/c$ substituted in Q' (Eq. (4.8)), we find the beam waist,

$$w_0 = 2\sqrt{a'} = 2\sqrt{a - b^2/c}. \tag{4.13}$$

We now instead set $Q(z_0)$ to be the waist and, thus, diagonal ($b = 0$), and using M^2 from Eq. (4.1), the a' element of Q' gives the beam width:

$$4a' = w(z) = w_0^2 + \left[\frac{M^2\lambda}{\pi w_0}\right]^2 (z - z_0)^2, \tag{4.14}$$

where $z' = z - z_0$. This gives the Rayleigh range,

$$z_R = \frac{\pi w_0^2}{M^2\lambda} = \frac{\lambda}{4\pi c}. \tag{4.15}$$

In other words, the propagation of an arbitrary paraxial beam is same as the one of a Gaussian beam magnified by M^2.

Equation (4.14) above is precisely the relation that the ISO standard uses to model the evolution of the beam width under spatial propagation. One first finds the intensity distributions in ten planes, then finds their widths, and then fits them to the hyperbolic function in Eq. (4.14). This fit yields values for M^2, the beam waist w_0 and its location z_0 [79]. A big disadvantage of this approach is that we already need to have rough values for the position of the beam waist as well as the M^2 parameter in order to compute the effective Rayleigh distance z_R and be able to follow the steps given in the protocol.

4.4 Recipe to compute M^2 and other beam parameters from the electric field

The goal of this section is to be self-contained and provide a straightforward set of steps to calculate M^2 and other key beam parameters directly from the transverse profile of the beam's scalar electric field $E \equiv E(x;z)$ that one has obtained in any plane z along the propagation direction. To do so, we use Eq. (4.9) and the covariance matrix Eq. (4.6), which was found under the assumption that $\langle x \rangle = \langle \theta \rangle = 0$. This assumption is valid for most optical simulations since the incoming beam is ideal and usually travelling centered on the x-axis and because the optical device (largely composed of spherical lenses and mirrors) usually is symmetric about $x = 0$. We give explicit formulae for elements a, b, and c for the more general case, offset beams, in Section 4.6. Reviewing, the beam must be paraxial with angles to z-axis much less than 1 rad; as usual, the width convention for the beam waist w_0 and angular spread $\Delta\theta$ is twice the intensity standard-deviation (e.g., $w_0 = 2\sigma_x$); and the beam quality is

$$M^2 \equiv \frac{\pi}{\lambda} w_0 \Delta\theta = \frac{4\pi}{\lambda}\sqrt{ac - b^2}. \tag{4.16}$$

To evaluate Eq. (4.16), one uses the covariance matrix elements, all of which are real-valued:

$$\begin{aligned} a &= \frac{1}{n}\int x^2 |E|^2 dx, \\ b &= -\frac{i\lambda}{4\pi}\left(1 + \frac{2}{n}\int x E^* \frac{\partial E}{\partial x} dx\right), \\ c &= -\left(\frac{\lambda}{2\pi}\right)^2 \frac{1}{n}\int E^* \frac{\partial^2 E}{\partial x^2} dx, \\ n &= \int |E|^2 dx, \end{aligned} \tag{4.17}$$

where we have used $\partial/\partial x\,(xE) = x\partial E/\partial x + E$ to simplify b.

From the a, b, and c matrix elements, given above we can also find other key beam parameters, such as the beam's angular width $\Delta\theta$, Rayleigh range z_R, beam waist w_0, and waist's

location relative to the plane of the measured field, Δz. In Eqs. (4.11, 4.15), we show that

$$\Delta z = \frac{b}{c}, \qquad w_0 = 2\sqrt{a - b^2/c}, \qquad \Delta\theta = 2\sqrt{c}, \qquad z_R = \frac{\lambda}{4\pi c}. \tag{4.18}$$

Thus, like the beam quality, these parameters can be determined from the electric field profile at one plane.

The formulas in this section are the main results of this work. The next section numerically verifies that they are correct, and also provides further details for the numerical evaluation of the electric field integrals in the a, b, and c parameters.

4.5 Numerical implementation and test of our covariance method

In this section, we compare the M^2 found from our covariance method to the ISO standard via a numerical simulation of propagation of two beams with nontrivial electric field profiles. We calculate $E(x)$ at an initial plane ($z = 0$) using the analytic form of the test beam $E(x)$ and then numerically propagate it to find $E(x; z)$ and the corresponding intensity profile $I(x; z) = |E(x; z)|^2$ at the ten requisite planes described below. This propagation is given by Eq. (1.38), which is accurate for all angles for a scalar field. That is, it is valid beyond the paraxial approximation. For the implementation, we used the native Fast Fourier Transform (FFT) in Python.

For the implementation of the ISO protocol, we calculate the standard deviation of the transverse intensity profile $I(x; z)$ to find the beam width $w(z)$ using Eq. (4.10). We do so at five planes within half a Rayleigh distance on either side of the beam waist and at five planes in the range of four to five Rayleigh distances. These widths are fit to Eq. (4.14) to find M^2.

At each of these planes, we calculate the entries, a, b, and c of the covariance matrix Q from $E(x; z)$ by Eq. (4.17) and use them to calculate M^2 using Eq. (4.9). We compare these ten values of M^2 from our covariance method, which should all be identical, to the single value from the ISO protocol.

We now briefly discuss some details of the numerical implementation for our covariance method. We use Python to evaluate Eq. (4.17). Derivatives were calculated with the Numpy gradient function, which uses the central difference method. Scipy integrate (Simpson's rule) was used for integrals. Code is provided in the Ref. [119].

The accuracy of our covariance method is mainly set by the x-position grid of the electric field profile $E(x)$. This directly sets the range and grid density in x and indirectly sets them in θ. In particular, since the two directions are related by a Fourier transform, the grid spacing $\delta\theta$ in the θ-direction is set by the range L in the x-direction, $\delta\theta \propto 1/(kL)$. Additionally, the range Θ in the θ-direction is set by the grid spacing δx in the x-direction, $\Theta \propto 1/(k\delta x)$.

The two goals are to have sufficient grid density to resolve the state and sufficient range to span the state in phase-space. That is, in both the x and θ directions, we aim for r points across the minimum width of the state and a range of s times the maximum width of the state. The

minimum and maximum widths in x are w_0 and $w(z)$, which set $\delta x < w_0/r$ and $L > w(z)\, s$. The beam width is always $\Delta\theta$ in θ, so $\delta\theta = \Delta\theta/r$ and $\Theta = s\Delta\theta$. By the Fourier relations from above, these set $L > \Delta\theta/(rk)$ and $\delta x < (s\Delta\theta)/k$, respectively. These four inequalities ensure sufficient resolution and range.

The inequalities determine the grid for a given beam state based on the chosen range and resolution parameters. For the latter, we suggest, $s = r = 10$. Since w_0 and $\Delta\theta$ are not known a priori, one should start with a trial grid $L > 10\ w(z)$ and then calculate w_0 and $\Delta\theta$ from Eqs. (4.17, 4.18). At this point, one would potentially need to revise the grid. Alternately, if $E(x)$ is from an optical simulation and computational power is not a constraint, setting $\delta x < \lambda/10$ will certainly be sufficient since it is five times better than the Nyquist limit. Satisfying the above conditions will ensure an accurate value of M^2.

Numerical comparison of the ISO and covariance methods

For the following test beams, $\lambda = 400\,\text{nm}$, we use position range of length $L = 200\,\text{mm}$ and a position grid density $\delta x = \lambda/2$. The ten planes range from $z = -7$ to $7\,\text{m}$.

For our first test beam, we use a linear combination of the first four Hermite-Gauss (HG) modes with complex coefficients. The electric field distribution at the initial plane for the m-th HG mode is taken to be

$$\begin{aligned} E_m(x) &= n_m\, H_m\left(\frac{\sqrt{2}x}{w_{0,m}}\right) e^{-\left(\frac{x}{w_{0,m}}\right)^2}, \\ n_m &= \left(\frac{2}{\pi}\right)^{1/4} \left(2^m m! w_{0,m}\right)^{-1/2}, \end{aligned} \tag{4.19}$$

where $w_{0,m}$ is the beam waist of the mode and H_m the m-th order Hermite polynomial. The set of all HG modes form an orthonormal basis, meaning that any other beam can be expressed as a superposition of HG modes. Thus, a superposition of HG modes can lead to intensity distributions that are not symmetric around $x = 0$ (as the ones shown in Figs. 4.1 and 4.2). Although each HG mode is centered the total beam will generally be offset from the $x = 0$ axis, so Eqs. (4.21, 4.22) in Section 4.6 must be used. In addition, the z location of the waist of a superposition will not generally be at the waist location of the individual HG modes. Furthermore, the intensity distribution of a general superposition will vary with the propagation distance z as each HG mode acquires a different phase upon free space propagation according to Eq. (1.14). Accommodating these possibilities, we derive a completely general theoretical value of M^2 for a superposition of HG modes with complex coefficients c_n in Appendix A, with Eq. (B.4) as the final result.

The test beam's electric field is given by

$$E(x) = (0.8 + 0.2i)\, E_0(x) + 0.3E_1(x) + (-0.099 - 0.01i)\, E_2(x) + 0.469E_3(x), \tag{4.20}$$

with $w_{0,m} = 642\,\mu\text{m}$ for $m = 0, 1, 2, 3$. This particular superposition happens to have little wavefront curvature and, thus, Eq. (85) for M^2 from [120] is approximately correct. Both the

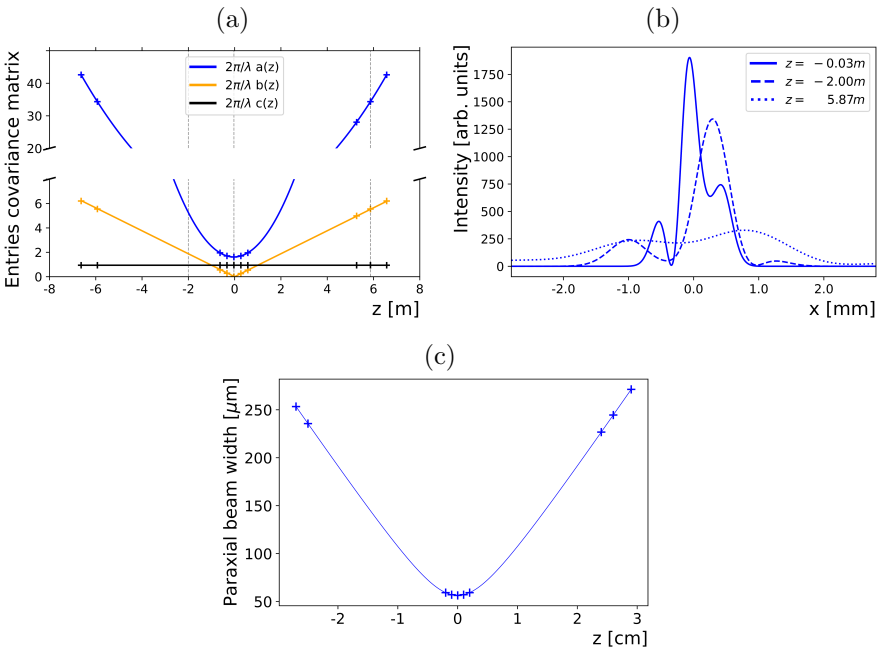


Figure 4.2: Analytic and numerical propagation of test beams. For a beam that is the superposition of the first four Hermite-Gauss modes given in the main text ($w_0 = 642\,\mu\text{m}$, $\lambda = 400\,\text{nm}$ $M^2 = 2.466$ given by Eq. (B.4)): (a) points are the scaled entries of the covariance matrix for the beam numerically propagated (by Eq. (1.38)) to each z distance and curves are the analytically propagated elements (by Eq. (4.8)) and (b) is the corresponding intensity profiles at three z-locations. For a paraxial supergaussian beam ($w_0 = 56\,\mu\text{m}$, order $= 50$, $\lambda = 400\,\text{nm}$, nominal $M^2 = 4.0466$): (c) points are the beam widths w_0 calculated from the numerically propagated beam and the curve is the analytical width from Eq. (4.14).

latter equation and our general formula, Eq. (B.4), predict a value of $M^2 = 2.466$ for the above superposition.

For this first test beam we plot the elements of the offset Q in Fig. 4.2(a). The element $a = w^2(z)/4$ (blue points) lies on the theoretical beam waist curve (Eq. (4.14), blue curve) and the other elements (yellow and black points) lie on the curves given by the matrix elements of the propagated Q matrix (Eq. (4.8), yellow and black curves). This agreement confirms that our numerical propagation of $E(x)$ is accurate and our calculation of the covariance matrix Q is correct.

With the correctness of our numerical calculations established, we now compare the ISO standard method to our covariance method for an offset beam to find M^2. The M^2 values found from our covariance method (Eqs. (4.16, 4.22)) at the ten planes have average value $M^2 = 2.470$ with a standard deviation of 0.006. Thus, as expected, our covariance method results in the same value of M^2 and it is independent of the plane z of the $E(x; z)$ used to calculate it. Moreover, it is in good agreement with the theoretical prediction of $M^2 = 2.466$.

A fit to the waists in the ten planes, the ISO method, gives $M^2 = 2.467$ with a fit error of 2.6×10^{-10}, again in good agreement. This suggests that using solely a single plane and the covariance method will agree with the ISO method up to the second decimal place of M^2.

Our second test beam is meant to approximate the top hat intensity-distribution that would be transmitted by a slit. In this way, it differs greatly from a basic Gaussian beam or even combinations of low-order HG modes. Moreover, an exact top hat has an angular intensity distribution that follows a sinc-squared function and thus has a $\Delta\theta$ that diverges and is not well-defined, which makes it a particularly challenging test case. A one-dimensional supergaussian, $E(x) = \exp\left[-\left(|x|/w_s\right)^N\right]$, of high order N approximates a top hat function of full-width $2w_s$. Using Eqs. (4.9, 4.17) we theoretically found $M^2 = N\sqrt{\Gamma(3/N)\,\Gamma(2-1/N)}/\Gamma(1/N) \approx \sqrt{N/3}$, where the approximation is valid for large N and Γ is the usual gamma function. (Note, this formula differs from formulae in the literature, which are for the "cylindrical M^2", e.g., [121].) The test beam is a supergaussian of order $N = 50$ with equivalent slit width of $2w_s = 0.1\,\text{mm}$ and a theoretical $M^2 = 4.0466$. Diffraction from such a slit has zero-to-zero angular width of 1.6×10^{-2} rad, putting it well within the paraxial regime.

