Far-field linear optical super-resolution imaging via Hermite-Gaussian microscopy.





Anastasiia Pushkina

Somerville College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2021

Abstract

The resolution of optical imaging devices is ultimately limited by the diffraction of light. To circumvent this limit, modern super-resolution microscopy techniques employ active interaction with the object by exploiting its optical nonlinearities, nonclassical properties of the illumination beam, or near-field probing. Thus, they are not applicable in areas where interaction is not possible, for example, in astronomy or non-invasive biological imaging. Far-field and linear-optical super-resolution technique based on passive analysis of light coming from the object would cover these gaps. This thesis presents the first proof-of-principle demonstration of such a technique for general objects. It works by accessing information about spatial correlations of the image optical field and, hence, about the object itself via measuring projections into Hermite-Gaussian transverse spatial modes. In theory, the resolution of the technique scales as the inverse square root of the number of modes in each transverse dimension. With a basis of 21 spatial modes in both dimensions ($21^2 = 441$ modes), this experimental work demonstrates two-dimensional imaging with twofold resolution enhancement beyond the diffraction limit.

Statement of Originality

I hereby declare that this thesis is my own work. I was the one responsible for the experimental realization of the project. I have arranged the laboratory and built the experimental setup. All measurements were performed by myself. Besides, I had a significant contribution to numerical simulations and data processing.

Acknowledgements

I am genuinely grateful to my supervisor Dr. Alexander Lvovsky for all chances and opportunities given to me during our collaboration. I enjoyed working on such an interesting and promising project, which, I believe, opens the path to my future career. I appreciate all the support and guidance during my work.

I am also grateful to my colleagues Jose Da Costa Filho and Giorgio Maltese. Jose provided helpful insights on the theoretical aspects of the project. Giorgio significantly facilitated my life in the lab. On his example, I have learnt the attitude and approach to follow in experimental work.

Contents

List of Figures viii

List of Abbreviations xi

1 Introduction 1
- 1.1 Classical imaging and resolution limit . . . 1
- 1.2 Super-resolution imaging . . . 6
- 1.3 Hermite-Gaussian microscopy . . . 8
- 1.4 Review of thesis chapters . . . 10

2 Comprehensive model and performance optimization of phase-only spatial light modulators 13
- 2.1 Operation and exploitation principles . . . 14
  - 2.1.1 Phase modulation . . . 14
  - 2.1.2 Simultaneous phase and amplitude modulation . . . 16
  - 2.1.3 Polarization modulation . . . 20
  - 2.1.4 Applications . . . 20
- 2.2 Imperfections of phase-only SLMs . . . 25
  - 2.2.1 Curvature of the back panel . . . 25
  - 2.2.2 Inner Fabry-Perot cavity . . . 26
  - 2.2.3 Pixel crosstalk . . . 28
  - 2.2.4 Non-linear and non-uniform response of liquid crystal to the grey level . . . 31
  - 2.2.5 Noise of the electric drive circuit . . . 32
- 2.3 Comprehensive characterization and compensation of the SLM imperfections . . . 33
  - 2.3.1 Experiment . . . 33
  - 2.3.2 Theoretical model . . . 35
  - 2.3.3 Fitting the response curves . . . 37
  - 2.3.4 Compensation of the spurious effects . . . 41
  - 2.3.5 Fidelity sensitivity to the spurious effects . . . 44

2.3.6 Compensation for the non-uniform profile of the illumination beam . . . 46
2.3.7 Conclusion . . . 46

3 Super-resolution imaging: a review **48**
3.1 Near-field techniques . . . 49
3.1.1 Tip-enhanced microscopy . . . 50
3.1.2 Aperture-SNOM . . . 53
3.1.3 Photon scanning tunnelling microscopy . . . 55
3.1.4 Feedback mechanisms . . . 56
3.1.5 Near-field optical random mapping microscopy . . . 58
3.2 Far-field techniques . . . 59
3.2.1 Image scanning microscopy . . . 59
3.2.2 Stimulated-emission-depletion microscopy . . . 62
3.2.3 Structured illumination microscopy . . . 65
3.2.4 Localization microscopy . . . 67
3.2.5 Rotating coherent scattering microscopy . . . 69

4 Hermite-Gaussian microscopy: background **72**
4.1 Introduction to the Fisher information . . . 73
4.1.1 Classical Fisher information . . . 73
4.1.2 Quantum Fisher information . . . 74
4.2 Super-resolution for two incoherent point sources . . . 74
4.2.1 Estimation of the separation . . . 74
4.2.2 Spatial-mode demultiplexing . . . 77
4.3 Implementation options of spatial-mode demultiplexing . . . 80
4.3.1 Mode-sorting . . . 80
4.3.2 Heterodyne detection . . . 86
4.4 Theory of Hermite-Gaussian microscopy . . . 91
4.4.1 Coherent case . . . 92
4.4.2 Incoherent case . . . 95
4.5 Review of experimental works . . . 96

5 Hermite-Gaussian microscopy: experiment **101**
5.1 Data to be acquired . . . 101
5.2 Experimental setup . . . 102
5.2.1 General description . . . 102
5.2.2 Laser . . . 104
5.2.3 Digital micromirror device . . . 105
5.2.4 Acousto-optic modulator . . . 106

5.2.5 Imaging system in the signal path . . . 107
5.2.6 Spatial light modulator . . . 109
5.2.7 Electronic system . . . 110
5.3 Setup alignment and data acquisition . . . 112
5.3.1 Optical setup for mode matching . . . 112
5.3.2 Mode matching procedure . . . 112
5.3.3 Automatic alignment procedure of the interferometer . . . 113
5.3.4 DMD datasets . . . 114
5.3.5 Acquisition of direct images . . . 117
5.3.6 Acquisition of photocurrents . . . 118
5.3.7 Noise and setup misalignment . . . 120
5.4 Data processing . . . 122
5.4.1 Pre-processing . . . 122
5.4.2 Neural network . . . 124
5.5 Results . . . 128
5.5.1 Comparison with direct imaging . . . 128
5.5.2 Comparison with deconvolution of direct images . . . 131
5.5.3 Alternative resolution improvement benchmark . . . 132

6 Conclusion 134
6.1 Summary . . . 134
6.2 Discussion. . . . 136
6.3 Future outlook . . . 137

Appendices

A Derivation of the quantum Fisher information for separation estimation 141

B Derivation of the maximum-likelihood estimator for the separation 144

C An optical Eratosthenes' sieve for large prime numbers 145
C.1 Motivation . . . 145
C.2 Theory . . . 146
C.2.1 Sieving mask. . . . 146
C.2.2 Displacement grating. . . . 149
C.3 Experiment . . . 150
C.4 Discussion . . . 155

List of Figures

1.1 Airy pattern 1
1.2 Resolution limits 3

2.1 LC layer of SLMs 14
2.2 Example of a hologram for simultaneous phase and amplitude encoding 18
2.3 Applications of SLMs in structured light. 21
2.4 Schematic cross-section of a phase-only SLM 25
2.5 Crosstalk effect 29
2.6 Field inversion schemes of the pixel voltage 32
2.7 Setup scheme for the collection of the intensity response curves. . . 34
2.8 Near-field images of the first and second diffraction orders. 35
2.9 Model of inner cavities of an SLM 36
2.10 Phase maps of the cavity thickness obtained from the fit and a wavefront camera 39
2.11 Comparison of the fit quality for the different estimation models. . . 40
2.12 Algorithm used to calculate an SLM hologram compensating for crosstalk. 41
2.13 Near- and far-field profiles of the $HG_{12,12}$ produced by the SLM addressed with non-, partially- and fully-corrected holograms 43
2.14 Fidelity sensitivity to the spurious effects 44
2.15 Simulated power efficiency as a function of the crosstalk point spread function width with and without the correction. 45

3.1 Tip-enhanced scattering microscopy. 50
3.2 Aperture near-field scanning microscopy. 53
3.3 Photon scanning tunnelling microscopy. 55
3.4 Feedback mechanisms of SNOM. 57
3.5 Near-field optical random mapping microscopy. 58
3.6 Image scanning microscopy. 60
3.7 Stimulated-emission-depletion microscopy. 63
3.8 Diagram explaining the Moirè pattern generation and structured illumination microscopy concept. 65

3.9 Diagram explaining the resolution enhancement of structured illumination microscopy. 65
3.10 Schematic of ROCS microscopy 69

4.1 Fisher information per photon and Cramer-Rao bound versus the separation between the two sources. 75
4.2 Illustration of a basic idea of spatial-mode demultiplexing. 77
4.3 Per-photon Fisher information of the SPADE for two partially coherent point sources. 80
4.4 Schematic of heterodyne detection for separation estimation. 89
4.5 Resolution enhancement of one-dimensional HGM beyond the Rayleigh limit versus the number of Hermite-Gaussian modes and absolute value of the electric field of two-dimensional Hermite-Gaussian modes. 94

5.1 Simplified schematic of the experimental setup for HGM. 103
5.2 Digital micromirror device 105
5.3 Schematic of an acousto-optic modulator. 106
5.4 Iris. 108
5.5 Schematic of the imaging system in the signal path. 108
5.6 Examples of SLM holograms used to produce Hermite-Gaussian modes 109
5.7 Electronics setup. 110
5.8 Test dataset for the neural network. 115
5.9 Train dataset for the neural network. 116
5.10 Transverse scanning of the University of Oxford logo for direct imaging 117
5.11 Data acquisition timing diagram. 118
5.12 Example of an oscilloscope trace saved during data acquisition. . . . 119
5.13 Photocurrents noise 121
5.14 Sketch of the data processing using the neural network. 125
5.15 Original, HGM-reconstructed and camera images of the University of Oxford logo and coat-of-arms test sets. 128
5.16 Original, HGM-reconstructed and camera images of the alphabet test set. 128
5.17 Quantitative assessment of the HGM resolution enhancement with respect to direct imaging. 129
5.18 Richardson-Lucy deconvolution applied to direct images of University of Oxford logo and coat of arms and alphabet datasets using the Airy PSF. 131
5.19 Resolution improvement evaluated through the MSE loss. 132

C.1 Light intensity profiles at near and far fields for an infinite grating of pinholes illuminated by a plane monochromatic wave over a unit length window 147
C.2 Optical sieve diffraction masks and the square root of the corresponding far field intensity profiles 148
C.3 Experimental implementation of the optical prime sieve 150
C.4 Experimental results for the optical prime sieve 151
C.5 Effect of the SLM's spatial resolution on the produced far field intensity profile 152
C.6 Number of misclassified primes and composites within the measured sieving range $139 < q_x < 22201$ as a function of the intensity threshold 154

List of Abbreviations

AOM . . . . . Acousto-optic modulator

AMC . . . . . Astigmatic mode converter

BS . . . . . Beamsplitter

CLSM . . . . . Confocal laser-scanning microscopy

DDS . . . . . Direct Digital Synthesizer

DI . . . . . Direct imaging

DMD . . . . . Digital micromirror device

FPALM . . . . Fluorescence photoactivation localization microscopy

FRFT . . . . . Fractional Fourier transform

FWM . . . . . Four-wave mixing

GL . . . . . Grey level

HD . . . . . Heterodyne detector

HGM . . . . . Hermite-Gaussian microscopy

HWP . . . . . Half-wave plate

ISM . . . . . Image scanning microscopy

LC . . . . . Liquid crystal

LCoS . . . . . Liquid crystal on silicon

LO . . . . . Local oscillator

LUT . . . . . Look-up table

MPLC . . . . . Multi-Plane Light Conversion

MSE . . . . . Mean squared error

NA . . . . . Numerical aperture

OAM . . . . . Orbital angular momentum

PALM . . . . . Photoactivated localization microscopy

PBS . . . . . Polarizing beamsplitter

PD . . . . . Photodiode

PSF . . . . . Point spread function

PSTM . . . . . Photon scanning tunnelling microscopy

QCRB . . . . . Quantum Cramer-Rao bound

QISM . . . . . Quantum image scanning microscopy

QKD . . . . . Quantum key distribution

RL . . . . . Richardson-Lucy

ROCS . . . . . Rotating coherent scattering microscopy

SHG . . . . . Second harmonic generation

SIM . . . . . Structured illumination microscopy

SLM . . . . . Spatial light modulator

SNOM . . . . Scanning near-field optical microscopy

SPADE . . . . Spatial-mode demultiplexing

SSIM . . . . . Saturated structured illumination microscopy

STED . . . . . Stimulated emission depletion microscopy

STORM . . . . Stochastic optical reconstruction microscopy

WSS . . . . . Wavelength selective switch

1 Introduction

1.1 Classical imaging and resolution limit

Perhaps, the most important component of a conventional imaging system such as a microscope or telescope is the objective lens. The objective lens collects light emitted by an object and defines the resolution of the imaging system, i.e the ability to distinguish fine features of the object. Due to the wave nature of light, resolution of any finite objective lens is limited by diffraction. Let us consider classical approaches to image formation.

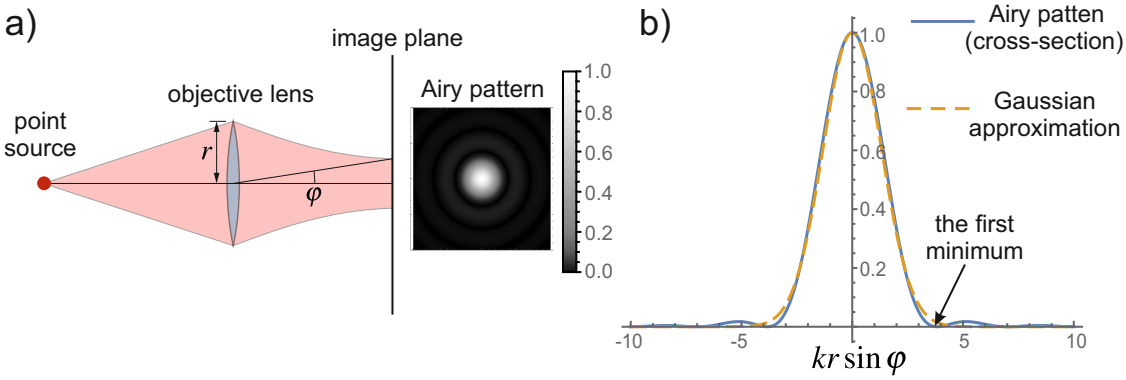


Figure 1.1: a) An image of a point source created by an objective lens of finite diameter is the Airy pattern. The grayscale of the Airy pattern is adjusted to provide contrast for the rings. b) A radial cross-section of the Airy pattern together with the Gaussian profile approximation.

Point spread function. As commonly accepted, we start the description of classical imaging with the simplest case when the object is a point source. Due to Fraunhofer diffraction, an image of the point source formed by an ideal lens with a circular aperture is the Airy pattern (the Airy disc) [2] with the intensity

distribution given by

$$I(\varphi) = I_0 \left[\frac{2J_1(kr \sin \varphi)}{kr \sin \varphi} \right]^2 \equiv I_0 \left[\frac{2J_1(q)}{q} \right]^2, \tag{1.1}$$

where I_0 defines the peak intensity, $J_1(q \equiv kr \sin \varphi)$ is the Bessel function of the first kind of order one, r is the radius of the lens, $k = \frac{2\pi}{\lambda}$ is the wavenumber, and φ is the observation angle. Fig. 1.1 illustrates the imaging process and the Airy pattern.

By definition, a function that describes the response of an imaging system to a point source is called a *point spread function* (PSF). Thus, the Airy disc represents the PSF of an objective lens with a circular aperture. The first minimum (the first zero) of the Airy disc occurs at

$$kr \sin \varphi \approx 3.83. \tag{1.2}$$

Hence, for small angles ($\sin \varphi \approx \varphi$), we obtain that the angle of first minimum is $\varphi \approx 3.83/(kr)$. We refer to the distance between the maximum intensity and the first minimum as the Airy disc radius.

It is common in imaging to approximate the Airy disc with a Gaussian profile to simplify calculations:

$$I(q) \approx I_0 \exp\left(-\frac{q^2}{2\sigma^2} \right). \tag{1.3}$$

The optimal Gaussian waist to fit the central lobe is approximately 0.21/0.61 of the Airy disc radius. Fig. 1.1(b) shows a radial cross-section of the Airy disc together with the optimal Gaussian fit ($\sigma \approx (0.21/0.61) \cdot 3.83$ in units of $q \equiv kr \sin \varphi$).

Angular and spatial resolution limits. Lord Rayleigh defined the resolution of an imaging system as the ability to distinguish two point sources separated by angular distance ϑ (see Fig. 1.2(a)) and established the angular resolution limit ϑ_R as follows. Two incoherent point sources of equal strength are barely resolved when the maximum of one Airy disc coincides with the first minimum of the other. In other words, the minimum resolvable angular distance is equal to the observation angle of first minimum of the Airy pattern PSF:

$$\vartheta_R \approx \frac{3.83}{kr} = 1.22 \frac{\lambda}{D}, \tag{1.4}$$

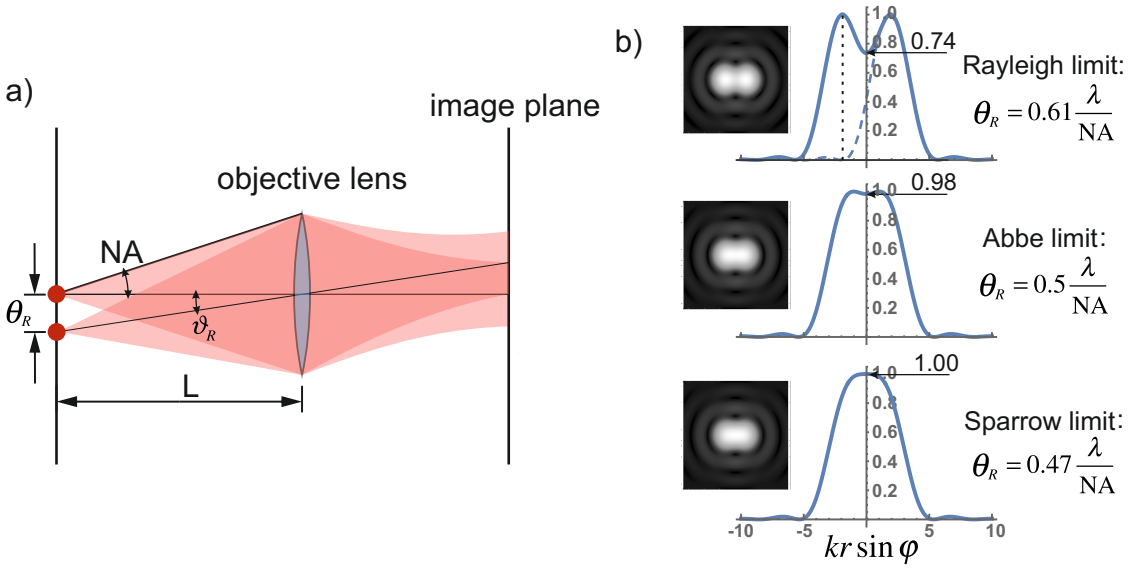


Figure 1.2: a) Imaging of two incoherent point sources. b) Illustration of heuristically defined resolution limits. Here, the horizontal axis of the intensity cross-section plots is shifted such that the curve is symmetric about zero.

where $D = 2r$ is the diameter of the lens. The visualization of the Rayleigh limit is shown on the top of Fig. 1.2(b).

The angular resolution can be converted into the spatial resolution θ_R. The minimum resolvable spatial separation between two incoherent sources is

$$\theta_R \approx L\vartheta_R = 1.22\frac{\lambda L}{2r} = 0.61\frac{\lambda}{\mathrm{NA}}, \tag{1.5}$$

where L is the distance between the lens and the object plane and $\mathrm{NA} \approx r/L$ is the numerical aperture of the imaging system (i.e the half-range of acceptance angles). The given expression for the NA is valid in the limit of $r \ll L$, otherwise the NA is defined as the sine of the maximum half-angle entering the lens.

The waist of the Gaussian approximation of the PSF can be expressed through the NA as follows. The Airy disc radius in the image plane is related to the resolution limit as $M\theta_R$, where M is the magnification of the imaging system. Let us assume unity magnification. Then, if the Gaussian approximation profile is written in spatial coordinates of the image plane instead of the angular coordinate q, we obtain that the waist of the Gaussian profile is $\sigma = (0.21/0.61)\cdot\theta_R \approx 0.21\lambda/\mathrm{NA}$.

Another way to see the Rayleigh resolution limit is by the ratio between the central intensity gap and the maximum intensity in the image plane. For incoherent

light, the Rayleigh limit expressed by Eq. (1.4) and Eq. (1.5) is equivalent to the statement that two point sources are barely resolved when the gap is 74% of the maximum. Defined via the gap relative intensity, the resolution limit becomes more general and can be applied to coherent light. For two coherent point sources in phase, the minimum resolvable separation, i.e the separation that results in the intensity gap of 74% in the image plane, is $0.82\lambda/\mathrm{NA}$. This resolution limit is obtained considering the interference of two Airy patterns in the image plane.

Fig. 1.2(b) compares the Rayleigh resolution limit with the Abbe and Sparrow limits, which have the gap of 98% ($0.5\lambda/\mathrm{NA}$ for incoherent light) and 100% ($0.47\lambda/\mathrm{NA}$ for incoherent light) respectively. As seen, the definition of resolution limits is heuristic, and the field of study dictates the choice of one. In astronomy, it is more common to use the Sparrow limit, while in microscopy, people apply the Rayleigh or Abbe limit.

Lateral and axial resolution. In the literature, terms 'lateral' and 'axial' resolution appear when considering 3D imaging. The lateral resolution stands for the ability to distinguish two point sources located in the plane, which is orthogonal to the optical axis of an imaging system. In other words, the lateral resolution is the resolution defined in the previous paragraph. The axial resolution refers to the ability of an imaging system, which is designed to produce 3D images, to distinguish two point sources separated along the optical axis or along a line parallel to the optical axis.

When a point source is imaged by an objective lens, the intensity distribution along the optical axis in the vicinity of the image plane is

$$I(\Delta z) = I_0 \left[\frac{\sin(\mu \Delta z)}{\mu \Delta z} \right]^2, \tag{1.6}$$

where

$$\mu = \frac{\pi \mathrm{NA}^2}{2\lambda} \tag{1.7}$$

and Δz is the distance from the image plane along the optical axis. The first minimum of this intensity distribution occurs at $\mu \Delta z' = \pi$, which results in the

following axial Rayleigh limit for incoherent light:

$$\theta_R^{Axial} = \Delta z' = \frac{2\lambda}{\mathrm{NA}^2}. \tag{1.8}$$

Here, the magnification of the imaging system is assumed to be unity.

In most cases in this thesis, we consider imaging of one-dimensional or two-dimensional objects with no extent along the optical axis. Thus, if the type of resolution is not specified, one should assume the lateral resolution.

Abbe theory of image formation. Ernst Abbe considered imaging of amplitude gratings [1]. The light diffracted by an object enters the imaging system, and the interference of diffraction orders forms an image. If the objective lens has infinite numerical aperture, no diffraction orders are lost, and the image precisely corresponds to the object. However, if the numerical aperture is finite, then it cuts high diffraction orders, which leads to image blurring.

The mathematical description of the Abbe theory is given by Fourier optics. If the electric field distribution in the object plane is $E(x)$, the electric field in the objective lens plane, located in the far field at distance L from the object, is given by the Fourier transform

$$E'(k_x) \propto \int_{-\infty}^{\infty} E(x) e^{-ik_x x} dx, \tag{1.9}$$

where k_x is the orthogonal component of the wave vector. The lens transmits spatial frequencies according to its optical transfer function $\psi(k_x)$. In the image plane, the electric field corresponds to the inverse Fourier transform of the field transmitted through the lens:

$$\begin{aligned} E_{img}(x') &\propto \int_{-\infty}^{\infty} E'(k_x)\psi(k_x) e^{ik_x x'} dk_x \\ &= \iint_{-\infty}^{\infty} E(x)\psi(k_x) e^{ik_x(x'-x)} dx dk_x. \end{aligned} \tag{1.10}$$

Let us denote the inverse Fourier transform of the transfer function as

$$\psi(x'-x) = \int_{-\infty}^{\infty} \psi(k_x) e^{ik_x(x'-x)} dk_x. \tag{1.11}$$

Combining Eq. (1.10) and Eq. (1.11), we obtain the following expression for the image field:

$$E_{img}(x') \propto \int_{-\infty}^{\infty} E(x)\psi(x'-x)dx = E(x) * \psi(x), \tag{1.12}$$

where $*$ denotes the convolution operation.

By definition, the inverse Fourier transform of the optical transfer function is the point spread function. Thus, $\psi(x)$ in Eq. (1.12) is the electric field PSF of the imaging system. In other words, blurring of the image can be modeled as a convolution of the object and the electric field PSF. This conclusion can be extended to objects of arbitrary shape.

Conclusion. Ernst Abbe and Lord Rayleigh published their research on optical imaging at the end of the 19th century [1, 128]. Since that, the diffraction resolution limit has been an ultimate bound for optical resolution until the 1980s. In absolute numbers, the spatial resolution of conventional microscopy for visible light is limited to ∼200 nm, since the lower bound of the visible wavelength range is about 380 nm, and the NA of dry objectives is 0.95 at most (1.45 for oil immersion objectives). In astronomy, attempts to enhance the angular resolution led to the effort to manufacture objective lenses of increasing diameters and cost.

However, the derivation of the diffraction limit is based on several assumptions. The diffraction resolution limit is valid for incoherent classical sources, linear imaging, and measurement of intensity in the far-field. Thus, the resolution of optical imaging can be enhanced beyond the diffraction limit by breaking at least one of these assumptions. Imaging techniques that achieve sub-diffraction (sub-Rayleigh) resolution define the domain of super-resolution microscopy.

1.2 Super-resolution imaging

Super-resolution imaging has emerged in the 1980s. All super-resolution techniques can be divided into two large categories: near-field and far-field techniques. Chapter 3 of the thesis provides the detailed description of operation principles

of various techniques from both categories. Here, we briefly review capabilities and limitations of the existing methods.

Near-field techniques achieve super-resolution by probing the evanescent field in the vicinity of the sample. The evanescent field is not affected by diffraction and is highly localized around fine features of the sample. The first idea of near-field microscope was proposed by Synge [147] in 1928. However, the first implementations [123, 21] were demonstrated only in the 1980s when technologies achieved the required level. Nowadays, near-field microscopy exists in various configurations and reaches the lateral resolution of up to 10 nm and beyond. The resolution is mainly limited by the sharpness of the probe.

Near-field microscopy is typically used for analysis of surface properties of solid samples. For instance, it provides access to the spatial distribution of optical refractive index and absorption coefficient [63], chemical composition analysis [151], and detection of physical and chemical defects [31]. The application area of near-field microscopy is limited to studies of surfaces or shallow subsurface regions due to the exponential decay of the evanescent field with the distance from the source. Another limiting factor is the complexity and long scanning time needed to obtain an image.

Requirements of biological studies in imaging of structures inside live cells in real time stimulated the development of far-field super-resolution techniques. The biology-inspired approach to super-resolution imaging is based on the widely used method of fluorescent labelling, which means that bright fluorophores can be selectively attached to cellular components such as proteins, peptides, antibodies, and nucleic acids using chemical methods. Utilizing nonlinear interactions between fluorophores and laser illumination, researchers can image cellular components, which fluorophores are bound to, with sub-Rayleigh resolution.

Due to the three-dimensional structure of biological samples, 3D imaging is popular in fluorescence microscopy. Various approaches to the axial and lateral resolution enhancement have been proposed since experimental demonstrations of fluorescence-based far-field super-resolution techniques have emerged in the early 1990s. The first techniques included standing-wave fluorescence microscopy [10],

image interference microscopy (I^nM) [67], and 4Pi microscopy [71]. The aforementioned techniques enhance the axial resolution up to seven times beyond the Rayleigh limit, however, do not provide noticeable improvement to the lateral resolution.

The breakthrough in far-field super-resolution imaging is associated with lateral techniques, which have appeared shortly after the axial ones. They are represented by confocal image scanning microscopy [112], structured illumination microscopy [66, 65], stimulated emission depletion microscopy [72] marked by the 2014 Nobel Prize in Chemistry, and single-molecule localization microscopy [20, 133, 75]. The lateral resolution ranges from 10 to 100 nm depending on the method.

Fluorescence microscopy provides tools for 3D imaging of biological samples with sub-Rayleigh resolution. However, this approach to far-field super-resolution imaging remains limited to fluorescent probes. The operation of far-field super-resolution microscopes is typically more complex and costly than conventional ones. Besides, labelling and excitation of fluorophores require interaction with the sample.

Conclusion. Existing super-resolution techniques are limited in applications and require active interaction with the sample. Thus, a universal far-field and linear-optical super-resolution technique would benefit scientific areas where super-resolution is not yet available, for example, astronomy and non-invasive biological imaging. Here, non-invasive imaging means imaging techniques that do not require labelling of samples with fluorescent probes or illuminating them with powerful and, hence, damaging laser beams to stimulate non-linear behaviour.

1.3 Hermite-Gaussian microscopy

In 2015, Tsang *et al.* [162] derived the fundamental resolution limit in measuring the separation between two weak incoherent point sources in the form of the quantum Cramér-Rao bound (QCRB), which establishes the ultimate resolution limit regardless of the measurement. Surprisingly, QCRB does not show the effect of the Rayleigh limit and remains constant versus separation. In other words, it means that such measurement exists that can distinguish arbitrary close sources equally well.

In the same paper, in addition to the derivation of QCRB, Tsang *et al.* proposed a measurement technique dubbed spatial-mode demultiplexing (SPADE) that achieves the fundamental resolution limit in separation estimation. SPADE is the decomposition of the optical field in the image plane into the Hermite-Gaussian modes basis with measurement of the coefficients. Intuitive explanation of SPADE for separation estimation is as follows. Let us assume a Gaussian PSF. A point source on the optical axis emits photons into the fundamental mode only, while a displaced source in the lateral direction has a probability to emit a photon into a higher order mode. Thus, the photon counting in each mode of the basis provides separation estimation with quantum optimal precision.

Shortly after Tsang's paper, our group has proposed an extension of SPADE to full imaging [175]. We named this extension Hermite-Gaussian microscopy (HGM). The idea beyond HGM is as follows. Projections into different modes access information about spatial correlations of the image optical field and, hence, about the object itself. This information is lost in conventional imaging. Theoretically, the resolution of HGM scales approximately as the inverse square root of the highest mode order in the basis. This follows from the inverse square root scaling of the features size of modes with their order.

Being linear-optical and operational in the far-field, SPADE and HGM are promising super-resolution techniques. However, existing proof-of-principle experimental studies, reviewed in detail in section 4.5, focus on separation estimation only.

This thesis presents the first experimental demonstration of full imaging using HGM [124]. This proof-of-principle experiment was performed with some simplifications. Namely, we used coherent macroscopic sources and an objective lens with low NA to simulate the sub-Rayleigh regime. Another simplification is that, instead of technically challenging demultiplexing into the Hermite-Gaussian basis, we performed heterodyne detection with a local oscillator prepared in each mode of the basis. An important feature of our experiment is a supervised neural network, which was trained to filter aberrations of the optical setup and other

systematic errors, as well as noise. Using a basis of 441 two-dimensional Hermite-Gaussian modes (21 modes in both vertical and horizontal directions), we achieved approximately two-fold resolution enhancement beyond the Rayleigh limit, which is close to the theoretical prediction.

It is fair to mention that rotating coherent scattering (ROCS) microscopy (see subsection 3.2.5) also can achieve two-fold resolution enhancement beyond the Rayleigh limit without relying on fluorescent labelling or near-field probing. However, such a result (the two-fold resolution enhancement) is the theoretical limit for ROCS microscopy, while the resolution of HGM is expected to improve further with the number of Hermite-Gaussian modes used for decomposition of the image field. As other existing super-resolution techniques, ROCS is a case-specific technique because it works only for objects under oblique coherent illumination. Although we demonstrated HGM for a certain scenario, namely, for imaging of coherent objects with a low-NA objective lens, our experiment is the first step towards implementing a universal super-resolution imaging technique, which would address imaging scenarios from microscopy to astronomy. Besides coherent objects, the setup presented in this work is suitable for imaging of quasi-monochromatic incoherent objects. In the future, we plan to demonstrate that our setup is compatible with high-NA microscope objectives. The final step will be to adapt our setup for imaging of broadband incoherent objects by implementing mode-sorting (see subsection 4.3.1).

1.4 Review of thesis chapters

Chapter 2 of the thesis presents a project, which has been done as a run-up for the HGM experiment. To produce Hermite-Gaussian modes for imaging, we chose to use a phase-only spatial light modulator (SLM). However, commercial SLMs commonly have various imperfections. To approach this issue, we have developed a novel calibration and performance optimization method of SLMs [125]. Chapter 2 describes in detail operation principles of SLMs and their imperfections, reviews existing compensation procedures and demonstrates the advantage of our procedure on the example of a high-order Hermite-Gaussian mode.

Chapter 3 reviews the existing super-resolution techniques both in the near-field and far-field. It focuses on the operating principles of techniques and compares different techniques in terms of resolution, application areas, and limitations. This chapter is necessary to define the place of HGM in the broad field of super-resolution imaging.

Chapter 4 starts with introducing concepts of the Fisher information and Cramér-Rao bound and their quantum counterparts. Then, it provides the derivation of QCRB for separation estimation and proves that SPADE achieves it. Chapter 4 compares two approaches to measuring coefficients of the decomposition into Hermite-Gaussian modes, namely mode-sorting and heterodyne detection, and justifies the choice of heterodyne detection in our experiment. Then, the chapter introduces Hermite-Gaussian microscopy by demonstrating how two-dimensional images can be reconstructed from heterodyne detector photocurrents and provides theoretically expected resolution enhancement versus the number of Hermite-Gaussian modes. The chapter ends with a review of experiments inspired by the theory of SPADE to show that, unlike the existing experiments focusing on separation estimation only, our work is the first experimental demonstration of full imaging with sub-Rayleigh resolution based on the SPADE approach.

Chapter 5 reports the original work on the experimental realization of HGM. It provides a comprehensive description of the experimental setup and data acquisition. The chapter describes the architecture and training of the neural network, which is dedicated to filter systematic errors from raw data and reconstruct HGM images in the presence of noise. Then, chapter 5 demonstrates our experimental results represented by HGM reconstructions of the University of Oxford logo and coat of arms and alphabet symbols. HGM reconstructions are compared with direct images, and the advantage of HGM is demonstrated. In this chapter, we experimentally estimate the resolution enhancement of HGM over direct imaging and investigate how it scales with the number of Hermite-Gaussian modes. Besides direct imaging, we compare HGM with the commonly used deconvolution algorithm and show that this comparison is in favour of HGM.

Chapter 6 provides a summary of our work and future outlook. It emphasizes the limitations of our current setup and discusses what changes have to be made for further development of HGM towards practical microscopy.