For this second test beam, we repeat the comparison of methods that was conducted on the first beam. The beam widths in the ten planes for the supergaussian are plotted in Fig. 4.2(c). The ten planes of our covariance method give an average value of M^2= 4.0462 with a standard deviation of 7.0×10^{-5}. The ISO method (based on a fit to the beam waists in multiple planes) gives $M^2 = 4.0506$ with a fit error of 1.4×10^{-9}. The ISO method's value only agrees with theory and covariance method up to two decimal places after rounding. In contrast, our covariance method agrees with the analytical theoretical value to four decimal places, suggesting it is a more reliable method for this particularly challenging test case.

The success of our covariance method for finding M^2 with these two paraxial test beams validates its correctness and demonstrates that it stands in good agreement with the ISO protocol even for the extreme case of highly non-Gaussian beams.

4.6 The covariance method for general beams

Until now we have focused on the simplest case, a paraxial coherent one-dimensional profile. Both the ISO protocol and our covariance method rely on the paraxial approximation. More problematically, the derivation of the M^2 from the uncertainty principle relies on the small-angle approximation, so it is not clear that the definition is completely general. We briefly examine the validity of both the ISO and covariance methods in the non-paraxial regime. For the following, $\lambda = 400\,\text{nm}$ and we use position range of length $L = 200\,\text{mm}$ and a position grid density $\delta x = \lambda/2$.

To examine the non-paraxial regime, we decrease the effective slit full-width of the order $N = 50$ supergaussian to 0.5 µm, which increases the zero-to-zero angular spread to 1.6 rad putting it outside the paraxial approximation. Fig. 4.3(b) should give an impression of the non-paraxiality of the supergaussian beam. The beam quality factor for a supergaussian is

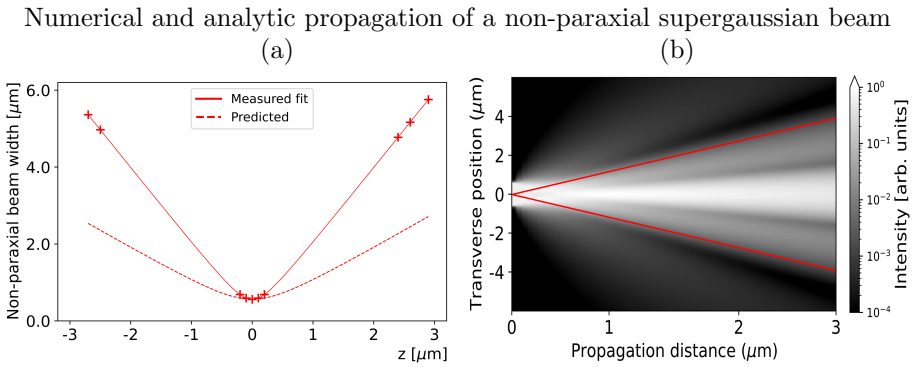


Figure 4.3: (a) Width evolution of a non-paraxial supergaussian ($w_0 = 0.56\,\mu\text{m}$, order $= 50$, $\lambda = 400\,\text{nm}$). Measured widths (indicated by points) do not lie on predicted curve for $M^2 = 4.05$. (b) Intensity plot of the same supergausssian. Red lines indicate the full angular spread of the beam $2\Delta\theta = 0.91\,\text{rad}$.

solely a function of order, so the theoretical beam quality should be the same as the $N = 50$ supergaussian from Section 4.5, $M^2 = 4.0466$. Applying the ISO method, the width measurements can still be fitted to a hyperbolic function (Fig. 4.3(a), top curve). However, the resulting M^2 factor more than doubles to a value of 9.28. While the covariance matrix gives the correct result for M^2 in the $z = 0$-plane, the values quickly diverge as soon as propagation starts. At $z = 2.9\,\mu\text{m}$, the covariance method gives, $M^2 = 43.5$, which is off by a factor of five at least. The limitation to paraxial beams might be lifted when defining the beam matrix through a non-paraxial version of the Wigner function [110], but a thorough investigation would be in order here. In summary, our covariance method and the ISO method fail for non-paraxial beams, as expected.

A one-dimensional description applies to simple coherent beams whose two-dimensional profile can be written as a product, $E(x)\,E(y)$, e.g., two dimensional HG modes. We now discuss the generalization of our covariance method to a wider variety of paraxial beams, including beams with general two-dimensional profiles and incoherent beams.

Offset beams

In some optical simulations and all laboratory measurements one cannot assume that the beam is perfectly centered on the beam axis, i.e., $\langle x\rangle \neq 0$ and $\langle\theta\rangle \neq 0$. In this case, the covariance matrix takes the more general form:

$$Q(z) = \begin{bmatrix} \langle x^2\rangle - \langle x\rangle^2 & \frac{1}{2}\langle x\theta + \theta x\rangle - \langle x\rangle\langle\theta\rangle \\ \frac{1}{2}\langle x\theta + \theta x\rangle - \langle x\rangle\langle\theta\rangle & \langle\theta^2\rangle - \langle\theta\rangle^2 \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}. \tag{4.21}$$

The corresponding generalized formulas for the matrix elements are

$$\begin{aligned} a &= \frac{1}{n}\int x^2|E|^2dx - \left(\frac{1}{n}\int x|E|^2dx\right)^2, \\ b &= \frac{i\lambda}{2\pi}\left[\left(\frac{1}{n}\int x|E|^2dx\right)\left(\frac{1}{n}\int E^*\frac{\partial E}{\partial x}dx\right) - \left(\frac{1}{2}+\frac{1}{n}\int xE^*\frac{\partial E}{\partial x}dx\right)\right], \\ c &= \left[\frac{\lambda}{2\pi}\right]^2\left[\left(\frac{1}{n}\int E^*\frac{\partial E}{\partial x}dx\right)^2 - \frac{1}{n}\int E^*\frac{\partial^2 E}{\partial x^2}dx\right], \\ n &= \int |E|^2dx. \end{aligned} \tag{4.22}$$

As before, these can be used to find M^2 and other beam parameters using Eqs. (4.16, 4.18). One can save some computational effort by noticing that the second and first integrals in a and c, respectively, also appear in b.

Simple astigmatic beams

So far, we have presented our method in one transverse dimension x for clarity. This one-dimensional analysis is straightforward to generalize to two-dimensional simple astigmatic beams; those that have circular and elliptical cross sections, respectively. To apply our covariance method to simple astigmatic beams, one first would find the principal axes of the elliptical transverse profile and denote these by x and y. Along these directions, the total field is separable, $E(x,y;z) = E_x(x;z)\,E_y(y;z)$, and each transverse direction can be separately analyzed by our covariance method. The result would be two values, M_x^2 and M_y^2, from which one can define an effective beam quality $M_{\text{eff}}^2 = \sqrt{M_x^2 M_y^2}$. See [79] for details.

General astigmatic beams

We now propose a route to adapt our covariance method to find the beam quality of more general two-dimensional paraxial fields, $E(x,y;z) \neq E_x(x;z)\,E_y(y;z)$. Unlike simple astigmatic beams, the principal axes of their transverse cross sections can rotate while the beam is propagating [122]. Part 2 of the ISO standard [123] derives an effective beam quality M_{eff}^2 in terms of moments of the Wigner function in the four-dimensional phase-space. Essentially, the state is a generalized ellipse in this four-dimensional space and M_{eff}^2 is the area of that state. Analogous to the generalized P matrix in Part 2 of the ISO standard, a generalized $Q_{xy}(z)$ matrix would be a 4×4 matrix:

$$Q_{xy}(z) \equiv \begin{bmatrix} Q_x & D_{xy} \\ D_{xy} & Q_y \end{bmatrix}. \tag{4.23}$$

Here, the block diagonals are the standard 2×2 Q matrices Eq. (4.6): $Q_x \equiv Q$ and Q_y is analogous. The x and y are arbitrary orthogonal transverse directions unconnected to the beam state. The two off-diagonal blocks are identical and given by

$$D_{xy}(z) = \begin{bmatrix} \langle xy \rangle & \langle x\theta_y \rangle \\ \langle y\theta_x \rangle & \langle \theta_x\theta_y \rangle \end{bmatrix}. \tag{4.24}$$

Note, the D blocks involve one variable from each direction and thus do not require the two orderings (e.g., $\langle x\theta + \theta x \rangle$) that are essential for the 2×2 Q matrix. With this generalized covariance matrix, $M^2_{\text{eff}} = \frac{4\pi}{\lambda} \sqrt[4]{\det Q_{xy}(z)}$. Unlike the P matrix in the ISO procedure, $Q_{xy}(z)$ is found directly from the electric field $E(x, y; z)$.

Incoherent beams

We now consider generalizing our covariance method to incoherent beams. We tested our covariance method with beams that are coherent superpositions of HG modes. If the beam is an incoherent mixture of beams, one would need to use a density matrix formalism, or equivalently, a coherence matrix to represent the beam in single z plane, $\rho(x, x') \equiv \overline{E(x) E^*(x')}$. Here, the bar indicates an ensemble or time average, as is standard in statistical optics. We leave the details for future work, but in short the expectation values in $Q_{xy}(z)$ would then be $\langle v \rangle = \text{Tr}[v\rho]$ and the beam quality M^2 would be found from Q as before, from Eq. (4.9). This approach would complete the presented covariance method. However previous work have shown how to obtain M^2 for an incoherent mixture of beams [76, 100], we refer the reader to such references for further details.

4.7 Conclusion

The work in this chapter has mostly focused on the simplest case, a paraxial coherent one-dimensional profile. In Section 4.6, we discussed the generalization of our covariance method to a wider variety of paraxial beams, including beams with general two-dimensional profiles and incoherent beams. We also showed that both our covariance method and the ISO method fail for non-paraxial beams, as expected. A one-dimensional description applies to simple coherent beams whose two-dimensional profile can be written as a product, $E(x) E(y)$, e.g., two dimensional HG modes.

In summary, we have demonstrated how one can compute the beam quality factor M^2 as well as angular width, waist width, and waist location for a beam from its electric field distribution in a single arbitrary plane. The conventional method, prescribed in the ISO standard [79], relies on finding the beam width in ten prescribed propagation planes. Since each width calculation requires three integrals of the intensity distribution, the ISO method requires thirty integrals in total. In comparison, our covariance method requires only six integrals (see Eq. 4.22) of the electric field distribution. Moreover, it eliminates the need to fit a function, a key step in the ISO method. Further, our covariance method eliminates the need for an initial estimate of the Rayleigh range and waist location.

Since there are now many tools to measure the electric field profile of a beam, our covariance method is amenable to use in laboratories and could be incorporated in optical test and measurement products. That said, we envision our method will be particularly useful for analyzing the electric field output of optical simulations and could be directly built into common simulation software. In this way, we expect our covariance method to calculate beam quality will streamline optical research and development.

Chapter 5

Theory and experiment for resource-efficient joint weak-measurement

This chapter is based on Martinez-Becerril, A. C., Bussières, G., Curic, D., Giner, L., Abrahao, R. A., and Lundeen, J. S. (2021). 'Theory and experiment for resource-efficient joint weak-measurement'. *Quantum*, 5, 599.

5.1 Weak measurement: a resource to obtain information on quantum incompatible observables

Modern quantum measurement techniques have pushed forward our understanding and ability to manipulate quantum particles. Often, fundamental and practical measurements involve the product of two or more observables of a quantum system. In particular, correlations of incompatible or non-commuting observables $\boldsymbol{A}$ and $\boldsymbol{B}$, defined by $[\boldsymbol{A},\boldsymbol{B}] \equiv \boldsymbol{AB} - \boldsymbol{BA} \neq 0$, are central to our understanding of entanglement [124, 125] and the Heisenberg uncertainty principle. A 'joint' measurement of $\boldsymbol{A}$ and $\boldsymbol{B}$ refers to the measurement process which outputs the expectation value of the product of the two observables $\langle \boldsymbol{BA} \rangle = \mathrm{tr}\,(\boldsymbol{BA\rho})$, where $\boldsymbol{\rho}$ is the density matrix of the system. The standard procedure to perform a joint measurement would be to measure observable $\boldsymbol{A}$, then measure $\boldsymbol{B}$. This fails for incompatible observables since the first measurement collapses the state of a particle into an eigenstate of $\boldsymbol{A}$, erasing the information about $\boldsymbol{B}$ and randomizing its value.

In contrast, weak measurement decreases the disturbance caused by the measurement process and thereby mostly preserves the quantum state of the system, thus allowing one to obtain correlations between any chosen set of general observables, including incompatible ones [43, 126–132]. To perform such a measurement, the observable is weakly coupled to a separate read-out system (the 'pointer') that indicates the average result of the measurement. Even though this approach refers to an individual system, weak measurement requires repeating the measurement on identically prepared systems, and averaging. This compensates for the little information that is extracted in a single trial. Weak measurement is a type of non-

destructive quantum measurement that minimizes disturbance of the measured system [46]. As shown in Ref. [133], as one decreases the disturbance caused by the measurement process, one also decreases the 'predictability' of the measurement. Weak measurement has a broad range of applications from amplifying tiny signals [134–136] to fundamental studies on the meaning of a quantum state [137, 138]. Particularly relevant to this work are Refs. [43–45, 127], which showed that if two observables are *weakly* measured, the average measurement outcome is simply the expectation value of the product of those two observables, $\langle \boldsymbol{BA} \rangle$. Remarkably, this holds even if $\boldsymbol{A}$ and $\boldsymbol{B}$ are incompatible, which would make $\boldsymbol{BA}$ non-Hermitian and, nominally, unobservable.

Joint weak-measurement

More recently the weak measurement formalism was expanded to deal with composite systems and performing a measurement of the product of two or more observables, known as a joint weak-measurement. Joint weak-measurement has proven to be useful, for example, in experimental realizations of the Cheshire cat [139, 140], the Hardy's paradox [141, 142], the study of quantum dynamics, and to give insight into the role of time ordering in the quantum domain [143, 144]. The ability to jointly measure incompatible observables has also shown to have many applications in the field of quantum metrology [145–150]. Joint weak-measurement of multiple observables enables sequentially probing a quantum system for characterizing its quantum evolution [131, 151]. Another example is the test of the Leggett-Garg inequalities for sequential measurements of multiple observables in a single system [152].

Known methods for the realization of a joint weak-measurement are resource-intensive. Specifically, they require either interactions that involve three or more particles or a separate read-out system for each observable. With a few exceptions [141], due to the absence of two-particle interactions, even single-observable weak measurement resorts to a strategy of using internal degrees of freedom (DOF) as the read-out systems [43–45, 127]. For example, one can measure the polarization of a photon by using its position DOF as a read-out [153]. For a joint measurement, this strategy is particularly limiting given that quantum particles have a limited number of DOF. For instance, for a photon there are just four DOF: polarization, and a three-dimensional wavevector (which, in turn, incorporates frequency-time and transverse position-momentum). Due to this limitation, joint weak-measurement experiments have never progressed beyond the product of two observables [129, 133]. To overcome this constraint, the present work theoretically introduces and experimentally demonstrates a technique to perform a joint weak-measurement of multiple observables using a single DOF as the read-out system.

We implement our technique to directly measure quantum states. This is a type of quantum state estimation where the state is fully determined by the shift of the pointer. Quantum state estimation has become an invaluable tool in the subject of quantum information, which requires verification of the quality (i.e., 'fidelity') of resource quantum states. The experimental demonstration of the direct measurement of the wave function opened up new research lines in quantum state estimation. The directness of the method means that one can obtain the complex amplitudes of a quantum state, in any chosen basis [137]. An important aspect of

direct state estimation is that no optimization or complicated inversion is involved. Solving such problems is one of the key goals of current research in quantum state estimation [154–156]. Further work demonstrated how to estimate a general quantum state by directly measuring the density matrix [45, 128], or by directly measuring phase-space quasiprobability distributions of states, such as the Dirac distribution [157, 158]. Similarly, we apply our single-pointer joint weak-measurement method to directly determine any chosen element of the density matrix. Specifically, we obtain the density matrix of photon polarization states using a single pointer for the two requisite observables.

The rest of the chapter is organized as follows. We start by describing weak measurement in terms of raising and lowering operators. Then, we outline the theory of our technique to perform a joint weak-measurement and introduce an important ingredient, the fractional Fourier transform. Next, we present the experimental demonstration of our technique and an application to quantum state estimation. Finally, we summarize our work and point out some future possible directions.

5.2 Joint weak-measurement with a single read-out system

For composite systems, one is interested in the average value of the product of observables such as $\langle \boldsymbol{BA} \rangle$. Universally, this involves correlations between two measurement outcomes (e.g., as in Bell's inequalities). In the von Neumann model, this corresponds to correlations between pointer distributions. This is true for both strong and weak measurements. In the latter case, the average outcome should be the joint weak-value,

$$\langle BA \rangle_w \equiv \frac{\langle F|\boldsymbol{BA}|I\rangle}{\langle F|I\rangle}. \tag{5.1}$$

A number of techniques have been proposed and demonstrated to observe the joint weak-value in pointer correlations. We now briefly review these techniques.

First, we review the case of compatible operators $\boldsymbol{A}$ and $\boldsymbol{B}$. These could be two different observables of a single particle or observables acting on two different particles. Ref. [127] proposed using a separate von Neumann interaction (i.e., Eq. (2.24)) and pointer for each observable (pointers 1 and 2 with creation operators $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ respectively). This was simplified in [43] assuming an equal coupling strength γ for measuring each observable. The resulting joint weak-value is given in terms of the product of the creation operators ($\boldsymbol{a}_1$ and $\boldsymbol{a}_2$) of both pointers: $\langle \boldsymbol{a}_1 \boldsymbol{a}_2 \rangle = \langle BA \rangle_w / \gamma^2$. This strategy of performing two separate weak measurements was experimentally demonstrated in [141].

A more challenging case, and the subject of this work, is the one in which $\boldsymbol{A}$ and $\boldsymbol{B}$ act on the same particle, but are incompatible e.g., two complementary observables such as position and momentum. Furthermore, the product $\boldsymbol{BA}$ is not Hermitian, thus it is not considered a valid observable in standard quantum mechanics. For example, naively replacing $\boldsymbol{A}$ with $\boldsymbol{BA}$ in the von Neumann Hamiltonian, Eq. (2.24), results in non-unitary time evolution.

However, in the weak regime, a measurement of $\boldsymbol{A}$ largely preserves the quantum state of the particle allowing a subsequent measurement of $\boldsymbol{B}$. The correlations between the outcomes of the two measurements give $\langle BA\rangle_w$. A technique along these lines was proposed in [44]. As with the compatible observable case above, it used a separate von Neumann interaction and pointer for each observable (pointer 1 for $\boldsymbol{B}$ and pointer 2 for $\boldsymbol{A}$). In [45], the required correlation between the pointers was shown to be $\langle \boldsymbol{a}_1 \boldsymbol{a}_2\rangle = \langle BA\rangle_w/\gamma^2$ and experimentally demonstrated in [128, 129]. In summary, for both compatible and incompatible observables, the same technique works. The drawback of the technique is that it requires one pointer for each observable.

In particular, this requirement of one pointer per observable is resource-intensive. In most implementations of weak measurement, pointers are internal DOF of the measured particle. For example, in [153] a photon's polarization is measured by coupling it to the same photon's transverse spatial DOF. In absence of inter-particle interactions, this facilitates the use of weak measurement, but quickly uses up all available internal DOF. In turn, this limits the number of observables in the product and the number of DOF that can be used in the measured system for other quantum information tasks. It is natural to ask: can we perform a joint weak-measurement with a single pointer?

The main contribution of the present work is to introduce and experimentally demonstrate such a technique. Our technique uses a sequence of two standard von Neumann interactions, each given by Eq. (2.24). Unlike the previous techniques, the two interactions couple the system to the same pointer. As in Section 2.5, the total initial state is $|I\rangle|0\rangle$. The first interaction $\boldsymbol{U}_A$ couples the pointer to $\boldsymbol{A}$, while the second $\boldsymbol{U}_B$ couples the *same pointer* to $\boldsymbol{B}$. The action of two von Neumann unitary transformations with equal interaction strength γ is $\boldsymbol{U}_B\boldsymbol{U}_A|I\rangle|0\rangle = e^{\gamma \boldsymbol{B}(\boldsymbol{a}^\dagger - \boldsymbol{a})} e^{\gamma \boldsymbol{A}(\boldsymbol{a}^\dagger - \boldsymbol{a})}|I\rangle|0\rangle = \sum_{m,n=0}^{\infty} \frac{\gamma^{n+m}}{n!m!}\boldsymbol{B}^m \boldsymbol{A}^n (\boldsymbol{a}^\dagger - \boldsymbol{a})^{n+m}|I\rangle|0\rangle$. This is the final state of the total system after the two interactions.

Motivated by the techniques outlined above, which used correlations between two different lowering operators $\langle \boldsymbol{a}_1 \boldsymbol{a}_2\rangle$, we will aim to find the expectation of the product of two identical lowering operators, $\langle \boldsymbol{a}^2\rangle$. Thus, we must expand the pointer state after the interaction to second order in the interaction strength γ. There are three second-order terms: $m = n = 1$; $m = 0$, $n = 2$; and $m = 2$, $n = 0$. Along with the zero and first order terms, this gives

$$\boldsymbol{U}_B\boldsymbol{U}_A|0\rangle|I\rangle = \Bigg(|0\rangle + \gamma(\boldsymbol{A}+\boldsymbol{B})|1\rangle + \frac{\gamma^2}{2}\big(2\boldsymbol{B}\boldsymbol{A} + \boldsymbol{A}^2 + \boldsymbol{B}^2\big)\big(\sqrt{2}|2\rangle - |0\rangle\big) + O\Big(\gamma^3\Big)\Bigg)|I\rangle\,. \quad (5.2)$$

Now we post-select the system on a final state $|F\rangle$. To second order in γ, the renormalized pointer's final state is

$$|\phi'\rangle = \frac{1}{\langle F|I\rangle}\Bigg(\langle F|I\rangle\,|0\rangle + \gamma\,\langle F|\,\boldsymbol{A}+\boldsymbol{B}\,|I\rangle\,|1\rangle + \frac{\gamma^2}{2}\,\langle F|\,2\boldsymbol{B}\boldsymbol{A}+\boldsymbol{A}^2+\boldsymbol{B}^2\,|I\rangle\,\big(\sqrt{2}\,|2\rangle - |0\rangle\big)\Bigg). \quad (5.3)$$

As per our aim, we now calculate the expectation value $\langle \boldsymbol{a}^2\rangle$ for $|\phi'\rangle$:

$$\langle \boldsymbol{a}^2\rangle = 2\gamma^2\,\langle BA\rangle_w + \gamma^2\Big(\langle A^2\rangle_w + \langle B^2\rangle_w\Big). \quad (5.4)$$

This equation contains the weak value of the product observable $\langle BA\rangle_w$ but also other nontrivial weak values, $\langle A^2\rangle_w$ and $\langle B^2\rangle_w$. However, if we limit the two observables to be projectors, then $\boldsymbol{A}^2=\boldsymbol{A}$ and $\boldsymbol{B}^2=\boldsymbol{B}$. This turns the nontrivial weak values into single-observable weak values, which we can replace with $\langle A+B\rangle_w = \langle \boldsymbol{a}\rangle/\gamma$. Using this and rearranging Eq. (5.4) to solve for $\langle BA\rangle_w$ we arrive at

$$\langle BA\rangle_w = \frac{1}{2\gamma^2}\Big(\langle \boldsymbol{a}^2\rangle - \gamma\,\langle \boldsymbol{a}\rangle\Big). \quad (5.5)$$

In this way, we have expressed the joint weak-value solely in terms of expectation values on the pointer's final state. However, an additional step is still necessary. While the expectation value of a single lowering operator is easily measured in an experiment by measuring $\boldsymbol{x}$ and $\boldsymbol{p}$ in separate trials, powers of lowering operators cannot be measured as easily. To solve this, we express $\langle \boldsymbol{a}^2\rangle$ using $\boldsymbol{a} = \frac{\boldsymbol{x}}{2\sigma_x} + i\frac{\boldsymbol{p}}{2\sigma_p}$, where we have used $\sigma_x\sigma_p = \frac{\hbar}{2}$ (which is valid since the pointer is in the minimum uncertainty state $|0\rangle$). Doing so, leads to the appearance of cross terms such as $\boldsymbol{x}\boldsymbol{p}+\boldsymbol{p}\boldsymbol{x}$, which do not correspond to a straightforwardly physical read-out system observable.

To overcome this problem, we can use the Hermitian observable $\boldsymbol{d}$ which is an equally weighted combination of $\boldsymbol{x}$ and $\boldsymbol{p}$: $\boldsymbol{d} = \frac{\sigma_d}{\sqrt{2}}\Big(\frac{\boldsymbol{x}}{\sigma_x} + \frac{\boldsymbol{p}}{\sigma_p}\Big)$. Here, σ_x, σ_p and σ_d are the standard deviations of the pointer in x, p and d spaces, respectively. The $\boldsymbol{d}$ observable naturally appears in a variety of quantum systems. In the Heisenberg picture in quantum optics, the $\boldsymbol{x}$ field quadrature rotates to $\boldsymbol{d}$ after an eighth of a period of oscillation; this is equivalent to an x-p phase-space rotation of $R\pi/2$, with $R=1/2$ where R is the rotation order. Similarly, $\boldsymbol{x}$ rotates to $\boldsymbol{p}$ after a quarter period ($R=1$). Just as the Fourier Transform links $\boldsymbol{x}$ and $\boldsymbol{p}$, the fractional Fourier Transform (FrFT) was introduced to calculate the effect of a rotation order R on a state in the Schrödinger picture [7]. In summary, there are established practical methods to physically implement FrFTs and measure $\boldsymbol{d}$.

The reason we have introduced this new observable is that the square of $\boldsymbol{d}$ will contain the desired cross terms. Calculating $\boldsymbol{d}^2$ and solving for the cross terms we find

$$\boldsymbol{x}\boldsymbol{p}+\boldsymbol{p}\boldsymbol{x} = \sigma_x\sigma_p\Big(2\frac{\boldsymbol{d}^2}{\sigma_d^2} - \frac{\boldsymbol{x}^2}{\sigma_x^2} - \frac{\boldsymbol{p}^2}{\sigma_p^2}\Big). \quad (5.6)$$

Upon substituting Eq. (5.6) in Eq. (5.5), we obtain an expression for the real and imaginary parts of $\langle BA \rangle_w$:

$$\mathrm{Re}\Big(\langle BA \rangle_w \Big) = \frac{1}{8\gamma^2} \left\langle \frac{\boldsymbol{x}^2}{\sigma_x^2} - \frac{\boldsymbol{p}^2}{\sigma_p^2} - \frac{gt\boldsymbol{x}}{\sigma_x^2} \right\rangle , \tag{5.7}$$

and

$$\mathrm{Im}\Big(\langle BA \rangle_w \Big) = \frac{1}{8\gamma^2} \left\langle 2\frac{\boldsymbol{d}^2}{\sigma_d^2} - \frac{\boldsymbol{x}^2}{\sigma_x^2} - \frac{\boldsymbol{p}^2}{\sigma_p^2} - \frac{gt}{\sigma_x}\frac{\boldsymbol{p}}{\sigma_p} \right\rangle . \tag{5.8}$$

Note that every term in Eqs. (5.7, 5.8) is a ratio of two variables with the same units, therefore each term is unitless. For the same reason, experimental scaling factors e.g., a magnification in the $\boldsymbol{x}$ domain, cancel out. Hence, characterization of experimental scaling factors is not required for the use of our technique.

In summary, Eqs. (5.7, 5.8) express the full complex joint weak-value for product observable $\boldsymbol{BA}$ in terms of Hermitian observables on the pointer's final state. As expected, the joint weak-value appears in second order powers of $\boldsymbol{x}$ and $\boldsymbol{p}$ and our new observable $\boldsymbol{d}$. This comprises our proposed technique to weakly measure the product of incompatible observables using only a single pointer.

5.3 Realization of a joint weak-measurement using a single read-out degree of freedom

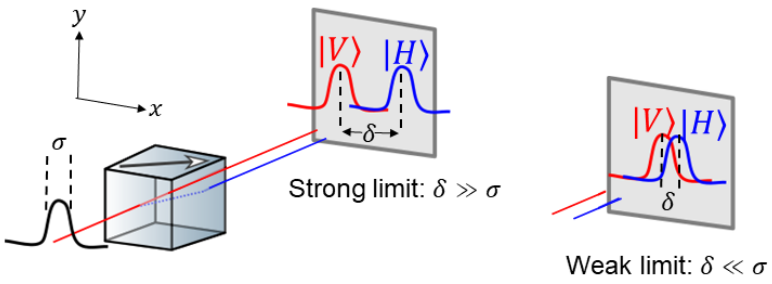


Figure 5.1: Experimental realization of a weak measurement of light polarization using the transverse spatial profile as read-out system. The beam displacer shifts the transverse position of vertically polarized light by δ. A weak measurement is characterized by $\delta \ll \sigma$, where σ is the beam size of the Gaussian transverse profile.