2 | Comprehensive model and performance optimization of phase-only spatial light modulators

We chose the preparation of high-quality Hermite-Gaussian modes as the very first task of our super-resolution imaging experiment for the following reason. The performance of Hermite-Gaussian microscopy was expected to be quite sensitive to systematic errors in higher-order Hermite-Gaussian modes because the impact of a mode on the final image is amplified via the factorial of its order (see theory in subsection 4.4.1 and Eq. (4.56) particularly). We used a phase-only spatial light modulator (SLM) to generate Hermite-Gaussian modes and noticed distortions in the produced optical fields, which triggered a comprehensive investigation of the SLM imperfections. Although we later came up with the idea of relying on a neural network to filter systematic errors, we decided to finish our study of the SLM spurious effects. This chapter presents our work and explains our motivation.

Phase-only liquid crystal on silicon spatial light modulators (LCoS SLMs) have been established as the most popular tool to generate spatially structured light beams by reshaping the wavefront of an incoming beam. By making use of high-yield CMOS backplanes (pixel pitch $\sim 10\mu$m, fill factor up to 98%), high diffraction efficiencies can be achieved, while the large electro-optic coefficients of liquid crystal materials enable significant modulation depths (up to several wavelengths) and real-time operation (millisecond response time) [182].

An ideal phase-only SLM should produce a predictable, linear and uniform phase response to the applied control voltage matrix. However, a few imperfections are

known to deteriorate SLM performance. We study for the first time the joint effect of the major ones: the curvature of the back panel, pixel crosstalk and the internal Fabry-Perot cavity. In addition, we introduce an effective compensation method. As a result, we generate high-order Hermite-Gaussian modes ($HG_{m,n}$) with improved fidelity and efficiency with respect to the currently adopted techniques.

2.1 Operation and exploitation principles

2.1.1 Phase modulation

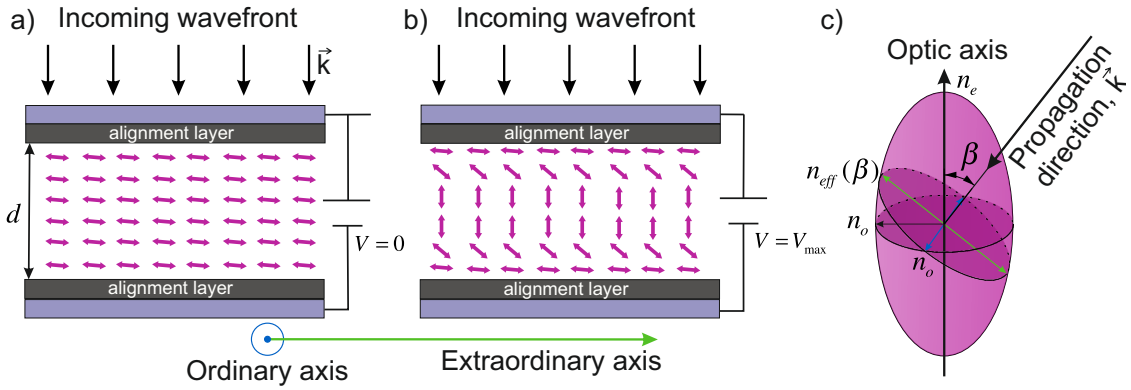


Figure 2.1: a), b) Orientation of LC molecules in dependence on the applied voltage. Magenta arrows represent the molecules' director (optic axis). The director gradient in the proximity of the alignment layers in figure (b) has a negligible impact on the overage orientation of molecules. The ordinary axis of the SLM is the axis that remains orthogonal to the molecules' director for any voltage. The extraordinary axis of the SLM is the axis orthogonal to the ordinary one. c) Refractive index ellipsoid, also called the optical indicatrix, which represents the refractive indices of the LC layer for light following different directions with respect to the optic axis. The blue and green arrows stand for the polarization states of the ordinary and extraordinary waves, respectively.

A layer of nematic liquid crystal (LC) possessing the electrically controlled birefringence provides modulating properties of phase-only SLMs. A voltage V applied to the LC layer defines a tilt angle of the molecules' director, which coincides with the optic axis. The voltage takes discrete values in a given range. Fig. 2.1(a) and (b) show a schematic of the LC layer for the utmost voltage values.

The LC layer is sandwiched between two alignment layers, the purpose of which is to define the initial orientation of LC molecules. The alignment layers keep molecules, which are adjacent to them, parallel to a certain direction and can be

represented, for example, by polymer films buffed along this direction [33]. To ensure that all molecules tilt uniformly (for example, all tilt clockwise) when a voltage is applied, the alignment layers provide the molecules with a small pre-tilt angle.

Controlling the tilt angle of molecules with voltage results in controlling the angle $\beta = \beta(V)$ between the propagation direction of the incident electromagnetic wave and the optic axis of the LC layer. The refractive index seen by the wave depends both on the polarization state of the wave and angle β, as illustrated in Fig. 2.1(c). If the polarization is orthogonal to the optic axis, then the wave experiences the ordinary refractive index n_o regardless of β. Such a wave is called an *ordinary wave*, and its polarization direction is indicated by the blue arrow in Fig. 2.1(c). An electromagnetic wave in the polarization state orthogonal to that of the ordinary wave (the green arrow) is called an *extraordinary wave*. It is governed by an *effective extraordinary refractive index* n_{eff}, which takes values between the ordinary n_o and extraordinary n_e refractive indices depending on β:

$$n_{eff}(\beta) = \frac{n_o n_e}{\sqrt{n_o^2 \sin^2(\beta) + n_e^2 \cos^2(\beta)}}. \tag{2.1}$$

Thus, the voltage-controlled LC layer can be seen as a variable waveplate. Unitary vectors of the polarization directions of the ordinary and extraordinary waves define the ordinary (fixed refractive index) and extraordinary (variable refractive index) axes of this waveplate, which are shown in Fig. 2.1(a)-(b). The phase retardation between the ordinary and extraordinary axes is given by

$$\Delta\Theta(V) = \frac{2\pi}{\lambda} L(n_{eff}(\beta(V)) - n_o), \tag{2.2}$$

where λ is the wavelength of the incident light and L is the length of the light path in the LC layer, which equals to the LC layer thickness d for a transmissive SLM and $2d$ for a reflective one.

To summarize, applying a voltage to the LC layer results in changing the refractive index and, hence, the optical path but only for the extraordinary wave. The SLM display accommodates an array of pixels. Each pixel applies a specific voltage across the LC layer defined via a computer-generated matrix, which contains

8-bit (from 0 to 255) numbers (10-bit for some SLM models) called *grey levels* $\mathrm{GL}(x, y)$. The grey level that induces 2π phase shift is called the 2π voltage, which we denote as $a_{2\pi}$. The pixel structure of the SLM allows spatially-dependent phase modulation of the extraordinary wave. Ideally, the SLM is calibrated to have a linear relation between the phase Θ and grey level: $\Theta(x, y) = \frac{2\pi}{a_{2\pi}}\mathrm{GL}(x, y)$. A phase mask imprinted on the SLM and given either in grey levels or radians, depending on the context, is called a *hologram*.

Besides straightforward phase modulation, researchers and engineers widely employ phase-only SLMs to achieve full control over polarization, amplitude, and temporal profiles of the optical field. The following subsections review existing approaches. If the polarization of the incident wave is not specified in the text, one should assume an extraordinary wave.

2.1.2 Simultaneous phase and amplitude modulation

Using two cascaded SLMs and polarization optics. An SLM combined with a polarizer can modulate the amplitude of an electromagnetic wave that is a superposition of the ordinary and extraordinary waves. For example, consider a plane wave of amplitude A_0 polarized at 45° with respect to the extraordinary axis that enters the SLM with a hologram $\Theta_1(x, y)$. After passing through the SLM, the extraordinary component becomes $E_e = \frac{\sqrt{2}}{2}A_0 e^{i\Theta_1(x,y)}$, while the ordinary one remains unaffected and is given by $E_o = \frac{\sqrt{2}}{2}A_0$. If then the wave passes through a 45° polarizer, its electric field becomes

$$E(x, y) = \frac{\sqrt{2}}{2}(E_o + E_e) = \frac{1}{2}A_0(e^{i\Theta_1(x,y)} + 1). \tag{2.3}$$

Both the amplitude and phase of this field are modulated by the hologram, however, not independently.

To control the amplitude and phase simultaneously and independently, work [187] proposes to use the second SLM with a hologram $\Theta_2(x, y)$. The optical setup is such that either a single lens or $4f$ system images the first SLM onto the second one. A half-wave plate rotates the polarization of the field created by the first

SLM and 45° polarizer to match the polarization with the extraordinary axis of the second SLM. After passing through the second SLM, the electric field becomes

$$E(x,y) = \frac{1}{2}A_0(e^{i\Theta_1(x,y)} + 1)e^{i\Theta_2(x,y)} = A_0 \cos\left(\frac{\Theta_1(x,y)}{2}\right) e^{i\frac{1}{2}(\Theta_1(x,y)+2\Theta_2(x,y))}. \quad (2.4)$$

As seen, $\Theta_1(x,y)$ defines the amplitude modulation and $\Theta_1(x,y) + 2\Theta_2(x,y)$ sets the phase modulation.

Via modulation of a blazed grating. The method of using two SLMs and polarization optics described in the previous paragraph seems to be simple, however, experimentally produced structured light commonly has significant distortions. The main reason for the deteriorated performance is a non-zero fraction of the unshaped extraordinary wave, which originates from inactive areas of the SLM and propagates collinear to the modulated beam. Another source of errors comes from the imperfect alignment of the centres of two SLMs.

Surprisingly, a single phase-only SLM is sufficient for the simultaneous encoding of the phase and amplitude profiles of a beam, and encryption methods are robust to the presence of unshaped light. This paragraph describes one of these methods, which utilizes diffractive holograms in the form of blazed grating patterns [43, 23]. The idea of the method is to produce the desired beam profile in the first diffraction order of the hologram, while the undesired light is diffracted to the zeroth order, which includes light passed through inactive SLM areas, and higher orders. The first diffraction order can be spatially separated from others using a spatial filter in the Fourier plane of the hologram.

The control over the amplitude profile of the first diffraction order is realized by spatially modulating the diffraction efficiency of a blazed grating pattern. Assuming an ideal SLM, consider a hologram of the form $\Theta(x,y) = M \bmod \left(\frac{2\pi x}{\Lambda}, 2\pi\right)$, where $\mathrm{mod}(a,b)$ stands for the modulo operation (remainder of the division of a by b). This is a blazed grating pattern of the period Λ, the diffraction efficiency of which depends on M. If $M = 1$, then the phase variation within one period is 2π, and the diffraction efficiency of the hologram is unity, i.e this hologram directs all the

incoming light to the first order. For any other M, some light is diffracted to other orders. Thus, if M has a slowly varying spatial profile, i.e $M = M(x, y)$, where $M(x, y)$ is a slowly varying function with respect to the grating period Λ, the hologram modulates the amplitude of the first order.

To control the phase profile of the first order, work [23] proposes to apply a spatially dependent phase shift to the blazed grating pattern. For the phase shift given by a slowly varying function $F(x, y)$, the resulting hologram is

$$\Theta(x, y) = M(x, y) \cdot \mathrm{mod}\left(F(x, y) + \frac{2\pi x}{\Lambda}, 2\pi\right). \tag{2.5}$$

Below we prove that for any desired optical field given by $A(x, y)e^{i\Phi(x,y)}$, where $A(x, y) \in [0, 1]$, there is a solution for $M(x, y) \in [0, 1]$ and $F(x, y) \in [0, 2\pi]$.

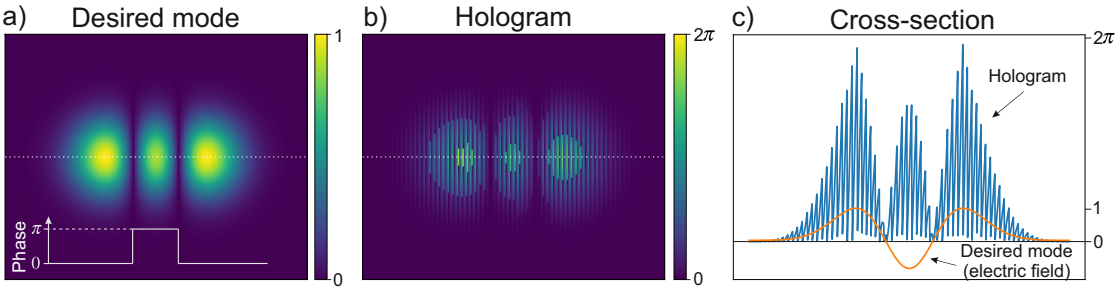


Figure 2.2: a) Amplitude profile of $HG_{2,0}$ mode. The white plot on the bottom schematizes the phase profile. b) Hologram given by Eq. (2.5) with $M(x, y)$ and $F(x, y)$ calculated for $HG_{2,0}$ mode. c) Comparison of the cross-sections marked by the white dashed lines in figures (a) and (b).

Assume that a plane wave of unity amplitude passes through the hologram given by Eq. (2.5). The electric field after the hologram is

$$E(x, y) = e^{iM(x,y)\cdot \mathrm{mod}\left(F(x,y)+\frac{2\pi x}{\Lambda}, 2\pi\right)}. \tag{2.6}$$

Since $M(x, y)$ and $F(x, y)$ are slowly varying functions, $E(x, y)$ can be considered periodic on short intervals multiple to the grating period and, hence, expanded into the Fourier series:

$$E(x, y) = \sum_{k=-\infty}^{\infty} E^{(k)}_{M,F}(x, y)e^{i2\pi kx/\Lambda}. \tag{2.7}$$

A calculation based on work [43] shows that the coefficients of the expansion are slowly varying functions of the transverse position given by

$$E^{(k)}_{M,F}(x,y) = (-1)^k \mathrm{sinc}[\pi M(x,y) - \pi k] e^{i[\pi M(x,y) + kF(x,y)]}. \tag{2.8}$$

Each term of the expansion represents a diffraction order that carries the amplitude and phase encoding. Selecting the first diffraction order (i.e $k = 1$) as the most efficient one to carry the desired optical field, we find that its amplitude and phase profiles are given by $\mathrm{sinc}[\pi M(x,y) - \pi]$ and $\pi M(x,y) + F(x,y)$, respectively (see Eq. (2.8)). Thus, we find that to obtain the desired amplitude $A(x,y)$ and phase $\Phi(x,y)$ in the first order, functions $M(x,y)$ and $F(x,y)$ must equal to

$$\begin{aligned} M(x,y) &= 1 + \frac{1}{\pi}\mathrm{sinc}^{-1}[A(x,y)] \\ F(x,y) &= \Phi(x,y) - \pi M(x,y), \end{aligned} \tag{2.9}$$

where $\mathrm{sinc}^{-1}(\xi)$ is the inverse function of the sinc in the domain of $[-\pi, 0]$.

We used the described approach to generate Hermite-Gaussian modes for the imaging experiment. As an example, Fig. 2.2 illustrates a hologram aimed to generate $HG_{2,0}$ mode. The encryption method via modulation of a grating reduces the shaping resolution of the SLM as the smallest size of the generated modes' features can not be smaller than the period of the grating. Also, such a method leads to different shaping resolutions in x and y directions if lines of the grating lie along one axis, as shown in Fig. 2.2(b). However, one can rotate the grating by 45° to avoid this issue. In our imaging experiment, the shaping resolution of the SLM was sufficient to produce Hermite-Gaussian modes up to $HG_{20,20}$ using a diagonal grating of a 15-pixel period. Reducing the grating period below 15 pixels affected the diffraction efficiency.

Via modulation of a sinusoidal grating. The option of using a hologram based on a modulated sinusoidal grating instead of a modulated blazed grating is discussed in [6]. The proposed form of the hologram is

$$\Theta(x,y) = M(x,y) \cdot \sin\left(F(x,y) + \frac{2\pi x}{\Lambda}\right). \tag{2.10}$$

Expanding the electric field $E(x,y) = e^{i\Theta(x,y)}$ after the hologram into the Fourier series, we obtain that the coefficient corresponding to the first diffraction order is

$$E^{(1)}_{M,F}(x,y) = J_1[M(x,y)]e^{iF(x,y)}, \tag{2.11}$$

where J_1 denotes the first-order Bessel function.

As seen, the maximum possible to encode amplitude $A_{max} \approx 0.5819$, which corresponds to the maximum of J_1, is significantly smaller than $A_{max} = 1$ for holograms based on a blazed grating. Though the hologram based on a sinusoidal grating is less efficient, it is advantageous when applied to a low-resolution SLM, as discussed in work [6]. According to this study, the hologram based on a sinusoidal grating results in a higher signal-to-noise ratio when non-signal diffraction orders, which are produced by the pixel structure of the SLM, are present.

2.1.3 Polarization modulation

Generation of beams with a spatial polarization structure extensively utilizes the idea of superposing two scalar (i.e having homogeneous polarization distribution) modes of orthogonal polarization [113, 107]:

$$E(x,y) = U_H(x,y)\hat{H} + U_V(x,y)\hat{V}, \tag{2.12}$$

where $\hat{H}$ and $\hat{V}$ denote unitary polarization basis vectors. Thus, phase-only SLMs enable full control over the phase, amplitude, and polarization distributions of the optical field.

2.1.4 Applications

The most diverse application area of SLMs is in *structured light*, which refers to the generation of customized light fields for certain purposes. A substantial example is an electromagnetic wave with a helical wavefront (see Fig. 2.3(a)) carrying orbital angular momentum (OAM). The phase pattern of *helical beams* contains a factor of the form $\exp(il\phi)$, where ϕ is the angular coordinate and l is an integer number. The magnitude of l defines the number of intertwined helices composing the wavefront

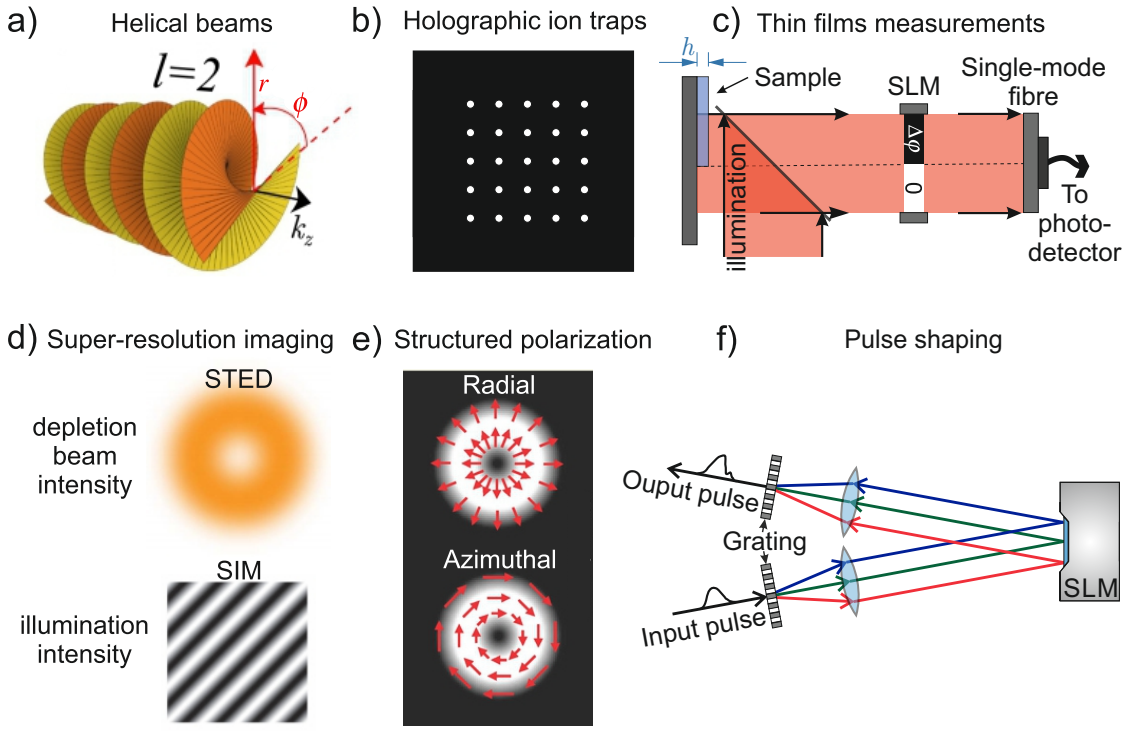


Figure 2.3: Applications of SLMs in structured light. a) A helical wavefront corresponding to the OAM state with $l = 2$. Reproduced from [140]. b) A 2D cross-section of an ion trap that consists of an array of focused beams, which are schematized by white dots. c) Experimental setup for measuring the nano-films' thickness. The beam reflected from the sample plane is imaged on an SLM that shifts the phase of one half of the beam by $\Delta\varphi$ with respect to the other half. After the SLM, the beam is coupled into a single-mode fibre. The SLM and fibre are equivalent to a projector onto a Gaussian mode with phase step $\Delta\varphi$. d) Examples of structured light in super-resolution imaging. The doughnut shape of the STED beam allows researchers to isolate an area of the sample that is in the central part of the beam from the surrounding and, hence, enhance the resolution. When the sinusoidal illumination of SIM is superposed with a sample, which has features with high spatial frequencies, a pattern carrying sub-Rayleigh information about the sample occurs. For more details, see subsections 3.2.3 and 3.2.2. Reproduced from [164]. e) Beam profiles with spatially variant polarization. Reproduced from [106]. f) Schematic of a basic pulse shaper.

and the sign of l defines their handedness. Such beams possess quantized OAM of $l\hbar$ per photon. The most common basis to represent helical beams is the basis of Laguerre-Gaussian modes. Another example is the high-order Bessel beams. SLMs enable both the generation and detection of light with OAM [57].

As a tool for the generation of beams with OAM, SLMs benefit a variety of quantum technologies. The OAM degree of freedom paves the path to high-dimensional quantum systems enabling single-photon qudits. The explicit profit of

the latter is in the increased information capacity of quantum channels [118, 139, 148]. OAM entangled states found applications in high-dimensional quantum key distribution (QKD) [103], quantum teleportation of multiple degrees of freedom [168] and quantum memories [47, 186]. High-dimensional states particularly advantage quantum cryptography due to their robustness to the quantum cloning [30]. As known, the no-cloning theorem forbids the exact copying of a quantum state. However, the creation of an approximate copy is possible and the upper bound of the fidelity depends on the dimension dim of the Hilbert space as $\mathcal{F} = \frac{1}{2} + \frac{1}{1+\dim}$. Moreover, qudit states enhance fault tolerance of QKD from the error threshold of 11% for two dimensions [137] to 19% for four dimensions [30] approaching 50% in the limit of $\dim \to \infty$.

Applications of SLM-generated helical beams are not limited by quantum protocols. Coupling of the OAM of light with mechanical momentum reveals information about moving objects. Optical metrology based on helical beams includes measurements of lateral velocity of motion [40], rotational frequency of spinning objects [95], and turbulence in media [59, 99]. In addition, helical beams have found implementation in biomedicine and chemistry due to their enhanced interaction with chiral molecules given by interplay of the chirality of light modes with that of a molecule [28, 188].

Although helical beams with OAM is an extensive subject, there are many other types of structured light that contribute to imaging, metrology, communication systems, etc. For example, in atomic physics, structured light is a base for optical traps and tweezers [64, 39] (see Fig. 2.3(b)). A hologram printed on the SLM produces multiple beams with individual propagation angle and degree of collimation. These beams are subsequently focused by an objective lens into a three-dimensional holographic optical trap.

In metrology, SLMs are used in advanced methods for measuring the nano-films' thickness [74]. An example of the setup is shown in Fig. 2.3(c). A Gaussian laser beam illuminates a sample of thickness h as shown in the figure. The beam reflected from the sample plane is projected onto a mode that can be described

as a Gaussian mode with the phase difference of $\Delta\varphi$ between its halves and of the same waist as the illumination beam. The normalized power $\mathcal{P}$ measured after the projection depends on $\Delta\varphi$ as

$$\mathcal{P}_{-\Delta\varphi} = \frac{1}{2}(1 + \cos(\delta + \Delta\varphi)), \tag{2.13}$$

where $\delta = 4\pi h/\lambda$. Thus, the thickness h of the sample, which is much smaller than the wavelength of light λ, can be estimated from $\mathcal{P}_{-\Delta\varphi} + \mathcal{P}_{\Delta\varphi} = \sin(\delta)\sin(\Delta\varphi)$.

An important application of structured light is in super-resolution microscopy such as structured illumination microscopy (SIM) and stimulated emission depletion (STED) microscopy (see subsections 3.2.3 and 3.2.2 for operation principles). SLMs generate illumination beams for these techniques (see Fig. 2.3(d)). Especially, SLMs benefit SIM. They can quickly change the orientation and other characteristics of illumination patterns while the sample and other optics remain stationary. Besides the generation of modes, SLMs play a key role in mode-sorting schemes. Implementations of mode-sorting are reviewed in subsection 4.3.1. Chapters 4 and 6 discuss the significance of mode-sorting for Hermite-Gaussian microscopy.

Subsection 2.1.3 has shown that SLMs can shape the polarization profile of a beam. Let us now consider prospects of structured polarization. Radially and azimuthally polarized beams (called *cylindrical vector beams* due to the cylindrical symmetry of the polarization pattern) are of particular interest (see Fig. 2.3(e)). When focused, radially and azimuthally polarized beams produce strong longitudinal electric field and magnetic field components, respectively. Moreover, they can be focused significantly sharper than scalar beams [50]. Properties of cylindrical vector beams make them of great utility for nano-probing [172], lithography [163], confocal microscopy [101, 108], and optical memories [181].

Another example of structured polarization are beams with polarization singularities. Such beams exhibit all possible polarization states described by ellipses. Singularity is a point where the ellipse orientation or ellipse handedness is not defined (circular and linear polarization, respectively). Topology studies use the ability of polarization singularities to form lines and surfaces for producing the

desired structures in space, such as knots [94] and Möbius strips [12]. Due to their robustness to external perturbations, these optical structures are being considered as information carriers in such applications as communication or encryption [93].

SLMs' contribution to applications that require temporal mode shaping is also significant. Fig. 2.3(f) illustrates a typical setup for the pulse shaping [170]. The input pulse enters a diffraction grating, which spatially splits wavelength components. After collimation by a lens, dispersed light passes through the SLM, which modulates the amplitude and phase of each spectral component independently. A lens and grating after the SLM recombine spectral components producing the output pulse, the temporal shape of which corresponds to the inverse Fourier transform of the pattern transferred by the SLM. Fibre communication systems use pulse shapers to perform compensation of chromatic dispersion effects [174]. In biomedical imaging, control of the temporal mode is necessary to ensure the shortest possible pulses [11]. In chemistry, tailored pulses enable fine control over some chemical reactions [8].

One more notable application of SLMs is in commercial telecommunication systems, where SLMs are used in wavelength selective switches (WSS) [132, 167]. WSS is a control element which routes and reconfigures optical signals depending on their bandwidth and central frequency within the fibre network.

In our laboratory, we used an SLM for the first experimental realization of an optical prime number sieve (see appendix C). The prime numbers distribution is encoded in the intensity zeros of the far field produced by a spatial light modulator hologram, which comprises a set of diffraction gratings whose periods correspond to all prime numbers below 149. To overcome the limited far field illumination window and the discretization error introduced by the SLM finite spatial resolution, we rely on additional diffraction gratings and sequential recordings of the far field. This strategy allows us to optically sieve all prime numbers below $149^2 = 22201$. This work connects prime numbers and physics in the field of classical optics.

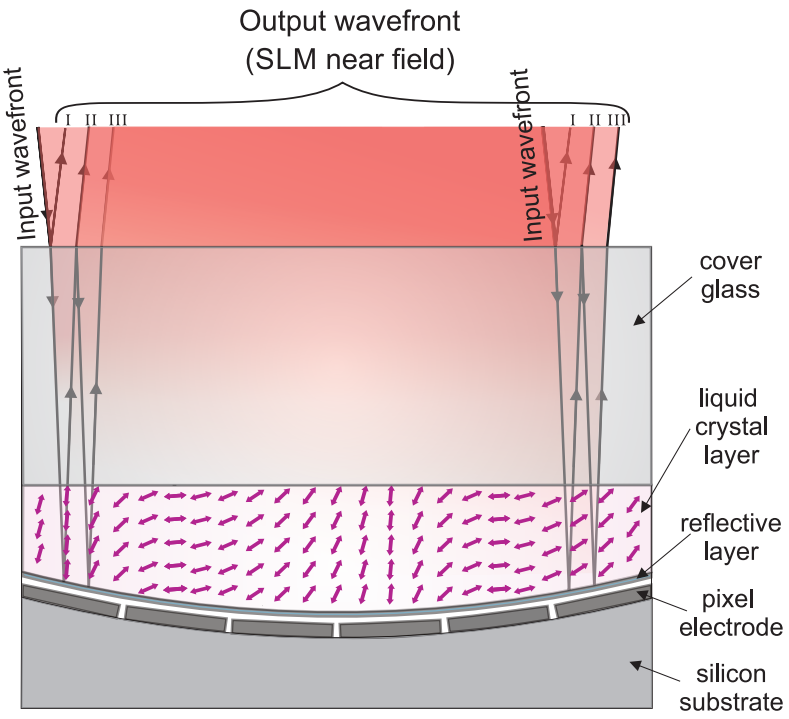


Figure 2.4: Schematic cross-section of a phase-only SLM illuminated by a wide laser beam. I, II, III outline the wavefronts produced by the Fabry-Perot cavity formed by the air-glass interface and the back panel. The orientation of the liquid crystal molecules (magenta arrows) corresponds to a situation when a blazed grating hologram of a 3-pixel period is printed on the SLM suffering from pixel crosstalk.

2.2 Imperfections of phase-only SLMs

The involvement of phase-only SLMs as the key component in the broad spectrum of applications leads to requirements for high accuracy of the device response. Therefore, the study of SLMs imperfections and elaboration of methods for their compensation is necessary.

2.2.1 Curvature of the back panel

Fig. 2.4 schematizes the layered structure of a phase-only reflective SLM to outline the origin of the imperfections. The lower part of Fig. 2.4 illustrates the curved SLM back panel leading to a non-uniform thickness of the liquid crystal layer. The back panel curvature occurs due to limitations in the manufacturing process and varies significantly from one SLM unit to another. A wavefront propagating in the liquid crystal layer experiences a phase retardation given by two contributions.

The first one depends on the liquid crystal molecules' orientation, determined by the pixel voltage matrix (i.e. the printed hologram). The second contribution, voltage-independent, is determined by the additional optical path associated with the non-uniform liquid crystal thickness.

Several methods exist for the characterization and compensation of phase errors related to the non-uniformity of the SLM cell. One approach to retrieve the undesired phase retardation is via interferometric techniques such as phase-shift interferometry [173] and Twyman-Green interferometry [180]. In phase-shift interferometry, the wavefront passed through the SLM addressed with a uniform pixels' voltage is overlapped with a plane reference beam controlled by a piezo mirror. A camera placed at the image plane of the SLM captures four intensity profiles I_1, I_2, I_3 and I_4 obtained for four positions of the piezo mirror spaced by $\frac{\lambda}{4}$. The phase at a specific pixel position is $\phi = \tan^{-1} \frac{I_4 - I_2}{I_1 - I_3}$. In opposite to phase-shift interferometry, the reference beam in Twyman-Green interferometer is slightly tilted so that the interference pattern consists of stripes. A single image is sufficient to retrieve the phase profile of the SLM cell by analyzing the stripes' spacing irregularity.

The retrieved information about the back panel curvature can be used for the curvature compensation. Let us denote the desired phase pattern as $\Theta_0(x, y)$, measured at a grey level GL_0, phase retardation due to the curvature as $\alpha(x, y)$, and the 2π voltage as $a_{2\pi}$. Assuming that the phase response of each SLM pixel to the grey level is linear with uniform slope $k = \frac{2\pi}{a_{2\pi}}$, the grey level pattern (hologram) compensating for the curvature is given by

$$\mathrm{GL}(x, y) = \mathrm{GL}_0 + \mathrm{int}\left[\frac{1}{k}(\Theta_0(x, y) - \alpha(x, y))\right], \tag{2.14}$$

where $\mathrm{int}\left[\frac{1}{k}\varphi\right]$ is the integer part of the expression inside the brackets with $\varphi \in [0, 2\pi]$.

2.2.2 Inner Fabry-Perot cavity

For the optimal performance of reflective phase-only SLMs, manufacturers recommend either normal or small oblique angles of incidence (maximum allowed angle

between the incident and reflected beams is about 5°) for the following reasons. The first reason is that the modulation curve (phase response of the SLM versus addressed grey level) is calibrated to be linear for normal incidence at the electronic driver level. This curve does not change significantly with small angles of incidence. However, when the angle is large, the modulation curve becomes different and varies with the angle of incidence, which follows from the non-linear behaviour of the effective refractive index versus the angle between the propagation direction of light and the optic axis of the LC layer (see Eq. (2.1)). Another reason to avoid oblique incidence is that it leads to a smaller range of phase modulation [116]. Finally, large angles of incidence intricate spatial phase modulation since each point of the incident wavefront can pass through the LC area controlled by several SLM pixels. The information provided in this paragraph is helpful for the description of the SLM's inner Fabry-Perot cavity given below.

Due to a refractive index step at the air-glass interface, an anti-reflection coating is usually applied to the coverglass surface. The resulting coverglass reflection coefficient is low (usually around 0.1 or less) but not negligible. This interface and the reflective layer of the back panel form a low-finesse Fabry-Perot cavity, which produces spurious reflections from the SLM surface [105]. In the following, we refer to this as the *cavity effect*. Since reflective SLMs operate in the parallel-aligned configuration (maximum allowed angle between the incident and reflected beams is about 5°), wavefronts produced by the cavity and marked as I, II, and III in Fig. 2.4 overlap and interfere with relative phases defined by the LC molecules orientation and the back panel curvature. Such interference leads to undesired phase and intensity modulation. Additional reflections (not shown in Fig. 2.4) may occur at the interface between the coverglass and the liquid crystal.

To estimate the reflection coefficients of the interfaces originating from the SLM layered structure, work [105] proposes to separate wavefronts I, II, and III by increasing the angle of incidence to 45°. When the angle of incidence is large enough, the thickness of the cover glass allows observers to distinguish the thee beams, if the incident beam is narrow enough. If the reflection coefficient of the

interface between the glass and LC is negligible, the incident beam intensity and intensities of the beams reflected by the SLM are sufficient to estimate the reflection coefficients of the remaining interfaces.

If the glass-LC reflection coefficient is not negligible, beams II and III are the results of interference of the reflections inside the LC layer, and their intensities oscillate versus grey level. The LC layer is too thin (~ 20 μm, which is significantly thinner than the cover glass) to make separation of the reflections inside it possible. Work [105] proposes a model that allows us to find the reflection coefficients of all interfaces by fitting intensities of beams II and III versus grey level and using the beam I intensity, which remains constant. The disadvantage of this method is that the reflection coefficients are estimated for a large incidence angle, and their values will change when the SLM is used in the standard configuration (i.e the SLM is illuminated under normal incidence).

The undesired modulation due to the cavity effect can be averaged out if the coherence length of the illumination is smaller than the thickness of the cavity. In this case, wavefronts I, II, and III interfere with random relative phases (a random spatially-independent phase is added to the spatial phase profile of each reflection at the output of the SLM), which results in averaging over different realizations of the output field at the measurement stage. However, the coherence length of our laser (~ 48 m) significantly exceeds the thickness of the inner cavity of the SLM. Later in this chapter, we propose a compensation method of the cavity effect for cases when it can not be neglected due to the short coherence length of the incident light.

2.2.3 Pixel crosstalk

Another imperfection to be addressed is crosstalk between neighbouring pixels. The crosstalk effect is that the orientation of liquid crystal molecules within a pixel is not uniform but affected by the adjacent pixels, as shown in Fig. 2.4. The crosstalk leads to a smoothing of the phase profile, which an incoming wavefront acquires passing through the SLM, compared to the desired phase profile. One reason underlying the crosstalk is the gradual change of the electric field amplitude and direction

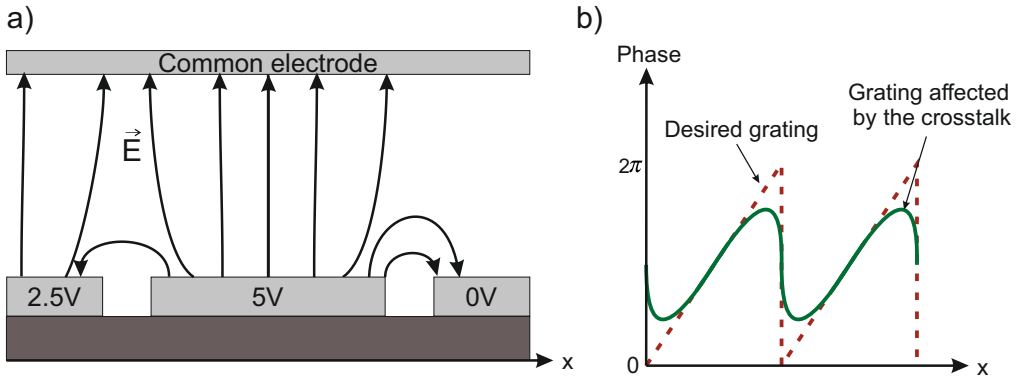


Figure 2.5: a) Fringing fields between pixels addressed with different voltages. Arrows represent electric field lines. b) Smoothing of the phase profile under the crosstalk effect. The dashed line shows the phase response of an ideal SLM addressed with a blazed grating hologram. The solid green line corresponds to the phase-response of an SLM suffering from pixel crosstalk.

between pixels addressed with different voltages (see Fig. 2.5(a)). For this reason, pixel crosstalk is also known as the fringing field effect. Another contribution to the effect comes from the viscosity of the LC material. The crosstalk primarily affects areas of sharp phase steps and decreases the diffraction efficiency (see Fig. 2.5(b)).

The smoothing of the phase profile can be considered as a low pass spatial filtering and modelled as the convolution of the ideal phase with a kernel called a *crosstalk point spread function* (PSF). Usually, a generalized Gaussian PSF

$$g(x,y) = \mathcal{N} \exp\left(-\left(\frac{x^2}{2w_x^2} + \frac{y^2}{2w_y^2}\right)^{\gamma}\right), \tag{2.15}$$

where $\mathcal{N}$ is the normalization factor, describes the crosstalk effect with satisfactory precision. The normalization factor is such that the integral of the crosstalk PSF is equal to one. The waist w ranges from a fraction of one pixel to several pixels, depending on the SLM model, and quantifies the strength of the effect, which, in general, is different for the x and y directions. To characterize the effect for the subsequent compensation, one have to determine the parameters of the crosstalk PSF: w_x, w_y and γ. The prevalent method is to fit the intensity of the first diffraction order of a binary grating (a grating of square fringes) versus the grating's amplitude (*the intensity response curve*) by the theoretical curve modelled using the convolution with the crosstalk PSF [121].

The described approach, being simple and computationally efficient, is preferable for the majority of applications. However, some researchers refer to the complexity of the crosstalk mechanisms and propose the detailed modelling of the LC molecules director behaviour [111]. The motivation for this complicated approach is in the asymmetry of the LC director configuration within a pixel even when a symmetric voltage pattern is applied around the pixel. The reason for the asymmetry is the pre-tilt angle of the LC molecules introduced by the alignment layers, which leads to the following. The electric field at the pixel borders tends to increase the director angle on the one side of the pixel and decrease on the other. Such behaviour can not be modelled precisely with the convolution method.

Another publication proposes to use the Jones matrices approach to the crosstalk problem [100]. In particular, the idea is to image the SLM screen with the magnification sufficient to resolve individual pixels and divide each pixel into sub-regions of a constant director orientation. Measuring Jones matrices of sub-regions for all possible grey level combinations of neighbouring pixels, one can create a huge look-up table which is used to predict the outcome of the imprinted diffraction grating and to find the optimal, in terms of efficiency, shape of the grating.

Knowledge of the crosstalk model allows one to quantify the discrepancy between the expected and real outcomes and develop iterative compensation algorithms. Algorithms differ by the cost function they aim to minimize. It can be the squared difference of the desired I_0 and modelled $\hat{I}$ intensity patterns produced at some distance from the hologram, for example, at the Fourier plane: $R = \sum (\hat{I} - I_0)^2 \rightarrow$ min. Other algorithms directly minimize the squared difference between the ideal and modelled SLM phase patterns: $R = \sum (\hat{\Theta} - \Theta_0)^2 \rightarrow$ min.

Alternatively to the modelling, the effect of crosstalk can be reduced by combining nearby pixels into square regions called super-pixels and assigning a uniform voltage within each of them. This approach reduces the relative area affected by crosstalk, compared to the total area of a super-pixel. However, it leads to the reduction of the available resolution [29].

2.2.4 Non-linear and non-uniform response of liquid crystal to the grey level

Ideally, the response of LC to the grey level, i.e $n_{eff}(\mathrm{GL})$, is linear and uniform. However, the curvature of the back panel not only causes aberrations but also affects the electric field applied to pixels, leading to the spatially dependent response of LC. The non-uniformity of the electric drive scheme also contributes to the spatial variation of $n_{eff}(\mathrm{GL})$. In addition, $n_{eff}(\mathrm{GL})$ can demonstrate spatially-dependent non-linear behaviour. In this case, the SLM requires additional calibration, the purpose of which is to obtain a look-up table (LUT) binding the grey level and phase. The global LUT is sufficient when the spatial variation of n_{eff} is negligible, otherwise, the position-dependent LUT is required.

Similar to one of the crosstalk characterization methods, phase calibration can be performed by displaying diffraction gratings on the SLM and measuring the intensity of diffraction orders. Ronchi gratings (binary phase gratings), which have the shape of a square wave function, are widely used for this purpose [117, 53]. The procedure is as follows. One level of the binary grating remains fixed (the reference grey level GL_{ref}) while another varies (GL_v). The intensity of the first diffraction order depends on the phase difference $\Delta\varphi = \varphi(\mathrm{GL}_v) - \varphi(\mathrm{GL}_{ref})$ between the two levels as $I_1 \propto 1 - \cos\Delta\varphi$. Dividing the SLM into small regions and measuring I_1 for each of them, one can construct position-dependent LUT that contains the mapping between the phase and grey level $\varphi(x, y, \mathrm{GL})$.

Some researchers use phase retrieval algorithms to characterize the SLM response [90]. Algorithms are based on simulations of the wave propagation. A wavefront emitted by the object (SLM) passes through an additional phase plate, which modulates it with a known phase profile. After the phase plate, the wavefront propagates to a camera, which records N intensity patterns corresponding to N phase plate positions. The post-processing algorithm iteratively retrieves the phase profile of the object as follows. First, the algorithm guesses it and simulates the propagation of the wave with this guessed phase profile to the camera for some phase plate position. Then, the algorithm generates the new estimate of the phase:

it 1) replaces the simulated amplitude of the propagated wave with the square root of the recorded by the camera intensity, 2) simulates the propagation of the electric field of the wave back to the plate position, and 3) removes the plate-induced phase from the result of step (2). The phase retrieval method is able to resolve even inner pixel phase variations of the hologram displayed on the SLM.

2.2.5 Noise of the electric drive circuit

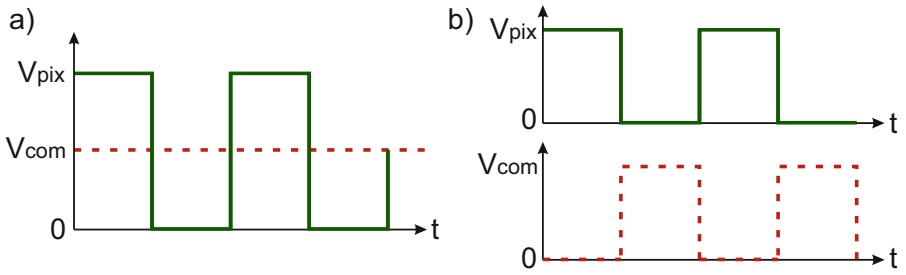


Figure 2.6: Field inversion schemes: a) flipping the pixel voltage around a constant voltage of the common electrode, b) sequentially flipping the voltage of both electrodes. Reproduced from [96].

The difference between the voltage of the common electrode V_{com} and the voltage of a pixel V_{pix} gives the voltage across the LC layer. However, the DC voltage would cause ions' drift and accumulation of the charge at inner surfaces, which would degrade the modulating properties of the SLM. For this reason, the LC layer is kept DC-balanced by the implementation of the periodic field inversion. The field inversion does not affect the tilt angle of LC molecules since the molecules' response is only sensitive to the magnitude of the electric field and not to the voltage polarity. Fig. 2.6 shows possible field inversion schemes. The switching frequency of the voltage is on a scale of several kHz.

Due to imperfections in the realization of the driving scheme such as, for example, the imprecise matching of V_{com} to the middle of V_{pix} (scheme Fig. 2.6(a)) the field inversion leads to the temporal modulation of the voltage magnitude and, hence, to the temporal phase noise [44]. In addition, pixel electrodes switch at slightly different times with respect to the common electrode (raw-based, column-based,

or with a specific pattern schemes), which contributes to the spatial variation of the phase response.

2.3 Comprehensive characterization and compensation of the SLM imperfections

The previous section reviews several SLM imperfections. Existing publications consider different types of imperfections individually, while usually several spurious effects act simultaneously, which significantly complicates their correct characterization. For example, popular methods of quantifying one chosen spurious effect are based on fitting the shape of intensity response curves of a diffraction grating. However, this shape is influenced by multiple effects, including crosstalk, cavity effect, curvature of the back panel, and non-linear response of the refractive index to the addressed grey level.

We propose a comprehensive model which takes into account three major spurious effects: the curvature of the back panel, the inner Fabri-Perot cavity and pixel crosstalk [125]. The non-linearity of the refractive index versus grey level can be omitted since manufacturers usually provide a built-in calibration. To estimate the model parameters with high accuracy, we generate blazed grating holograms and acquire the intensity response curves of the first and second diffraction orders. In addition to considering the joint action of three spurious effects, simultaneous fitting of two diffraction orders is another difference between our method and others. The quantitative model is used to generate compensating holograms, which can produce optical modes with high fidelity.

2.3.1 Experiment

Fig. 2.7 illustrates the optical setup used to characterize the SLM. We illuminate the whole screen of a reflective phase-only LCoS-SLM (Hamamatsu X13138-02) using a continuous laser beam at 785 nm. The SLM operates in a parallel-aligned configuration (the angle between the incoming and reflected beams does not exceed 5°). Resolution of the display is 1280×1024 pixels and the pixel pitch is 12.5 μm.

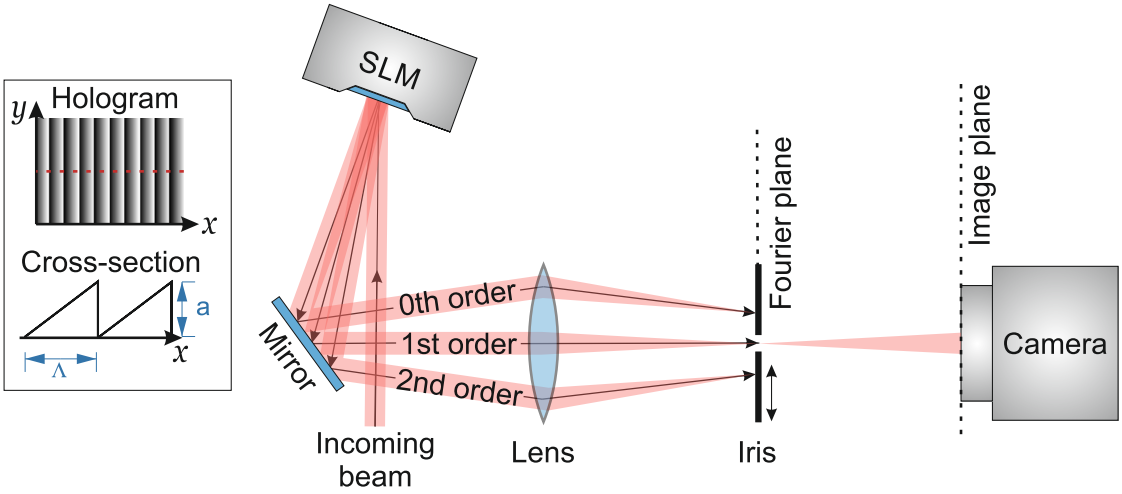


Figure 2.7: Setup scheme for the collection of the intensity response curves. The SLM hologram containing a horizontal blazed grating diffracts the incoming beam. An iris in the Fourier plane selects either the first or second diffraction order and a camera in the image plane acquires the intensity profile corresponding to the near-field of the SLM.

To characterize the spurious effects, we print on the SLM screen a blazed grating hologram of a 20-pixel period, varying its amplitude from 0 to 248 grey levels in steps of 2. The reflected beam is focused by a 2-inch aperture lens ($f = 250$ mm) and subsequently re-imaged by another lens onto a CCD camera (UI-2140SE). An iris diaphragm in the focal plane of the first lens selects either the first or the second diffraction order.

Fig. 2.8(a) shows these images for three different values of grating amplitude. We notice an annular structure in the intensity profiles, which is a sign of the previously mentioned Fabry-Perot cavity effect. Specifically, it is a result of interference of the fields reflected from the SLM after different number of passes through the liquid crystal layer of spatially varying thickness, marked as I, II, and III in Fig. 2.4. For high grating amplitudes, this structure is much more pronounced in the second order than in the first one.

Intuitively this can be understood as follows. Field I is always emitted into the zeroth order. Because of the low reflectivity of the air-glass interface, beam II is generally much stronger than beam III. But when the grating voltage is 2π, the effect of the grating upon the single-pass beam (II) becomes equivalent to that of a plane mirror directing the entire optical field towards the first order. By the same

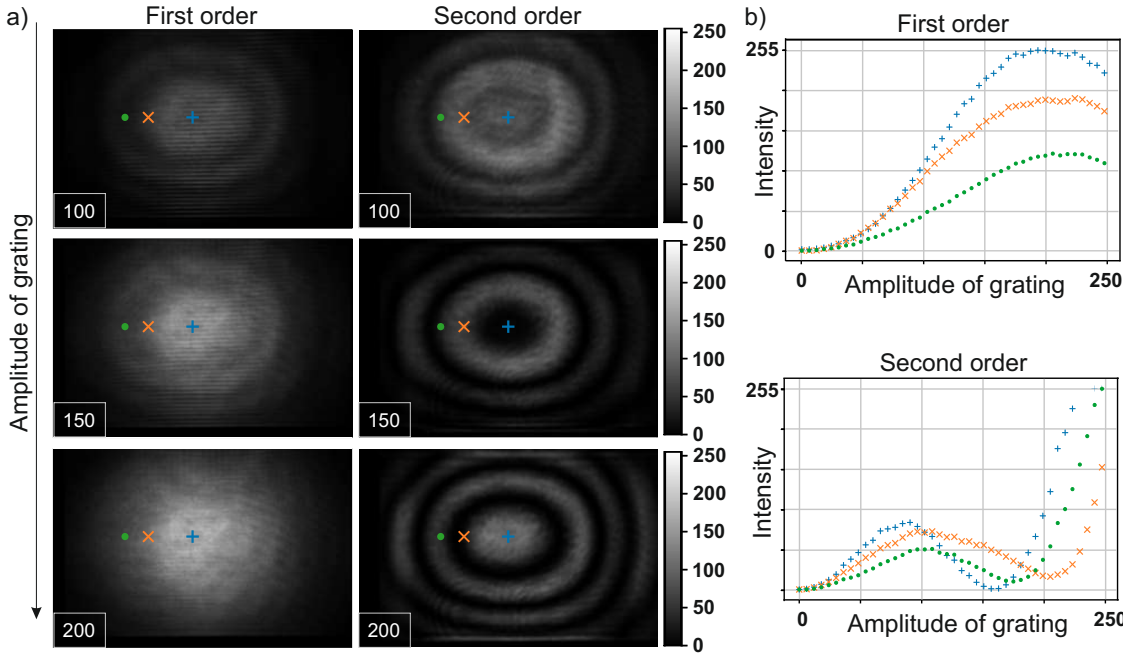


Figure 2.8: a) Near-field images of the first and second diffraction orders for three different values $a = 100, 150, 200$ of grating amplitude. b) Diffracted intensity versus grating amplitude for the three SLM sections marked in (a).

token, the double-pass beam (III) experiences a grating of amplitude 4π, so all of its energy goes into the second order. As a result, in the neighborhood of that voltage the amplitudes of wavefronts II and III in the second order become comparable with their relative phase dependent on the liquid crystal layer thickness.

This effect is further visible in Fig. 2.8(b), where we plot the integrated intensity response for the three 20×20 sections of the SLM screen centred as marked in Fig. 2.8(a). The second order intensity is low, with the behaviour strongly dependent on the cavity thickness, for the grating amplitudes around the 2π voltage (~ 200 grey levels for our wavelength), resulting in well-defined rings.

2.3.2 Theoretical model

Let us start with modelling the inner Fabry-Perot cavity. The optical field emitted by the SLM is given by the sum over all possible reflections occurred at three interfaces between the materials: the air-glass (r_1, t_1), the glass-LC (r_2, t_2), and the LC-reflective-layer (r_3, t_3), where r_i, t_i denote the reflection and transmission coefficients of the corresponding interface (see Fig. 2.9(a)). Phases accumulated

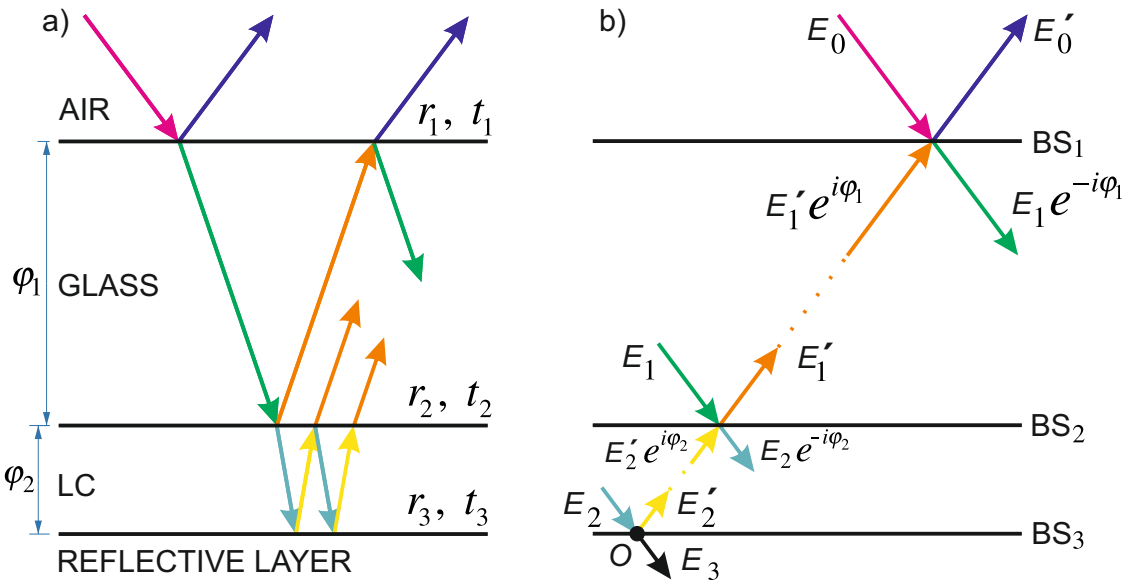


Figure 2.9: a) Reflections occurring inside the SLM. b) The corresponding beamsplitter (BS) model.

due to the propagation through the glass and LC layers are denoted as φ_1 and φ_2, respectively. Such three-interface structure is equivalent to three cascaded beamsplitters, as shown in Fig. 2.9(b). Point O is the phase reference. Colours indicate the input and output channels of each beamsplitter. Using beamsplitters input-output relations significantly simplifies the calculation of the resulting output field E_0', which comes down to solving the following system of equations:

$$\begin{pmatrix} E_0' \\ E_1 e^{-i\varphi_1} \end{pmatrix} = \begin{bmatrix} t_1 & -r_1 \\ r_1 & t_1 \end{bmatrix} \begin{pmatrix} E_1' e^{\varphi_1} \\ E_0 \end{pmatrix} \tag{2.16}$$

$$\begin{pmatrix} E_1' \\ E_2 e^{-i\varphi_2} \end{pmatrix} = \begin{bmatrix} t_2 & -r_2 \\ r_2 & t_2 \end{bmatrix} \begin{pmatrix} E_2' e^{\varphi_2} \\ E_1 \end{pmatrix} \tag{2.17}$$

$$\begin{pmatrix} E_2' \\ E_3 \end{pmatrix} = \begin{bmatrix} t_3 & -r_3 \\ r_3 & t_3 \end{bmatrix} \begin{pmatrix} 0 \\ E_2 \end{pmatrix}. \tag{2.18}$$

Taking into account that $t_i = \sqrt{1 - r_i^2}$, we obtain the following expression for the output electric field:

$$E_0' = -E_0 \frac{r_1 + r_1 r_2 r_3 e^{i2\varphi_2} + r_2 e^{i2\varphi_1} + r_3 e^{i2(\varphi_1+\varphi_2)}}{1 + r_2 r_3 e^{i2\varphi_2} + r_1 r_2 e^{i2\varphi_1} + r_1 r_3 e^{i2(\varphi_1+\varphi_2)}}. \tag{2.19}$$

Let us now turn to the modelling of the spatially-dependent phase φ_2 induced by the LC layer. This phase is given by

$$\varphi_2(x,y) = \Theta(x,y) * g(x,y) + \alpha(x,y). \tag{2.20}$$

The first term is due to the liquid crystal's response to the applied voltage, with $\Theta(x,y)$ being the phase shift in the absence of crosstalk. The crosstalk effect is modelled by convolving $\Theta(x,y)$ with a crosstalk PSF given by Eq. (2.15). The second, voltage independent term, $\alpha(x,y)$, is associated with the spatially variable thickness of the LC layer cavity.

If the reflection coefficient of the glass-LC interface is negligible ($r_2 \approx 0$), Eq. (2.19) reduces to

$$E_0' = -E_0 \frac{r_1 + r_3 e^{i\varphi}}{1 + r_1 r_3 e^{i\varphi}}, \tag{2.21}$$

where $\varphi = 2(\varphi_1 + \varphi_2)$ is the phase accumulated by a wavefront via the round trip through the SLM. In this case, the expression given by Eq. (2.20) models the phase φ (instead of φ_2) with $\alpha(x,y)$ representing the spatially variable thickness of the cavity formed by the air-glass interface and the reflective layer. An alternative way to obtain the expression given by Eq. (2.21) is the summation over multiple reflections produced by the two surfaces:

$$E_0' = -E_0 \left(r_1 + t_1^2 r_3 e^{i\varphi} \sum_l \left[-r_1 r_3 e^{i\varphi} \right]^l \right). \tag{2.22}$$

2.3.3 Fitting the response curves

The voltage applied to the SLM in our experiment corresponds to a blazed grating with its lines along the y axis (see Fig. 2.7):

$$\Theta(x,y) = \frac{a}{a_{2\pi}} \cdot \mathrm{mod}\left(\frac{2\pi x}{\Lambda}, 2\pi \right), \tag{2.23}$$

where a and Λ are the grating amplitude and period, respectively, and $a_{2\pi}$ is the 2π voltage. We use the simplified model of the cavity given by Eq. (2.21) to simulate the response curves of the grating. We assume the reflection coefficient of the

highly-reflective back panel to equal 1 (i.e. $r_3 = 1$). We use a simplified crosstalk PSF, compared to that given by Eq. (2.15). Namely, we assume a Gaussian PSF with the direction-independent waist, i.e our PSF is $g(x, y) = \mathcal{N} e^{-\frac{x^2+y^2}{2w^2}}$.

Thus, the field distribution (2.21) depends upon the following set of parameters: $\{r_1, w, a_{2\pi}(x, y), \alpha(x, y)\}$. We assume that r_1 and w are spatially independent, while $\alpha(x, y)$ and the 2π voltage $a_{2\pi}(x, y)$ are functions of transverse position. We numerically calculate the Fourier transform of Eq. (2.21) and obtain the theoretical response curve of the first (second) diffraction order $I^{\text{th}}_{1^{\text{st}}(2^{\text{nd}})}(a, x, y, \{r, w, a_{2\pi}, \alpha\})$, which is a function of the grating amplitude and model parameters. We estimate the parameters by fitting the acquired experimental curves $I^{\text{exp}}(a, x, y)$ for each 20×20 section of the SLM screen (examples of experimental curves are shown in Fig. 2.8(b)). We search for the optimal set in the following intervals: $r \in [0, 0.15]$ in steps of 0.005, $w \in [0, 2]$ pixels in steps of 0.05, $\alpha \in [0, 2\pi]$ in steps of 0.1, and $a_{2\pi} \in [190, 220]$ in steps of 1.

We fit both diffraction orders simultaneously by considering the joint response curves $I = I_{1^{\text{st}}} \cup I_{2^{\text{nd}}}$. To evaluate the quality of the fit, we use chi-squared

$$\chi^2(x, y) = \sum_{i=0}^{n} \frac{\left(I_i^{\text{exp}} - I_i^{\text{th}}\right)^2}{I_i^{\text{th}}}, \tag{2.24}$$

where $n = 125$ is the number of voltage values for which the data were acquired. The following optimization procedure was used: first, for each pair of r and w we find the optimal spatial distribution of $\alpha(x, y)$ and $a_{2\pi}(x, y)$ by minimizing $\chi^2(x, y)$ of the joint curves for each SLM section; then, we choose the pair of r and w which gives the lowest value of the spatially averaged chi-squared $\chi^2_{\text{avg}} = \langle \chi^2(x, y) \rangle_{x,y}$.

Fitting both orders together significantly improves the accuracy of the parameter estimation. For our SLM, the results were as follows: coverglass reflection coefficient $r_1 = 0.055$, crosstalk Gaussian point spread function width $w = 0.75$ pixels, and the 2π voltage map dropping from 206 in the central part of the SLM to around 197 towards the borders.

Fig. 2.10(a) reports $\alpha(x, y)$ resulting from the fit. As seen, $\alpha(x, y)$ resembles an elliptical paraboloid, with the peak-to-valley value of about 12 rad (1.9λ). To

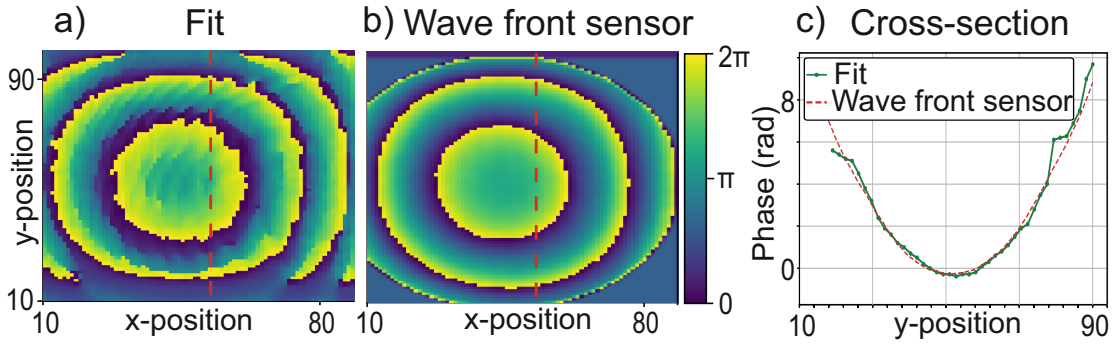


Figure 2.10: Phase map $\alpha(x, y)$ a) calculated by fitting the experimental data (intensity response curves across the SLM screen) with our model and b) measured with a wavefront camera. The profiles of the cross-sections marked by dashed red lines are shown in (c).

verify the accuracy of the reconstructed $\alpha(x, y)$, we have measured it using a commercial wavefront sensor camera (Phasics SID4). The fitted (Fig. 2.10(a)) and measured (Fig. 2.10(b)) phase maps are very similar, as evidenced further by comparing one of their cross-sections (Fig. 2.10(c)). The noise visible on the fitted phase map originates from the CCD camera used to take the near-field images. The camera also has an inner cavity, which results in the weak modulation of the images by curved fringes.

To demonstrate that all spurious effects included in the model are essential to describe shapes of the experimental response curves, we compare the accuracy of our model with three simpler models, in which the crosstalk and/or cavity effects are neglected, i.e. w and/or r are set to zero (see Fig. 2.11). The model neglecting both spurious effects (blue curve) is the least accurate (χ^2 is the highest). The quality of the fit improves if the model includes either cavity (orange curve) or crosstalk (purple curve), and is maximum if both effects are taken into account (green curve).

The extended model of three-layer cavity given by Eq. (2.19) has not refined the fit, which proves the validity of our assumptions about reflection coefficients r_2 and r_3. We tested different deviations of the crosstalk PSF from a Gaussian shape, including asymmetric shapes, as well. None of these PSF provided a better result. Thus, we conclude that convolution with a Gaussian kernel accurately models crosstalk.

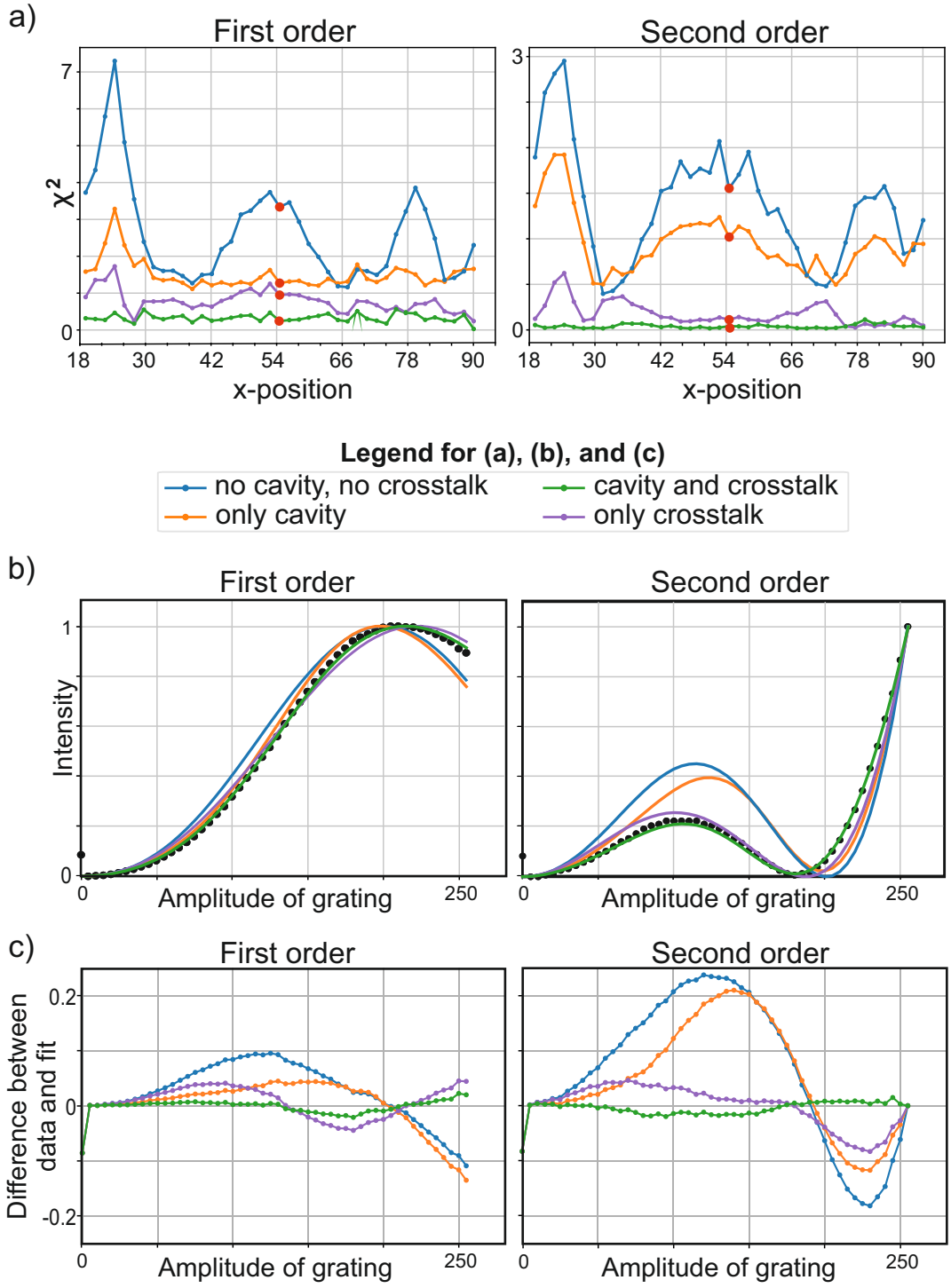


Figure 2.11: Comparison of the fit quality for the different estimation models. a) Fit quality parameter $\chi^2(I^{\text{th}}, I^{\text{exp}})$ calculated along an SLM cross-section. b) Measured (black dot) and fitted (coloured line) intensity response curves for an arbitrary SLM screen position (red dot in (a)). c) Difference between the fitted and measured intensity response curves. For example, the blue line in (c) is the difference between the blue line and the black dotted line in (b).