In this section, we present the experimental demonstration of our proposed technique using photons. Specifically we perform a joint weak-measurement of incompatible polarization projectors. The experimental setup is shown in Fig. 5.2. The measured observable will be in the photon's polarization DOF. The pointer is the photon's transverse $\boldsymbol{x}$ position with probability-distribution given by the absolute square of the wave function in Eq. (2.22) with $\sigma_x = 403\,\mu\text{m}$. The photon source is a He:Ne laser at 633 nm with a power of 1.19 mW. The setup can be divided into state preparation, weak measurements, strong measurement stages, and a read-out apparatus section. In order to test our technique, we prepare a range of polarization states $|I\rangle = \alpha\,|H\rangle + \beta\,|V\rangle$, where $|H\rangle$ ($|V\rangle$) is the horizontal (vertical) polarization. For state preparation, we use a polarizing beam splitter (PBS) followed by a half-wave plate (HWP), set at an angle of $\theta/2$ with respect to the $|H\rangle$ polarization, and a quarter-wave plate (QWP) (see the caption in Fig. 5.2 for setting details).

A von Neumann measurement of polarization can be performed with a birefringent crystal (e.g., a BBO crystal) acting as a beam displacer, see Fig. 5.1. This optical component transversely shifts the photon by $\Delta x = gt = 150\,\mu\text{m}$ if the photon is in the $|H\rangle$ polarization state and leaves it unshifted if it is in $|V\rangle$. In this way, the crystal couples the polarization observable $\boldsymbol{A} = |H\rangle\langle H|$ to the photon's transverse spatial position $\boldsymbol{x}$ that plays the role of the pointer. The strong measurement regime is characterized by Δx greater than σ_x, in which the eigenstates of $\boldsymbol{A}$ are fully separated. Our experiment is performed in the weak measurement regime where Δx is less than σ_x.

In order to measure the product of two observables with our technique, the setup performs two weak measurements in a row. Each von Neumann interaction (i.e., Eq. (2.24)) is achieved with a separate BBO crystal. Both crystals are aligned such that they shift the transverse profile of horizontally polarized photons in the horizontal direction x, leaving the transverse profile in the y direction unchanged. Thus, they couple to the same pointer, the x DOF. The first BBO implements a measurement of $\boldsymbol{A} = \boldsymbol{\pi}_H = |H\rangle\langle H|$. Before the second BBO, there is a HWP oriented at 22.5°. This effectively rotates the second measured observable to $\boldsymbol{B} = \boldsymbol{\pi}_{45^\circ} = |45^\circ\rangle\langle 45^\circ|$, with $|45^\circ\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}$. These two measurements and their read-out constitute an experimental application of our joint weak-measurement technique that uses a single pointer. Lastly, a strong measurement of polarization observable $\boldsymbol{\pi}_j$ ($j = H$ or V) is performed.

In our experiment, we need the ability to measure three incompatible observables of the pointer, $\boldsymbol{x}$, $\boldsymbol{p}$, and $\boldsymbol{d}$. This is the read-out of the result of the weak measurement. As we will explain, lens transformations will allow us to switch between these spatial observables, transforming them to a final transverse position x' on a camera. We measure the probability distribution of the observables in Eqs. (5.7, 5.8) on a monochrome 8 bit CMOS camera with a pixel width in x' of $2.2\,\mu\text{m}$. To make room for the optical lengths required for the lens transformations we add a 4f lens-pair to the imaging system ($f_1 = 100$ cm and $f_2 = 120$ cm). The 4f is positioned such that f_1 is 100 cm after the crystals. This ensures that the spatial wave function at the exit surface of the second crystal is recreated 120 cm after the f_2 lens. Our goal is to leave the camera fixed in place while different lenses are inserted in order to

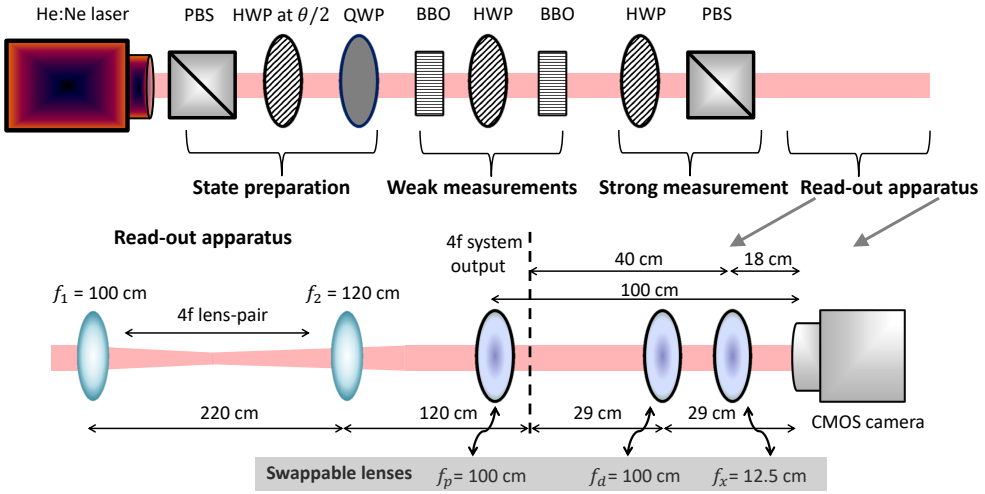


Figure 5.2: Experimental setup for performing a joint weak-measurement of a photon's polarization state using a single pointer, the photon's transverse x position. We work with three sets of pure polarization states $|\psi_1\rangle = \cos\theta\,|H\rangle + \sin\theta\,|V\rangle$, $|\psi_2\rangle = \cos\theta\,|H\rangle + i\sin\theta\,|V\rangle$ and $|\psi_3\rangle = \frac{1}{\sqrt{2}}\big(|H\rangle - ie^{2i\theta}\,|V\rangle\big)$. **State preparation** : To produce such states, we use a polarizing beam splitter (PBS), a half-wave plate (HWP) set at $\theta/2$ and a quarter-wave plate (QWP). The QWP is removed for preparing $|\psi_1\rangle$, and it is set at $0°$ and $45°$ for $|\psi_2\rangle$ and $|\psi_3\rangle$, respectively. **Weak measurements** : A first walk-off crystal (BBO) implements a weak measurement of $\boldsymbol{\pi}_j$ where j can be $|H\rangle$ or $|V\rangle$. A HWP at $22.5°$ and a second BBO effectively perform a weak measurement of $\boldsymbol{\pi}_{45°} = |45°\rangle\langle 45°|$, with $|45°\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}$. Both crystals have their optical axes aligned to create walk-off along the x-axis, the read-out DOF. **Strong Measurement** : A final HWP and a PBS implement a strong measurement in the $\{|H\rangle, |V\rangle\}$ basis. **Read − out apparatus** : A 4f lens-pair ($f_1 = 100$ cm and $f_2 = 120$ cm) is required to obtain the probability distributions involved in Eqs. (5.7, 5.8) i.e., to determine $\langle BA\rangle_w$. For $\boldsymbol{p}$, we use a Fourier transform lens of focal length $f_p = 100$ cm, for $\boldsymbol{d}$, a Fractional Fourier Transform (FrFT) lens of focal length $f_d = 100$ cm and, for $\boldsymbol{x}$, an imaging lens of focal length $f_x = 12.5$ cm. Each lens is set at the specified distance from a fixed CMOS camera, the obtained images are used to calculate the required expectation values as described in the text.

measure $\boldsymbol{x}$, $\boldsymbol{p}$, and $\boldsymbol{d}$.

To measure $\boldsymbol{p}$, and $\boldsymbol{d}$ we use an optical FrFT of the spatial DOF. The special case of rotation order $R = 1$ (a standard Fourier Transform), is already widely used; the transverse position x' at one focal length after lens $f_p = 100$ cm is proportional to p at any distance before the lens. Hence, lens f_p can be placed at any distance after the 4f lens pair as long as it is f_p distance from the camera. Less common is the optical spatial FrFT, which was introduced in [13,16,159] (more details were given in Section 1.4). At a distance z after lens $f_d = 100$ cm, x' will be proportional to the $\boldsymbol{d}$ observable a distance z before the lens. Here, $z = f_d \tan\big(\frac{R\pi}{4}\big)\sin\big(\frac{R\pi}{2}\big)$. For $\boldsymbol{d}$, the phase-space rotation parameter R equals $1/2$ making $z = 29$ cm. This d lens

transformation fixes the distance of the camera from the 4f lens pair. Lastly, a single lens ($f_x = 12.5$ cm) placed 160 cm after f_2 relays the image from the 4f lens pair. This ensures that 18 cm after f_x, x' on the camera is proportional to x at the crystals. The values of f_x, f_p, and f_d were chosen so that each measured x' distribution spans many pixels. By switching in one lens at a time, f_x, f_p, or f_d, the camera effectively measures the corresponding observable.

A key experimental simplification is that we do not need to experimentally or theoretically determine the proportionality constants between x' at the camera and $\boldsymbol{x}$, $\boldsymbol{p}$, and $\boldsymbol{d}$. The imaging magnification between x and x' is an example of such a proportionality constant. Since they depend on the focal lengths and lens-camera distances, these constants are difficult to experimentally determine precisely. Instead, Eqs. (5.7, 5.8) show that each observable is divided by the width of the pointer's initial distribution in that observable, e.g. $\boldsymbol{p}/\sigma_p$. Consequently, the units cancel and all calculations can be conducted directly in terms of x', i.e. camera pixel index.

The data acquisition consisted of taking five camera images per pointer observable (i.e., per lens configuration). A background image, taken with the laser blocked, was subtracted from each. The resulting image was integrated along the vertical direction y', and normalized to the brightest image obtained in that configuration. The resulting one dimensional probability distribution $P(x')$, corresponds to the probability of detecting a photon in position x' with final polarization $|H\rangle$ or $|V\rangle$. With the f_x lens in place, this is effectively an $\boldsymbol{x}$ read-out. The expectation values required for the joint weak-value from Eqs. (5.7, 5.8) can be obtained as $\langle \boldsymbol{x} \rangle / \sigma_x = \langle \boldsymbol{x'} \rangle / \sigma_{x'}$, where $\langle \boldsymbol{x'} \rangle = \int P(x') x' dx'$. For $\boldsymbol{p}$ and $\boldsymbol{d}$, a similar procedure is followed.

As our first demonstration of the technique, we weakly measure the non-Hermitian product observable $\boldsymbol{\pi}_{45^\circ}\boldsymbol{\pi}_V$ for a range of input states, $|I\rangle$. Specifically, we set the state preparation HWP at an angle of $\theta/2$ and the QWP at 45° in order to produce the state: $|I\rangle = \frac{1}{\sqrt{2}}\big(|H\rangle - ie^{2i\theta}|V\rangle\big)$. We increment θ from 0° to 180° in steps of 4°. For each input state and each image, the expectation values for the joint weak-value were evaluated according to Eqs. (5.7, 5.8). The uncertainties were estimated by the standard deviation in the joint weak-value across the five recorded images.

Curves for the real and imaginary parts of the joint weak-value are shown in Fig. 5.3. The experimental values closely follow the expected curves calculated from the nominal input state $|I\rangle$. However, they do not agree within error. These deviations are likely due to imperfections in the waveplates. These imperfections will also propagate to the alignment of the displacement axes of the BBO crystals since the waveplates are used in the alignment process. Such imperfections have been shown to be the dominant source of systematic error in similar past experiments [128, 160]. Nonetheless, the results demonstrate the validity of our proposed technique using a single pointer.

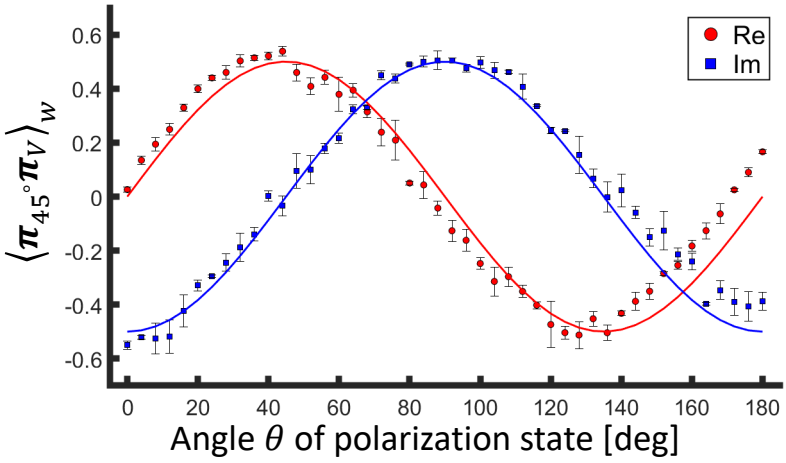


Figure 5.3: Joint weak-value of the product of incompatible observables $\boldsymbol{\pi}_{45^\circ}\boldsymbol{\pi}_V$ with a post-selection on the state $|H\rangle$. The input state is $|I\rangle = \frac{1}{\sqrt{2}}\big(|H\rangle - ie^{2i\theta}|V\rangle\big)$ obtained by setting the preparation HWP at $\theta/2$ and the QWP at 45°. The real and imaginary parts of the weak value are displayed with markers, while solid lines correspond to the joint weak-value, Eq. (5.1). The error bars are calculated solely from measurement statistics and correspond to the standard deviation. The standard deviation of the error bars can be decreased by increasing the number of trials for each measurement as well as mechanically stabilizing the setup upon swapping different lenses.

5.4 Direct measurement of a quantum state

We now move to a more sophisticated demonstration of our technique, the direct measurement of each element of the density matrix of polarization states. Such a direct measurement was introduced in [45,128]. An important advantage of this direct state estimation approach is the number of measurement bases it requires. To obtain a given element of the density matrix, this method requires joint weak-measurements in two complementary bases independently of the dimension of the quantum system. This contrasts with standard quantum state tomography, which requires $O(\mathrm{m})$ bases for an m-dimensional system. The direct estimation approach determines the density matrix of a quantum system element-by-element. One can envision a scenario where the off-diagonal elements of a system's density matrix (known as coherences) are monitored (via direct estimation) as a way to detect decoherence [161]. Unlike in Refs. [45,128], which used two pointers for the joint weak-measurements, here we use only a single pointer (measured in three different bases) for the same task.