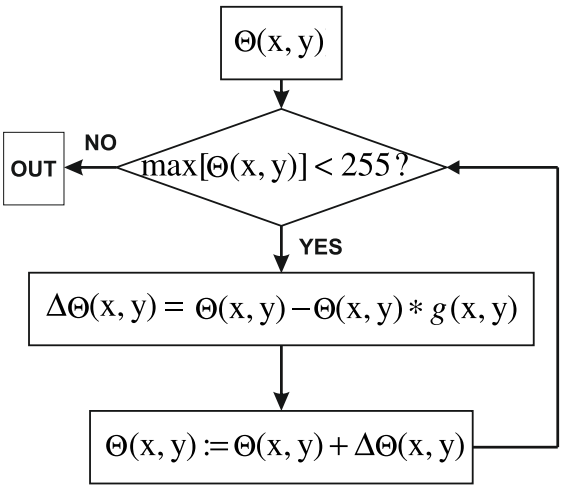


Figure 2.12: Algorithm used to calculate an SLM hologram compensating for crosstalk. Here, 255 is the maximum available grey level of the SLM.

2.3.4 Compensation of the spurious effects

We now address the question of how the above characterized spurious effects can be taken into account when using the SLM to produce arbitrary optical fields of amplitude $A(x, y)$ and phase $\Phi(x, y)$. We base our approach on the widely used encoding proposed by Bolduc *et al.* [23] (see subsection 2.1.2 for the detailed description), in which the pattern printed on the SLM is a modulated blazed grating of the form

$$\Theta(x, y) = M(x, y) \cdot \mathrm{mod}\left(F(x, y) + \frac{2\pi x}{\Lambda}, 2\pi\right), \tag{2.25}$$

where $M(x, y) \in [0, 1]$ and $F(x, y) \in [0, 2\pi]$ are slowly varying functions on the scale of the grating period Λ. Similarly to our experimental scheme, the optical field after the SLM is subjected to a direct and then inverse Fourier transform by means of two lenses. In the Fourier plane, spatial filtering is implemented to select the first diffraction order. The goal is to choose the functions $M(x, y)$ and $F(x, y)$ such that the field obtained in the image plane is the desired $A(x, y)e^{i\Phi(x,y)}$.

For an ideal SLM modelled in Ref. [23], the transverse profile of the field after reflection from the SLM surface is given by $E(x, y) = E_0 e^{i\Theta(x,y)}$, where E_0 is the incident field. In this case, the amplitude of $E(x, y)$ is constant, but the phase is modulated with the period Λ, with the modulation depth and offset determined by

the slowly varying $M(x, y)$ and $F(x, y)$. In an SLM with a cavity effect, however, $E(x, y)$ is given by Eq. (2.21), so small amplitude modulation is also present. Importantly, the function $\alpha(x, y)$ is also slowly varying, so the reflected amplitude can still be considered periodic on short intervals multiple to the grating period.

A function of this kind can be expanded into the Fourier series

$$E(x, y) = \sum_{k=-\infty}^{\infty} E^{(k)}_{M,F,\alpha}(x, y) e^{i2\pi kx/\Lambda}, \tag{2.26}$$

where the components

$$E^{(k)}_{M,F,\alpha}(x, y) = \frac{1}{\Lambda} \int_{x-\Lambda/2}^{x+\Lambda/2} E(x', y) e^{-i2\pi kx'/\Lambda} \mathrm{d}x' \tag{2.27}$$

are slowly varying functions of the transverse position. After the spatial filtering of the first diffraction order (i.e selecting $k = 1$), the field in the image plane of the SLM is given by $E^{(1)}_{M,F,\alpha}(x, y)$. We wish this field to match the desired profile:

$$E^{(1)}_{M,F,\alpha}(x, y) = A(x, y) e^{i\Phi(x,y)}. \tag{2.28}$$

Eq. (2.28) can be numerically solved to obtain $M(x, y)$ and $F(x, y)$ for the given $A(x, y)$ and $\Phi(x, y)$, and thereby construct the cavity corrected hologram. To that end, for each point (x, y) we set up a grid of values $(M(x, y), F(x, y))$, for each of which we evaluate $E(x, y)$ according to Eq. (2.21) using $\alpha(x, y)$ and r_1 known from the fit (r_3 and E_0 are equal to 1). Note that at this stage it is convenient to neglect the crosstalk, simplifying Eq. (2.20) to $\varphi(x, y) = \Theta(x, y) + \alpha(x, y)$. Subsequently we apply Eq. (2.27) to evaluate $E^{(1)}_{M,F,\alpha}(x, y)$ for each point on the grid, assuming $M(x, y)$ and $F(x, y)$ to be constant within the integration limits. In this way, we obtain a look-up table that allows us to find the pair $(M(x, y), F(x, y))$ for any desired complex output at each point. More details about this procedure can be found in our Python implementation [126].

To compensate for the crosstalk effect, we modify the hologram $\Theta(x, y)$ applying the iterative algorithm shown on Fig. 2.12, where $g(x, y)$ is the fitted crosstalk Gaussian kernel. The iterations have to be stopped when the new hologram values are about to exceed the available range of phase modulation. Compared

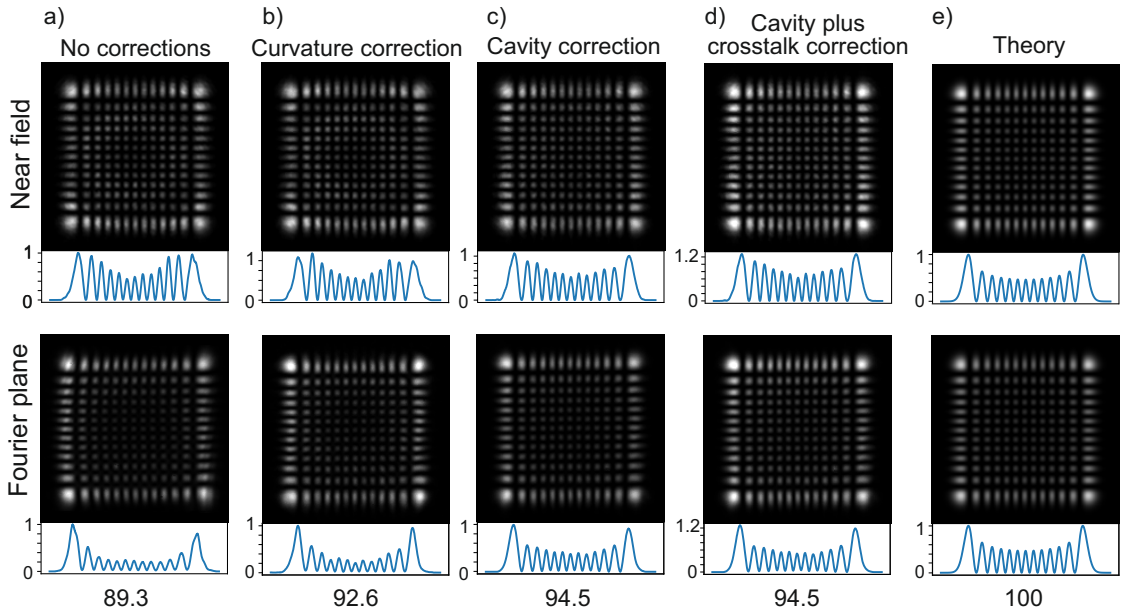


Figure 2.13: Near- and far-field profiles of the $HG_{12,12}$ as produced by the SLM and measured by a camera (a-d) and expected theoretically (e). Each panel indicates the type of correction applied to the SLM hologram and the evaluated fidelity. For each image, the horizontal cross-sections through the mode centre is shown in blue line.

to other crosstalk compensation methods, this approach is not restricted to a specific type of holograms [121], does not reduce the spatial resolution [29] or involve complicated modelling [111].

We evaluate the proposed compensation encoding by generating high-order Hermite-Gaussian (HG) modes. As a figure of merit, we evaluate the fidelity of the produced mode with respect to the ideal. For this purpose, we acquire with a camera the near- and far-field intensity images and apply an iterative maximum-likelihood reconstruction method [102] to obtain the first-order normalized coherence profile

$$g^{(1)}(x,y,x',y') = \frac{\langle E(x,y)E^*(x',y')\rangle}{\int\langle E(x,y)E^*(x,y)\rangle \mathrm{d}x\mathrm{d}y}. \tag{2.29}$$

The fidelity of the experimental mode is then calculated as

$$F = \int E^*_{\mathrm{th}}(x,y)E_{\mathrm{th}}(x',y')g^{(1)}(x,y,x',y')\mathrm{d}x\mathrm{d}y\mathrm{d}x'\mathrm{d}y',$$

where $E_{\mathrm{th}}(x,y)$ is the normalised theoretical profile of the ideal mode.

An example for $HG_{12,12}$ is shown in Fig. 2.13, comparing the theoretical mode profile with those generated by non-, partially- and fully-corrected holograms. The

non-compensated hologram (a) produces a mode whose intensity profile has a ring-shaped modulation in the near field (especially visible in the marginal plot below the 3D photograph), due to the Fabry-Perot cavity effect, and distorted in the far field, as a consequence of the curved back panel. The fidelity of this mode with the ideal one is 89.3%. The curvature-corrected hologram (b) generates a mode with a higher fidelity, 92.6%. The far-field profile is no longer distorted, but the near-field intensity distribution is still modulated by the cavity interference fringes. Hologram (c) compensates both the curvature and cavity effects, producing a mode which best approximates the ideal mode profile (e), with a fidelity of 94.5.

To correct for the crosstalk, we apply three iterations according to Fig. 2.12 to the previously calculated curvature- and cavity-compensated hologram. While the fidelity of the produced mode is unchanged, the diffraction efficiency increases by 28%, as seen in Fig 2.13(d). The crosstalk compensation primarily modifies the areas of discontinuity in the phase profile $\Theta(x, y)$ leading to a sharper blazed grating, whereas the smooth modulation functions $M(x, y)$ and $F(x, y)$ are not affected, which explains the improvement.

2.3.5 Fidelity sensitivity to the spurious effects

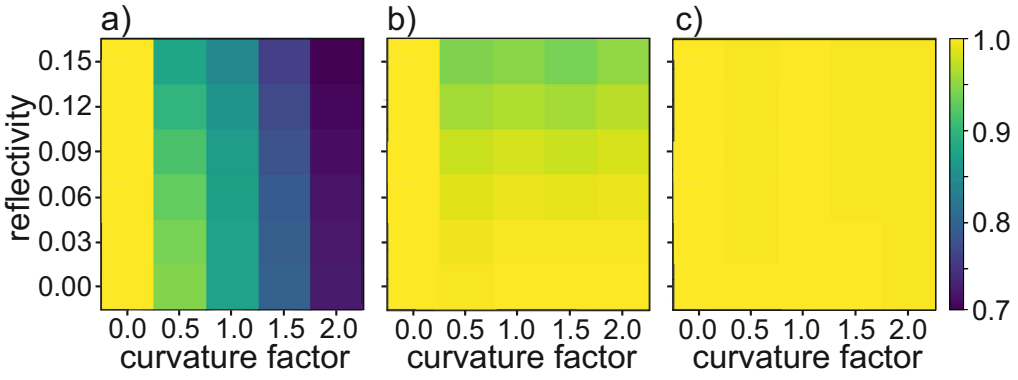


Figure 2.14: Simulated $HG_{12,12}$ fidelity as a function of the reflection coefficient r_1 and curvature multiplying factor, for a) uncorrected, b) curvature corrected and c) cavity corrected holograms.

To show the sensitivity of the proposed method to the three spurious effects, we performed additional numerical simulations. To analyse the cavity plus curvature

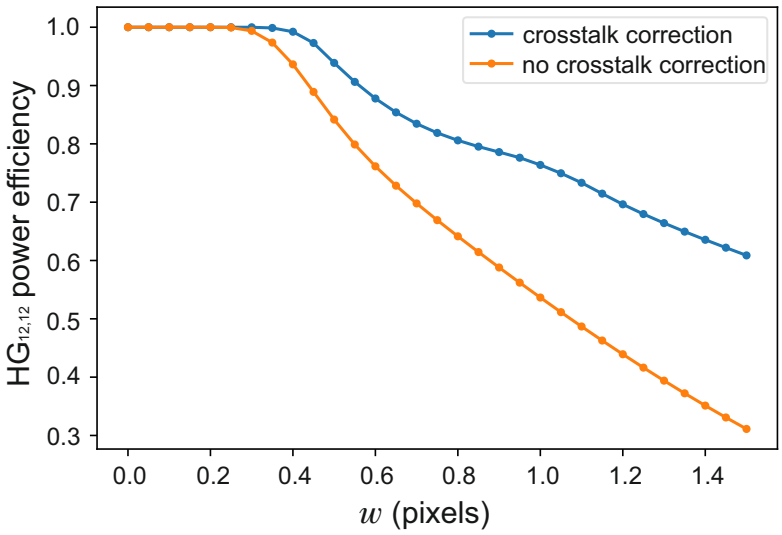


Figure 2.15: Simulated power efficiency as a function of the crosstalk point spread function width with and without the correction.

correction, we model an SLM that has a glass-air interface reflection coefficient r_1 that varies from 0 to 0.15, and the back plane curvature that is similar to our SLM (Fig.2.10), but multiplied by a factor within the interval $[0, 2.0]$. In Fig.2.14(a), we calculate the fidelity of $HG_{12,12}$ without any compensation, i.e. we use the algorithm of reference [23] to compute the SLM hologram directly. The fidelity decreases as the SLM curvature and reflection coefficient increase. In Fig.2.14(b), we compensate only for its curvature, neglecting the cavity effect. For $r_1 = 0$ (no cavity), we notice a significant improvement in the calculated fidelity, which however decreases with increasing r_1 as the cavity effect becomes more significant. Finally, in Fig.2.14(c), we use an SLM hologram that accounts for both the curvature and cavity effect. We see that our algorithm is capable of generating a mode with high fidelity, even when these two effects are significant.

To characterize the crosstalk correction, we model the diffraction efficiency of the mode produced by the hologram corrected for the crosstalk effect and compare it with a hologram with no corrections. We plot the efficiency versus the width of the crosstalk point spread function in Fig.2.15. As seen, the crosstalk correction significantly improves the efficiency for high widths. Both the reflectivity of the air-glass interface and curvature of the back panel are set to zero for this simulation.

As mentioned previously, the influence of the crosstalk on the fidelity is insignificant.

2.3.6 Compensation for the non-uniform profile of the illumination beam

The non-uniform illumination of the SLM screen is not among imperfections of the SLM itself, however, it deteriorates the generated optical field. In the above description, the illumination was assumed to be a plane wave. To take into account the incoming beam profile, the procedure is as follows. We acquire an image of the near field of the SLM screen to obtain the intensity distribution $I_{illu}(x, y)$ of the illumination on the screen. Then, we take the square root of the intensity profile to get the distribution of the absolute value of the electric field. Assuming that the desired amplitude of the generated field is $A(x, y)$, we replace it by

$$A'(x, y) = \mathcal{N} \frac{A(x, y)}{\sqrt{I_{illu}(x, y)}} \tag{2.30}$$

before calculating holograms to compensate for the non-uniform illumination. In Eq. (2.30), $I_{illu}(x, y)$ is normalized so that $I_{illu}(x, y) \in [0, 1]$, analogously to $A(x, y)$, and $\mathcal{N}$ is the normalization factor to get $A'(x, y) \in [0, 1]$. In our experiment, the illumination had non-zero values everywhere across the screen. Otherwise, one can replace zeros with a small non-zero number.

Thus, if the compensation for the non-uniform illumination is needed, one should calculate holograms for $A'(x, y)e^{i\Phi(x,y)}$ using the same methods as described before in this chapter. All experimental images of $HG_{12,12}$ in Fig. 2.13 are obtained after correcting for the incoming beam profile.

2.3.7 Conclusion

We optimize the performance of a LCoS SLM by accounting for all its major spurious effects: the curvature of the back panel, the pixel crosstalk and the low-finesse internal cavity. The model parameters are evaluated by measuring the intensity images of the first and second diffraction orders versus the SLM

blazed grating amplitude. Our compensation method allows us to produce modes of significantly higher fidelity.

The presented model, characterization and compensation method can be readily applied to any phase-only LCoS SLM, with potential benefits for a vast range of applications, such as free space mode division multiplexing [57, 166] and maskless interference lithography [14, 171]. Moreover, the ability to generate and manipulate high-order, high-fidelity Hermite-Gaussian modes may have an important impact in super-resolution imaging, paving the way towards the experimental demonstration of super-resolution techniques based on mode-sorting [162].

3 | Super-resolution imaging: a review

This thesis presents a super-resolution technique dubbed Hermite-Gaussian microscopy. However, the field of super-resolution imaging and microscopy exists since the 1980s and includes various approaches to achieve the resolution beyond the Rayleigh limit. All the existing super-resolution techniques have their individual advantages and limitations, operation principles, and areas of applicability. Thus, a comprehensive review of the most significant techniques is required to identify a place of Hermite-Gaussian microscopy in the field of super-resolution and underline its novelty. This chapter provides such a review and familiarizes a reader with the available methods to break the diffraction limit.

A general classification divides super-resolution techniques into two broad parts: near-field and far-field microscopy. This chapter starts with near-field microscopy based on the intuitive idea that the diffraction limit is not imposed until the electromagnetic field propagates away from the sample. However, it is not possible to probe the electromagnetic field in the vicinity of the sample in many scenarios, which has motivated the development of the far-field techniques reviewed after the near-field ones in this chapter. As will be shown, the existing far-field techniques are based on the non-linear properties of samples and require interaction with them. The motivation behind Hermite-Gaussian microscopy is to develop a far-field and linear-optical super-resolution technique, the applicability of which is not limited to near-field probing or specific samples.

As was mentioned in the introduction (see paragraph "Lateral and axial resolution" in section 1.1), the resolution can be enhanced in the lateral (associated with

the sample plane) or axial (associated with the depth of a sample) directions. Since Hermite-Gaussian microscopy is a two-dimensional imaging technique that enhances the lateral resolution, this chapter focuses on the lateral super-resolution techniques.

3.1 Near-field techniques

Near-field, also known as evanescent wave, is a non-propagating electromagnetic field spatially concentrated at the distances shorter than a wavelength of light from a source. The evanescent wave carries high-frequency spatial information which vanishes with free propagation away from the source. Near-field imaging techniques are based on the detection of such non-radiating waves. Scanning near-field optical microscopy (SNOM) is the most representative example of near-field techniques. In common, all SNOM apparatuses use a sharply pointed probe of a sub-wavelength size in the vicinity of the sample that converts the evanescent wave into a propagating wave preserving the information. An image with sub-Rayleigh details is obtained by scanning the probe over the sample surface. The diversity of SNOM techniques starts with physical effects used to detect evanescent waves, shape of the probe, and illumination mode.

In general, there are two types of SNOM probes: aperture (i.e. with a hole at the tip of the probe) and apertureless (i.e. solid tip). Probes can either be used as a source of the evanescent wave illuminating a sample (illumination mode) or as a collector of light, while the evanescent field around the sample is created by the far-field illumination (collection mode). In some cases, a probe serves both purposes simultaneously.

Near-field microscopy provides a high spatial resolution, which approaches 10 nm level. However, it suffers from the following disadvantages: extremely short working distances, long scanning times, and complications to interpret an image. This subsection provides a review of near-field techniques and highlights underlying physical processes and existing limitations.

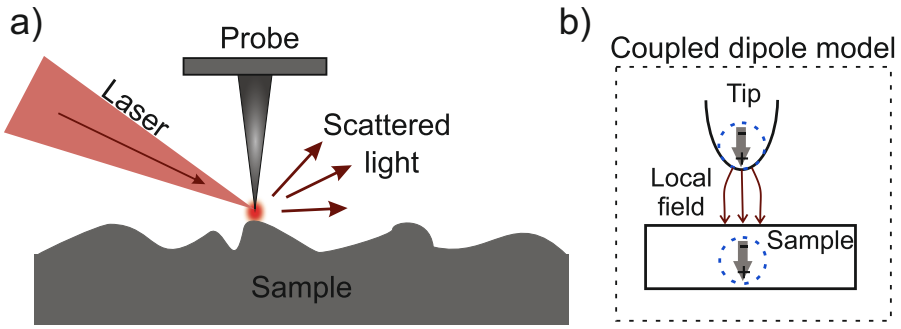


Figure 3.1: a) Schematic of tip-enhanced scattering microscopy. The evanescent field is created around the apex of a sharply pointed metallic probe. The light scattered due to the local tip-sample interaction is detected in the far field and forms a signal for imaging. b) The coupled dipole model of the tip-sample interaction.

3.1.1 Tip-enhanced microscopy

Tip-enhanced microscopy, also known as apertureless SNOM due to usage of an opaque sharply pointed probe, achieves lateral resolution well beyond 50 nm and approaching 10 nm in some cases [69, 150, 86]. This type of near-field microscopy is based on the detection of elastically or inelastically (tip-enhanced Raman spectroscopy) scattered light. Fig. 3.1(a) shows the typical setup. An electromagnetic wave is focused on a sharp metallic tip, which serves two purposes. First, it creates a strong tip-enhanced evanescent field above the sample surface. Second, the tip interacts with the sample resulting in the coupled tip-sample system that acts as a scattering centre for the evanescent field [127]. A detector captures the scattered light in the far-field. The scattered light delivers information about topography and optical properties of the sample to the detector.

The main phenomenon underlying tip-enhanced microscopy is the concentration of incident light around the tip apex. There are three effects known to contribute to the local field enhancement: lightning rod effect [7], optical antenna resonances [79], and localized surface plasmon resonances [46]. *Lightning rod effect* is the field enhancement due to highly localized surface charge densities at geometric singularities. The strength of the effect depends on the conductivity of the tip material at a given wavelength. *Optical antenna resonances* are analogous to their radio and microwave counterparts. When electrons, moving under the action of

the electric field of light, reflect from the antenna apex, and the reflected current is in phase with the optical field at the apex, the local field enhancement occurs. Differences between optical antenna resonances and their radio counterparts arise in resonance conditions on the antenna geometry because metals are not perfect conductors at optical frequencies, unlike radio frequencies, and behave like a plasma. A *localized surface plasmon* is a collective oscillation of electron clouds, stimulated by incident light, around nuclei at the interface between spherical metallic nano-particles (apex of the tip) and air. When the frequency of light approaches the plasmon resonance frequency (i.e. the eigenfrequency at which the displaced electron clouds tend to oscillate around nuclei under the restoring Coulombic force), dependent on the geometry of the tip, near-field at the tip surface is greatly enhanced.

The simplest way to describe the mechanism of the information transfer into the scattered light is a coupled dipole model [41] (see Fig. 3.1(b)). The model considers a spearhead of the tip as a spherical dipole located at the distance d from a half-space representing the sample. The interaction of such system is commonly quantified by assuming a mirror dipole at the distance $-d$ from the boundary of the half-space akin to the method of mirror images in electrostatics. This two-dipole system scatters light according to its effective polarizability, which depends on the distance d and dielectric constants of the tip and sample. Given that geometrical and optical properties of the tip are known, scattered light contains information about sample properties.

Tip-enhanced microscopy has found applications in surface analysis [22], strain visualization in semiconductors [73], imaging of sub-surface structures [156, 135], surface plasmon imaging [32], and biochemical analysis [114, 26]. However, it suffers from the following limitations.

In the case of elastic scattering microscopy, the main challenge is large background noise since illuminating and scattered light have the same frequency. To distinguish the useful signal from one produced by the background scattering, the tip oscillates in the direction orthogonal to the sample surface, modulating the distance d. The demodulated signal is used to form an image of the sample.

Notwithstanding the signal modulation, the poor signal-to-noise ratio remains a drawback of the technique since the spurious scattering of the illumination field exceeds the informative scattering by orders of magnitude. In the case of Raman scattering, the main limitation is an extremely low scattering cross-section.

One way to improve the signal-to-noise ratio is the elaboration of the tip design. Material and geometrical shape of the tip substantially influence the local field enhancement at the tip apex and quantitative characterization of the tip-sample interactions [87, 27].

An alternative approach is to utilize nonlinear properties of the sample (nth order susceptibilities), where this is possible. Tip-enhanced nonlinear imaging combines the sub-Rayleigh spatial resolution of the near-field probing and sensitivity of the nonlinear processes to the structure and symmetry of the medium [178, 25]. The tip acts as a local excitation source for the sample nonlinear response. Though signal detection in near-field nonlinear imaging is still based on the scattering from the tip, it avoids problems related to high background scattering since the signal has a wavelength different from the illumination wavelength and can be spectrally filtered.

A good example of tip-enhanced nonlinear imaging is the investigation of surfaces with topological defects via second harmonic generation (SHG) [179]. SHG is forbidden in media with inversion symmetry. So, the only source of the second-order non-linearity are defects that break the symmetry. The SHG method provides a particularly high contrast when imaging metal surfaces [179]. In this case, a defect acts as a source for an electromagnetic wave travelling along the surface. This wave, called a surface plasmon polariton, causes a significant local enhancement of the second harmonic. Besides imaging of surfaces with defects, near-field SHG measurements are used to determine the local poling direction of ferroelectric materials with nanometric precision and image magnetic-domain structures [149, 88]. Another example of a nonlinear process used in tip-enhanced imaging is four-wave mixing (FWM). The FWM method has been used to image DNA networks [81] and determine the size of gold nano-particles [42].

Limitations of apertureless SNOM based on nonlinear processes are as follows. The use of metallic tips imposes a problem of image interpretation due to nonlinear processes in the tip itself. The tip becomes a source of a frequency-converted signal that overlaps with the signal from the sample. Non-metallic tips do not provide significant local field enhancement at their apex (yet, there is some enhancement), which leads to low signal from the sample because the frequency conversion is generally a weak effect. In this case, a long signal averaging time is required, which has a negative impact on the resolution of the method due to drifts in the apparatus during acquisition.

3.1.2 Aperture-SNOM

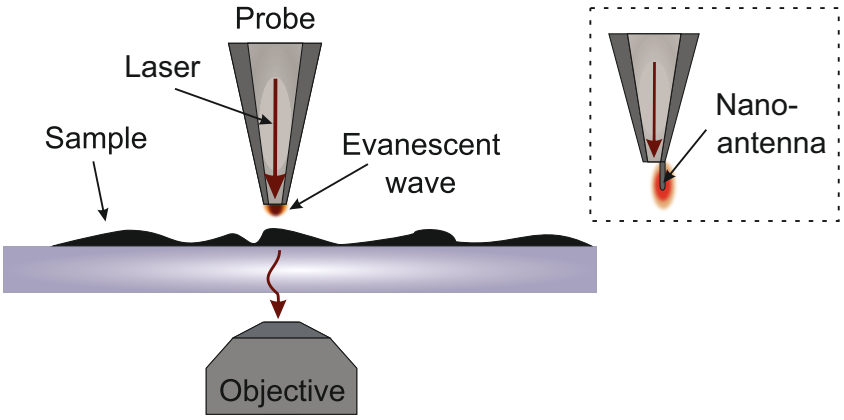


Figure 3.2: Schematic of aperture near-field scanning microscopy. The laser light focused through a nanoscopic aperture at the apex of a tapered metal-coated fibre produces an evanescent field on the exit side of the aperture, which is used to interact with the sample locally. The inset demonstrates a tip-on-aperture probe designed to improve the resolution.

Due to shortcomings of apertureless SNOM such as background scattering and the necessity of modelling tip-sample interactions to extract information from the scattered light, aperture-SNOM is a more popular type of SNOM. A typical probe of aperture-SNOM is a tapered metal-coated fibre with a nanoscopic aperture at the apex, as shown in Fig. 3.2. The aperture acts as a source of evanescent field, which emerges when an electromagnetic wave coupled into the fibre propagates towards its tapered end. The portion of light transmitted through the aperture

depends on the probe diameter. If the diameter is significantly smaller than the wavelength of light, no propagating light passes through the aperture (it is either back-reflected or absorbed), and only the evanescent field remains.

The amount of light that couples through the aperture is associated with the sensitivity of aperture-SNOM. The sensitivity drops while the resolution improves as the aperture size decreases. The tradeoff between resolution and sensitivity is the aperture size of about 50-100 nm. The typical transmittance through the aperture ranges from 10^{-5} to 10^{-3}.

The probe scans over the sample surface, which emits propagating signal in response to the evanescent wave due to the dipole radiation or fluorescence. The typical aperture size of 50-100 nm provides highly localized interaction with the sample and defines the resolution of the method [70]. The emitted signal is then collected in the far-field by a conventional microscope objective.

Aperture-SNOM has a wide range of applications due to the variety of available image formation mechanisms such as amplitude, phase, and fluorescence contrast. Contrast refers to the ability to distinguish between the sample and substrate. Image formation using variations in the amplitude of light emerging from the interaction zone of the evanescent field and absorptive sample refers to the *amplitude-contrast* mechanism. The *phase-contrast* mechanism, which is used for transparent samples, stands for the image formation by converting phase shifts induced by the sample into brightness variations using interferometry.

The *fluorescence-contrast* mechanism is the first and most used image formation mechanism of aperture-SNOM, the application area of which is imaging of biological samples [60, 51]. By locally exciting the sample and detecting the emitted light (which defines the image formation mechanism), aperture-SNOM reveals the density distribution of fluorescent molecules and molecules' dipole moment orientation. The latter follows from the sensitivity of the fluorescent transition to the polarization of the optical field. Namely, the transition occurs when a molecule dipole moment is parallel to the polarization of light. This dependence can be used conversely

to determine the field distribution around the aperture by using a single molecule of known orientation.

Raman spectroscopy is among applications of aperture-SNOM [83], notwithstanding difficulties related to low energy transmission through the probe combined with low efficiency of Raman scattering.

The main advantage of aperture-SNOM over the tip-enhanced microscopy is excellent optical contrast due to the absence of background illumination. However, the resolution of aperture-SNOM is typically lower due to limitations on the aperture size, which is hard to reduce below 50 nm because of dramatic decrease of transmission through the aperture. To combine advantages of the aperture and apertureless configurations, tip-on-aperture generation of probes has been developed [58]. Such probes have a metallic nano-antenna formed at the apex of the probe, which concentrates light transmitted trough the aperture (see the inset of Fig. 3.2). Tip-on-aperture probes enhance resolution down to 25 nm.

3.1.3 Photon scanning tunnelling microscopy

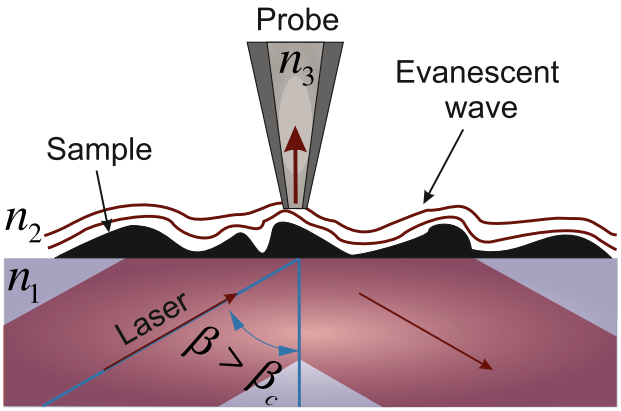


Figure 3.3: Schematic of photon scanning tunnelling microscopy. The evanescent field created on the surface of total internal reflection and modulated by the sample interacts with the probe, which leads to the coupling of light into the probe due to the frustrated total internal reflection effect.

Photon scanning tunnelling microscopy (PSTM) relies on the *frustrated total internal reflection effect*. As is known, total internal reflection occurs when a wave arrives at the interface between two media from the medium with a higher

refractive index under an angle exceeding the critical one. Let us denote the refractive indices of the two media as n_1 and n_2, assuming that $n_1 > n_2$. The critical angle is then given by

$$\beta_c = \arcsin \frac{n_2}{n_1}. \tag{3.1}$$

In the case of total internal reflection, there is no refracted beam and, hence, no energy transfer to the medium with a lower refractive index. However, the evanescent wave is present at the vicinity of the interface. The frustrated total internal reflection effect occurs when there is a third medium of refractive index $n_3 > n_1$ at a sufficiently short distance to interact with the evanescent wave. In this case, one can observe some light transmitted into the third medium. The frustrated total internal reflection has been found similar to the quantum tunnelling effect, when a particle penetrates through a potential barrier. For this similarity, the effect is also called *photon tunnelling*.

In a photon tunnelling microscope, a dielectric or semiconductor sample of a wavelength-order thickness is placed on the surface of total internal reflection and modulates the evanescent field (see Fig. 3.3). The evanescent field is scanned by a sharp probe, which is usually represented by a tapered uncoated fibre. The degree of coupling of the evanescent wave into the probe strongly depends on the distance from the sample. Thus, the microscope obtains a topographical image of the sample by counting photons transmitted into the probe. A feedback circuit prevents the collision between the probe and sample. Since the probe is used for collection in PSTM, there is no need for the aperture surrounded by metal coating, as in aperture-SNOM, where the probe is a source of the evanescent field. For this reason, PSTM surpasses aperture-SNOM and achieves the lateral resolution of 10 nm and vertical (axial) resolution of 1 nm [152].

3.1.4 Feedback mechanisms

A feedback mechanism is an essential part of SNOM that keeps tip-sample working distance in the sub-wavelength domain to assure efficient sensing of rapidly decaying

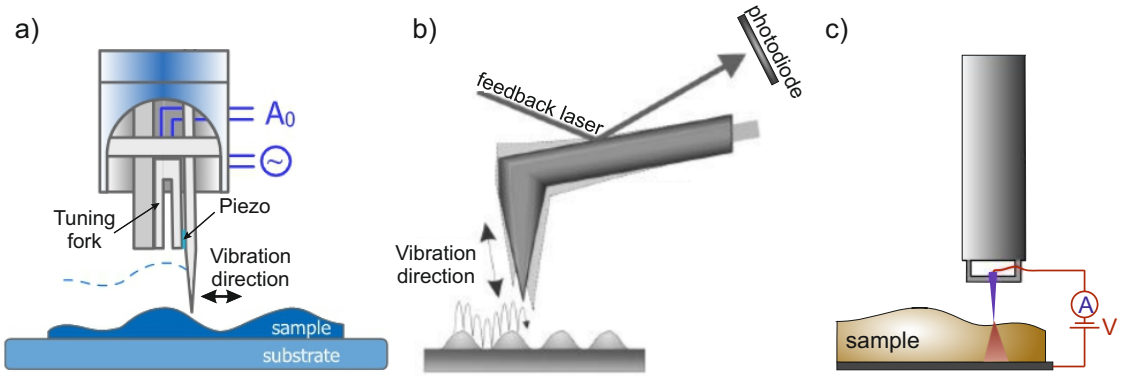


Figure 3.4: Feedback mechanisms of SNOM. a) Shear-force feedback. Reproduced from [142]. b) Tapping mode feedback. Reproduced from [45]. c) Tunnelling current feedback. Reproduced from [159].

evanescent waves. Information obtained from feedback loops supplements optical information in SNOM images. The three main feedback systems are shear-force feedback, tapping mode feedback, and tunnelling current detection feedback [97].

Fig.3.4(a) schematizes the shear-force feedback system. A probe attached to a quartz tuning fork oscillates in the transverse direction at the fork resonant frequency. Oscillations are induced by an external piezo element under AC voltage. In response, the tuning fork produces a voltage output that is dependent on the amplitude of mechanical oscillations. When the tip of the probe nears the sample surface, atomic forces between tip and sample damp the probe oscillations and change the voltage response of the tuning fork.

In the tapping mode feedback system, shown in Fig. 3.4(b), a cantilever probe oscillates in the vertical direction with respect to the sample. A laser beam reflected from the cantilever to the position-sensitive detector tracks changes in amplitude and frequency of oscillations due to the interaction with the sample. Damping forces consist of atomic forces and capillary action of the thin layer of water that covers both the tip of the probe and the sample surface in standard atmospheric conditions [189]. A drag force appears when the tip is lowered into the layer of water on the sample surface. The tapping mode is preferable for biological samples since the shear-force system damages cells. Furthermore, the resolution of the shear-force system is lower than that of the tapping mode due to lateral oscillations.

Fig. 3.4(c) illustrates the tunnelling current feedback. This feedback type is applicable to the conductive tip and conductive sample configuration. In this case, a bias voltage is applied between the tip and sample. When the tip approaches the sample surface, tunnelling current occurs, which can be used as a signal.

3.1.5 Near-field optical random mapping microscopy

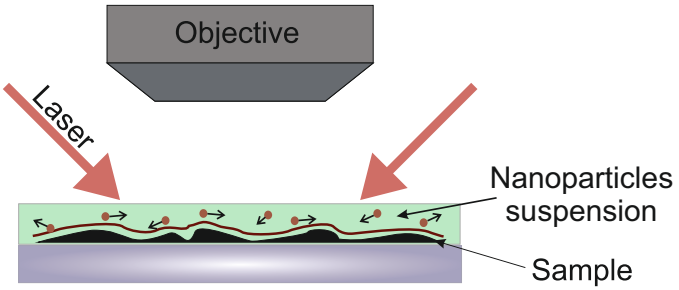


Figure 3.5: Schematic of near-field optical random mapping microscopy based on the scattering of the evanescent field over the sample by nanoparticles.

This type of near-field microscopy does not use a tip scanning over a sample. Instead, the sample is placed into a scattering medium, such as suspension of nanoparticles, and observed by an ordinary microscope (see Fig. 3.5). The brightness of light scattered by a nanoparticle depends on the local near-field intensity over the sample. Thus, if the distribution of nanoparticles is known, a single frame captured by a camera contains information about the near-field intensity at location points of nanoparticles. Particles move in suspension according to the Brownian motion. By collecting multiple frames that correspond to different distributions of particles, the whole near-field can be reconstructed. Nanoparticles appear in the microscope like round spots, lateral positions of which can be determined with the precision much higher than the resolution of the objective.

The constructed image requires further Fourier filtering due to the following reason. Nanoparticles move at different distances from the sample surface. More distant particles can image only low spatial frequencies, which results in the disproportional contribution of the latter. Fourier filtering corrects the associated blurring by damping low frequencies.

The resolution of the method is defined by the size of nanoparticles and resolution of the objective used to detect them. The original paper reports 96 nm resolution by using nanoparticles of 120 nm size [109], microscope objective with $NA = 0.4$, where NA is numerical aperture, and illumination at 520 nm wavelength. The achievable resolution capability of the method is estimated to be 10 nm.

3.2 Far-field techniques

Applications of near-field microscopy are limited mainly to the study of surfaces due to short working distances. The motivation for developing far-field super-resolution techniques has been initiated by biological studies. For this reason, far-field techniques rely on fluorescent properties of samples.

3.2.1 Image scanning microscopy

Image scanning microscopy (ISM) is a fluorescence-based technique that has emerged from confocal laser-scanning microscopy (CLSM) [155]. In the confocal microscope shown in Fig. 3.6(a), laser light is focused by an objective lens to excite a particular point at a certain depth within a fluorescent sample, unlike in a conventional wide-field microscope where the whole sample is illuminated. The light emitted by the excited point is then collected by the same objective lens, decoupled from the illumination path by a dichroic mirror and imaged into a bucket detector (i.e. a detector with no spatial resolution). A pinhole in the conjugate focal plane in front of the detector cuts the signal from out-of-focus points. By scanning the sample in the transverse direction, a sample plane at the chosen depth can be imaged. Closing the pinhole not only eliminates the out-of-focus signal and improves the axial resolution, but makes it possible to increase the lateral resolution by up to a factor of $\sqrt{2}$ beyond the diffraction limit. The latter statement is derived below.

CLSM would achieve the $\sqrt{2}$-fold resolution enhancement in the limit of a fully closed pinhole. For this reason, it suffers from low signal-to-noise ratio. This issue has been overcome in ISM, schematic of which is shown in Fig. 3.6(b), by replacing the pinhole and bucket detector by an array detector, which consists of many pixels

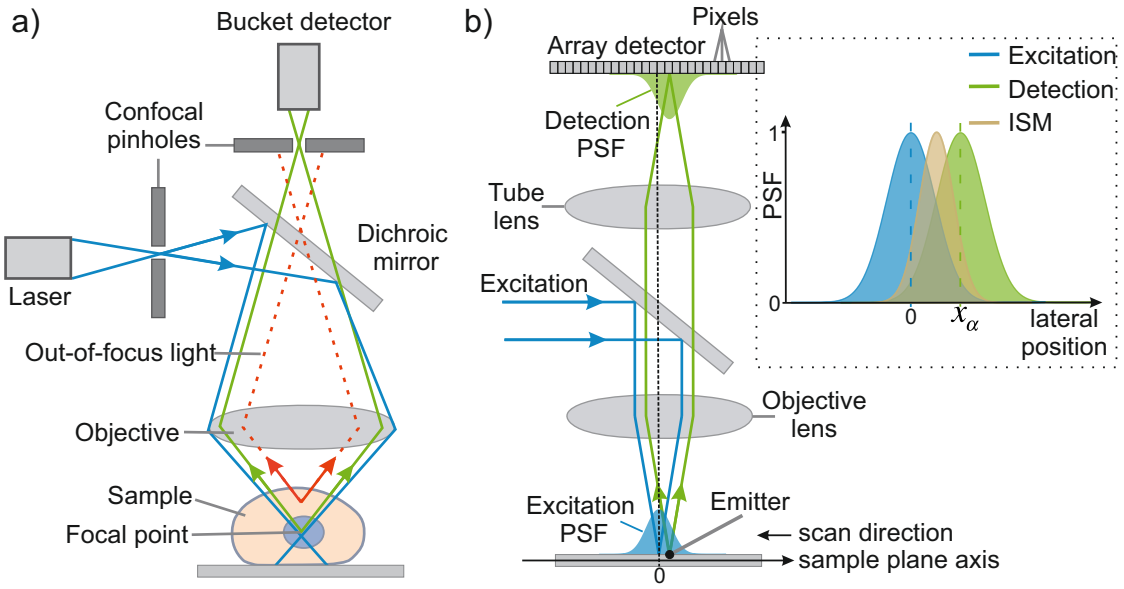


Figure 3.6: a) Schematic of confocal laser scanning microscopy. The bucket detector records a scan image of the sample with the $\sqrt{2}$-fold resolution enhancement in the lateral direction, compared to wide-field microscopy, in the limit of a zero-size pinhole. Reproduced from [36] b) Schematic of image scanning microscopy. Each pixel of the detector records a scan image with approximately the $\sqrt{2}$-fold resolution enhancement, which results in an improved signal-to-noise ratio, compared to CLSM, of the final image obtained after post-processing. The black dashed line indicates the optical axis. The sample plane axis is the axis used to describe the positions of emitters forming the sample. The origin of this axis is defined by the intersection with the optical axis. For simplicity, the figure shows the emission from a single emitter in the sample plane. The inset illustrates the formation of the effective PSF for the scan image recorded by the pixel located at position x_α. All PSF profiles in the figure are intensity profiles.

[112, 143]. Each pixel records a signal versus scan position of the sample, called a scan image. Since the pixel size is much smaller than the waist of PSF of the imaging system, which images the sample plane into the detector, the scan image of each pixel is analogous to that of the bucket detector combined with a near zero-size pinhole in CLSM. Images from different pixels are shifted to compensate for a pixel viewing angle and summed into a single image. Thus, ISM provides resolution comparable to that of the confocal microscope with a zero-size pinhole, but with a significantly increased signal-to-noise ratio since there are no light losses.

To describe the image formation and derive the lateral resolution of ISM, let us treat elements of the sample plane as single-photon emitters located at positions x_i with respect to the optical axis. The excitation laser is focused to a diffraction-

limited spot centred at $x = 0$, the intensity distribution of which is given by the excitation PSF $\psi_{exc}(x)$ (see Fig. 3.6(b)). Positions of pixels in the image plane are x_α with respect to the optical axis. Suppose that emitters shift by x_s due to scanning in the direction opposite to the sample plane axis. Thus, ith emitter moves from position x_i to $x_i - x_s$.

The excitation probability of this emitter is proportional to the illumination intensity, which is given by $\psi_{exc}(x_i - x_s)$. The probability of a pixel located at position x_α to detect the emitted photon is proportional to $\psi_{det}(x_\alpha - (x_i - x_s))$, where $\psi_{det}(x)$ is the detection PSF. Thus, the pixel located at x_α records a scan image $G_\alpha(x_s)$ given by

$$G_\alpha(x_s) \propto \sum_i \psi_{exc}(x_i - x_s) \cdot \psi_{det}(x_\alpha - (x_i - x_s)) \equiv \sum_i \psi_\alpha(x_i - x_s), \tag{3.2}$$

where $\psi_\alpha(x) = \psi_{exc}(x) \cdot \psi_{det}(x_\alpha - x)$ represents the effective PSF of a pixel image.

In general, $\psi_{det}(x)$ is slightly wider than $\psi_{exc}(x)$ due to the Stokes shift of the fluorescence emission and linear scaling of the PSF width with the wavelength of light. However, we neglect the Stokes shift for simplicity and assume that $\psi_{exc}(x) = \psi_{det}(x) \equiv \psi(x)$. If $\psi(x)$ is represented by a Gaussian function of width σ, then $\psi_\alpha(x)$ is a Gaussian function of width $\frac{\sigma}{\sqrt{2}}$ centred around $\frac{x_\alpha}{2}$, as illustrated in the inset of Fig. 3.6(b). As seen, the lateral resolution enhancement of $\sqrt{2}$ times below the diffraction limit is achieved due to the multiplication of the excitation and detection PSFs. In CLSM, the origin of the $\sqrt{2}$-fold resolution enhancement is analogous since the bucket detector and near zero-size pinhole can be seen as a pixel positioned on the optical axis.

To form the final image, scan images from individual pixels are centred around $x = 0$ and added together to increase the signal-to-noise ratio:

$$G_{ISM}(x_s) = \sum_\alpha G_\alpha \left(x_s - \frac{x_\alpha}{2} \right). \tag{3.3}$$

After this procedure, the effective PSF of the method becomes

$$\psi_{ISM}(x) = \sum_\alpha \psi_\alpha \left(x - \frac{x_\alpha}{2} \right), \tag{3.4}$$

which remains, however, about $\sqrt{2}$ times narrower than the PSF of conventional imaging.

A photon correlation measurement has been proposed for further enhancement of the ISM resolution [157]. The idea is as follows. Each emitter in the sample plane emits at most one photon when excited by a pulse that is shorter than the fluorescence lifetime (photon antibunching). However, there is a probability that two emitters inside the excitation spot get excited by the same pulse. Due to antibunching, the probability to detect two photons simultaneously (i.e. with a delay not exceeding the pulse duration) is lower than the probability to detect two photons at a delay of one or more pulses. The difference between the two probabilities is proposed as a signal that forms an image.

For a pair of pixels α and β, the signal is given by

$$\begin{aligned}\Delta G_{\alpha\beta}(x_s) &= G_\alpha(x_s)\cdot G_\beta(x_s) - G_{\alpha\beta}(x_s)\\ &= \sum_i \psi_\alpha(x_i - x_s)\cdot\psi_\beta(x_i - x_s) \equiv \sum_i \psi_{\alpha\beta}(x_i - x_s),\end{aligned} \quad (3.5)$$

where $G_{\alpha\beta}(x_s) \propto \sum_{i\neq j}\psi_\alpha(x_i - x_s)\cdot\psi_\beta(x_j - x_s)$ is the probability of simultaneous detection of a photon pair from two different sources with two different pixels, $G_\alpha(x_s)$ and $G_\beta(x_s)$ are given by Eq. (3.2). As seen, $\psi_{\alpha\beta}(x)$ is $\sqrt{2}$ times narrower than $\psi_\alpha(x)$ and hence 2 times narrower than the PSF of conventional microscopy. To obtain the final image, which we denote as $G_{Q-ISM}(x_s)$, with a high signal-to-noise ratio, signals from each pixel pair are shifted and summed:

$$G_{Q-ISM}(x_s) = \sum_{\alpha\neq\beta}\Delta G_{\alpha\beta}\left(x_s - \frac{x_\alpha + x_\beta}{4}\right). \quad (3.6)$$

This method is called quantum ISM (Q-ISM).

Fourier reweighting of ISM images, which amplifies high spatial frequencies, allows one to improve the lateral resolution even further. About 2 and 2.3 times enhancement of resolution with respect to conventional imaging has been demonstrated for ISM and Q-ISM, respectively, after Fourier reweighting.

3.2.2 Stimulated-emission-depletion microscopy

Stimulated-emission-depletion microscopy (STED) is one of the fundamental far-field super-resolution techniques [72, 89], which was marked by the Nobel Prize

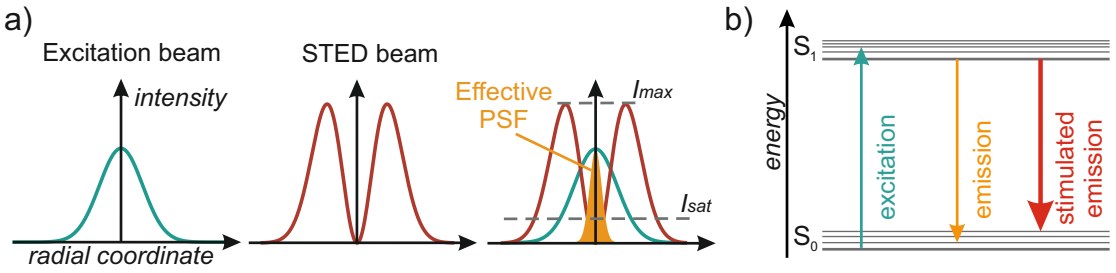


Figure 3.7: Reproduced from [84]. a) Spatial intensity distribution of the excitation and STED beams and the effective PSF of STED. b) Energy levels diagram and optical transitions of STED.

in Chemistry in 2014. STED is a fluorescence-based technique that achieves sub-Rayleigh resolution by depleting fluorescence emission in the outer region of a diffraction-limited excitation spot.

The depletion of fluorophores is implemented by overlapping the excitation beam in the object plane with a doughnut-shaped beam, which is called a STED beam. Fig. 3.7(a) shows intensity profiles of the beams versus radial coordinate. Fig. 3.7(b) shows a diagram of the energy levels involved in STED. A pulse of the excitation beam excites fluorophores from the lowest vibrational sublevel of the ground state S_0 to a higher vibrational sublevel of the exited state S_1. Molecules decay to the lowest sublevel of S_1 due to vibrational relaxation within few ps. The lifetime of this sublevel is of ns order.

A STED pulse arrives immediately after the excitation pulse under the control of a delay-time optimization circuit. In the central low-intensity region of the STED beam, transition from S_1 to S_0 occurs due to spontaneous emission. However, in the region where the intensity of the STED beam is significant, stimulated emission at a wavelength of the STED laser dominates. The STED beam is red-shifted with respect to the fluorescence wavelength chosen to be detected. Photons at the STED wavelength are filtered out and the detector sees only the narrow spot where the fluorescence is unaffected by the STED beam (the effective PSF in Fig. 3.7(a)).

The resolution of the method depends on the width of the spot seen by the detector. The probability to detect a photon from location r (i.e the effective

PSF of STED) is

$$\psi_{STED}(r) \propto \psi_{exc}(r) \cdot \eta(r), \tag{3.7}$$

where $\psi_{exc}(r)$ is the intensity profile of the excitation spot and $\eta(r)$ is the fluorescence probability after the STED pulse [68]:

$$\eta(r) = \exp\left[-\frac{I_{STED}(r)}{I_{sat}}\right]. \tag{3.8}$$

$I_{STED}(r)$ denotes intensity profile of the STED beam and I_{sat} is the saturation intensity. Assuming a Gaussian excitation spot, i.e. $\psi_{exc}(r) \propto \exp\left[-\frac{r^2}{2\sigma^2}\right]$ and approximating the central part of the STED beam intensity profile by a parabola

$$I_{STED}(r) \approx 4 I_{max} s^2 r^2, \tag{3.9}$$

where I_{max} is the maximum intensity of the STED beam and s is the steepness parameter, we obtain the following expression for the effective PSF of STED using Eq. (3.7) and Eq. (3.8):

$$\psi_{STED}(r) \propto \exp\left[-\frac{r^2}{2\sigma^2}\right] \cdot \exp\left[-\frac{4 I_{max} s^2 r^2}{I_{sat}}\right] = \exp\left[-\frac{r^2}{2\sigma_{STED}^2}\right]. \tag{3.10}$$

The width of $\psi_{STED}(r)$ is given by

$$\sigma_{STED} = \frac{\sigma}{\sqrt{1 + 8\sigma^2 s^2 \frac{I_{max}}{I_{sat}}}}. \tag{3.11}$$

For $I_{max} = 0$, i.e. in the absence of the STED beam, σ_{STED} coincides with the width σ of the diffraction-limited excitation spot. For $I_{max} > 0$, the resolution of STED overcomes the diffraction limit.

Theoretically, infinitely high resolution can be achieved by increasing the power of the STED beam. In practice, however, the available power is limited. Besides, extremely high power of the STED beam leads to unwanted effects such as photobleaching, multiphoton absorption, sample heating or destruction. An alternative way to enhance the resolution is to reduce I_{sat} by optimizing the wavelength of the STED beam and increasing the cross-section of the stimulated emission. Experimentally achieved resolution of STED for biological samples ranges from 20 to 100 nm. The highest reported resolution is 6 nm for nitrogen vacancy centres in diamonds [131].

3.2.3 Structured illumination microscopy

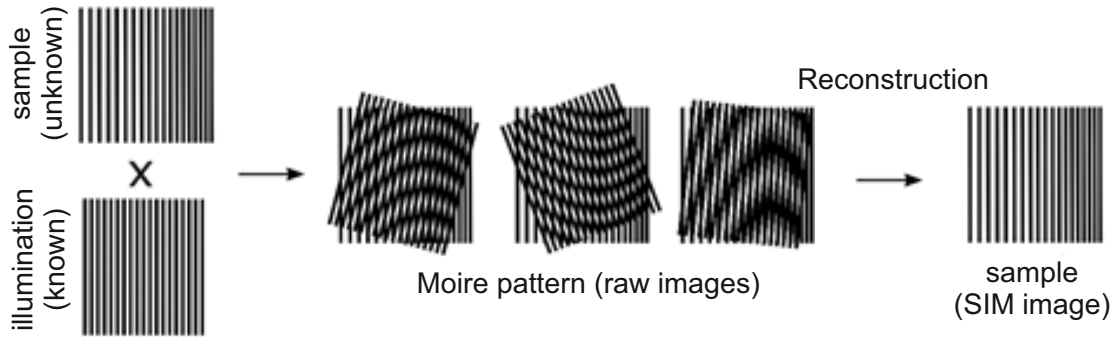


Figure 3.8: Diagram explaining the Moirè pattern generation and SIM concept. Reproduced from [134].

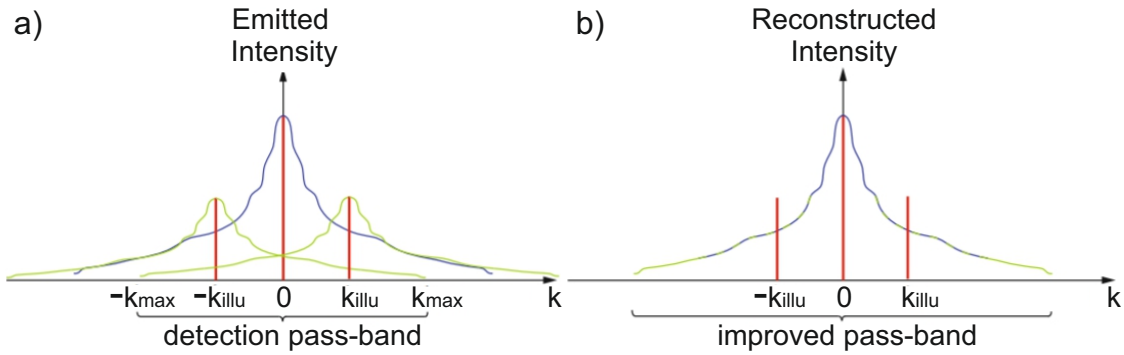


Figure 3.9: Diagram explaining the resolution enhancement of SIM. Reproduced from [4]. a) Schematic illustration of the Fourier transform of the Moirè pattern. b) Combination of all Fourier components properly reweighted and placed in correct positions.

In structured illumination microscopy (SIM) [66], a fluorescent sample is typically illuminated by a sinusoidal fringes pattern, which is the result of interference of two beams at the focal plane of the objective. The interfering beams are diffraction orders of a grating, which is generated by a spatial light modulator. An SLM controls the parameters of the pattern such as periodicity, orientation, and phase shift.

A signal emitted by the sample is given by

$$I_s(r) = S(r) \cdot I_{illu}(r), \tag{3.12}$$

where $S(r)$ is the sample density and $I_{illu}(r)$ is the sinusoidal illumination pattern. Fig. 3.8 illustrates the concept of SIM that is based on the Moirè effect. When a fine pattern of the sample density is superposed with the illumination fine pattern,

a third pattern of lower periodicity than the first two is produced. The latter is the Moirè pattern. After the Moirè pattern passes through the imaging system, some information about high spatial frequencies of $S(r)$ can be restored from the image, given that $I_{illu}(r)$ is known. Multiple Moirè patterns obtained with different illumination patterns allow to reconstruct a full image of the sample with up to twofold resolution enhancement beyond the diffraction limit.

Fig. 3.9 explains the resolution enhancement of SIM [136]. The Fourier transform of the Moirè pattern is given by

$$\tilde{I}_s(k) = \tilde{S}(k) * \tilde{I}_{illu}(k), \tag{3.13}$$

where $*$ denotes the convolution operation. The imaging system acts as a low pass filter and cuts spatial frequencies higher than $k_{max} = \frac{2NA}{\lambda}$, where λ is the wavelength of light. The Fourier transform of the illumination pattern $\tilde{I}_{illu}(k)$ corresponds to three delta peaks. The convolution of $\tilde{S}(k)$ and $\tilde{I}_{illu}(k)$ is represented by the sum of three objects shown in Fig. 3.9(a). The central object (blue) corresponds to $\tilde{S}(k)$ itself, the two side objects (green) correspond to $\tilde{S}(k)$ rescaled and displaced by $\pm k_{illu}$. Due to the displacement of the side objects, higher spatial frequencies of the sample are shifted inside the passband of the imaging system. By processing multiple Moirè patterns, these frequencies can be properly reweighted and placed in correct positions in the Fourier space. Thus, observable spatial frequencies are extended from k_{max} to $k_{max} + k_{illu}$ (see Fig. 3.9(b)). However, spatial frequencies of the illumination pattern are limited by the diffraction limit as well. Hence, SIM is able to image spatial frequencies up to 2 times higher than conventional imaging.

The twofold resolution enhancement limit has been surpassed by saturated structured illumination microscopy (SSIM) [65]. In SSIM the peak intensity of the illumination pattern exceeds the saturation intensity, which results in a non-linear response of the emission rate on the illumination near maximums. Since the response of a single fluorophore is no longer sinusoidal, the Fourier transform of the effective illumination pattern contains spatial frequencies nk_{illu}, where n is an integer number. In principle, n can be infinite. In practice, however, a number of

spatial harmonics, which are strong enough to contribute to imaging, is limited. Experimentally, SSIM achieves the lateral resolution of about 50 nm.

3.2.4 Localization microscopy

Localization microscopy is a branch of fluorescence-based techniques that achieve super-resolution by isolating individual emitters and measuring their positions with high precision. Localization microscopy includes photoactivated localization microscopy (PALM) [20], fluorescence photoactivation localization microscopy (FPALM) [75], and stochastic optical reconstruction microscopy (STORM) [133]. Principles of localization microscopy have been awarded by the 2014 Nobel Prize in Chemistry together with STED.

All three techniques (PALM, FPALM, and STORM) were published in 2006 and share a common operation principle. Generally, localization microscopy relies on specific labels (molecules attached to a sample) that change their fluorescent properties after absorption of a photon at a certain wavelength: they switch from their dark (off-) state to fluorescent (on-) state. A wide-field illumination of a sample consists of two lasers. One is a weak activation laser that stochastically changes the fluorescent state of a small portion of molecules from off- to on-state. Another laser illuminates the fluorescent transition of the on-state. The activated molecules fluoresce for some time and then return to the dark state or bleach.

The basic idea is that only sparse emitters are activated at a time so that the distance between emitters in the on-state exceeds the Rayleigh distance. Each emitter appears in the microscope like a diffraction-limited spot, which, however, can be distinguished from others. The precision for determining the centre of an isolated diffraction-limited spot by fitting the spot to a Gaussian function significantly exceeds the precision for distinguishing two close spots, which is bounded by the Rayleigh limit. The localization precision σ_{loc} scales as $\sigma_{loc} \approx \sigma/\sqrt{N}$, where N is the number of detected photons from the emitter and σ is the width of the spot in the image plane.

A full image of the sample is reconstructed from several thousands of activation and localization cycles using one of the following approaches. The first approach is to plot Gaussian intensity distributions of width that corresponds to the localization precision at the determined positions. Another approach is to define a virtual pixel grid of the desired pixel size and assign each pixel with the intensity value equal to the number of localized emitters inside the pixel area. The aforementioned approaches result in smoother and more accurate images than images obtained by drawing a dot of a certain size and uniform intensity at each position.

The resolution of localization microscopy depends not only on the localization precision, which sets the best achievable resolution, but also on the number of points sampling the object. According to the Nyquist-Shannon theorem, the average distance between the detected emitters must be smaller than half the desired resolution. Thus, higher resolution of images requires longer acquisition times since only a few emitters are imaged per one activation-localization cycle. In practice, localization microscopy achieves the lateral resolution from 10 to 60 nm. The resolution of 60 nm can be obtained at a rate of 30 sec per image [146]. An imaging session for higher resolution often takes more than 5 min and sometimes more than 30 min [183].

The localization microscopy techniques differ in the type of labels they are based on. PALM and FPALM use photo-switchable protein fluorophores, which bleach after activation, while STORM relies on photo-switchable synthetic dyes, which can be switched between on- and off-states many times. The two main features of labels that affect the resolution of localization microscopy are brightness (i.e. the number of photons per molecule in the on-state) and contrast (i.e. the ratio between the brightness of the on-state and background signal from the off-state). Each type of labels has its own advantages and weaknesses. Protein fluorophores provide better contrast but are less bright than synthetic dyes. Synthetic dyes have limited applicability, imposing difficulties in labelling specific samples, however, they are promising tools for multi-colour imaging due to their diversity, which implies flexibility in the activation and emission wavelengths.

Among the general limitations of localization microscopy is the tradeoff between the resolution and duration of the imaging session. Long sessions make the imaging process vulnerable to sample drifting and require expensive microscope stages, the usage of which reduces the drift significantly, however, does not eliminate it. A shortcoming of localization microscopy is the photo-damage of biological samples caused by the powerful excitation laser. This shortcoming, however, is common for all super-resolution techniques based on fluorescence [80].

3.2.5 Rotating coherent scattering microscopy

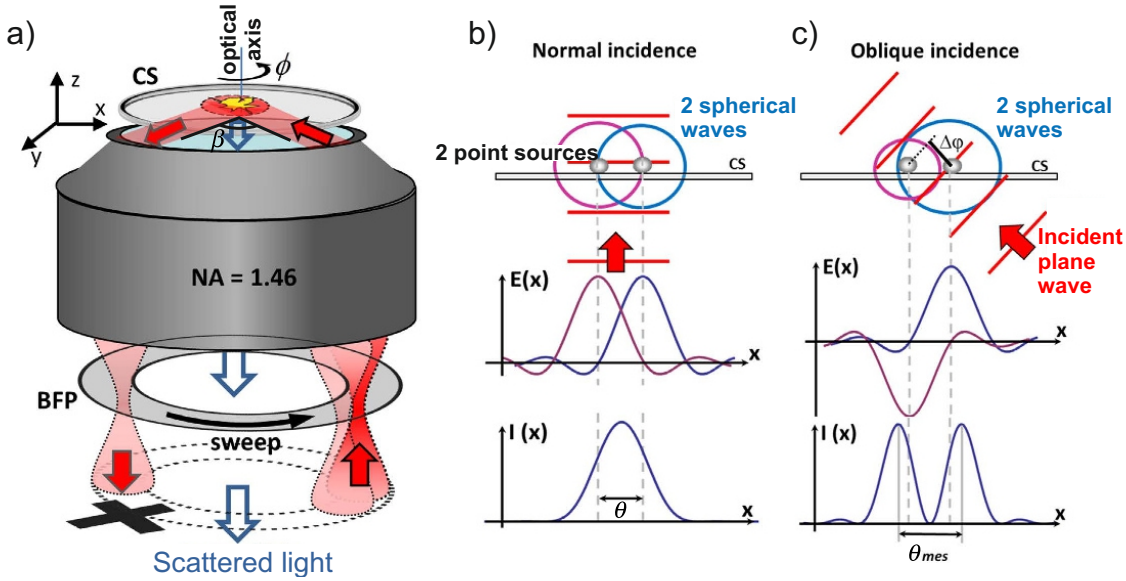


Figure 3.10: Schematic of ROCS microscopy. Reproduced from [85]. a) A coverslip (CS) with the sample (yellow spot) is located in the back focal plane of a microscope objective. The laser beam focused by the objective experiences total internal reflection from the coverslip, providing oblique illumination of the sample. The reflected light is blocked to enable dark-field imaging. A camera collects light that is scattered by the sample. While the polar angle β of illumination remains constant, the azimuthal angle ϕ sweeps values from 0 to 2π, such that the sample is illuminated from all possible directions during the integration time of the camera. b) The sample is illuminated under the normal incidence. In this case, image electric fields $E(x)$ of the two points of the sample, which are separated by distance θ, interfere constructively, and these two points can not be resolved on the camera. c) Due to oblique illumination, a phase delay of $\Delta\varphi \approx \pi$ is introduced between the two points, which makes them resolvable.

Far-field super-resolution imaging is widely represented but not limited to fluorescence-based techniques. Rotating coherent scattering (ROCS) microscopy

achieves resolution enhancement by illuminating a non-fluorescent sample with oblique coherent illumination [85], as shown in Fig. 3.10(a). The light scattered downwards by the sample forms an image on the camera.

The super-resolution of ROCS microscopy is based on the fact that the ability to distinguish two coherent point sources in the image plane depends not only on the distance between them but also on the relative phase. Indeed, when two coherent sources are in phase (see Fig. 3.10(b)), they interfere constructively in the image plane and become less distinguishable than incoherent sources. However, if a relative phase between the coherent sources is π, the destructive interference makes them more distinguishable than incoherent sources, for which the Rayleigh limit is defined. By using oblique illumination, one introduces phase shifts between adjacent points of the sample, such that the two points separated by distance $\theta \sim \theta_R$, where θ_R is the Rayleigh limit for incoherent light, are exited with a relative phase of π and become resolvable (see Fig. 3.10(c)).

To introduce phase shifts for each direction in the sample, either a coverslip with the sample or illumination rotates around the optical axis (see Fig. 3.10(a)). The angle of rotation ϕ is scanned from 0 to 2π in a few milliseconds. The time of the scan is matched with the integration time of the camera. Therefore, an image produced by the camera is an incoherent superposition of coherent images of the sample obtained for each illumination direction. While a single coherent image is full of speckles, the final ROCS image is almost free of artefacts and of sub-Rayleigh resolution.

The mathematical expression for the ROCS image is as follows. The electric field image of jth point source for a given illumination direction ϕ is $E_j(r,\phi) = \psi_j(r)e^{-i\varphi_j(\phi)}$, where $\psi(r)$ is the PSF, $\psi_j(r) = A_j\psi(r - r_j)$, r_j is the position of jth source, A_j is the amplitude, and $\varphi_j(\phi)$ is the phase of the source due to oblique illumination. Thus, the intensity distribution measured by the camera is

$$\begin{aligned} I(r) &= \frac{1}{2\pi}\int_0^{2\pi}\left|\sum_j E_j(r,\phi)\right|^2 d\phi \\ &= \sum_j |\psi_j|^2 + \frac{1}{\pi}\sum_{j\neq l}|\psi_j||\psi_l|\int_0^{2\pi}\cos\left[\varphi_j(\phi) - \varphi_l(\phi)\right]d\phi. \end{aligned} \tag{3.14}$$

The first term in the second line of Eq. (3.14) is equivalent to the incoherent image of the object, while the second term contains a super-resolving interference pattern. By taking the Fourier transform of the ROCS intensity image, we obtain the image spectrum:

$$\widetilde{I}(k) = \mathrm{AC}[F(k - k_{illu}) \cdot \psi(k)], \tag{3.15}$$

where AC[·] is the autocorrelation function, $F(k)$ is the object spectrum, k_{illu} is the illumination wave vector, and $\psi(k)$ is the optical transfer function of the microscope objective.

From Eq. (3.15) follows that, at best, ROCS can achieve twofold resolution enhancement beyond the Rayleigh limit. The explanation for this is analogous to that provided for SIM. Oblique illumination shifts observable spatial frequencies of the object (i.e spatial frequencies selected by the optical transfer function) by the incident wave vector k_{illu}. Thus, ROCS microscopy can retrieve spatial frequencies up to $k_{max} + k_{illu} = 2k_{max}$, where k_{max} is the Rayleigh limit on spatial frequencies, which is defined by the waist of the optical transfer function. The experimentally reported resolution is about 150 nm [85].

4 | Hermite-Gaussian microscopy: background

The review of super-resolution techniques in the previous chapter demonstrates that all of them are powerful imaging tools that have, however, constrained applicability. The main limitation of near-field techniques is the short working distance. Far-field techniques rely on nonlinear properties of samples, which narrows down their application area to biomedical studies. Most of the far-field techniques require labelling of the sample with bright fluorophores before imaging. Thus, a far-field non-invasive super-resolution imaging technique that can recover information about fine features of the sample from the emitted light is desired. Such a technique would contribute to the development of areas where super-resolution is not yet available, for example, astronomy.