A joint weak-measurement of the product $\boldsymbol{\pi}_i\boldsymbol{\pi}_{45^\circ}\boldsymbol{\pi}_j$, with $i,j = H$ or V (with no post-selection) gives the element $\rho(i,j)$ of the density matrix. This is schematically shown in Fig. 5.4. As shown in [45], the average outcome of a weak measurement without post-selection is the 'weak average' (rather than the weak value), which is equal to the expectation value of the measured observable $\langle \boldsymbol{C}\rangle = \mathrm{tr}\,[\boldsymbol{C}\boldsymbol{\rho}]$. Thus, the direct measurement procedure results in a joint

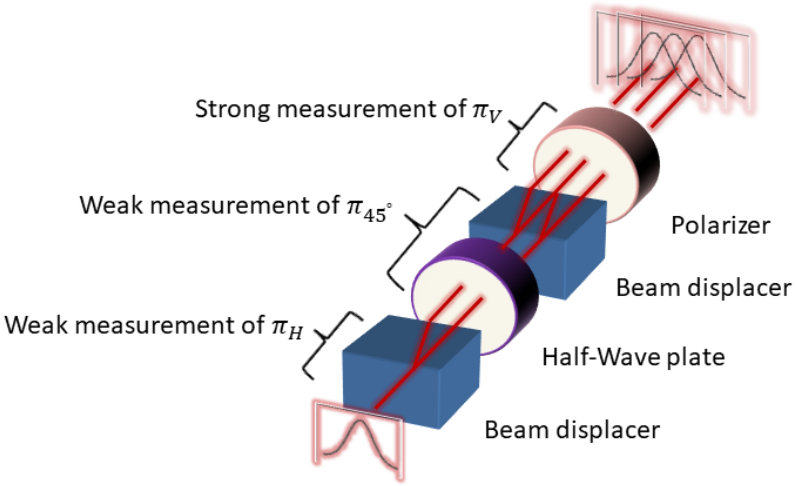


Figure 5.4: Schematic of the use of a joint weak-measurement for the direct measurement of the density matrix of photon polarization.

weak-average, $\rho(i,j) = 2\,\mathrm{tr}\left[\boldsymbol{\pi}_i \boldsymbol{\pi}_{45^\circ} \boldsymbol{\pi}_j \boldsymbol{\rho}\right]$. Therefore, by varying the first and last projectors, the density matrix can be directly determined element-by-element.

To measure the density matrix experimentally, we changed the HWPs settings to scan over the projectors $\boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_j$. As shown in [45], the final observable in the product can be measured either weakly or strongly. The last PBS implements a strong measurement. For each pair of projectors, we measure the required expectation values in Eqs. (5.7, 5.8).

We test our direct measurement with three sets of pure polarization states. A general polarization state $|\psi\rangle = \alpha\,|H\rangle + \beta\,|V\rangle$ has density matrix $\boldsymbol{\rho} \equiv |\psi\rangle\langle\psi| = |\alpha|^2\,|H\rangle\langle H| + \alpha\beta^*\,|H\rangle\langle V| + \beta\alpha^*\,|V\rangle\langle H| + |\beta|^2\,|V\rangle\langle V|$. The states sets can be found by use of Jones matrices of half and quarter-wave plates (which can be found for example in [3]). Such states sets are $|\psi_1\rangle = \cos\theta\,|H\rangle + \sin\theta\,|V\rangle$, $|\psi_2\rangle = \cos\theta\,|H\rangle + i\sin\theta\,|V\rangle$ and $|\psi_3\rangle = \frac{1}{\sqrt{2}}\big(|H\rangle - ie^{2i\theta}\,|V\rangle\big)$. For all cases, the first HWP varies the parameter θ scanning the interval $[0^\circ, 180^\circ]$. The QWP is removed for $|\psi_1\rangle$, and it is set at 0° and 45° for $|\psi_2\rangle$ and $|\psi_3\rangle$, respectively. The polarization states are visualized in the Poincaré sphere in Fig. 5.5a. For achieving such a representation one can visualize some states and interpolate continuous trajectories. For example, for $|\psi_1\rangle$ and $\theta = 0, \pi/4, \pi/2, 3\pi/4$, the corresponding states are $|H\rangle, |45^\circ\rangle, |V\rangle, |-45^\circ\rangle$ respectively. These states can easily be identified on the Poincaré sphere leading to a trajectory of states along the great circle passing for such states. For $|\psi_3\rangle$ and $\theta = 0, \pi/4, \pi/2, 3\pi/4$, the corresponding

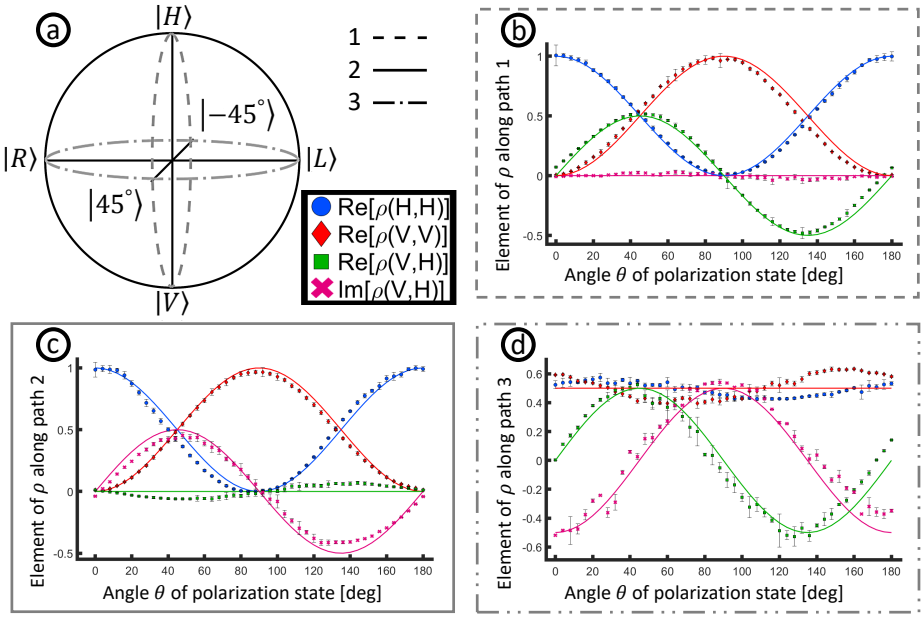


Figure 5.5: The polarization states used to test the method can be visualized in frame (a). These states are located on the three great circles (labeled 1, 2 and 3) in the Poincaré sphere passing through the states $|H\rangle, |V\rangle, |45°\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}, |-45°\rangle = \frac{|H\rangle-|V\rangle}{\sqrt{2}}, |R\rangle = \frac{|H\rangle-i|V\rangle}{\sqrt{2}}$ and $|L\rangle = \frac{|H\rangle+i|V\rangle}{\sqrt{2}}$. In frames (b) to (d), we show experimental elements of the density matrix $\boldsymbol{\rho}$ of the polarization states $|\psi_1\rangle = \cos\theta\,|H\rangle + \sin\theta\,|V\rangle$, $|\psi_2\rangle = \cos\theta\,|H\rangle + i\sin\theta\,|V\rangle$ and $|\psi_3\rangle = \frac{1}{\sqrt{2}}\big(|H\rangle - ie^{2i\theta}\,|V\rangle\big)$, respectively. Solid lines correspond to the theory. States following path 1 corresponds to linear polarization (no QWP in the setup), states along paths 2 and 3 were obtained by setting the QWP at 0° and 45°, respectively. Error bars are calculated using the standard deviation in the joint weak-value across the five recorded images, as previously employed in Fig. 5.3.

states are $|L\rangle, |45°\rangle, |R\rangle, |-45°\rangle$ respectively.

Our results are shown in Fig. 5.5b - d. The solid lines correspond to the real and imaginary parts of the elements of the theoretical density matrix and the points are the corresponding experimental joint weak-values. The latter should be equal to the real and imaginary parts of the density matrix elements and indeed follow the expected curve for each element. As before, deviations are thought to be the result of systematic errors in the polarization optics. This direct determination of the density matrix demonstrates the utility of our technique for weak measurement applications in quantum information.

5.5 Towards a general joint measurement with one read-out system

Before summarizing, we discuss some special cases and extensions of our technique. First, the special case where $\boldsymbol{B} = \boldsymbol{A}$ and the observable is general, i.e., not necessarily a projector. In this case, the unitary evolution (Eq. (5.2)) of the two von Neumann interactions (i.e., Eq. (2.24)) is equivalent to a single unitary given by $\boldsymbol{U}_A\boldsymbol{U}_A = e^{\gamma \boldsymbol{A}(\boldsymbol{a}^\dagger+\boldsymbol{a})}e^{\gamma \boldsymbol{A}(\boldsymbol{a}^\dagger+\boldsymbol{a})} = e^{2\gamma \boldsymbol{A}(\boldsymbol{a}^\dagger+\boldsymbol{a})}$. This corresponds to a unitary of a single von Neumann interaction of the $\boldsymbol{A}$ observable with a doubled interaction strength 2γ. By measuring $\boldsymbol{x}$, $\boldsymbol{p}$, and $\boldsymbol{d}$ on the pointer and using Eqs. (5.7, 5.8), we can then find the weak value $\langle A^2 \rangle_w$. This behaviour can be generalized so that a single von Neumann interaction can be used to measure $\langle A^N \rangle_w$. To do so, one will need to measure corresponding powers of observables on the pointer, e.g., $\boldsymbol{x}^N$ as well as the N-th power of hybrid observables such as $\boldsymbol{d}$.

In the derivation of our technique, we focused on the case that $\boldsymbol{A}$ and $\boldsymbol{B}$ are projectors. However, for general $\boldsymbol{A}$ and $\boldsymbol{B}$, measuring the product of projectors is enough to obtain the joint weak-value $\langle BA \rangle_w$. This can be seen if we express $\boldsymbol{A}$ in its spectral decomposition $\boldsymbol{A} = \sum_\alpha \alpha \boldsymbol{\pi}_\alpha$. Here α is an eigenvalue corresponding to the eigenstate $|\alpha\rangle$, and $\boldsymbol{\pi}_\alpha = |\alpha\rangle\langle\alpha|$ [18]. Analogously, for $\boldsymbol{B}$ we have $\boldsymbol{B} = \sum_\beta \beta \boldsymbol{\pi}_\beta$. Therefore, $\boldsymbol{BA} \equiv \sum_{\alpha,\beta} \alpha\beta \boldsymbol{\pi}_\beta \boldsymbol{\pi}_\alpha$. Thus, by measuring each of the products $\boldsymbol{\pi}_\beta\boldsymbol{\pi}_\alpha$ and adding the results, the joint weak-value $\langle BA \rangle_w$ can be obtained. In summary, our method can be used to weakly measure the product of general incompatible observables $\boldsymbol{A}$ and $\boldsymbol{B}$.

Our technique is also applicable to observables on separate quantum systems, e.g., $\boldsymbol{A}$ is measured on a first particle and $\boldsymbol{B}$ is measured on a second particle. The standard procedure for weak measurement would couple $\boldsymbol{BA}$ to a single pointer using $\boldsymbol{H} = g\boldsymbol{BAp}$, which is Hermitian now. This, however, requires a three-particle interaction which is challenging. Our method uses a two-particle interaction on each system while still only using a single pointer. One would first use the standard von Neumann interaction Eq. (2.24), to couple the pointer to $\boldsymbol{A}$, then couple the same pointer to $\boldsymbol{B}$. This double coupling can be challenging to implement. Particularly, in photons one would need an optical nonlinear effect at the single-photon level. As our technique largely preserves the initial quantum state, potential applications include demonstrating contextuality [162–165], and tests of Leggett-Garg inequalities [152, 166–171]. Motivated by the direct measurement of entangled systems [172, 173] and entanglement witnesses [174, 175], a potential future research direction is the use of this technique for characterizing entanglement.

A possible extension of our technique is measuring the product of m observables of a quantum system, using a single pointer. One would perform subsequent couplings between each of the m observables, and the pointer. The product of the m observables will appear in the expectation value of the m power of the lowering operator of the pointer $\langle \boldsymbol{a}^m \rangle$. Performing such a read-out potentially requires full-tomography on the pointer. The advantage of this approach is that only two-particle interactions are employed, and a single pointer is required for measuring the product of m observables.

As an outline of future work, the technique introduced in this chapter can also be extended to general types of pointers such as spin pointers. Indeed, previous work showed that the lowering operator formalism can be extended to spin pointers (e.g., polarization) and spin lowering operators [43]. To measure the product of N observables, one would need to sequentially couple the N observables to the spin pointer. The product of the N observables will appear as a coefficient of the N^{th} excited state of the pointer (just as the product $\boldsymbol{BA}$ appears as a coefficient of the second excited state in Eq. (5.2)). Thus the key requirement will be that one needs N spin levels (i.e., $N = 2S + 1$, where S is the spin) to measure the product of N observables. This shows that the method trades the resource of N dimensions in the pointer for an N-particle interaction (Eq. (2.24)) or, alternately, N separate pointer systems as in Ref. [43]. Instead of encoding the measurement information in many separate pointers as in Ref. [43,128,141], we encode the information from the measurement in a single high-dimensional pointer.

5.6 Conclusions

As the main contribution of the work on this chapter, we theoretically derived and experimentally demonstrated a method to perform a joint weak-measurement of two incompatible observables using a single pointer. We then employed this method to directly and individually measure each element of a system's density matrix. Since product observables are ubiquitous in quantum information processing, our technique may be useful for probing and characterizing such processors *in situ* without substantially disturbing them. Our work optimizes the use of resources needed to perform a joint-weak measurement freeing degrees of freedom of a quantum particle for quantum information tasks. We hope our technique facilitates the use of weak measurement in complex dynamics and new studies in the quantum realm.

Chapter 6

Conclusions

The spatial mode of a photon is probably the richest degree of freedom a photon possesses. Complex high-dimensional states can be created in the spatial mode as well as transformations on them. In this thesis, we have presented three projects which have in common the spatial mode of a photon as underlying degree of freedom. Each result contributes in different areas of classical and quantum information processing.

In Chapter 3, we have focused on a Multi-plane light converter (MPLC) as a system to achieve modal transformations. We have experimentally demonstrated an MPLC can be reconfigured in time to implement an arbitrary unitary transformation in a two-level system. An MPLC system is a linear optical network. As suggested from boson sampling, such a network is complex enough to achieve quantum advantage in the case of much more modes than photons in the network. This could be a possible future project as it has been demonstrated that MPLC systems can implement high-dimensional transformations [1, 48].

We have focused on the state of a single photon as input state. When multiple photons are used as inputs, multiple-photon interference will come into play. This presumably would make the system harder to classically simulate, and harder to design. However, this is the ideal type of systems that an experimentalist could have in order to further explore and use quantum effects. In such a scenario, the design approach we have used would no longer suffice. It would be needed a different approach where an MPLC system be treated as a quantum simulator.

In Chapter 4, we have demonstrated how one can compute the beam quality factor M^2 as well as the other beam properties using solely the complex electric field in a single plane. Such a method highlights M^2 as a fundamental quantity that is conserved under free space propagation. This clarifies the concept of such a parameter and explains why it is used to characterize the propagation of a laser beam.

In Chapter 5, we theoretically and experimentally demonstrated a method to perform a joint weak-measurement of two incompatible observables using a single read-out system. Our work optimizes the resources for a joint-weak measurement. In our experimental implementation of the method, the read-out system was a beam array, but this time with highly overlapping beams. We used our joint weak-measurement method for reconstructing the quantum state of a photon's polarization. More specifically, we directly and individually obtained every element of the density matrix of a photon's polarization state.

Appendix A

Derivation of the wave equation for light

The starting point in Chapter. 1 was the wave equation for light. We now derive such an equation starting from Maxwell's equations. The derivation is based on Ref. [3].

Maxwell's equations consist of four vectorial equations, two of such being for the curl and two for the divergence of the electric E and magnetic fields H. For a dielectric, linear, nondispersive, homogeneous and isotropic medium Maxwell equations read as follows:

$$\nabla \times H = \epsilon \frac{\partial E}{\partial t}, \tag{A.1}$$

$$\nabla \times E = -\mu \frac{\partial H}{\partial t}, \tag{A.2}$$

$$\nabla \cdot E = 0, \tag{A.3}$$

$$\nabla \cdot H = 0. \tag{A.4}$$

Here ϵ is the electric permittivity of the medium and μ is its magnetic permeability.

Let's apply the curl operator to Eq. A.2 and use the following vector identity $\nabla \times (\nabla \times) = \nabla (\nabla \cdot) - \nabla^2$. We obtain the following for the left hand side

$$\nabla \times (\nabla \times E) = \nabla (\nabla \cdot E) - \nabla^2 E \tag{A.5}$$

$$= -\nabla^2 E. \tag{A.6}$$

Where Eq. A.3 was used. Now, for the right hand side we obtain

$$\nabla \times \mu \frac{\partial H}{\partial t} = \mu \frac{\partial}{\partial t} \nabla \times H \tag{A.7}$$

$$= \epsilon \mu \frac{\partial^2 E}{\partial t^2} \tag{A.8}$$

where we noticed that the time and spatial derivatives commute i.e., $\nabla \times \frac{\partial}{\partial t} = \frac{\partial}{\partial t} \nabla$ and also used Eq. A.1 to obtain an expression in terms of the electric field E only. Finally we equate

the left and right hand sides,

$$\nabla^2 E - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = 0, \tag{A.9}$$

which is the wave equation we started from in Chapter. 1! Here $c = 1/\sqrt{\epsilon\mu}$ is the velocity of light in the medium.

Appendix B

Derivation of M2 for a general superposition of HG beams

We now derive a general formula for M^2 for a beam $E(x)$ that is a superposition of HG beams $E_m(x)$. This result is more general than the equations obtained in Refs. [99, 100, 176]. Mathematically $E_m(x)$ also describes the wave function of the energy states of a simple harmonic oscillator in quantum mechanics. Thus we use the *bra-ket* mathematical formalism to calculate expectation values as is standard in quantum mechanics. Consider a general state $|\psi\rangle$ described in the HG beam basis i.e., $|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$ where $|n\rangle$ is the state of the HG mode of order n and c_n are complex coefficients satisfying $\sum_{n=0}^{\infty} |c_n|^2 = 1$.

In the main text we found Eq. (4.9) as the basis of the covariance method. We use the general form for the covariance matrix elements in Section 4.6, Eq. (4.22) to find M^2 from $|\psi\rangle$. It is useful to use dimensionless position X and momentum P (rather than angle) operators defined as follows $X = x/2\sigma_x$ and $P = p\sigma_x/\hbar = 2\pi\sigma_x\theta/\lambda$. In terms of these dimensionless operators, Eq. (4.9) is rewritten as

$$\begin{aligned} M^2 &= 4\sqrt{\det Q} \qquad \text{(B.1)} \\ &= 4\sqrt{(\Delta X)^2 (\Delta P)^2 - (\langle XP + PX\rangle/2 - \langle X\rangle\langle P\rangle)^2} \\ &= 4\sqrt{\left(\langle X^2\rangle - \langle X\rangle^2\right)\left(\langle P^2\rangle - \langle P\rangle^2\right) - (\langle XP + PX\rangle/2 - \langle X\rangle\langle P\rangle)^2} \\ &= 4\sqrt{\langle X^2\rangle\langle P^2\rangle - \langle X^2\rangle\langle P\rangle^2 - \langle X\rangle^2\langle P^2\rangle + 2\langle X\rangle\langle P\rangle \operatorname{Re}\langle XP\rangle - (\operatorname{Re}\langle XP\rangle)^2}, \end{aligned}$$

where we have used $\frac{1}{2}\langle XP + PX\rangle = \operatorname{Re}\langle XP\rangle$.

The calculation of the expectation values appearing in above equation can be done using the ladder operators a and $a^\dagger$ operators commonly used in quantum mechanics. These operators are related to X and P according to the following equations $X = \left(a + a^\dagger\right)/2$ and $P = i\left(a^\dagger - a\right)/2$. The action of these operators on the $|n\rangle$ states is given by $a|n\rangle = \sqrt{n}\,|n-1\rangle$ and $a^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle$. Expectation values can now be easily obtained, for example $\langle P\rangle$ is calculated

as follows:

$$\begin{aligned}\langle P\rangle &= \left\langle i\left(a^{\dagger}-a\right)/2\right\rangle \\ &= \frac{i}{2}\sum_{n=0}^{\infty}\left(c_n c_{n+1}^{*}-c_n^{*}c_{n+1}\right)\sqrt{n+1} \\ &= -\operatorname{Im}\left(B\right), \end{aligned} \tag{B.2}$$

where constants A,B, C will be defined explicitly at the end. While $\left\langle P^2\right\rangle$ is given by

$$\begin{aligned}\left\langle P^2\right\rangle &= -\left\langle a^2+\left(a^{\dagger}\right)^2-2\langle n\rangle-1\right\rangle/4 \\ &= -\frac{1}{2}\sum_{n=0}^{\infty}\operatorname{Re}\left(c_n c_{n+2}^{*}\right)\sqrt{(n+1)(n+2)}+1/2\sum_{n=0}^{\infty}n|c_n|^2+1/4 \\ &= -\frac{1}{2}\operatorname{Re}\left(A\right)+\frac{C}{4}. \end{aligned} \tag{B.3}$$

Following a similar procedure, one obtains $\langle X\rangle=\operatorname{Re}\left(B\right)$, $\left\langle X^2\right\rangle=\frac{1}{2}\operatorname{Re}\left(A\right)+\frac{C}{4}$, and $\operatorname{Re}\left(XP\right)=-\frac{1}{2}\operatorname{Im}\left(A\right)$.

Substituting these into Eq. (B.1) we obtain the following general expression of M^2 for an arbitrary superposition of HG beams:

$$M^2=\sqrt{C^2+8\operatorname{Re}\left(B^2A^{*}\right)-4|A|^2-4|B|^2C}, \tag{B.4}$$

where A^{*} is the complex conjugate of A and

$$A = \sum_{n=0}^{\infty}c_n c_{n+2}^{*}\sqrt{(n+1)(n+2)}, \tag{B.5}$$

$$B = \sum_{n=0}^{\infty}c_n c_{n+1}^{*}\sqrt{n+1}, \tag{B.6}$$

$$C = 1+2\sum_{n=0}^{\infty}n|c_n|^2. \tag{B.7}$$

Bibliography

[1] Nicolas K Fontaine, Roland Ryf, Haoshuo Chen, David T Neilson, Kwangwoong Kim, and Joel Carpenter. Laguerre-Gaussian mode sorter. *Nature Communications*, 10(1):1–7, 2019.

[2] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994.

[3] B.E.A. Saleh and M.C. Teich. *Fundamentals of Photonics*. Wiley Series in Pure and Applied Optics. Wiley, 2020.

[4] Okan K Ersoy. *Diffraction, Fourier optics and imaging*. John Wiley & Sons, 2006.

[5] Eugene Hecht. *Optics*. Pearson Education India, 2012.

[6] EU Condon. Immersion of the Fourier transform in a continuous group of functional transformations. *Proceedings of the National academy of Sciences of the United States of America*, 23(3):158, 1937.

[7] Victor Namias. The Fractional Order Fourier Transform and its Application to Quantum Mechanics. *IMA Journal of Applied Mathematics*, 25(3):241–265, 1980.

[8] Miguel A. Alonso. Wigner functions in optics: describing beams as ray bundles and pulses as particle ensembles. *Adv. Opt. Photon.*, 3(4):272–365, Dec 2011.

[9] H.M. Ozaktas, D. Mendlovic, M.A. Kutay, and Z. Zalevsky. *The Fractional Fourier Transform: With Applications in Optics and Signal Processing*. Wiley Series in Pure and Applied Optics. Wiley, 2001.

[10] P. Kunche and N. Manikanthababu. *Fractional Fourier Transform Techniques for Speech Enhancement*. Springer Briefs in Speech Technology. Springer International Publishing, 2020.

[11] Naveen Kumar Nishchal, Joby Joseph, and Kehar Singh. Securing information using fractional Fourier transform in digital holography. *Optics Communications*, 235(4):253–259, 2004.

[12] Pierre Pellat-Finet. Fresnel diffraction and the fractional-order Fourier transform. *Optics Letters*, 19(18):1388–1390, 1994.

[13] Adolf W. Lohmann. Image rotation, Wigner rotation, and the fractional Fourier transform. *J. Opt. Soc. Am. A*, 10(10):2181–2186, 1993.

[14] David Mendlovic and Haldun M Ozaktas. Fractional Fourier transforms and their optical implementation: I. *JOSA A*, 10(9):1875–1881, 1993.

[15] Haldun M Ozaktas and David Mendlovic. Fractional Fourier transforms and their optical implementation. II. *JOSA A*, 10(12):2522–2531, 1993.

[16] Steffen Weimann, Armando Perez-Leija, Maxime Lebugle, Robert Keil, Malte Tichy, Markus Gräfe, René Heilmann, Stefan Nolte, Hector Moya-Cessa, Gregor Weihs, et al. Implementation of quantum and classical discrete fractional Fourier transforms. *Nat. Commun.*, 7:11027, 2016.

[17] C. Fabre and N. Treps. Modes and states in quantum optics. *Reviews of Modern Physics*, 92(3), September 2020.

[18] E. Merzbacher. *Quantum Mechanics*. Wiley, 1998.

[19] Jun John Sakurai and Eugene D Commins. Modern quantum mechanics, revised edition, 1995.

[20] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73:58–61, Jul 1994.

[21] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[22] William R. Clements, Peter C. Humphreys, Benjamin J. Metcalf, W. Steven Kolthammer, and Ian A. Walmsley. Optimal design for universal multiport interferometers. *Optica*, 3(12):1460–1465, Dec 2016.

[23] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martín-López, Nicholas J. Russell, Joshua W. Silverstone, Peter J. Shadbolt, Nobuyuki Matsuda, Manabu Oguma, Mikitaka Itoh, Graham D. Marshall, Mark G. Thompson, Jonathan C. F. Matthews, Toshikazu Hashimoto, Jeremy L. O'Brien, and Anthony Laing. Universal linear optics. *Science*, 349(6249):711–716, 2015.

[24] David P DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, 48(9-11):771–783, 2000.

[25] Seth Lloyd. Almost any quantum logic gate is universal. *Physical Review Letters*, 75(2):346, 1995.

[26] Pieter Kok, Hwang Lee, and Jonathan P Dowling. Single-photon quantum-nondemolition detectors constructed with linear optics and projective measurements. *Physical Review A*, 66(6):063814, 2002.

[27] Emanuel Knill, Raymond Laflamme, and Gerald J Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, 2001.

[28] Norbert Lütkenhaus, John Calsamiglia, and K-A Suominen. Bell measurements for teleportation. *Physical Review A*, 59(5):3295, 1999.

[29] Lev Vaidman and Nadav Yoran. Methods for reliable teleportation. *Physical Review A*, 59(1):116, 1999.

[30] Pieter Kok, William J Munro, Kae Nemoto, Timothy C Ralph, Jonathan P Dowling, and Gerard J Milburn. Linear optical quantum computing with photonic qubits. *Reviews of Modern Physics*, 79(1):135, 2007.

[31] Joachim von zur Gathen. Permanent and determinant. *Linear Algebra and its Applications*, 96:87–100, 1987.

[32] Jan Van Leeuwen. *Handbook of theoretical computer science (vol. A). Algorithms and Complexity*. MIT Press, 1991.

[33] Matthew A Broome, Alessandro Fedrizzi, Saleh Rahimi-Keshari, Justin Dove, Scott Aaronson, Timothy C Ralph, and Andrew G White. Photonic boson sampling in a tunable circuit. *Science*, 339(6121):794–798, 2013.

[34] Justin B Spring, Benjamin J Metcalf, Peter C Humphreys, W Steven Kolthammer, Xian-Min Jin, Marco Barbieri, Animesh Datta, Nicholas Thomas-Peter, Nathan K Langford, Dmytro Kundys, et al. Boson sampling on a photonic chip. *Science*, 339(6121):798–801, 2013.

[35] Max Tillmann, Borivoje Dakić, René Heilmann, Stefan Nolte, Alexander Szameit, and Philip Walther. Experimental boson sampling. *Nature Photonics*, 7(7):540–544, 2013.

[36] Andrea Crespi, Roberto Osellame, Roberta Ramponi, Daniel J Brod, Ernesto F Galvao, Nicolo Spagnolo, Chiara Vitelli, Enrico Maiorino, Paolo Mataloni, and Fabio Sciarrino. Integrated multimode interferometers with arbitrary designs for photonic boson sampling. *Nature Photonics*, 7(7):545–549, 2013.

[37] Jacques Carolan, Jasmin DA Meinecke, Peter J Shadbolt, Nicholas J Russell, Nur Ismail, Kerstin Wörhoff, Terry Rudolph, Mark G Thompson, Jeremy L O'brien, Jonathan CF Matthews, et al. On the experimental verification of quantum complexity in linear optics. *Nature Photonics*, 8(8):621–626, 2014.

[38] Nicolo Spagnolo, Chiara Vitelli, Marco Bentivegna, Daniel J Brod, Andrea Crespi, Fulvio Flamini, Sandro Giacomini, Giorgio Milani, Roberta Ramponi, Paolo Mataloni, et al. Efficient experimental validation of photonic boson sampling against the uniform distribution. *arXiv preprint arXiv:1311.1622*, 2013.

[39] Hui Wang, Jian Qin, Xing Ding, Ming-Cheng Chen, Si Chen, Xiang You, Yu-Ming He, Xiao Jiang, L You, Z Wang, et al. Boson sampling with 20 input photons and a 60-mode interferometer in a 10^{14} dimensional hilbert space. *Physical Review Letters*, 123(25):250503, 2019.

[40] Aram W Harrow and Ashley Montanaro. Quantum computational supremacy. *Nature*, 549(7671):203–209, 2017.

[41] Dominik Hangleiter and Jens Eisert. Computational advantage of quantum random sampling. *Reviews of Modern Physics*, 95(3):035001, 2023.