In 2015, Tsang *et al.* proposed a theory of far-field linear-optical super-resolution technique for distance estimation between two weak incoherent point sources [162]. He approached the problem by using statistical analysis and the concept of quantum Fisher information. Using the latter, he derived the fundamental quantum limit of the distance estimation precision and demonstrated that this limit surpasses the conventional imaging precision infinitely, in theory. In other words, he proved that the estimation technique exists that is not subject to Rayleigh curse. Moreover, he proposed a method of spatial-mode demultiplexing (SPADE) that achieves the fundamental quantum limit.

This chapter starts with the review of Tsang's theory, considering that his ideas have inspired our work. Further, the chapter describes generalizations of this theory including Hermite-Gaussian microscopy (HGM), which is the theoretical base of

our experimental work. The theory of HGM has been first described in the paper [175] of our group, which is a predecessor of the present experiment. The chapter ends with the review of experimental works based on SPADE.

4.1 Introduction to the Fisher information

4.1.1 Classical Fisher information

The goal of parameter estimation problems in statistics is to determine unknown parameter θ, which defines the distribution of a random variable X, using a finite set of outcomes of X. The *Fisher information* is a measure of the amount of information that each outcome carries about the parameter. It is given by

$$\mathcal{K}(\theta) = E_\theta \left[\left(\frac{\partial}{\partial \theta} \log f(X|\theta) \right)^2 \right] = \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial \theta} \log f(x|\theta) \right]^2 f(x|\theta) dx, \tag{4.1}$$

where $f(X|\theta)$ is the probability density function of random variable X (i.e. such a function that $f(x)dx$ gives the probability of $X \in [x, x+dx]$) conditioned on θ and log is the natural logarithm. Intuitively, the Fisher information measures the sensitivity of the probability distribution to changes in θ. A higher value of the Fisher information for a sharply peaked f to θ reflects the fact that, for the sharper f, the estimation of θ with a given precision requires less samples of X, *i.e* each sample carries more information.

The inverse of the Fisher information defines another important quantity: the *Cramer-Rao bound*. Let us denote a rule applied to the observed data to produce an estimate as $\check{\theta}$. In statistics, such a rule is called *an estimator*. If the mathematical expectation of an estimator corresponds to the true value of the parameter, this estimator is *unbiased*. The Cramer-Rao bound expresses the lower bound on the variance of any unbiased estimator of the parameter:

$$\mathrm{Var}(\check{\theta}) \geq \frac{1}{N\mathcal{K}(\theta)}, \tag{4.2}$$

where N is the number of observations. If an unbiased estimator achieves the Cramer-Rao bound, this estimator is *efficient*.

Note that the classical Fisher information and Cramer-Rao bound set the upper bound on the precision for a given measurement (the best precision among estimators can not be higher than the Cramer-Rao bound). However, they depend on the choice of measurement, which defines the probability density function $f(x|\theta)$.

4.1.2 Quantum Fisher information

In the quantum formalism of parameter estimation, the probability density function is replaced by the density matrix $\hat{\rho}(\theta)$. The quantum Fisher information is defined as

$$\mathcal{K}_Q = Tr[\hat{\rho}\mathcal{L}_\theta^2], \tag{4.3}$$

where $\mathcal{L}_\theta$ is the symmetric logarithmic derivative given by the equation

$$\frac{\partial}{\partial\theta}\hat{\rho} = \frac{1}{2}(\hat{\rho}\mathcal{L}_\theta + \mathcal{L}_\theta\hat{\rho}). \tag{4.4}$$

Writing the density matrix in the eigenbasis as

$$\hat{\rho} = \sum_i p_i \left|e_i\right\rangle \left\langle e_i\right|, \tag{4.5}$$

the symmetric logarithmic derivative takes the following form:

$$\mathcal{L}_\theta = \sum_{i,j} \frac{2}{p_i + p_j} \left\langle e_i\right| \frac{\partial\hat{\rho}}{\partial\theta} \left|e_j\right\rangle \left|e_i\right\rangle \left\langle e_j\right| \equiv \sum_{i,j} \mathcal{L}_\theta^{ij} \left|e_i\right\rangle \left\langle e_j\right|. \tag{4.6}$$

Unlike the classical Fisher information, which is a function of measurement, the quantum Fisher information is optimized over all possible measurements. The inverse of the quantum Fisher information defines the quantum Cramer-Rao bound, which provides the fundamental limit on the achievable variance in parameter estimation.

4.2 Super-resolution for two incoherent point sources

4.2.1 Estimation of the separation

The paper of Tsang *et al.* considers the Rayleigh limit in the context of parameter estimation problem. The distance θ between the two incoherent point sources is the parameter to be estimated. Since the estimation of the centroid is independent from

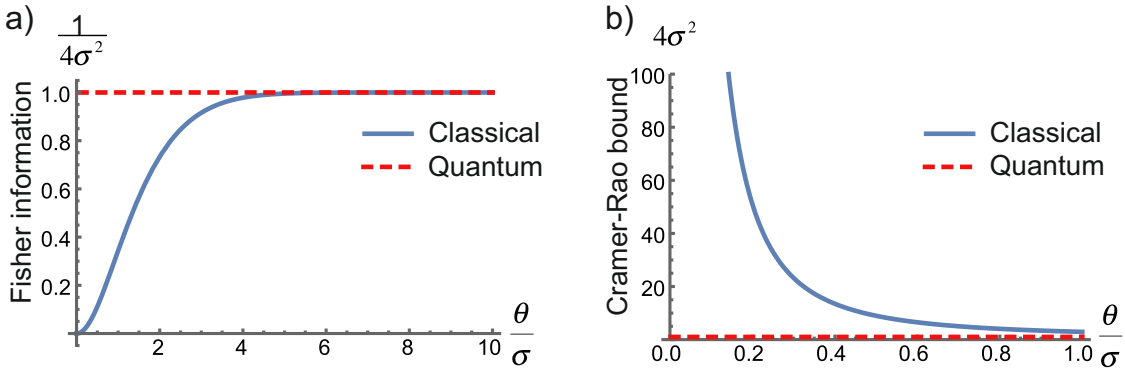


Figure 4.1: Fisher information per photon (a) and Cramer-Rao bound (b) versus the separation between two incoherent sources for a Gaussian PSF.

the separation estimation and is not affected by the Rayleigh curse, we assume the centroid to be known and equal to zero. Let us consider the separation estimation task in the classical and quantum domain.

Classical imaging can be viewed as spatially resolved photon counting in the image plane. Position x of a photon is the observable random variable. The probability density function of x is given by

$$\Lambda(x) = \frac{1}{2}\left(\left|\psi\left(x - \frac{\theta}{2}\right)\right|^2 + \left|\psi\left(x + \frac{\theta}{2}\right)\right|^2\right), \tag{4.7}$$

where $\psi(x)$ is the electric field PSF of the imaging system and $\pm\frac{\theta}{2}$ are positions of the sources. In the photon-counting representation, functions

$$\psi_1 \equiv \psi\left(x - \frac{\theta}{2}\right), \ \psi_2 \equiv \psi\left(x + \frac{\theta}{2}\right) \tag{4.8}$$

have the meaning of image-plane wave functions from each source. Using Eq. (4.1) and Eq. (4.7), we obtain the classical Fisher information per photon:

$$\mathcal{K}(\theta) = \int_{-\infty}^{\infty} \frac{1}{\Lambda(x)}\left[\frac{\partial\Lambda}{\partial\theta}\right]^2 dx. \tag{4.9}$$

The quantum formalism requires the quantum optical description of quasi-monochromatic incoherent weak point sources. Due to incoherence, light arriving on the image plane is represented by a multimode thermal state, the density matrix of which is given by

$$\hat{\rho} = \sum_{n=0}^{\infty} \pi_n \hat{\rho}_n, \tag{4.10}$$

where $\hat{\rho}_n$ is n-photon multimode Fock state and π_n is the probability of this state. Note that $\hat{\rho}$ describes the image field generated by both sources. Since the two sources are weak, the average number of photons ϵ arriving on the image plane within a short coherence time interval is much smaller than 1. This means that multiphoton events can be neglected and the density matrix within each coherence time interval becomes

$$\hat{\rho} = (1-\epsilon)\hat{\rho}_0 + \epsilon\hat{\rho}_1 + \mathcal{O}(\epsilon^2). \tag{4.11}$$

The single-photon state contains all the information about the unknown parameter θ. The density matrix of the single-photon state in the image plane can be represented as

$$\hat{\rho}_1 \approx \frac{1}{2}(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|), \tag{4.12}$$

where

$$|\psi_s\rangle = \int_{-\infty}^{\infty} \psi_s(x)|x\rangle\, dx,\ s = 1, 2, \tag{4.13}$$

$|x\rangle$ is the one-photon position state in the image plane, and $\psi_s(x)$ is defined in Eq. (4.8).

Given that the vacuum state does not contain any information and multiphoton events are neglected, from Eq. (4.3), Eq. (4.11), and Eq. (4.12) follows that the quantum Fisher information per photon is (see appendix A for the derivation)

$$\mathcal{K}_Q \approx Tr[\hat{\rho}_1\mathcal{L}_\theta^2] = \Delta k^2, \tag{4.14}$$

where

$$\Delta k^2 = \int_{-\infty}^{\infty} \left[\frac{\partial\psi(x)}{\partial x}\right]^2 dx. \tag{4.15}$$

If the PSF is a Gaussian function given by

$$\psi(x) = \frac{1}{\sqrt[4]{2\pi\sigma^2}} e^{-\frac{x^2}{4\sigma^2}}, \tag{4.16}$$

the quantum Fisher information per photon becomes

$$\mathcal{K}_Q \approx \frac{1}{4\sigma^2}. \tag{4.17}$$

Fig. 4.1 compares the classical and quantum Fisher information and Cramer-Rao bound. The integral in Eq. (4.9), which gives the classical Fisher information

(blue line), has been calculated numerically assuming a Gaussian PSF given by Eq. (4.16). As seen, for well-separated sources ($\theta \gg \sigma$) classical imaging is quantum optimal, while in the limit of $\frac{\theta}{\sigma} \to 0$ the classical Fisher information vanishes and the Cramer-Rao bound diverges revealing the Rayleigh curse. Remarkably, the quantum Fisher information and Cramer-Rao bound remain constant even for infinitesimally small separations, which means that the Rayleigh limit can be avoided for separation estimation.

4.2.2 Spatial-mode demultiplexing

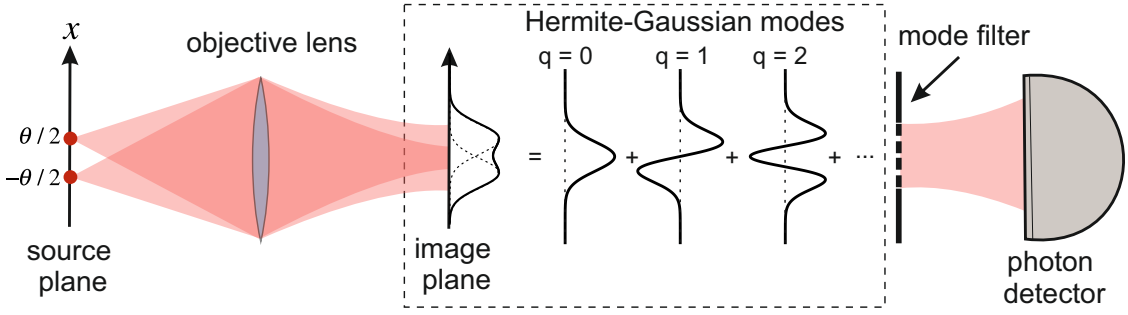


Figure 4.2: Illustration of a basic idea of spatial-mode demultiplexing. The central part shows the electric field profiles of few one-dimensional Hermite-Gaussian modes.

In their work, Tsang *et al.* proposed a quantum optimal measurement for separation estimation, named spatial-mode demultiplexing (SPADE). The idea of the method is to extract additional information from spatial correlations of the optical field in the image plane via demultiplexing this field into the Hermite-Gaussian modes basis. In contrast to SPADE, classical imaging ignores spatial correlations and measures only the intensity distribution of the field in the image plane.

Fig. 4.2 schematizes the idea of SPADE. Eigenvectors of the Hermite-Gaussian modes basis are given by

$$|\phi_q\rangle = \int_{-\infty}^{\infty} \phi_q(x)\,|x\rangle\,dx, \tag{4.18}$$

where

$$\phi_q(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/4} \frac{1}{\sqrt{2^q q!}} H_q\left(\frac{x}{\sqrt{2}\sigma}\right) \exp\left(-\frac{x^2}{4\sigma^2}\right),\ q = 0, 1, 2... \tag{4.19}$$

H_q is the Hermite polynomial. The positive-operator valued measure, which represents demultiplexing into the Hermite-Gaussian modes basis, can be expressed as projections

$$E_0 = |\text{vac}\rangle \langle \text{vac}|, \; E_q = |\phi_q\rangle \langle \phi_q|. \tag{4.20}$$

Considering that the density matrix of the single-photon state in the image plane is given by Eq. (4.12), the probability to detect a photon in qth mode is

$$P(q) = \frac{1}{2}(|\langle \phi_q|\psi_1\rangle|^2 + |\langle \phi_q|\psi_2\rangle|^2). \tag{4.21}$$

Functions $\phi_q(x)$ describing Hermite-Gaussian modes (see Eq. (4.19)) are mathematically equivalent to position wave functions of Fock states. For a Gaussian PSF (see Eq. (4.16)), $\psi_1(x) \equiv \psi(x - \theta/2)$ and $\psi_2(x) \equiv \psi(x + \theta/2)$ are equivalent to wave functions of coherent states with displacements $\pm\theta/(4\sigma)$. Utilizing the equivalence of $|\phi_q\rangle$ and $|\psi_s\rangle$ for $s = 1, 2$ to quantum optics states, we evaluate the overlaps in Eq. (4.21) and obtain that $P(q)$ becomes

$$P(q) = \exp(-Q)\frac{Q^q}{q!}, \; Q \equiv \frac{\theta^2}{16\sigma^2}. \tag{4.22}$$

Finally, the classical Fisher information per photon for the SPADE measurement is

$$\mathcal{K}(\theta) = \sum_{q=0}^{\infty} P(q)\left[\frac{\partial}{\partial\theta}\log P(q)\right]^2 = \frac{1}{4\sigma^2}, \tag{4.23}$$

which is equal to the quantum Fisher information given by Eq. (4.17). Thus, SPADE achieves the quantum Cramer-Rao bound of precision for the separation estimation between two incoherent point sources.

Suppose that L is the total amount of photons in the image plane and $\{m_q, q = 0, 1, ...\}$ is the set of photon counts in each Hermite-Gaussian mode. The maximum-likelihood estimator, which one needs to apply to this data for the separation estimation, is (see appendix B for the derivation)

$$\check{Q} = \frac{1}{L}\sum_q q m_q, \; \check{\theta} = 4\sigma\sqrt{\check{Q}}. \tag{4.24}$$

The estimator asymptotically approaches the Cramer-Rao bound in the limit of a large number of trials to detect a photon.

Since the Rayleigh curse occurs only for small $\frac{\theta}{\sigma}$, and only low-order Hermite-Gaussian modes are exited in this case, SPADE can be simplified to counting photons only in the first mode ($q = 1$). The Fisher information for this measurement is

$$\mathcal{K}(\theta) = \frac{1}{4\sigma^2} \exp(-Q)(Q-1)^2, \tag{4.25}$$

where Q is defined in Eq. (4.22). The maximum-likelihood estimator becomes

$$\check{Q} = -\log \frac{m_0}{L}, \; \check{\theta} = 4\sigma\sqrt{\check{Q}}. \tag{4.26}$$

In this case, the Fisher information is not constant and decreases when the sources are far apart. However, this measurement overcomes the Rayleigh limit for small separations.

The advantage of SPADE over direct imaging has the following intuitive interpretation. Two close sources emit primarily to the fundamental Gaussian mode, which is superposed with higher-order modes in the image plane. Thus, for any small separation, the intensity distribution in the image plane measured by direct imaging is almost indistinguishable from a Gaussian intensity profile. In SPADE, however, photons are counted in every mode separately enabling a sensitive *null measurement*: it is significantly simpler to distinguish between the absence of photons in higher-order modes, which corresponds to zero separation, and a small number of photons than between small deviations from a powerful zeroth-order background.

The above analysis is provided assuming a Gaussian PSF. Work [130] shows that the quantum Fisher information is constant for any spatially symmetric PSF. Moreover, for a real symmetric PSF, demultiplexing into any complete basis of real transverse modes with well-defined parity, which means that $\langle -x|\phi_q\rangle = \pm\langle x|\phi_q\rangle$, achieves the quantum Cramer-Rao bound.

For a more comprehensive review of the problem of distinguishing the two point sources with SPADE, let us consider the case of coherent and partially coherent light. The influence of the coherence between the two sources on the SPADE performance was investigated in [161]. The degree of coherence between the sources is defined as $\gamma = \langle E_1^* E_2\rangle / I_0$, where E_1 and E_2 are random variables representing the electric

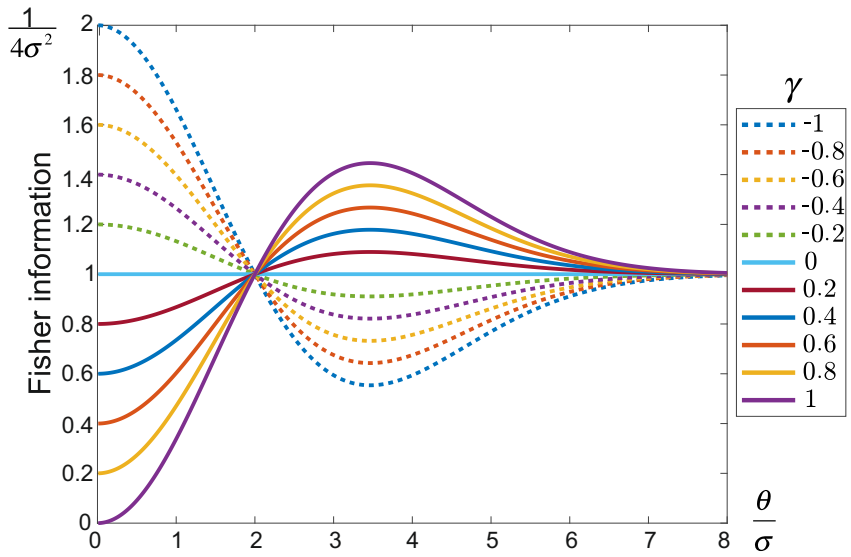


Figure 4.3: Per-photon Fisher information of the SPADE measurement in Hermite-Gaussian modes up to $q = 20$ for two partially coherent point sources. The legend specifies the degree of coherence. Reproduced from [161].

field amplitudes of the point sources, $\langle|E_1|^2\rangle = \langle|E_2|^2\rangle = I_0$. Thus, $\gamma = 0$ for two incoherent sources, $\gamma = 1$ for two coherent sources in phase, and $\gamma = -1$ for two coherent sources out of phase. Fig. 4.3 shows the classical Fisher information for demultiplexing into Hermite-Gaussian modes up to $q = 20$ for various degrees of coherence γ. As seen, for anticorrelated sources ($\gamma < 0$), the Fisher information does not show a sign of the Rayleigh curse. Moreover, there is an information enhancement for close anticorrelated sources, which can be intuitively explained by the enhanced coupling into the first Hermite-Gaussian mode ($q = 1$). For correlated sources ($\gamma > 0$), the Fisher information vanishes only when $\gamma = 1$, while for all other values of γ SPADE overcomes the Rayleigh limit.

4.3 Implementation options of spatial-mode demultiplexing

4.3.1 Mode-sorting

The decomposition of the optical signal into a transverse modes basis is called *mode-sorting* and this is the most challenging part of SPADE. In their work Tsang *et al.* proposed, for example, to use a multimode waveguide with a grating coupler, which

couples each mode of the image field into a free-space plane wave with a specific propagation angle. Another idea was to use evanescent coupling into single-mode waveguides fabricated to have propagation constants equal to those of modes in the multimode waveguide. However, both ideas are unfeasible with the current precision of lithography and etching technology used to fabricate waveguides and Bragg grating couplers. There are more practical ways to implement mode-sorting.

Sorting of modes with AOM using fork-shaped holograms. The first experimental implementations were demonstrated for the sorting of modes carrying orbital angular momentum (OAM). Methods relied on diffractive phase plates or holograms of a so called forked shape [177]. A forked-shape hologram appears when a radial phase gradient $l\varphi$, where φ is the angular coordinate and l is an integer number, is added to a blazed-grating phase pattern and the modulo 2π is applied. If a plane wave enters such a hologram, the hologram produces a desired mode with OAM of l in the first diffraction order. The scheme also works in reverse. If a mode carrying OAM of l enters the hologram of $-l$, the hologram produces a Gaussian beam. If the output of the hologram is coupled to a single-mode fibre, only the mode with OAM of l couples efficiently, which enables mode-sorting. Phase holograms can be either fabricated or generated via a SLM. However, such holograms test just one state a time.

Sorting of modes with AOM using the Cartesian to log-polar coordinate transformation. A technique based on the Cartesian to log-polar coordinate transformation has improved the mode-sorting efficiency for modes carrying AOM [17]. The idea is to transform a circular phase gradient into a transverse phase gradient. The latter can be discriminated by a lens, which maps a transverse phase gradient into a position in the focal plane. An increment in the angular phase change of 2π guarantees low cross-talk between different AOM states for the following reason. An additional phase change of 2π across the lens aperture results in a difference in the focal plane position comparable with the Rayleigh resolution limit.

The technique requires two phase elements (SLMs). The first SLM followed by a Fourier-transforming lens of focal length f performs the following mapping $(x, y) \rightarrow (u, v)$:

$$v = a \arctan\left(\frac{y}{x}\right), \; u = -a \log\left(\frac{\sqrt{x^2+y^2}}{b}\right), \tag{4.27}$$

where (x, y) and (u, v) are the Cartesian coordinate systems in the input and output plane, which is the Fourier plane of the first SLM, parameters a and b are explained below. The phase profile of the SLM is given by

$$\Theta_1(x, y) = \frac{2\pi a}{\lambda f}\left[y \arctan\left(\frac{y}{x}\right) - x \log\left(\frac{\sqrt{x^2+y^2}}{b}\right) + x\right]. \tag{4.28}$$

The mapping $(x, y) \rightarrow (u, v)$ corresponds to the transformation of an input image that consists of concentric circles into an output image of parallel lines. The parameter a (in units of length) scales the transformed image, and b translates the image in the u direction. The image transformation causes a phase distortion due to the difference in optical paths of rays composing the transforming beam. The second SLM, which is located in the output (Fourier) plane of the first one, with a phase profile given by

$$\Theta_2(u, v) = -\frac{2\pi ab}{\lambda f} \exp\left(-\frac{u}{a}\right) \cos\frac{v}{a} \tag{4.29}$$

corrects for the distortion. Thus, the two SLMs transform beams with helical phase gradients to beams with linear transverse gradients, which can be further spatially separated by a lens. The technique was demonstrated for mode-sorting of eleven OAM states.

Sorting of the Hermite-Gaussian modes using a mode converter. The methods described above are not directly applicable to mode-sorting of the Hermite-Gaussian modes. However, in 2018 Robert Boyd's group implemented a Hermite-Gaussian mode sorter [184] using an *astigmatic mode converter* (AMC), which consists of two cylindrical lenses spaced by $\sqrt{2}f$, where f is focal length. AMC performs the conversion of a two-dimensional Hermite-Gaussian mode $HG_{m,n}(x, y) \propto$

$\phi_m(x)\phi_n(y)$ (see Eq. (4.19) for $\phi_q(x)$), which is characterized by the indices m and n along x and y directions, into a Laguerre-Gaussian mode LG_p^l with the radial p and azimuthal l indices given by $p = \min(m, n)$ and $l = m - n$. The index l defines AOM of the mode.

The operating principle of AMC can be understood from the following description [15]. The decomposition of a Laguerre-Gaussian mode $LG_{\min(m,n)}^{(m-n)}$ into the Hermite-Gaussian modes basis is

$$LG_{\min(m,n)}^{(m-n)}(x, y) = \sum_{k=0}^{n+m} i^k b(n, m, k) HG_{m+n-k,k}(x, y), \tag{4.30}$$

where $b(n, m, k)$ are real coefficients. It has been noticed that a Hermite-Gaussian mode rotated by 45° to (x, y) axes can be decomposed into the same set of Hermite-Gaussian modes:

$$HG_{m,n}\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right) = \sum_{k=0}^{n+m} b(n, m, k) HG_{m+n-k,k}(x, y) \tag{4.31}$$

with the same coefficients $b(n, m, k)$ as $LG_{\min(m,n)}^{(m-n)}$. The only difference is that the components of the decomposition in Eq. (4.30) are summed with relative phases $\frac{\pi}{2}k$, which follows from the phase factor i^k.

If a 45°-oriented Hermite-Gaussian mode enters AMC, AMC rephases the terms in the decomposition (Eq. (4.31)) by making the beam astigmatic in a confined region between the two cylindrical lenses, axes of which are aligned with (x, y) axes. Due to the astigmatism, Gouy phase accumulates differently for each transverse direction. The resulting phase accumulated by a mode $HG_{s,j}(x, y)$ can be expressed as $\left(s + \frac{1}{2}\right)\beta_x + \left(j + \frac{1}{2}\right)\beta_y$, where β_x and β_y are the phase contributions from the corresponding directions. In AMC, $\beta_y - \beta_x = \frac{\pi}{2}$, which is defined by the spacing between the lenses. Thus, components $HG_{m+n-k_1,k_1}(x, y)$ and $HG_{m+n-k_2,k_2}(x, y)$ in Eq. (4.31) acquire a phase difference of $\frac{\pi}{2}(k_2 - k_1)$ passing trough AMC, which stands for the proper rephasing. In practice, instead of rotating the input Hermite-Gaussian modes by 45°, one can rotate axes of the cylindrical lenses by 45° and obtain the conversion $HG_{m,n}(x, y) \to LG_{\min(m,n)}^{(m-n)}(x, y)$, after which the modes can be sorted by the index l with the above methods for modes with AOM.

The sorting by l is not sufficient for two-dimensional modes, and an additional sorting module is required. Work [184] utilizes a fractional Fourier transform (FRFT) module, which is placed before AMC, as an additional sorting mechanism. The module consists of two SLMs that modulate only the horizontal polarization and two spherical lenses. The two lenses implement FRFT of order α_1 to vertical polarization, while the horizontal polarization experiences FRFT of order α_2 passing through the module. A Hermite-Gaussian mode acquires a mode-dependent Gouy phase under FRFT of order α given by $-(m+n)\alpha$. Thus, if a diagonally polarized Hermite-Gaussian mode enters the module, a phase difference of $-(m+n)\Delta\alpha$, where $\Delta\alpha = \alpha_1 - \alpha_2$, accumulates between the vertical and horizontal polarization components. By measuring the polarization state at the output of the module, one can sort incoming modes by the value of $m+n$.

Work [184] demonstrated mode-sorting of 16 Hermite-Gaussian modes using the FRFT module, AMC, and OAM sorter based on the Cartesian to log-polar coordinate transformation.

Sorting transverse-spatial modes of any type using Multi-Plane Light Conversion. Mode-sorting methods based on Multi-Plane Light Conversion have been developing in parallel with methods designed for or reducing to the sorting of modes with OAM, such as Laguerre-Gaussian modes. In 2010, work [110] showed that a system of j phase elements $\{\Theta_s(x,y), s = 1, ..., j\}$ and free-space propagations between them, called a Multi-Plane Light Conversion (MPLC) system, can perform a unitary transformation of any transverse-spatial mode into a fundamental Gaussian mode ($HG_{0,0}$). The MPLC system allows one to implement mode-sorting, relying on the idea that only a fundamental mode couples efficiently into a single-mode fibre.

In a mode-sorter with one output channel, the MPLC system transforms a selected input mode into a Gaussian output mode. To filter another transverse-spatial mode from the input field, phase elements should be updated. A configuration with multiple spatially-separated output channels (an array of single-mode fibres) allows one to separate many input transverse-spatial modes simultaneously (each

input mode couples into an individual fibre after passing through the MPLC system) and make the mode-sorting procedure instantaneous. If such a mode-sorter is used in the reverse direction, it acts as a multiplexer, i.e., it takes N spatially separated fundamental Gaussian beams as the input and turns them into a superposition of N transverse modes. MPLC-based mode-sorters and multiplexers have been demonstrated on different types of transverse-spatial modes, for instance, the LP modes of a few-mode fibre [92, 61, 48], where [92] is the first demonstration, Hermite-Gaussian modes [9, 56, 76], and Laguerre-Gaussian modes [55].

As an example, let us consider a numerical procedure for calculating phase profiles for a mode-sorter with one output channel [76]. Phases of all elements are set to zero in the beginning. The input mode is numerically propagated forward to the last phase element, and the complex amplitude of the mode $A_s(x, y)$ is calculated at the plane of each element. Then, the output fundamental Gaussian mode is propagated backwards through the same system of phase elements. Starting with jth element, phase profiles of all elements are updated according to the following procedure. First, the overlap between the complex amplitudes of the forward and backwards propagating beams is calculated for the jth element:

$$o_j(x, y) = A_j^*(x, y)G_j(x, y)e^{i\Theta_j(x,y)}, \tag{4.32}$$

where $G_j(x, y)$ is the complex amplitude of the propagating backwards beam. Then, the phase profile of the jth element is adjusted by

$$\Delta\Theta_j(x, y) = -\arg[o_j(x, y)e^{-i\beta}] \tag{4.33}$$

to match phase distributions of the two beams up to the offset term β, which is the mean value of the transverse phase of $o_j(x, y)$. Phase $\Delta\Theta_j(x, y)$ is added to the phase of $G_j(x, y)$, the complex field $G_j(x, y)$ is propagated to the element $j - 1$, and the procedure is repeated (and so on until the first element). Due to the free space propagation, amplitudes of beams are also slowly adjusted between the phase elements.

For a large enough number of phase elements, one optimization run is sufficient for adjusting the set of phase elements to perform the transformation of the input mode into the fundamental mode. If the number of elements is limited, one can repeat the procedure for the forward propagation. Work [76] experimentally demonstrated mode-sorting of 9 Hermite-Gaussian modes using $j = 3$ phase elements displayed by 3 separate regions of a single SLM.

Conclusion. As seen from the given examples, Hermite-Gaussian mode-sorting is challenging to implement. The complexity rises with the number of modes to be distinguished. Schemes require several SLMs, which makes the setup expensive and rises the demand of fine calibration. As will be seen further, Hermite-Gaussian microscopy requires mode-sorting of a significantly higher number of modes (we used 441 modes) than has been demonstrated by the above experiments. We used heterodyne detection (the description follows) instead of mode-sorting for Hermite-Gaussian microscopy because of the simplicity of the former. However, we keep in mind that mode-sorting is a promising tool for future experiments.

MPLC-based mode-sorters attract special attention for the following reasons. A mode-sorter with multiple output channels provides the possibility of measuring all coefficients of the decomposition of the optical field into the Hermite-Gaussian basis simultaneously using photodetectors. With such a data acquisitions scheme, Hermite-Gaussian microscopy could become a high-frame-rate super-resolution technique. MPLC-based mode-sorters and multiplexers are under continuous development. Nowadays, commercial mode-sorters are available, for instance, Cailabs PROTEUS-C, which can sort up to 45 modes simultaneously. Recently, work [54] reported an MPLC-based multiplexer supporting 1035 Hermite-Gaussian modes.

4.3.2 Heterodyne detection

An alternative to mode-sorting is heterodyne detection with the local oscillator (LO) prepared in one of the Hermite-Gaussian modes. A heterodyne detector measures an overlap between the signal and LO mode, and the preparation of the desired

mode is significantly easier than its extraction from the total signal field. To prove the former statement, let us consider a basic theory of heterodyne detection.

Basics of heterodyne detection. Consider two beams, the signal and LO, of slightly different optical frequencies ω_{sig} and ω_{LO} and the same polarization that interfere on a 50:50 beamsplitter (BS). Describing the optical fields of the two beams as traveling waves $\tilde{E}_{sig} = A_{sig}\cos(\omega_{sig}t)$ and $\tilde{E}_{LO} = A_{LO}\cos(\omega_{LO}t + \beta)$, we obtain that optical fields at the outputs of the beamsplitter are

$$E_1 = \frac{1}{\sqrt{2}}(\tilde{E}_{sig} - \tilde{E}_{LO}),\ E_2 = \frac{1}{\sqrt{2}}(\tilde{E}_{sig} + \tilde{E}_{LO}). \tag{4.34}$$

Each output optical field is directed to the corresponding photodiode of the heterodyne detector that measures an electronic signal J_s for $s = 1, 2$ proportional to the power of the field: $J_1 \propto E_1^2$ and $J_2 \propto E_2^2$. An electronic circuit of the detector subtracts the two photocurrents and sends the result to the output, which corresponds to the output electronic signal given by $J \propto \tilde{E}_{sig}\tilde{E}_{LO} = A_{sig}A_{LO}\cos(\omega_{sig}t)\cos(\omega_{LO}t + \beta)$. The product of the cosines gives two components at the difference $\omega_{sig} - \omega_{LO}$ and sum $\omega_{sig} + \omega_{LO}$ frequencies. Since the sum frequency can not be resolved by the detector and its average is zero, the output electronic signal (also called the output photocurrent) is

$$J \propto A_{sig}A_{LO}\cos[(\omega_{sig} - \omega_{LO})t - \beta] \propto E_{sig}E_{LO}^* + c.c., \tag{4.35}$$

where the right-hand side is written in the form of complex electric fields $E_{sig} = A_{sig}e^{i\omega_{sig}t}$ and $E_{LO} = A_{LO}e^{i\omega_{LO}t+\beta}$, and $c.c$ denotes the complex conjugate.

If $\omega_{sig} = \omega_{LO}$, the measurement is called *homodyne* detection. Such a measurement is commonly applied in quantum optics to characterize quantum states of light. Let us provide a quantum optical description of homodyne detection for the future comparison of homodyne (heterodyne) detection with direct imaging and mode-sorting in terms of Fisher information.

In quantum optics, the complex amplitude of a quantum optical field is represented by the creation and annihilation operators denoted as $\hat{a}$ and $\hat{a}^\dagger$, respectively.

Assuming the LO to be in a high-amplitude coherent state described classically using its complex amplitude $\alpha_{LO} = |\alpha_{LO}|e^{i\beta}$ and describing the signal as a quantum field, we rewrite Eq. (4.35) as

$$J \propto \alpha_{LO}\hat{a}^{\dagger}_{sig} + \alpha^{*}_{LO}\hat{a}_{sig} \propto |\alpha_{LO}|(\hat{\mathrm{X}}\cos\beta + \hat{\mathrm{P}}\sin\beta). \tag{4.36}$$

$\hat{\mathrm{X}}$ and $\hat{\mathrm{P}}$ are quadratures of the signal field: $\hat{\mathrm{X}} = \frac{1}{2}(\hat{a}^{\dagger}_{sig} + \hat{a}_{sig})$ and $\hat{\mathrm{P}} = \frac{i}{2}(\hat{a}^{\dagger}_{sig} - \hat{a}_{sig})$. As seen, a homodyne detector measures the observable $\hat{X}_\beta = \hat{X}\cos\beta + \hat{P}\sin\beta$ with the amplification defined by the absolute value of the LO amplitude.