[42] Howard M. Wiseman and Gerard J. Milburn. *Quantum Measurement and Control*. Cambridge University Press, 2009.

[43] J.S. Lundeen and K.J. Resch. Practical measurement of joint weak values and their connection to the annihilation operator. *Physics Letters A*, 334(5):337–344, 2005.

[44] Graeme Mitchison, Richard Jozsa, and Sandu Popescu. Sequential weak measurement. *Phys. Rev. A*, 76:062105, 2007.

[45] Jeff S. Lundeen and Charles Bamber. Procedure for Direct Measurement of General Quantum States Using Weak Measurement. *Phys. Rev. Lett.*, 108:070402, 2012.

[46] Yakir Aharonov, David Z. Albert, and Lev Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.*, 60:1351–1354, 1988.

[47] Hailong Zhou, Jianji Dong, Junwei Cheng, Wenchan Dong, Chaoran Huang, Yichen Shen, Qiming Zhang, Min Gu, Chao Qian, Hongsheng Chen, Zhichao Ruan, and Xinliang Zhang. Photonic matrix multiplication lights up photonic accelerator and beyond. *Light-Science & Applications*, 11(1), Feb 3 2022.

[48] Florian Brandt, Markus Hiekkamäki, Frédéric Bouchard, Marcus Huber, and Robert Fickler. High-dimensional quantum gates using full-field spatial modes of photons. *Optica*, 7(2):98–107, Feb 2020.

[49] Ohad Lib, Kfir Sulimany, and Yaron Bromberg. Processing entangled photons in high dimensions with a programmable light converter. *Phys. Rev. Appl.*, 18:014063, Jul 2022.

[50] Zhi-Qiang Jiao, Jun Gao, Wen-Hao Zhou, Xiao-Wei Wang, Ruo-Jing Ren, Xiao-Yun Xu, Lu-Feng Qiao, Yao Wang, and Xian-Min Jin. Two-dimensional quantum walks of correlated photons. *Optica*, 8(9):1129–1135, Sep 2021.

[51] Chiara Esposito, Mariana R Barros, Andrés Durán Hernández, Gonzalo Carvacho, Francesco Di Colandrea, Raouf Barboza, Filippo Cardano, Nicolò Spagnolo, Lorenzo Marrucci, and Fabio Sciarrino. Quantum walks of two correlated photons in a 2D synthetic lattice. *npj Quantum Information*, 8(1):34, 2022.

[52] Andreas Schreiber, Aurél Gábris, Peter P Rohde, Kaisa Laiho, Martin Štefaňák, Václav Potoček, Craig Hamilton, Igor Jex, and Christine Silberhorn. A 2D quantum walk simulation of two-particle dynamics. *Science*, 336(6077):55–58, 2012.

[53] Yichen Shen, Nicholas C Harris, Scott Skirlo, Mihika Prabhu, Tom Baehr-Jones, Michael Hochberg, Xin Sun, Shijie Zhao, Hugo Larochelle, Dirk Englund, et al. Deep learning with coherent nanophotonic circuits. *Nature Photonics*, 11(7):441–446, 2017.

[54] Ying Zuo, Bohan Li, Yujun Zhao, Yue Jiang, You-Chiuan Chen, Peng Chen, Gyu-Boong Jo, Junwei Liu, and Shengwang Du. All-optical neural network with nonlinear activation functions. *Optica*, 6(9):1132–1137, 2019.

[55] Johannes Feldmann, Nathan Youngblood, C David Wright, Harish Bhaskaran, and Wolfram HP Pernice. All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature*, 569(7755):208–214, 2019.

[56] Chaoran Huang, Shinsuke Fujisawa, Thomas Ferreira de Lima, Alexander N Tait, Eric C Blow, Yue Tian, Simon Bilodeau, Aashu Jha, Fatih Yaman, Hsuan-Tung Peng, et al. A silicon photonic–electronic neural network for fibre nonlinearity compensation. *Nature Electronics*, 4(11):837–844, 2021.

[57] Aashu Jha, Chaoran Huang, and Paul R. Prucnal. Reconfigurable all-optical nonlinear activation functions for neuromorphic photonics. *Opt. Lett.*, 45(17):4819–4822, Sep 2020.

[58] Jean-François Morizur, Lachlan Nicholls, Pu Jian, Seiji Armstrong, Nicolas Treps, Boris Hage, Magnus Hsu, Warwick Bowen, Jiri Janousek, and Hans-A. Bachor. Programmable unitary spatial mode manipulation. *J. Opt. Soc. Am. A*, 27(11):2524–2531, Nov 2010.

[59] Guillaume Labroille, Bertrand Denolle, Pu Jian, Philippe Genevaux, Nicolas Treps, and Jean-François Morizur. Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion. *Opt. Express*, 22(13):15599–15607, Jun 2014.

[60] Seiji Armstrong, Jean-François Morizur, Jiri Janousek, Boris Hage, Nicolas Treps, Ping Koy Lam, and Hans-A Bachor. Programmable multimode quantum networks. *Nature Communications*, 3(1):1–8, 2012.

[61] Xing Lin, Yair Rivenson, Nezih T. Yardimci, Muhammed Veli, Yi Luo, Mona Jarrahi, and Aydogan Ozcan. All-optical machine learning using diffractive deep neural networks. *Science*, 361(6406):1004–1008, 2018.

[62] Hugo Larocque and Dirk Englund. Universal linear optics by programmable multimode interference. *Opt. Express*, 29(23):38257–38267, Nov 2021.

[63] Caterina Taballione, Reinier van der Meer, Henk J Snijders, Peter Hooijschuur, Jörn P Epping, Michiel de Goede, Ben Kassenberg, Pim Venderbosch, Chris Toebes, Hans van den Vlekkert, et al. A universal fully reconfigurable 12-mode quantum photonic processor. *Materials for Quantum Technology*, 1(3):035002, 2021.

[64] ZI Borevich and SL Krupetskii. Subgroups of the unitary group that contain the group of diagonal matrices. *Journal of Soviet Mathematics*, 17(4):1951–1959, 1981.

[65] Víctor López Pastor, Jeff Lundeen, and Florian Marquardt. Arbitrary optical wave evolution with fourier transforms and phase masks. *Opt. Express*, 29(23):38441–38450, Nov 2021.

[66] MG Moharam and L Young. Criterion for Bragg and Raman-Nath diffraction regimes. *Applied Optics*, 17(11):1757–1759, 1978.

[67] Yohei Sakamaki, Takashi Saida, Toshikazu Hashimoto, and Hiroshi Takahashi. New optical waveguide design based on wavefront matching method. *Journal of lightwave technology*, 25(11):3511–3518, 2007.

[68] Nicolas K. Fontaine, Roland Ryf, Haoshuo Chen, David Neilson, and Joel Carpenter. Design of high order mode-multiplexers using multiplane light conversion. In *2017 European Conference on Optical Communication (ECOC)*, pages 1–3, 2017.

[69] Jeffrey A. Davis, Don M. Cottrell, Juan Campos, María J. Yzuel, and Ignacio Moreno. Encoding amplitude information onto phase-only filters. *Appl. Opt.*, 38(23):5004–5013, Aug 1999.

[70] Eliot Bolduc, Nicolas Bent, Enrico Santamato, Ebrahim Karimi, and Robert W. Boyd. Exact solution to simultaneous intensity and phase encryption with a single phase-only hologram. *Opt. Lett.*, 38(18):3546–3549, Sep 2013.

[71] John M. Khosrofian and Bruce A. Garetz. Measurement of a Gaussian laser beam diameter through the direct inversion of knife-edge data. *Appl. Opt.*, 22(21):3406–3410, Nov 1983.

[72] P. Picart and J. Li. *Digital Holography*. ISTE. Wiley, 2013.

[73] T. Kreis. *Handbook of Holographic Interferometry: Optical and Digital Methods*. Wiley, 2006.

[74] Salvador Elías Venegas-Andraca. Quantum walks: a comprehensive review. *Quantum Information Processing*, 11(5):1015–1106, 2012.

[75] Orad Reshef, Michael P DelMastro, Katherine KM Bearne, Ali H Alhulaymi, Lambert Giner, Robert W Boyd, and Jeff S Lundeen. An optic to replace space and its application towards ultra-thin imaging systems. *Nature Communications*, 12(1):1–8, 2021.

[76] Anthony E. Siegman. New developments in laser resonators. In Dale A. Holmes, editor, *Optical Resonators*, volume 1224, pages 2 – 14. International Society for Optics and Photonics, SPIE, 1990.

[77] Anthony E Siegman. Defining, measuring, and optimizing laser beam quality. *Laser Resonators and Coherent Optics: Modeling, Technology, and Applications*, 1868:2–12, 1993.

[78] Anthony E Siegman. How to (maybe) measure laser beam quality. In *Diode Pumped Solid State Lasers: Applications and Issues*, page MQ1. Optica Publishing Group, 1998.

[79] Lasers and laser-related equipment - Test methods for laser beam widths, divergence angles and beam propagation ratios - Part 1. ISO 11146-1:2021(E).

[80] J. W. Goodman and R. W. Lawrence. Digital image formation from electronically detected holograms. *Applied Physics Letters*, 11(3):77–79, 1967.

[81] S. Grilli, P. Ferraro, S. De Nicola, A. Finizio, G. Pierattini, and R. Meucci. Whole optical wavefields reconstruction by digital holography. *Opt. Express*, 9(6):294–302, Sep 2001.

[82] R.W. Boyd, S.G. Lukishova, and V.N. Zadkov. *Quantum Photonics: Pioneering Advances and Emerging Applications*. Springer Series in Optical Sciences. Springer International Publishing, 2019.

[83] Johannes Hartmann. *Objektivuntersuchungen*. Springer, 1904.

[84] Roland V Shack. Production and use of a lenticular Hartmann screen. In *Spring Meeting of Optical Society of America, 1971*, volume 656, 1971.

[85] Bernd Schäfer and Klaus Mann. Determination of beam parameters and coherence properties of laser radiation by use of an extended Hartmann-Shack wave-front sensor. *Applied Optics*, 41(15):2809–2817, 2002.

[86] Richa Sharma, J Solomon Ivan, and CS Narayanamurthy. Wave propagation analysis using the variance matrix. *JOSA A*, 31(10):2185–2191, 2014.

[87] Thomas Kreis. *Digital Recording and Numerical Reconstruction of Wave Fields*, chapter 3, pages 81–183. John Wiley & Sons, Ltd, 2004.

[88] U. Schnars, C. Falldorf, J. Watson, and W. Jüptner. *Digital Holography and Wavefront Sensing: Principles, Techniques and Applications*. EBL-Schweitzer. Springer Berlin Heidelberg, 2014.

[89] Stephan Wielandy. Implications of higher-order mode content in large mode area fibers with good beam quality. *Optics Express*, 15(23):15402–15409, 2007.

[90] S Liao, M Gong, and H Zhang. Theoretical calculation of beam quality factor of large-mode-area fiber amplifiers. *Laser Physics*, 19:437–444, 2009.

[91] Fabian Stutzki, Florian Jansen, Hans-Jürgen Otto, Cesar Jauregui, Jens Limpert, and Andreas Tünnermann. Designing advanced very-large-mode-area fibers for power scaling of fiber-laser systems. *Optica*, 1(4):233–242, Oct 2014.

[92] H. Zhou, A. Jin, Z. Chen, B. Zhang, X. Zhou, S. Chen, J. Hou, and J. Chen. Combined supercontinuum source with > 200 W power using a 3×1 broadband fiber power combiner. *Opt. Lett.*, 40(16):3810–3813, Aug 2015.

[93] N Hodgson and T Haase. Beam parameters, mode structure and diffraction losses of slab lasers with unstable resonators. *Optical and quantum electronics*, 24:S903–S926, 1992.

[94] Cesar Jauregui, Christoph Stihler, and Jens Limpert. Transverse mode instability. *Adv. Opt. Photon.*, 12(2):429–484, Jun 2020.

[95] Gerald F Marshall and Glenn E Stutz. *Handbook of optical and laser scanning*. Taylor & Francis, 2012.

[96] Andrew Forbes. *Laser beam propagation: generation and propagation of customized light*. CRC Press, 2014.

[97] Shimon Lavi, Ron Prochaska, and Eliezer Keren. Generalized beam parameters and transformation laws for partially coherent light. *Applied Optics*, 27(17):3696–3703, 1988.

[98] Anthony E Siegman. Defining the effective radius of curvature for a nonideal optical beam. *IEEE Journal of Quantum Electronics*, 27(5):1146–1148, 1991.

[99] H Weber. Some historical and technical aspects of beam quality. *Optical and Quantum Electronics*, 24:S861–S864, 1992.

[100] K M Du, G Herziger, P Loosen, and F Rühl. Coherence and intensity moments of laser light. *Optical and quantum electronics*, 24:S1081–S1093, 1992.

[101] G. Nemes and A. E. Siegman. Measurement of all ten second-order moments of an astigmatic beam by the use of rotating simple astigmatic (anamorphic) optics. *J. Opt. Soc. Am. A*, 11(8):2257–2264, Aug 1994.

[102] C Gao, H Weber, and M Gao. Characterization of laser beams by using intensity moments. In *ICO20: Lasers and Laser Technologies*, volume 6028, pages 428–435. SPIE, 2005.

[103] Hidehiko Yoda, Pavel Polynkin, and Masud Mansuripur. Beam quality factor of higher order modes in a step-index fiber. *J. Lightwave Technol.*, 24(3):1350, Mar 2006.

[104] Oliver A. Schmidt, Christian Schulze, Daniel Flamm, Robert Brüning, Thomas Kaiser, Siegmund Schröter, and Michael Duparré. Real-time determination of laser beam quality by modal decomposition. *Opt. Express*, 19(7):6741–6748, Mar 2011.

[105] Daniel Flamm, Christian Schulze, Robert Brüning, Oliver A. Schmidt, Thomas Kaiser, Siegmund Schröter, and Michael Duparré. Fast M^2 measurement for fiber beams based on modal analysis. *Appl. Opt.*, 51(7):987–993, Mar 2012.

[106] Yong-zhao Du, Guo-ying Feng, Hong-ru Li, Zhen Cai, Hong Zhao, and Shou-huan Zhou. Real-time determination of beam propagation factor by Mach-Zehnder point diffraction interferometer. *Optics Communications*, 287:1–5, 2013.

[107] Yongzhao Du, Yuqing Fu, and Lixin Zheng. Complex amplitude reconstruction for dynamic beam quality M^2 factor measurement with self-referencing interferometer wavefront sensor. *Applied Optics*, 55(36):10180–10186, 2016.

[108] Mengqi Du, Lars Loetgering, Kjeld S. E. Eikema, and Stefan Witte. Measuring laser beam quality, wavefronts, and lens aberrations using ptychography. *Opt. Express*, 28(4):5022–5034, Feb 2020.

[109] Julio Serna, Fernando Encinas-Sanz, and George Nemes. Complete spatial characterization of a pulsed doughnut-type beam by use of spherical optics and a cylindrical lens. *J. Opt. Soc. Am. A*, 18(7):1726–1733, Jul 2001.

[110] M. Alonso. Wigner functions in optics: describing beams as ray bundles and pulses as particle ensembles. *Adv. Opt. Photon.*, 3:272–365, 2011.

[111] M Crampin and F.A.E. Pirani. *Applicable Differential Geometry*. Lecture note series. Cambridge University Press, 1986.

[112] M. Hazewinkel. *Encyclopaedia of Mathematics*. Number v. 1 in Encyclopaedia of Mathematics. Springer Netherlands, 2012.

[113] VV Dodonov and OV Man'ko. Universal invariants of paraxial optical beams. *Computer Optics*, 1(1):65–68, 1989.

[114] R Simon and N Mukunda. Optical phase space, Wigner representation, and invariant quality parameters. *JOSA A*, 17(12):2440–2463, 2000.

[115] Victor V Dodonov and Olga V Man'ko. Universal invariants of quantum-mechanical and optical systems. *JOSA A*, 17(12):2403–2410, 2000.

[116] A.J. Pettofrezzo. *Matrices and Transformations*. Dover books on advanced mathematics. Prentice-Hall, 1966.

[117] Martin J. Bastiaans. Second-order moments of the Wigner distribution function in first-order optical systems. In *Optical Society of America Annual Meeting*, page FC2. Optica Publishing Group, 1991.

[118] Lasers and laser-related equipment - Test methods for laser beam widths, divergence angles and beam propagation ratios - Part 3 (ISO 11146-3: 2004). ISO 11146-3:2004(E).

[119] `https://figshare.com/s/a329aba5e892d1d29b6a`. Code available in Optica Publishing Group figshare platform.

[120] H Weber. Propagation of higher-order intensity moments in quadratic-index media. *Optical and Quantum Electronics*, 24:S1027–S1049, 1992.

[121] Shirong Luo and Baida Lü. M2 factor and kurtosis parameter of super-gaussian beams passing through an axicon. *Optik*, 114(5):193–198, 2003.

[122] Andreas Letsch and Adolf Giesen. Characterization of a general astigmatic laser beam by measuring its ten second order moments. In Laurent Mazuray and Rolf Wartmann, editors, *Optical Design and Engineering II*, volume 5962. International Society for Optics and Photonics, SPIE, 2005.

[123] Lasers and laser-related equipment — Test methods for laser beam widths, divergence angles and beam propagation ratios — Part 2: General astigmatic beams (ISO 11146-2: 2021). ISO 11146-2:2021(E).

[124] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 47:777–780, 1935.

[125] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, 2009.

[126] Yakir Aharonov and Lev Vaidman. Properties of a quantum system during the time interval between two measurements. *Phys. Rev. A*, 41:11–20, 1990.

[127] K. J. Resch and A. M. Steinberg. Extracting Joint Weak Values with Local, Single-Particle Measurements. *Phys. Rev. Lett.*, 92:130402, 2004.

[128] G. S. Thekkadath, L. Giner, Y. Chalich, M. J. Horton, J. Banker, and J. S. Lundeen. Direct Measurement of the Density Matrix of a Quantum System. *Phys. Rev. Lett.*, 117:120401, 2016.

[129] F. Piacentini, A. Avella, M. P. Levi, M. Gramegna, G. Brida, I. P. Degiovanni, E. Cohen, R. Lussana, F. Villa, A. Tosi, F. Zappa, and M. Genovese. Measuring Incompatible Observables by Exploiting Sequential Weak Values. *Phys. Rev. Lett.*, 117:170402, 2016.

[130] M.A. Ochoa, W. Belzig, and A. Nitzan. Simultaneous weak measurement of non-commuting observables: a generalized Arthurs-Kelly protocol. *Sci. Rep.*, 8(1):1–8, 2018.

[131] Y. Kim, YS Kim, SY Lee, Sang-Wook Han, Sung Moon, Yoon-Ho Kim, and Young-Wook Cho. Direct quantum process tomography via measuring sequential weak values of incompatible observables. *Nat. Commun.*, 9(1):1–6, 2018.

[132] Jiang-Shan Chen, Meng-Jun Hu, Xiao-Min Hu, Bi-Heng Liu, Yun-Feng Huang, Chuan-Feng Li, Can-Guang Guo, and Yong-Sheng Zhang. Experimental realization of sequential weak measurements of non-commuting Pauli observables. *Opt. Express*, 27(5):6089–6097, 2019.

[133] G S Thekkadath, F Hufnagel, and J S Lundeen. Determining complementary properties using weak-measurement: uncertainty, predictability, and disturbance. *New Journal of Physics*, 20(11):113034, 2018.

[134] Onur Hosten and Paul Kwiat. Observation of the Spin Hall Effect of Light via Weak Measurements. *Science*, 319(5864):787–790, 2008.

[135] David J. Starling, P. Ben Dixon, Andrew N. Jordan, and John C. Howell. Optimizing the signal-to-noise ratio of a beam-deflection measurement with interferometric weak values. *Phys. Rev. A*, 80:041803, 2009.

[136] Nicolas Brunner and Christoph Simon. Measuring Small Longitudinal Phase Shifts: Weak Measurements or Standard Interferometry? *Phys. Rev. Lett.*, 105:010405, 2010.

[137] Jeff S. Lundeen, Brandon Sutherland, Aabid Patel, Corey Stewart, and Charles Bamber. Direct measurement of the quantum wavefunction. *Nature*, 474(7350):188–191, 2011.

[138] Justin Dressel, Mehul Malik, Filippo M. Miatto, Andrew N. Jordan, and Robert W. Boyd. Colloquium: Understanding quantum weak values: Basics and applications. *Rev. Mod. Phys.*, 86:307–316, 2014.

[139] Yakir Aharonov, Sandu Popescu, Daniel Rohrlich, and Paul Skrzypczyk. Quantum Cheshire Cats. *New Journal of Physics*, 15(11):113015, 2013.

[140] T Denkmayr, H Geppert, S Sponar, H Lemmel, A Matzkin, J Tollaksen, and Y Hasegawa. Observation of a quantum Cheshire cat in a matter-wave interferometer experiment. *Nat. Commun.*, 5(1):1–7, 2014.

[141] J. S. Lundeen and A. M. Steinberg. Experimental Joint Weak Measurement on a Photon Pair as a Probe of Hardy's Paradox. *Phys. Rev. Lett.*, 102:020404, 2009.

[142] Kazuhiro Yokota, Takashi Yamamoto, Masato Koashi, and Nobuyuki Imoto. Direct observation of Hardy's paradox by joint weak measurement with an entangled photon pair. *New Journal of Physics*, 11(3):033011, 2009.

[143] Justin Dressel, Areeya Chantasri, Andrew N. Jordan, and Alexander N. Korotkov. Arrow of Time for Continuous Quantum Measurement. *Phys. Rev. Lett.*, 119:220507, 2017.

[144] D. Curic, M. C. Richardson, G. S. Thekkadath, J. Flórez, L. Giner, and J. S. Lundeen. Experimental investigation of measurement-induced disturbance and time symmetry in quantum physics. *Phys. Rev. A*, 97:042128, 2018.

[145] Holger F. Hofmann. Uncertainty limits for quantum metrology obtained from the statistics of weak measurements. *Phys. Rev. A*, 83:022106, 2011.

[146] Lee A. Rozema, Ardavan Darabi, Dylan H. Mahler, Alex Hayat, Yasaman Soudagar, and Aephraim M. Steinberg. Violation of Heisenberg's Measurement-Disturbance Relationship by Weak Measurements. *Phys. Rev. Lett.*, 109:100404, 2012.

[147] Shengshi Pang, Justin Dressel, and Todd A. Brun. Entanglement-Assisted Weak Value Amplification. *Phys. Rev. Lett.*, 113:030401, 2014.

[148] Andrew N. Jordan, Julián Martínez-Rincón, and John C. Howell. Technical Advantages for Weak-Value Amplification: When Less Is More. *Phys. Rev. X*, 4:011031, 2014.

[149] Jérémie Harris, Robert W. Boyd, and Jeff S. Lundeen. Weak Value Amplification Can Outperform Conventional Measurement in the Presence of Detector Saturation. *Phys. Rev. Lett.*, 118:070802, 2017.

[150] Binke Xia, Jingzheng Huang, Chen Fang, Hongjing Li, and Guihua Zeng. High-Precision Multiparameter Weak Measurement with Hermite-Gaussian Pointer. *Phys. Rev. Applied*, 13:034023, 2020.

[151] Nicole Yunger Halpern, Brian Swingle, and Justin Dressel. Quasiprobability behind the out-of-time-ordered correlator. *Phys. Rev. A*, 97:042105, Apr 2018.

[152] Marco Barbieri. Multiple-measurement Leggett-Garg inequalities. *Phys. Rev. A*, 80:034102, Sep 2009.

[153] N. W. M. Ritchie, J. G. Story, and Randall G. Hulet. Realization of a measurement of a "weak value". *Phys. Rev. Lett.*, 66:1107–1110, 1991.

[154] S Glancy, E Knill, and M Girard. Gradient-based stopping rules for maximum-likelihood quantum-state tomography. *New Journal of Physics*, 14(9):095017, 2012.

[155] Eliot Bolduc, George C Knee, Erik M Gauger, and Jonathan Leach. Projected gradient descent algorithms for quantum state tomography. *npj Quantum Information*, 3(1):1–9, 2017.

[156] Rajveer Nehra, Miller Eaton, Carlos González-Arciniegas, M. S. Kim, Thomas Gerrits, Adriana Lita, Sae Woo Nam, and Olivier Pfister. Generalized overlap quantum state tomography. *Phys. Rev. Research*, 2:042002, 2020.

[157] J Z Salvail, M Agnew, Allan S Johnson, Eliot Bolduc, Jonathan Leach, and Robert W Boyd. Full characterization of polarization states of light via direct measurement. *Nat. Photonics*, 7(4):316–321, 2013.

[158] Charles Bamber and Jeff S. Lundeen. Observing Dirac's Classical Phase Space Analog to the Quantum State. *Phys. Rev. Lett.*, 112:070405, 2014.

[159] Haldun M. Ozaktas and David Mendlovic. Fractional Fourier optics. *J. Opt. Soc. Am. A*, 12(4):743–751, 1995.

[160] A. Hariri, D. Curic, L. Giner, and J. S. Lundeen. Experimental simultaneous readout of the real and imaginary parts of the weak value. *Phys. Rev. A*, 100:032119, 2019.

[161] Gregory D Scholes, Graham R Fleming, Lin X Chen, Alán Aspuru-Guzik, Andreas Buchleitner, David F Coker, Gregory S Engel, Rienk Van Grondelle, Akihito Ishizaki, David M Jonas, et al. Using coherence to enhance function in chemical and biophysical systems. *Nature*, 543(7647):647–656, 2017.

[162] Arthur Fine. Hidden Variables, Joint Probability, and the Bell Inequalities. *Phys. Rev. Lett.*, 48:291–295, Feb 1982.

[163] Matthew F. Pusey. Anomalous Weak Values Are Proofs of Contextuality. *Phys. Rev. Lett.*, 113:200401, Nov 2014.

[164] Ravi Kunjwal, Matteo Lostaglio, and Matthew F. Pusey. Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts. *Phys. Rev. A*, 100:042116, Oct 2019.

[165] Valeria Cimini, Ilaria Gianani, Fabrizio Piacentini, Ivo Pietro Degiovanni, and Marco Barbieri. Anomalous values, Fisher information, and contextuality, in generalized quantum measurements. *Quantum Science and Technology*, 5(2):025007, Mar 2020.

[166] A. J. Leggett and Anupam Garg. Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Phys. Rev. Lett.*, 54:857–860, Mar 1985.

[167] Alessio Avella, Fabrizio Piacentini, Michelangelo Borsarelli, Marco Barbieri, Marco Gramegna, Rudi Lussana, Federica Villa, Alberto Tosi, Ivo Pietro Degiovanni, and Marco Genovese. Anomalous weak values and the violation of a multiple-measurement Leggett-Garg inequality. *Phys. Rev. A*, 96:052123, Nov 2017.

[168] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan. Experimental violation of Two-Party Leggett-Garg Inequalities with Semiweak Measurements. *Phys. Rev. Lett.*, 106:040402, Jan 2011.

[169] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O'Brien, A. G. White, and G. J. Pryde. Violation of the Leggett–Garg inequality with weak measurements of photons. *Proceedings of the National Academy of Sciences*, 108(4):1256–1261, 2011.

[170] Justin Dressel and Alexander N. Korotkov. Avoiding loopholes with hybrid Bell-Leggett-Garg inequalities. *Phys. Rev. A*, 89:012125, Jan 2014.

[171] Theodore C White, JY Mutus, Justin Dressel, J Kelly, R Barends, E Jeffrey, D Sank, A Megrant, B Campbell, Yu Chen, et al. Preserving entanglement during weak measurement demonstrated with a violation of the Bell–Leggett–Garg inequality. *npj Quantum Information*, 2(1):1–5, 2016.

[172] Wei-Wei Pan, Xiao-Ye Xu, Yaron Kedem, Qin-Qin Wang, Zhe Chen, Munsif Jan, Kai Sun, Jin-Shi Xu, Yong-Jian Han, Chuan-Feng Li, and Guang-Can Guo. Direct Measurement of a Nonlocal Entangled Quantum State. *Phys. Rev. Lett.*, 123:150402, Oct 2019.

[173] Ming-Cheng Chen, Yuan Li, Run-Ze Liu, Dian Wu, Zu-En Su, Xi-Lin Wang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan. Directly Measuring a Multiparticle Quantum Wave Function via Quantum Teleportation. *Phys. Rev. Lett.*, 127:030402, Jul 2021.

[174] Mohamed Bourennane, Manfred Eibl, Christian Kurtsiefer, Sascha Gaertner, Harald Weinfurter, Otfried Gühne, Philipp Hyllus, Dagmar Bruß, Maciej Lewenstein, and Anna Sanpera. Experimental Detection of Multipartite Entanglement using Witness Operators. *Phys. Rev. Lett.*, 92:087902, Feb 2004.

[175] Nicolai Friis, Giuseppe Vitagliano, Mehul Malik, and Marcus Huber. Entanglement certification from theory to experiment. *Nature Reviews Physics*, 1(1):72–87, 2019.

[176] Klaus Floettmann. Coherent superposition of orthogonal Hermite–Gauss modes. *Optics Communications*, 505:127537, 2022.