In quantum optics, a state of light in phase space spanned on quadratures $\hat{\mathrm{X}}$ and $\hat{\mathrm{P}}$ is described by the Wigner function $W(\mathrm{X}, \mathrm{P})$. Thus, measuring $\hat{X}_\beta$ for different relative phases between the signal and LO with homodyne detection, one can reconstruct the Wigner function of the optical signal.

If the signal and LO have a spatial transverse structure, the product $E_{sig}E^{*}_{LO}$ in Eq. (4.35) should be replaced by the overlap integral between the signal and LO electric field profiles, which follows from the fact that an electronic signal of each photodiode is proportional to the total optical power given by the integral over a spatial distribution of $|E_s|^2$ for $s = 1, 2$ (see Eq. (4.34) for E_s). In the quantum optical description, a homodyne detector measures $\hat{X}_\beta$ of the optical signal in the spatial mode that matches the mode of LO.

Heterodyne detection for separation estimation. Fig. 4.4(a) illustrates heterodyne detection for the separation estimation task in the simplest case when just the first Hermite-Gaussian mode is used. For this task, heterodyne detection is advantageous over homodyne detection due to lower flicker noise. Hermite-Gaussian modes are eigenmodes of a confocal cavity, which gives an option to prepare the LO in a given mode. The switch between different modes can be done by tuning the cavity length. Another option, which we chose in our experiment, is to use an SLM to generate the LO. See chapter 2 for details.

Consider a single point source located at position x_p. Then, the electric field distribution in the object plane is $E(x) \propto \delta(x - x_p)$. In the image plane, the electric

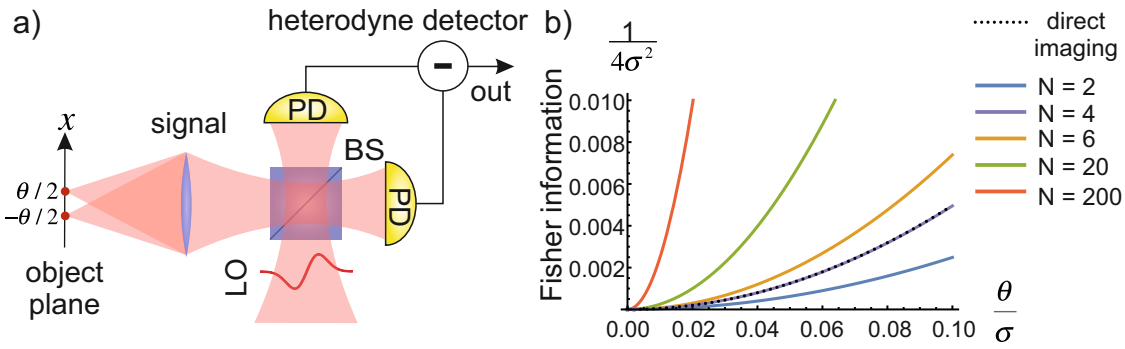


Figure 4.4: a) Schematic of heterodyne detection for separation estimation. The optical signal generated by two point sources is collected by an objective lens. The image field is overlapped with LO prepared in the first Hermite-Gaussian mode. The two beams interfere on a 50:50 beamsplitter (BS). Two photodiodes (PD) of the heterodyne detector measure photocurrents induced by the outputs of the beamsplitter. The output subtraction photocurrent is proportional to the overlap integral between the signal and LO. b) Per-photon Fisher information for separation estimation using heterodyne detection for different average photon numbers of each thermal source. Reproduced from [176].

field is given by the convolution with the imaging system PSF $\psi(x)$:

$$E'(x') = \int_{-\infty}^{\infty} E(x)\psi(x - x')dx = \psi(x' - x_p), \tag{4.37}$$

where x' is the spatial coordinate of the image plane. The output photocurrent generated by the heterodyne detector is proportional to the overlap integral between the signal and LO fields:

$$J(x_p) \propto \int_{-\infty}^{\infty} E'(x')E^*_{LO}(x')dx' = \int_{-\infty}^{\infty} \psi(x' - x_p)E^*_{LO}(x')dx'. \tag{4.38}$$

Assuming a Gaussian PSF and the LO prepared in the first Hermite-Gaussian mode, which means that $E_{LO}(x') = \phi_1(x')$ (see Eq. (4.19) for $\phi_1(x)$), the heterodyne detector photocurrent becomes

$$J(x_p) \propto \frac{1}{2\sigma} x_p e^{-\frac{x_p^2}{8\sigma^2}}. \tag{4.39}$$

The corresponding electronic power is

$$\mathcal{P}(x_p) = J^2(x_p) = \frac{1}{4\sigma^2} x_p^2 e^{-\frac{x_p^2}{4\sigma^2}}. \tag{4.40}$$

The signal power from two incoherent point sources located at $x = \pm\theta/2$ is

$$\mathcal{P}\left(\frac{\theta}{2}\right) + \mathcal{P}\left(-\frac{\theta}{2}\right) = \frac{\theta^2}{8\sigma^2} e^{-\frac{\theta^2}{16\sigma^2}} \approx \frac{\theta^2}{8\sigma^2}, \tag{4.41}$$

where the approximation is calculated using the Taylor expansion for small θ.

For comparison, direct imaging measures the intensity in the image plane given by

$$I(x') \propto \left|\psi\left(x' - \frac{\theta}{2}\right)\right|^2 + \left|\psi\left(x' + \frac{\theta}{2}\right)\right|^2. \tag{4.42}$$

Approximating $\left|\psi\left(x' \pm \frac{\theta}{2}\right)\right|^2 = |\psi(x')|^2 \pm \frac{\theta}{2}\frac{\partial}{\partial x'}|\psi(x')|^2 + \mathcal{O}(\theta^2)$, we get

$$I(x') \propto 2|\psi(x')|^2 + \mathcal{O}(\theta^2). \tag{4.43}$$

The difference between the heterodyne detector signal power and the direct intensity measurement given by Eq. (4.41) and Eq. (4.43), respectively, is as follows. In direct imaging, a term proportional to θ^2 is superposed with a strong zeroth-order background, while the measurement of the heterodyne detector signal in the first Hermite-Gaussian mode is a sensitive null measurement, which is a feature of SPADE.

Fisher information for separation estimation using heterodyne detection. In the previous paragraph, the optical signal coming from the object plane is assumed to be classical, and photon counting is not a part of the measurement. Our experiment is also performed with a classical signal. However, to demonstrate that heterodyne detection can provide the advantage over direct imaging in terms of Fisher information and, hence, to prove that the choice of heterodyne detection is relevant for our experiment, let us describe separation estimation with heterodyning in the quantum domain. Such a description can be found in work [176].

For a thermal state, which represents each incoherent point source, with average photon number N, the Wigner function is

$$W(\mathrm{X}, \mathrm{P}) = \frac{1}{\pi(2N+1)} \exp\left[-\frac{\mathrm{X}^2 + \mathrm{P}^2}{2N+1}\right]. \tag{4.44}$$

The average photon number in the first Hermite-Gaussian mode from one source is $\frac{\theta^2}{16\sigma^2}N$, which intuitively follows from the relative power of the mode in the total signal field given by the overlap integral between the signal and LO squared (Eq. (4.40) in which $x_p = \frac{\theta}{2}$ and the exponent is neglected). Hence, the average

photon number from both sources is $N' = \frac{\theta^2}{8\sigma^2}N$. First, let us suppose that we perform homodyne detection and measure X-quadrature. Substituting N' into Eq. (4.44) and integrating over P, we get the probability distribution of X in the first Hermite-Gaussian mode:

$$f(\mathrm{X}|\theta) = \frac{1}{\sqrt{\pi\left(\frac{\theta^2}{4\sigma^2}N + 1\right)}} \exp\left[-\frac{\mathrm{X}^2}{\frac{\theta^2}{4\sigma^2}N + 1}\right]. \tag{4.45}$$

Thus, the overall Fisher information for homodyne detection is

$$\mathcal{K}(\theta) = \left\langle \left(\frac{\partial}{\partial\theta}\log f(\mathrm{X}|\theta)\right)^2 \right\rangle = \frac{2N^2\theta^2}{\left(N\theta^2 + 4\sigma^2\right)^2}. \tag{4.46}$$

Heterodyne detection is equivalent to the splitting of the signal light by a 50:50 beamsplitter and performing homodyne detection of orthogonal quadratures in the two outputs. Therefore, one can obtain the overall Fisher information for heterodyne detection by multiplying Eq. (4.46) by 2 to account for two quadratures and replacing N with $\frac{N}{2}$ to account for the splitting of photons. Finally, dividing the overall Fisher information by the total number of photons $2N$ from both sources, we obtain the per-photon Fisher information of heterodyne detection:

$$\mathcal{K}(\theta) = \frac{\frac{1}{2}N\theta^2}{\left(\frac{1}{2}N\theta^2 + 4\sigma^2\right)^2}. \tag{4.47}$$

Fig. 4.4(b) illustrates the Fisher information given by Eq. (4.47) for different N and small separations between the sources. Heterodyne detection does not reach the quantum Cramer-Rao bound as mode-sorting due to the shot noise of the LO. However, it overcomes direct imaging when the average photon number from each source exceeds 4. Considering the implementation simplicity of heterodyne detection with comparison to mode-sorting, we chose to use heterodyne detection in our experiment.

4.4 Theory of Hermite-Gaussian microscopy

Heterodyne detection in the first Hermite-Gaussian mode is sufficient for separation estimation when the object consists of two point sources. Measuring photocurrents

in higher-order Hermite-Gaussian modes enables the full image reconstruction of an arbitrary object with sub-Rayleigh resolution [175]. This super-resolution imaging technique is dubbed Hermite-Gaussian microscopy (HGM). This section shows how to obtain the full image of an object via HGM both for coherent and incoherent objects.

4.4.1 Coherent case

The heterodyne detector photocurrent in response to a single point source located at position x is given by Eq. (4.38). For the LO prepared in qth Hermite-Gaussian mode, this photocurrent becomes

$$J_q(x) \propto \int_{-\infty}^{\infty} \psi(x'-x)\phi_q(x')dx' \propto \frac{1}{\sqrt{q!}}\left(\frac{x}{2\sigma}\right)^q e^{-\frac{x^2}{8\sigma^2}}. \tag{4.48}$$

Let us now consider an arbitrary object with the electric field distribution in the object plane $E(x)$. The heterodyne detector photocurrent measured for this object is given by

$$\mathcal{I}_q \propto \int_{-\infty}^{\infty} E(x)J_q(x)dx. \tag{4.49}$$

For an object of sub-Rayleigh size, this photocurrent can be approximated as

$$\mathcal{I}_q \propto \frac{1}{\sqrt{q!}}\int_{-\infty}^{\infty} E(x)\left(\frac{x}{2\sigma}\right)^q e^{-\frac{x^2}{8\sigma^2}}dx \simeq \frac{1}{(2\sigma)^q\sqrt{q!}}\int_{-\infty}^{\infty} E(x)x^q dx. \tag{4.50}$$

Here, we neglect the exponent, i.e. $e^{-\frac{x^2}{8\sigma^2}} \approx 1$, assuming that the object occupies an interval $\Delta x \ll \sigma$.

Thus, the heterodyne detector photocurrent in qth Hermite-Gaussian mode gives the qth geometric moment of the object field. The geometric moments of a distribution, which has nonzero values only in a finite region, provide extensive information about the distribution. For example, they can be used to determine the centre of mass or the principal axes (the expression for the two-dimensional moments will be given further) [78]. Besides, the geometric moments are widely used to derive invariants with respect to specific transformations. It was proved that the moments are uniquely determined by the pattern, and conversely, a pattern

is uniquely determined by its moments [77]. The purpose of HGM is the usage of photocurrents, which carry the geometric moments, to reconstruct the object.

The proof that the geometric moments possess the reconstruction ability is provided in work [62]. Below, we demonstrate how the field of the object can be reconstructed from a set of geometric moments (photocurrents). Consider the following decomposition of the object field into the Hermite-Gaussian basis:

$$E(x) \propto \sum_{k=0}^{\infty} \frac{\beta_k H_k\left(\frac{x}{2\sigma}\right) e^{-\frac{x^2}{8\sigma^2}}}{2^{k+1} k! \sqrt{\pi}\sigma}, \tag{4.51}$$

where β_k are coefficients of the decomposition. Hermite polynomials $H_k(x)$ can be written as

$$H_k(x) = \sum_{q=0}^{k} \alpha_{kq} x^q, \tag{4.52}$$

where α_{kq} are coefficients of the kth-order Hermite polynomial.

Knowing all values of β_k, one can calculate $E(x)$ using Eq. (4.51). The idea of HGM is to extract β_k from the heterodyne detector photocurrents measured for the object. Starting with the definition of β_k and then using Eq. (4.52), Eq. (4.48), and Eq. (4.49), we derive the following expression that links each coefficient with a set of photocurrents:

$$\begin{aligned} \beta_k &:= \int_{-\infty}^{\infty} E(x) H_k\left(\frac{x}{2\sigma}\right) e^{-\frac{x^2}{8\sigma^2}} dx = \int_{-\infty}^{\infty} E(x) \sum_{q=0}^{k} \alpha_{kq} \left(\frac{x}{2\sigma}\right)^q e^{-\frac{x^2}{8\sigma^2}} dx \\ &= \sum_{q=0}^{k} \sqrt{q!}\alpha_{kq} \int_{-\infty}^{\infty} E(x) J_q(x) dx = \sum_{q=0}^{k} \sqrt{q!}\alpha_{kq} \mathcal{I}_q. \end{aligned} \tag{4.53}$$

In theory, acquiring photocurrents for all Hermite-Gaussian modes allows image reconstruction with arbitrarily high resolution. In practice, however, the available number of modes is limited. Fig.4.5(a) shows the numerically calculated resolution enhancement of HGM beyond the Rayleigh limit versus the number of modes involved in the reconstruction. The resolution scales approximately as the inverse square root of this number.

The above description considers the one-dimensional case. HGM for two-dimensional objects is analogous. Measuring photocurrents in two-dimensional

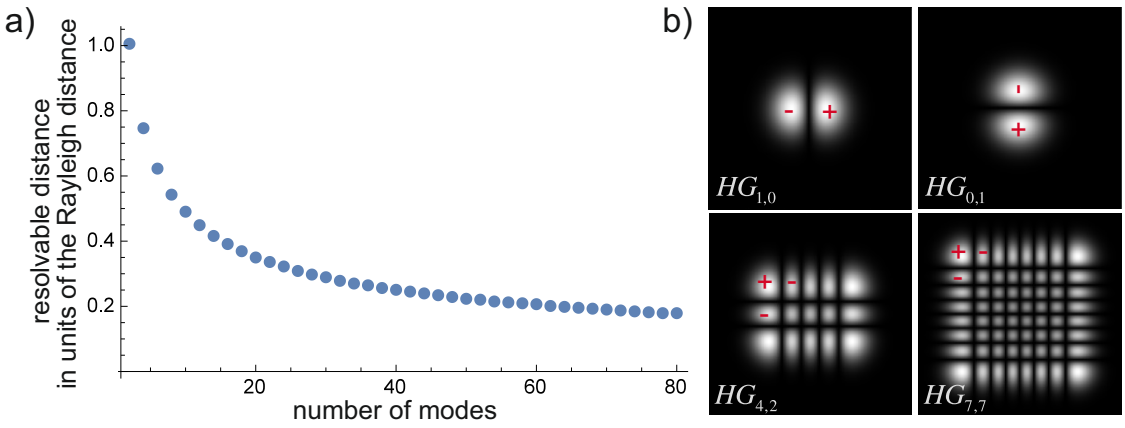


Figure 4.5: a) Resolution enhancement of one-dimensional HGM beyond the Rayleigh limit versus the number of Hermite-Gaussian modes. For two-dimensional HGM, the plot is analogous with the highest mode order in each dimension instead of the number of modes along the horizontal axis. The resolution is defined as the distance between the point objects such that, in the reconstructed image, the midpoint intensity is at 75% of the maximum intensity. Reproduced from [175]. b) Absolute value of the electric field of two-dimensional Hermite-Gaussian modes. The alternation of sign of the electric field inside lobes is marked by the red colour.

Hermite-Gaussian modes

$$\phi_{m,n}(x,y) = \frac{1}{\sqrt{\pi}\sigma}\frac{1}{\sqrt{2^{m+n+1}m!n!}}H_m\left(\frac{x}{\sqrt{2}\sigma}\right)H_n\left(\frac{y}{\sqrt{2}\sigma}\right)e^{-\frac{x^2+y^2}{4\sigma^2}}, \tag{4.54}$$

one can reconstruct the electric field distribution of the object using

$$E(x,y) = \sum_{m,n=0}^{\infty}\beta_{m,n}\frac{H_m\left(\frac{x}{2\sigma}\right)H_n\left(\frac{y}{2\sigma}\right)}{2^{m+n+2}m!n!\pi\sigma^2}e^{-\frac{x^2+y^2}{8\sigma^2}}, \tag{4.55}$$

where coefficients $\beta_{m,n}$ are related to the measured photocurrents $\mathcal{I}_{m,n}$ as

$$\beta_{m,n} = \sum_{k=0}^{m}\sum_{l=0}^{n}\sqrt{m!n!}\alpha_{mk}\alpha_{nl}\mathcal{I}_{m,n}. \tag{4.56}$$

Analogously to Eq. (4.50), photocurrents $\mathcal{I}_{m,n}$ carry the two-dimensional geometric moments of the object field $E(x,y)$:

$$\mathcal{I}_{m,n} \propto \frac{1}{(2\sigma)^{m+n}\sqrt{m!n!}}\iint_{-\infty}^{\infty}E(x,y)x^m y^n dxdy. \tag{4.57}$$

The resolution of two-dimensional HGM scales as the inverse square root of the highest mode order in each dimension (see Fig. 4.5(a)).

In the text, we use the abbreviation $HG_{m,n}$ to refer to two-dimensional Hermite-Gaussian modes given by Eq. (4.54). Fig. 4.5(b) shows the electric field distribution

of several modes. The features size of modes reduces as the inverse square root of the mode order, which explains the scaling of the HGM resolution. Projections into higher-order modes access increasingly finer details of the spatial distribution of the signal field correlations and, therefore, the sub-Rayleigh information they carry. Note that for coherent objects photocurrents are complex numbers. Both amplitudes and phases of photocurrents are needed for imaging since, without the phases, information about the antisymmetric features of the object would be lost.

4.4.2 Incoherent case

If the object is incoherent, we can only observe the output power of the heterodyne detector $\mathcal{P} = |J|^2$, where J is the output photocurrent, averaged over all realizations of the field in the image plane $E'(x')$:

$$\langle \mathcal{P} \rangle \propto \left\langle \left| \int_{-\infty}^{\infty} E'(x') E^*_{LO}(x') dx' \right|^2 \right\rangle = \iint_{-\infty}^{\infty} \langle E'(x') E'^*(x'') \rangle \, E^*_{LO}(x') E_{LO}(x'') dx' dx'', \quad (4.58)$$

where

$$\langle E'(x') E'^*(x'') \rangle = \iint_{-\infty}^{\infty} \langle E(x_1) E^*(x_2) \rangle \, \psi(x' - x_1) \psi^*(x'' - x_2) dx_1 dx_2. \quad (4.59)$$

Since the object is incoherent, we get

$$\langle E(x_1) E^*(x_2) \rangle \propto I(x_1) \delta(x_1 - x_2), \quad (4.60)$$

where $I(x)$ is the intensity distribution in the object plane. Hence,

$$\langle \mathcal{P} \rangle \propto \iiint_{-\infty}^{\infty} I(x) \psi(x' - x) \psi^*(x'' - x) E_{LO}(x') E^*_{LO}(x'') dx dx' dx''. \quad (4.61)$$

For a one-dimensional object, the heterodyne detector power in qth mode is

$$\langle \mathcal{P}_q \rangle \propto \frac{1}{q!} \int_{-\infty}^{\infty} I(x) \left(\frac{x}{2\sigma} \right)^{2q} e^{-\frac{x^2}{4\sigma^2}} dx \simeq \frac{1}{(2\sigma)^{2q} q!} \int_{-\infty}^{\infty} I(x) x^{2q} dx. \quad (4.62)$$

As seen, HGM retains only even moments of $I(x)$ in the incoherent case. This would result in the reconstruction given by $\frac{1}{2}[I(x) + I(-x)]$. In the two-dimensional case, the reconstruction would be $\frac{1}{4}[I(x,y) + I(-x,y) + I(x,-y) + I(-x,-y)]$. One

can avoid the overlap of the actual image with the collateral images by placing the object to a single quadrant of the reference frame, the centre of which is defined by the Hermite-Gaussian modes axis [175]. Another approach to incoherent imaging is to recover antisymmetric components of the object intensity by performing heterodyne detection with the LO prepared in superpositions of Hermite-Gaussian modes [160]. This approach is briefly reviewed below.

Consider the following superpositions of Hermite-Gaussian modes:

$$\phi_+(x') = \frac{1}{\sqrt{2}}\left[\phi_q(x') + \phi_{q'}(x')\right], \ \phi_-(x') = \frac{1}{\sqrt{2}}\left[\phi_q(x') - \phi_{q'}(x')\right], \tag{4.63}$$

where $\phi_q(x')$ is the electric field distribution of the qth Hermite-Gaussian mode in the image plane. Let us denote the qth moment of $I(x)$ as Y_q, i.e. $Y_q = \int_{-\infty}^{\infty} I(x)x^q dx$. According to work [160], heterodyne detection with the LO prepared in $\phi_+(x')$ and $\phi_-(x')$ yields photocurrents of power given by

$$\langle \mathcal{P}_+ \rangle = B_{q,q'} + \mathcal{N}_{q,q'} Y_{q+q'}, \ \langle \mathcal{P}_- \rangle = B_{q,q'} - \mathcal{N}_{q,q'} Y_{q+q'}, \tag{4.64}$$

where

$$B_{q,q'} = \frac{1}{2}\left(\mathcal{N}_{q,q} Y_{2q} + \mathcal{N}_{q',q'} Y_{2q'}\right), \ \mathcal{N}_{q,q'} = \frac{1}{2^{q+q'}\sqrt{q!q'!}}. \tag{4.65}$$

As seen from Eq. (4.64), the heterodyne detector power contains the term $\mathcal{N}_{q,q'} Y_{q+q'}$ that gives access to odd moments of $I(x)$.

4.5 Review of experimental works

Being a far-field linear-optical super-resolution technique, SPADE has attracted the attention of the imaging community. Shortly after the theoretical proposal of SPADE for separation estimation, several experimental groups performed proof-of-principle experiments. This section reviews methodology and results of these experiments.

The first experimental work has been published in September 2016 [154]. It was an experiment on separation estimation between the two Gaussian beams of $\sigma = 95.5$ μm of orthogonal polarizations, which mimic the two incoherent sources. A translation stage with a step of 1μm controlled the separation. Neither an objective lens nor Hermite-Gaussian mode-sorting was involved in the experiment. The

measurement of the fraction of the first Hermite-Gaussian mode in the signal field was replaced by the image inversion interferometry. The basic idea of the method is as follows. A signal from the sources enters a Mach-Zehnder interferometer with an image inverter in one arm. A detector measures the optical power in the interferometer output associated with the destructive interference for on-axis light. Thus, if the displacement of both sources from the optical axis is zero (i.e they coincide), the measured power is zero. The power rises with the separation. The experiment demonstrated separation estimation with finite error even when the sources completely overlap.

The next experimental realization [119] was closer to the original theoretical proposal. Two point sources were imaged by an objective lens and the image field was demultiplexed into the zeroth and the first Hermite-Gaussian modes. The experiment was performed in the following settings. The two sources were generated by a digital micromirror device (DMD), which was illuminated by a He-Ne laser, with a pixel size of 7.6μm. The numerical aperture of the objective lens was reduced by an iris to establish the sub-Rayleigh regime. The separation between the sources was varied in steps of 0.2 of σ of the PSF. To simulate incoherence, the DMD displayed only one source at a time switching between the two with a high frequency. Projections onto $HG_{0,0}$ and $HG_{0,1}$ modes were performed using a hologram displayed on an amplitude SLM. Remarkably, this experiment demonstrated the separation estimation precision close to the quantum Cramer-Rao bound.

Almost simultaneously with the experiment just described, the work on separation estimation using heterodyne detection was published [175]. $HG_{0,1}$ mode for the LO was prepared using a cavity. Pairs of slits illuminated by the laser light, which passed through a moving white paper library card in front of the slits to average over the incoherent light statistics, played a role of two incoherent sources. Such a way to create incoherence is valid when the distance between the slits significantly exceeds the typical coherence length of the light passed through the card. In the experiment, four pairs of slits had a separation of 0.25, 0.5, 0.75, and 1.00 mm. The separation of the closest slits was equal to the 0.28 of the Rayleigh limit of

the imaging system. This separation was estimated with a precision of 0.019 of the Rayleigh limit proving the practicality of heterodyne detection for separation estimation. More importantly, this paper has proposed an extension of heterodyne detection to Hermite-Gaussian microscopy.

In 2017, paper [158] proposed to perform the projection of the signal field onto the function

$$\eta(x,y) = \exp\left[-\left(x^2 + \frac{y^2}{4\sigma^2}\right)\right]\mathrm{sgn}(x) \tag{4.66}$$

instead of $HG_{0,1}$ arguing that this projection is significantly simpler and cheaper to implement and that such a measurement still surpasses direct imaging. Here, $\mathrm{sgn}(x)$ is the signum function. Function $\eta(x,y)$ is constructed to give zero overlap with HG_{00} so that the contribution of the bright fundamental mode, which carries no information, to the signal is eliminated. The projection onto $\eta(x,y)$ was achieved as follows. First, the signal from the sources at zero separation was coupled into a single-mode $HG_{0,0}$ fibre. Then, a phase plate was inserted in front of the fibre such that, for zero separation, it shifts the phase of one semicircular cross-section of the beams by π with respect to another semicircular cross-section. The phase plate consisted of two glass slips under an adjustable obtuse angle centred with the optical axis of the fibre. This measurement is not quantum optimal, however, it is not affected by the Rayleigh curse and achieves comparable precision with the quantum Cramer-Rao bound. For example, for $\theta < 0.47\sigma$, the standard deviation in the measured separation was approximately twice the quantum limit.

One of the latest works on separation estimation between the incoherent sources performed SPADE into several Hermite-Gaussian modes, namely $HG_{0\leq m,n\leq 2}$ (9 modes in total) [24]. A mode-sorter in this experiment, Cailabs PROTEUS-C, was based on the Multi-Plane Light Conversion technique (see paragraph "Sorting transverse-spatial modes of any type using Multi-Plane Light Conversion" in subsection 4.3.1). The output modes of the mode-sorter were spatially separated $HG_{0,0}$ beams coupled into individual fibres, which allowed performing intensity measurement on all Hermite-Gaussian modes simultaneously. The experiment

was conducted in the high-intensity regime and demonstrated a good agreement between the experimental and theoretical intensities of all Hermite-Gaussian modes versus separation. The inclusion of higher-order modes allowed estimating the separation in both directions and extend the sensitivity of the method outside the range of very small displacements.

One research group stepped aside from the incoherent sources and performed an experiment on two partially coherent sources with variable degree of coherence [165]. Parity sorting using image inversion interferometry was chosen as the separation-sensitive measurement. The results of the experiment confirmed that partial anticorrelation of the two sources increases the Fisher information of separation estimation.

Some experiments switched from lateral super-resolution to the quantum-limited axial localization of point sources and axial separation estimation between them [129, 185]. Particularly, work [185] showed both theoretically and experimentally that axially-separated incoherent sources can be resolved with precision approaching the quantum Cramer-Rao bound in the limit of zero separation by decomposing the signal into the radial Laguerre–Gaussian modes. The experiment was based on a binary mode-sorter that has two output channels: all odd-order radial modes are directed to one output channel, while even-order modes are directed to another channel.

The SPADE approach was extended to the time domain. Work [49] demonstrated separation estimation between frequency- and time-shifted incoherent pulses at the single-photon level. Separations as small as one-tenth of the pulses' optical bandwidth and temporal duration were measured with the improvement in precision ten times over the precision of a conventional method based on detecting spectral and temporal intensity profiles. To achieve this spectral and temporal resolution, the signal was projected into the first two Hermite-Gaussian time-frequency modes (the zeroth and first mode) using a mode-selective sum-frequency process in a nonlinear waveguide: a weak input signal is mixed with a spectrally shaped pump, and the power of the upconverted signal is measured. The same scientific group covered a

multiparameter estimation problem and demonstrated the simultaneous estimation of the timing centroid, offset, and intensity imbalance of the two incoherent pulses [3].

5 | Hermite-Gaussian microscopy: experiment

"An experiment is a question which science poses to Nature, and a measurement is the recording of Nature's answer."

Max Planck

All existing experiments based on SPADE focus on estimating just one or several parameters of the object, such as the separation between two point sources. This chapter reports the first experimental implementation of Hermite-Gaussian microscopy, which is the extension of SPADE to full imaging. Our work [124] is a proof-of-principle experiment performed for two-dimensional coherent objects in the high intensity regime using heterodyne detection.

5.1 Data to be acquired

In our experiment, we image coherently illuminated objects. According to the theoretical proposal (see subsection 4.4.1), the data to be acquired are HD photocurrents, which are proportional to the overlap between the signal (the light emitted by an object and passed through the imaging system) and Hermite-Gaussian modes (the LO). The explicit expression for HD photocurrents is

$$\begin{aligned} \mathcal{I} &\propto \left[\iint_{-\infty}^{\infty} E_{sig}(x,y)E_{LO}^{*}(x,y)dxdy\right] e^{i(\omega_{sig}-\omega_{LO})t+\beta} + c.c \\ &\equiv [Ae^{i\varphi}]e^{i(\omega_{sig}-\omega_{LO})t+\beta} + c.c, \end{aligned} \tag{5.1}$$

where $E_{sig}(x,y)$ and $E_{LO}(x,y)$ are transverse electric field profiles of the signal and LO, respectively, ω_{sig} and ω_{LO} are optical frequencies of the two beams, β

is the relative phase between the two beams due to the optical path difference, and $c.c$ means the complex conjugate.

As seen from Eq. (5.1), the only term that contains information about the object is $Ae^{i\varphi} \equiv \iint_{-\infty}^{\infty} E_{sig}(x,y)E^*_{LO}(x,y)dxdy$ (the overlap), while $e^{i(\omega_{sig}-\omega_{LO})t+\beta}$ is defined by settings of the experimental setup. For this reason, photocurrents in the theoretical recipe for reconstructing the electric field of the object, which is given by Eq. (4.55) and Eq. (4.56), are assumed to be complex numbers equal to the overlap only. Thus, for a given object and LO mode, one should extract $Ae^{i\varphi}$ from $\mathcal{I}$ (see Eq. (5.1)).

If $\omega_{sig} = \omega_{LO}$, a photocurrent produced by the detector is $\mathcal{I} = [Ae^{i\varphi}]e^{i\beta} + c.c \propto A\cos(\phi+\beta)$. To determine the quantities of interest, A and φ (called the amplitude and phase of the photocurrent in the text), it is necessary to modulate β (i.e. the optical path difference). The modulation can be done with a piezo mirror. If β changes as $\beta \propto t$, where t is time, then $\mathcal{I}(t)$ is a sinusoidal signal, the amplitude of which is A and the phase is φ (phases of all photocurrents should be measured with respect to a known position of the piezo mirror). The disadvantage of the described method is that piezo mirrors are slow (the driving frequency is ~ 1 kHz). Electronic signals at low frequencies are known to be affected by the flicker noise. Besides the noise, the speed of the piezo mirror would become a limiting factor for the entire experiment. To overcome these issues, we modulate photocurrents by shifting the optical frequency of the signal with respect to the LO using an acousto-optic modulator (the typical frequency shift is from 65 to 95 MHz) instead of modulating β.

5.2 Experimental setup

5.2.1 General description

Before going into details, let us start with a general description of the experimental setup. A schematic of the setup is shown in Fig. 5.1. A laser beam from a continuous wave diode laser operating at 785 nm is split into two paths: the signal and local

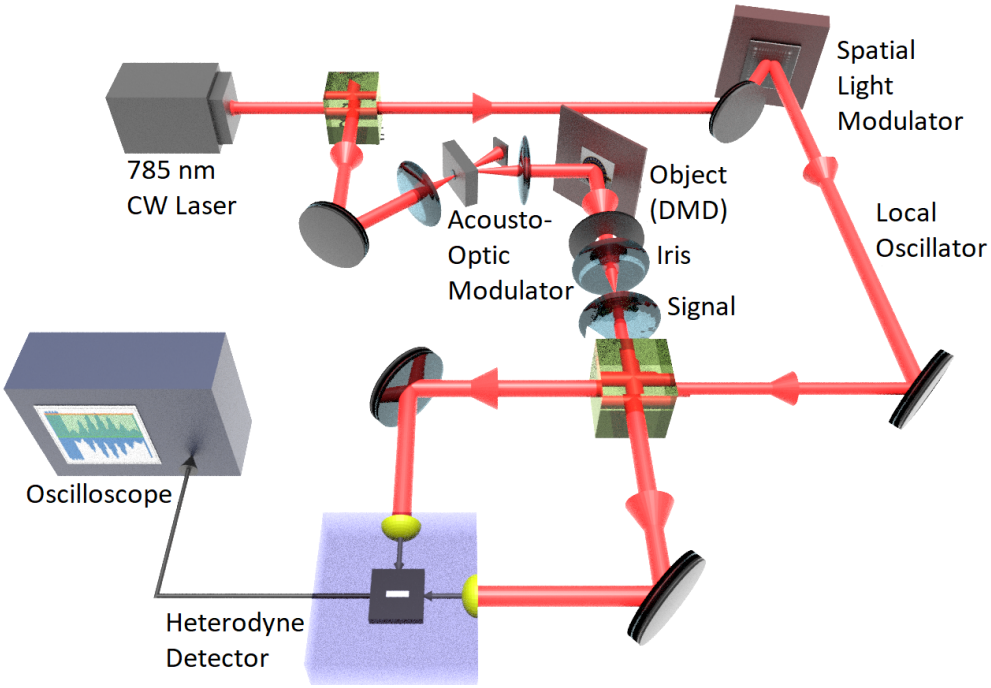


Figure 5.1: Simplified schematic of the experimental setup.

oscillator (LO). A diffractive phase hologram printed on a SLM produces Hermite-Gaussian modes $HG_{0\leq m,n\leq 20}$ in the first diffraction order. Since both indices m and n run from 0 to 20, our basis consists of 441 modes.

The central element of the signal path is a digital micromirror device (DMD) that displays two-dimensional binary amplitude patterns. These patterns play a role of objects for imaging. The beam, which illuminates the DMD, is frequency-shifted by 92.05 MHz via an acousto-optic modulator (AOM) with respect to the LO. After the DMD, the signal propagates towards an objective lens with a numerical aperture small enough to provide the sub-Rayleigh regime.

After the signal and LO are recombined on a beamsplitter, a heterodyne detector (HD) produces a photocurrent according to the result of their interference. This HD photocurrent is sinusoidal with a frequency of 92.05 MHz, which corresponds to the frequency difference between the signal and LO. The output of the heterodyne detector is connected to an oscilloscope, which acquires photocurrents for different objects and Hermite-Gaussian modes. After pre-processing, photocurrents are stored as complex numbers.

Theoretically, after all photocurrents for 441 Hermite-Gaussian modes are obtained for a given object, one can reconstruct an image of the object using Eq. (4.55) and Eq. (4.56). However, we have found that experimental imperfections make it impossible to use the theoretical equation. Systematic errors include non-unity fidelity of Hermite-Gaussian modes, phase aberrations, imperfect mode matching between the signal and LO, deviations of the PSF from a Gaussian shape, etc. Random errors due to mechanical vibrations and drift of optics, air fluctuations, electronic noise, and optical power fluctuations make photocurrents unstable in time.

The contribution of systematic errors and noise leads to the discrepancy between theory and experiment. To overcome this issue, we rely on a supervised feedforward neural network. The neural network filters systematic errors by finding a mapping between experimental photocurrents and modelled ideal image reconstructions, which would be obtained in the absence of imperfections, and increases the robustness of experimental image reconstructions, which we obtain using the trained network, to noise. To train the neural network, we use 26501 samples produced by the DMD.

Data acquisition for the training of the neural network takes approximately 9 hours. To avoid significant mode matching misalignment between the signal and LO, we perform an automatic alignment procedure every 3 hours using the SLM. The procedure adjusts a transverse position and wave vector of the LO.

5.2.2 Laser

A continuous wave laser diode (Eagleyard EYP-DFB0785), which operates at a wavelength $\lambda = 785$ nm, is controlled using a laser diode driver (Arroyo 4205-DR Series LaserSource) in the constant current mode. The current through the diode is set to produce about 84 mW of optical power (the maximum output optical power is 110 mW). The laser output passes through an optical isolator and then is coupled into a single-mode optical fibre (Thorlabs P3-780PM-FC-1) to obtain a beam with a Gaussian spatial profile. The isolator transmission and fibre coupling efficiency are approximately 77% and 60%, respectively.

The laser diode was chosen for the combination of its characteristics. Considering significant light losses on the SLM and DMD in our experiment, and taking into account that the signal-to-noise ratio of heterodyne detection depends on the power of LO, a laser with a lower output power is not desirable for the experiment. Besides the power considerations, the chosen diode is a distributed feedback diode, which means that its active region contains a diffraction grating providing optical feedback for the diode. The optical feedback results in a relatively narrow bandwidth of 2 MHz (full width at half maximum) of the diode, compared to the 0.5-2 nm bandwidth of diodes without feedback.

5.2.3 Digital micromirror device

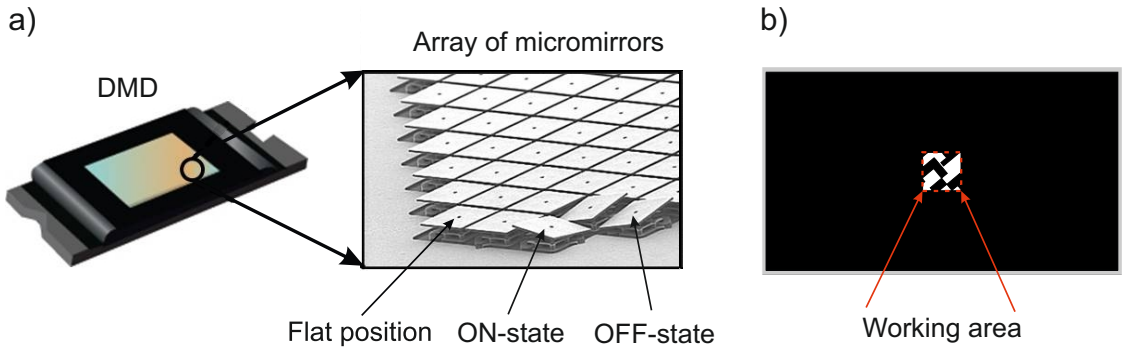


Figure 5.2: a) DMD screen and possible micromirror states. b) Example of a bitmap image that defines the state of each pixel. White and black colours correspond to the ON and OFF states, respectively.

The DMD (DLP LightCrafter 6500) consists of 1920×1080 micromirrors (pixels). Each of them can be set to the ON or OFF-state. ON and OFF-state correspond, respectively, to a tilt by $+12°$ or $-12°$ along the diagonal axis of mirrors and relative to their flat position, as shown in Fig. 5.2(a). ON-state defines the signal path, while light reflected from the OFF-state is blocked.

The pixel pitch of the DMD is 7.56 μm. In our experiment, we used only the central 210×210 pixel area to reduce aberrations due to the intrinsic DMD curvature and for the signal stability. Thus, the total working area is 1.588 mm wide. We generate bitmap images on a computer and upload them to the DMD via USB. Fig. 5.2(b) shows an example of such an image.

The DMD has various operation modes. In our experiment, we worked with the 'pattern-on-the-fly' mode. In this mode, the DMD can store up to 400 bitmap images in its internal memory. These images make up the *sequence*. The DMD displays images one by one. Each image stays active for a user-defined exposure time, which can be adjusted from 105 μs to few seconds. In the 'pattern-on-the-fly' mode, the DMD generates two output trigger signals of a square shape. One of them indicates the beginning of the sequence (DMD sequence trigger) and another marks the switch of an image inside the sequence (DMD frame trigger).

5.2.4 Acousto-optic modulator

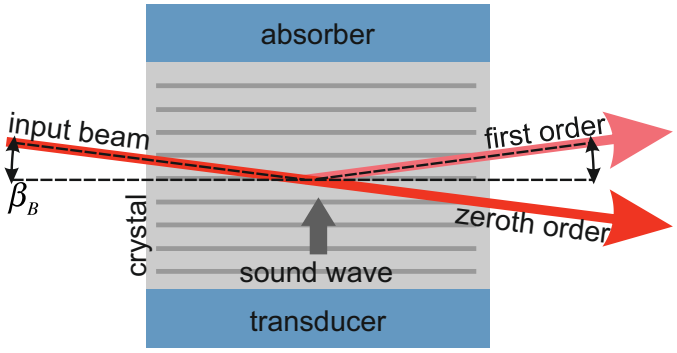


Figure 5.3: Schematic of an acousto-optic modulator.

The basic operation principle of an AOM is shown in Fig.5.3. A transparent crystal is placed between a piezoelectric transducer and acoustic absorber. The transducer, powered by an RF signal, produces a travelling sound wave through the crystal, which is then absorbed by the absorber. The sound wave modulates the refractive index of the crystal. Passing through the AOM, an input beam experiences Bragg diffraction on the travelling sound wave (unlike conventional Bragg diffraction on the crystal lattice), which leads to the following effect. The diffracted beams experience Doppler frequency shifts given by

$$\Delta\gamma_m = m\Upsilon, \tag{5.2}$$

where $\Delta\gamma_m$ is the frequency difference between the input beam and diffracted beam of mth order, and Υ is the frequency of the sound wave.

To concentrate the optical power in one of the frequency-shifted beams, most AOMs, including ours (Isomet 1205C-1), operate in the Bragg regime. In this regime, light enters the AOM under the Bragg angle, and most of the power is diffracted into the first order. The Bragg angle depends on the wavelength of the sound λ_s as

$$\beta_B = \frac{\lambda}{2n_c\lambda_s}, \tag{5.3}$$

where n_c is the refractive index of the crystal in the absence of modulation and λ is the wavelength of light. The incident and diffraction angles are identical. The frequency of the RF signal controls both the frequency shift and direction of the first order, whereas the power of the RF signal defines the diffracted optical power.

In our experiment, the AOM is placed at the focal point of a lens of $f = 250$ mm, where f is the focal distance. All diffraction orders except the first one are blocked by an iris. After passing through the iris, the first order is magnified by a four-lens telescope ($f_1 = 50$ mm, $f_2 = 50$ mm and $f_1' = 100$ mm, $f_2' = 250$ mm), where the first pair of lenses controls the divergence and the second pair collimates the beam to a certain size, to illuminate the working area of the DMD. The RF signal at 92.05 MHz operating the AOM is generated using a Direct Digital Synthesizer (DDS, AD9959 EVAL MODULE). This signal is amplified using a voltage-controlled amplifier to maximize the optical power of the first diffraction order.

5.2.5 Imaging system in the signal path

After the DMD, the amplitude-modulated signal propagates for $L = 245.5 \pm 1$ cm before it is collected by an objective lens. To adjust the numerical aperture of the imaging system, we placed an iris right before the lens of and set its diameter to $D = 3.5 \pm 0.03$ mm. The iris has a shape shown in Fig. 5.4. To estimate its diameter, we placed a metal ruler next to it and imaged both objects with a high-resolution camera. From the camera image, we compared the number of pixels corresponding to the iris diameter and the ruler scale and found the iris diameter in millimeters.

Thus, we obtained that the numerical aperture of the imaging system is NA $= D/(2L) = (7.1 \pm 0.07) \times 10^{-4}$. The corresponding (theoretical) coherent light

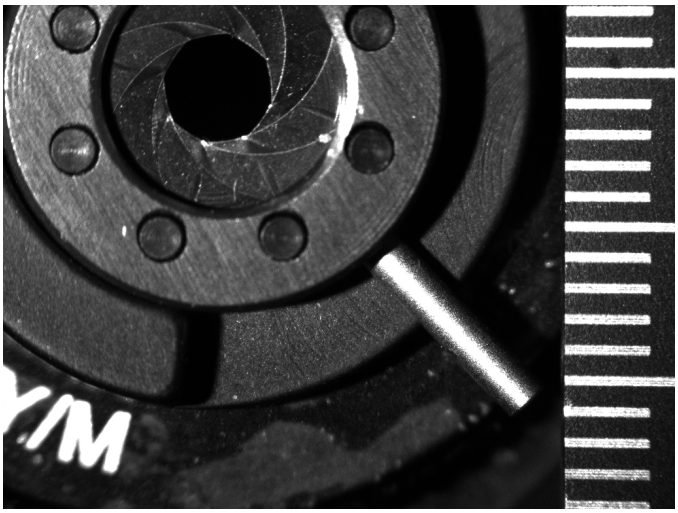


Figure 5.4: Iris.

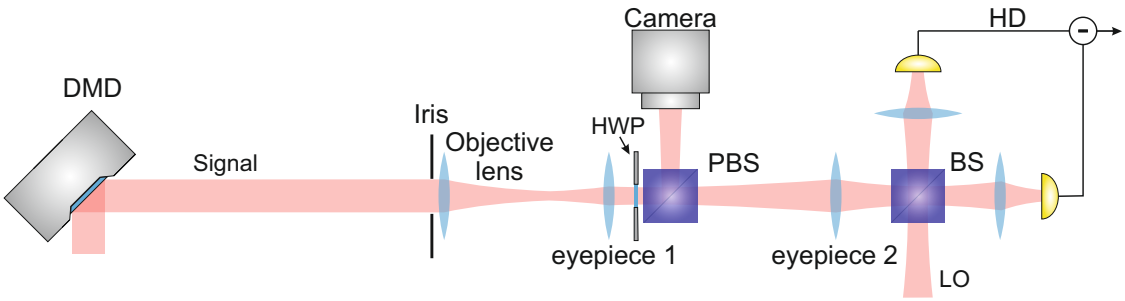


Figure 5.5: Schematic of the imaging system in the signal path that consists of a direct imaging part and HGM part, which are separated by a polarizing beamsplitter (PBS) and half-wave plate (HWP).

Rayleigh limit is $0.82\lambda/\mathrm{NA} = 906\mu$m (120 DMD pixels) and, for comparison, the classical incoherent light limit is $0.61\lambda/\mathrm{NA} = 674\mu$m (89 DMD pixels). Here, we define the Rayleigh limit for coherent light as such distance between the two sources in phase that the gap between them in the image plane is at 75% of the maximum intensity (see paragraph "Angular and spatial resolution limits" in section 1.1).

The objective lens, which has a focal distance of 100 mm, and an "eyepiece 1" lens of 50 mm focal distance image DMD bitmaps onto a conventional camera (see Fig. 5.5). We acquire direct images for comparison with Hermite-Gaussian microscopy. A half-wave plate and a polarizing beamsplitter, which are placed after the "eyepiece 1" lens, are used to reflect a small fraction of the signal to the camera. In the transmitted path, the signal is overlapped with LO and directed to the HD photodiodes. The divergence of the transmitted signal is

controlled by another "eyepiece" lens ("eyepiece 2", $f = 50$ mm) on an optical rail. After the mixing point of the signal and LO, two lenses ($f = 50$ mm each) focus the beams onto the photodiodes. The "eyepiece 2" lens is set such that photodiodes are located approximately in the image plane to avoid the formation of the defocused Airy pattern on them.

5.2.6 Spatial light modulator

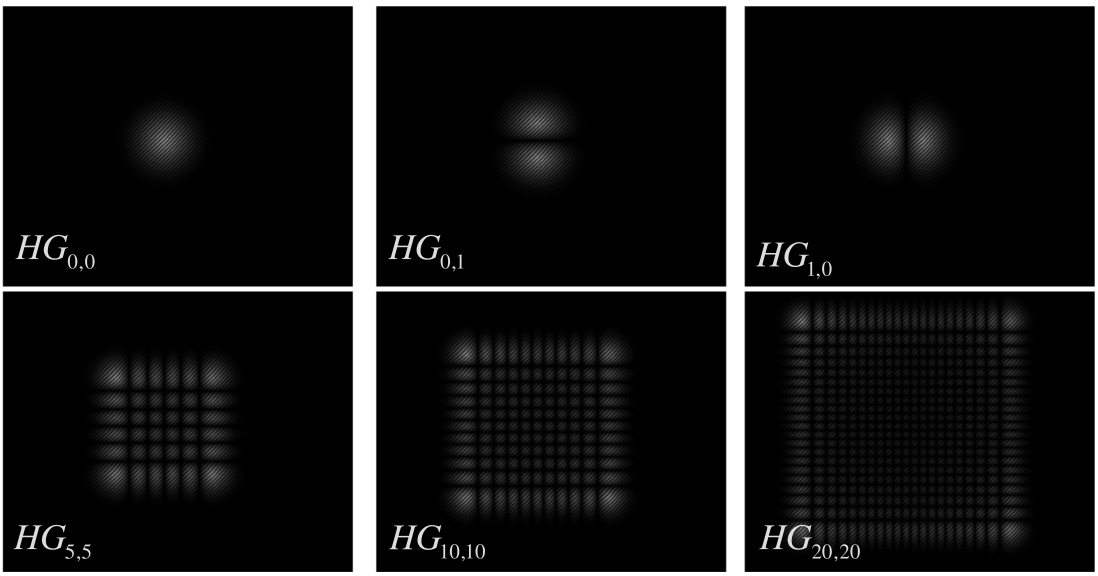


Figure 5.6: Examples of SLM holograms used to produce modes specified in the bottom left corner of each hologram. Dimensions of holograms correspond to the resolution of the SLM. Modes are centred with the physical curvature of the SLM back panel. The grating is diagonal of 15 pixel period.

A spatial light modulator (Hamamatsu X13138-02) is illuminated by a laser beam preliminary magnified by a telescope ($f_1 = 50$ mm, $f_2 = 250$ mm) to cover the whole screen. The incident beam hits the SLM screen almost perpendicularly, so that the angle between the incoming and reflected light is smaller than 5°. The polarization of the incident beam is parallel to the extraordinary axis of the SLM. Hermite-Gaussian modes are generated in the first diffraction order of modulated-grating holograms. The light reflected by the SLM is collected by a telescope ($f_1 = 250$ mm, $f_2 = 100$ mm) with an iris in its focal plane that selects only the first order.

Examples of holograms are shown in Fig. 5.6. These holograms compensate for the back panel curvature of the SLM and non-uniform profile of the incoming

beam. We omitted the cavity correction and crosstalk correction in this experiment to speed up the hologram generation during the automatic alignment procedure of the interferometer, which will be described further. The decrease of the modes' fidelity due to this decision can be considered a systematic error compensated by the neural network.

5.2.7 Electronic system

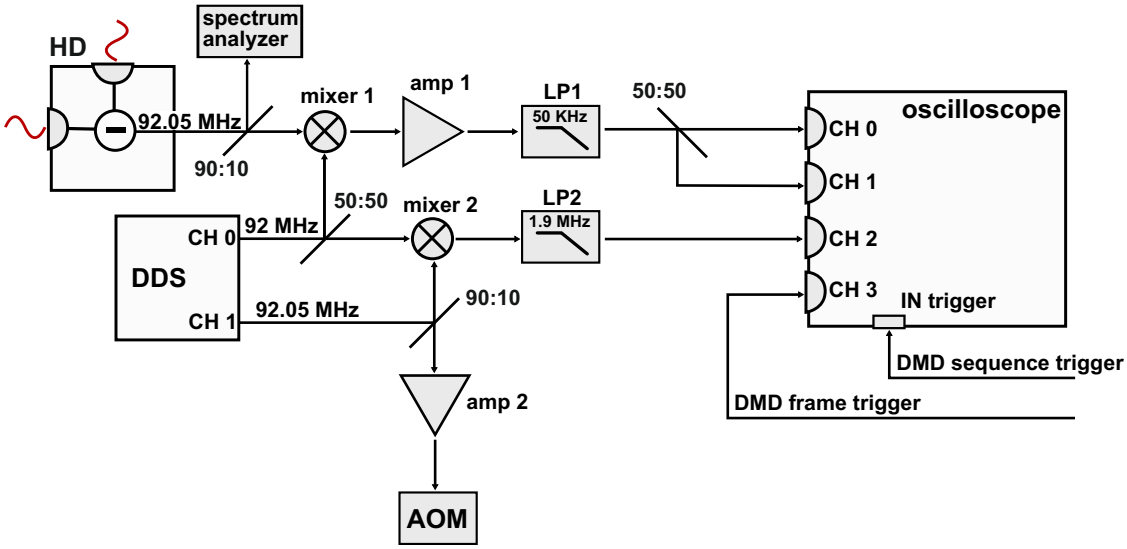


Figure 5.7: Electronics setup.

Fig. 5.7 shows a scheme of the electronic system. A DDS (Direct Digital Synthesizer AD9959 EVAL MODULE for generating analog waveforms) board generates two sinusoidal signals at 92 and 92.05 MHz. Each of them splits into two paths. 90% of the 92.05 MHz signal drives the acousto-optic modulator after the amplification. The 92 MHz signal is split 50 : 50. One half of it is mixed with the HD detector signal on frequency mixer 1, another with 10% of the DDS 92.05 MHz signal on frequency mixer 2. The output of mixers is a superposition of the sum and difference frequencies. Low pass filters (LP1 and LP2) after the mixers transmit only the difference frequency at 50 kHz. The down-converted HD signal is amplified and divided between two oscilloscope channels (CH0 and CH1), which will be discussed further in the data acquisition section. The oscilloscope model is Tektronix MSO5104. This is a digital oscilloscope (DAQ system) that allows

us to synchronize data acquisition with the DMD, control data acquisition with the PC, and receive data in a convenient format for analysis.

We reduce frequency of the HD signal to decrease the size of data and speed up data transfer and saving. As was mentioned before, DMD can update images not faster than every 105 μs. Without the down conversion to 50 kHz, HD photocurrent for each DMD frame would contain thousands of periods of 92.05 MHz signal, while only a few periods are needed for the amplitude and phase measurement.

The difference frequency signal of the two DDS channels enters CH2 of the oscilloscope. This is our *reference signal* that, like the down-converted HD signal, oscillates at 50 kHz. The reference signal is necessary for the phase measurement of photocurrents. CH3 and IN trigger channel of the oscilloscope are connected to the trigger outputs of the DMD. 10% of the HD signal is sent to a spectrum analyzer before down conversion for the automatic alignment procedure of the interferometer.

The list of components in Fig. 5.7 is

- mixer 1 and mixer 2: Mini-circuits ZX05-1L-S+,
- amp 1: Mini-Circuits ZFL-500-BNC and ZFL-1000LN+ in series,
- amp 2: Mini-Circuits ZX60-8008E and ZHL-32A in series,
- LP1: Mini-circuits BLP 1.9+ and Thorlabs EF 124 in series,
- LP2: Mini-circuits BLP 1.9+,
- 90 : 10 power splitter: Mini-circuits ZFDC-10-1-75,
- 50 : 50 power splitter: Mini-circuits ZMSC-2-1+,
- spectrum analyzer: Rigol DSA800,
- oscilloscope: Tektronix MSO 5104,
- HD: homemade heterodyne detector [91].

5.3 Setup alignment and data acquisition

5.3.1 Optical setup for mode matching

Two telescopes in the LO path ($f_1 = 250$ mm, $f_2 = 150$ mm and $f_1' = 100$ mm, $f_2' = 100$ mm), placed after the telescope that collects light after the SLM, control the waist and divergence of Hermite-Gaussian modes, respectively. "Eyepiece" lenses of the telescopes are mounted on translation stages for fine alignment. Two mirrors, located near the mixing point with the signal, are used to adjust the propagation direction of the LO and match it with the signal.

5.3.2 Mode matching procedure

The goal of the mode matching procedure is to maximize the overlap between $HG_{0,0}$ and PSF of the imaging system. Ideally, we should image one DMD pixel, which represents a point source, at the centre of the screen to obtain the PSF. However, a single pixel gives a low signal-to-noise ratio. For this reason, we display a 20×20 pixel square for the waist and divergence matching and a 100×100 pixel square for the propagation direction matching.

Before data acquisition, we align our setup manually in two steps. In the first step, we generate LO in $HG_{0,0}$ mode and match the size and divergence and propagation direction of the LO and signal using telescopes and mirrors in the LO path (see subsection 5.3.1). The criterion for optimization is the maximized power of the HD detector photocurrent.

In the second step, we perform fine matching of the propagation direction. We prepare LO in $HG_{0,1}$ and $HG_{1,0}$ modes for the alignment along the vertical and horizontal directions, respectively. The goal now is to find a local minimum of the photocurrent power since we expect the PSF to be orthogonal to $HG_{0,1}$ and $HG_{1,0}$ when the signal and LO are perfectly centred relative to each other. Being a null measurement, this procedure is significantly more sensitive to the relative beam displacement. We characterize the alignment by calculating the suppression ratio

as $10\log\left(\frac{\mathcal{P}_{0,1/1,0}}{\mathcal{P}_{0,0}}\right)$, where $\mathcal{P}_{n,m}$ is the photocurrent power in $HG_{m,n}$ mode. The alignment is finished when we achieve the suppression of -40 dB and beyond.

5.3.3 Automatic alignment procedure of the interferometer

We automated the fine alignment procedure to repeat it periodically (every 3 hours) during the data acquisition. The DMD displays a 100×100 pixel square at the centre. The SLM starts with the horizontal direction and displays $HG_{1,0}$ hologram. The auto-alignment program iteratively repeats the following two algorithms until the suppression ratio reaches -40 dB. The first algorithm looks for the optimal horizontal displacement of the mode on the SLM screen. It scans several positions around the current location of the mode with a step of 1 pixel and fixes the position x^{opt} where the photocurrent power is minimized. The second algorithm is designed to find the optimal propagation angle. It generates holograms centred at x^{opt} with various linear phase gradients along the horizontal direction and finds the optimal gradient. Thus, ith iteration of the program can be symbolically written as

$$\begin{aligned}
&x_i^{opt} = \min_{\mathcal{P}_{HD}} M(x-x_i,y)\cdot \mathrm{mod}\left[\Phi(x,y)+\mathrm{grid}(x,y),2\pi\right], \text{ where}\\
&\Phi(x,y) = F(x-x_i,y)-\alpha(x,y)+\kappa_{i-1}^{opt}(x-x_i),\ x_i\in[x_{i-1}^{opt}-\delta_x, x_{i-1}^{opt}+\delta_x];\\
&\kappa_i^{opt} = \min_{\mathcal{P}_{HD}} M(x-x_i^{opt},y)\cdot \mathrm{mod}\left[\Phi(x,y)+\mathrm{grid}(x,y),2\pi\right], \text{ where}\\
&\Phi(x,y) = F(x-x_i^{opt},y)-\alpha(x,y)+\kappa_i(x-x_i^{opt}),\ \kappa_i\in[\kappa_{i-1}^{opt}-\delta_\kappa, \kappa_{i-1}^{opt}+\delta_\kappa].
\end{aligned} \tag{5.4}$$

$\mathcal{P}_{HD}$ is the HD detector photocurrent power, $\mathrm{grid}(x,y)$ is the diffraction grating phase pattern, $\alpha(x,y)$ is the SLM back panel curvature, κ is the phase gradient slope, δ_x and δ_κ are search bounds, and $M(x,y)$ and $F(x,y)$ are grating modulation functions calculated according to Eq. (2.9) to produce $HG_{1,0}$. Search bounds are $\delta_x = 10$ pixels in the first iteration (5 pixels in subsequent iterations) and $\delta_\kappa = \frac{3}{1280}\frac{\mathrm{rad}}{\mathrm{px}}$ in the first iteration ($\frac{1}{1280}\frac{\mathrm{rad}}{\mathrm{px}}$ in subsequent iterations), where 1280 pixels (px) is the horizontal resolution of the SLM. The slope κ changes in steps of $0.1\delta_\kappa$ within the search bounds. After the horizontal direction, the program aligns the vertical one identically. The corresponding code can be found in [104].

Note that we only shift functions $M(x, y)$ and $F(x, y)$ but not the entire hologram to keep $\alpha(x, y)$ matching the physical SLM curvature. After the alignment procedure is finished, the program generates holograms for all Hermite-Gaussian modes with the optimal displacement and propagation angle (defined by the phase gradient slope). To speed up the program, we do not apply the cavity correction in our experiment. Due to this simplification, we avoid the time-consuming re-calculation of $M(x, y)$ and $F(x, y)$ (see the cavity correction procedure given in subsection 2.3.4) when the position of modes and phase gradient are updated. Instead, we shift the pre-calculated modulation functions and substitute them into Eq. (5.4). It takes about 3 minutes to find the optimal position and propagation angle and save new holograms.

5.3.4 DMD datasets

Our DMD sample set can be divided into two main groups: a dataset to train the neural network and a dataset to be imaged, called a test set. In total, we have 26803 samples, where 26501 is the train set and 302 is the test set.

Fig. 5.8 shows objects of the test set, which we aim to image. University of Oxford logo and coat of arms (Fig. 5.8(a) and (d)) are split into smaller areas of 210×210 DMD pixels. An example of the corresponding DMD bitmap is shown in Fig. 5.8(a) beneath the logo. We apply HGM reconstruction to each area and join obtained images to get full objects. This procedure is equivalent to the transverse scanning of the object (logo or coat of arms). As seen in Fig. 5.8(b) and (c), each symbol of the alphabet and each pair of lines are imaged without the scanning because they fit the working area.

The training/cross-validation dataset consists of

- 20000 frames that are tiled with black and white rectangles. The tile size is randomly chosen from 10 to 50 pixels in each dimension and the “colour” (ON-state of micromirrors corresponds to white colour, and OFF-state corresponds to black) of each tile is also random, sampled from a uniform distribution between 0.2 and 0.8; the entire matrix is then randomly rolled along both axes and randomly oriented, Fig. 5.9(a);

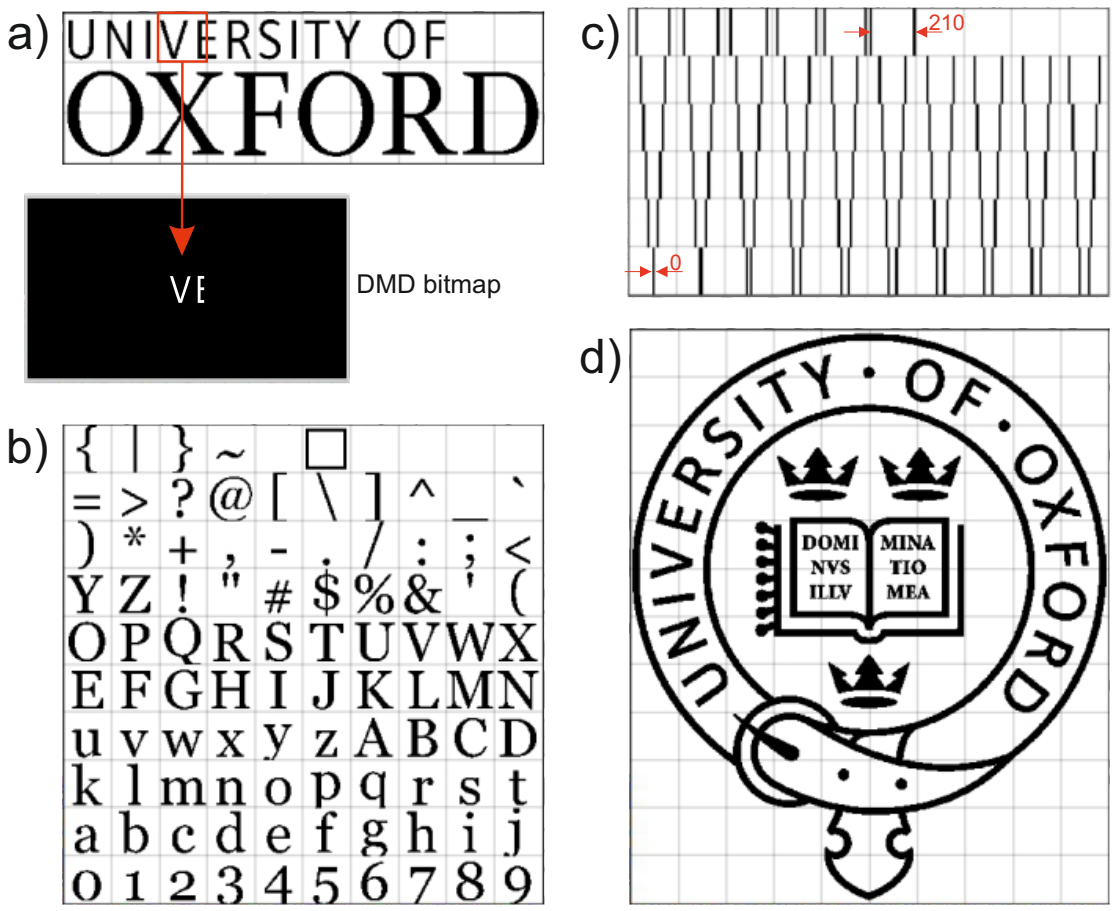


Figure 5.8: Test set. Squares of the grey grid correspond to the DMD working area size (210×210 pixels). a) University of Oxford logo, 10×3 frames. b) Alphabet, 96 text symbols. c) 56 pairs of vertical lines of 10 pixel thickness and separation varying from 0 to 210 pixels. d) University of Oxford coat of arms, 10×12 frames.

- 3000 random sets of 1 to 5 lines with random orientation, position and width (10 to 70 pixels), Fig. 5.9(b);
- 3000 random sets of 1 to 5 segments of elliptic rings with randomly variable eccentricity and width (minor and major axis from 40 to 120 pixels, width from 10 to 15 pixels), Fig. 5.9(c);
- 58 single squares of sizes 30 pixels and 70 pixels at various positions within the 210 × 210 frame;
- 441 squares of 30 pixel size at variable positions within the 210 × 210 frame, from which a smaller square of 10 pixel size is subtracted, with variable positions of the smaller square with respect to the larger one, Fig. 5.9(d);

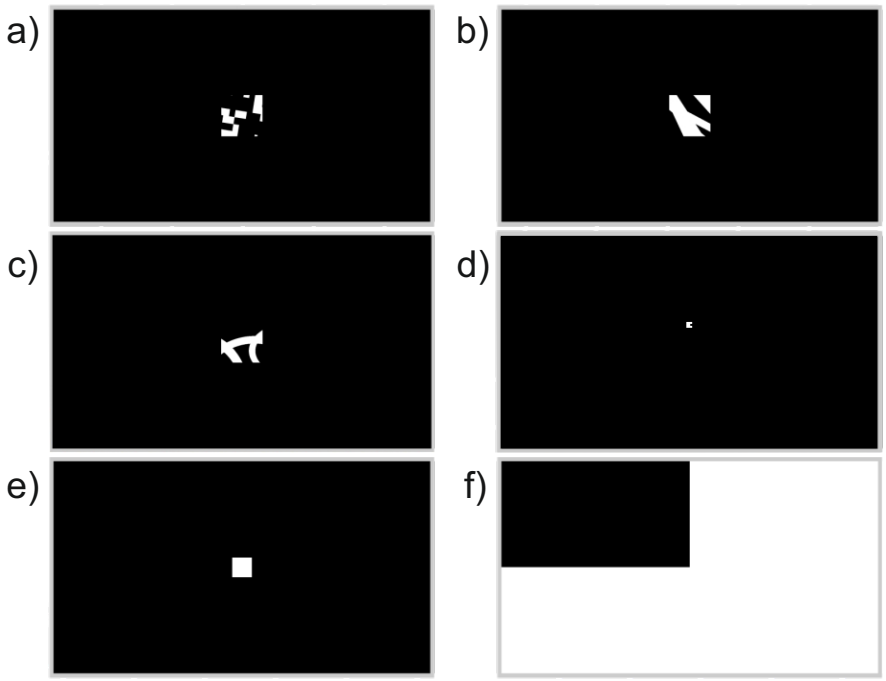


Figure 5.9: Train dataset and special objects ((e) and (f)). a) Random matrix. b) Superposed random lines. c) Random parts of an ellipse. d) 30×30 square with subtracted 10×10 square. e) Alignment square. f) Phase reference object (phase reference bitmap).

- a blank image;
- a square occupying the entire 210×210 pixel frame.

During the training time, the train/cross-validation dataset is randomly split into 90% for training and 10% for validation.

Besides the train and test set, we have two special objects: alignment square and phase reference object (Fig. 5.9(e) and (f)). The alignment square is a square of 100×100 pixels at the centre of the DMD working area. It is used to keep track of the relative alignment between LO and signal beams during data acquisition. An asymmetric shape of the phase reference bitmap has nonzero overlap with all LO modes, so this object always generates a measurable photocurrent. A phase of the photocurrent produced by the phase reference object with a given LO mode is used as a reference for all other DMD objects.

DMD datasets can be downloaded at [104].



Figure 5.10: Transverse scanning of the University of Oxford logo for direct imaging. a) The scanning procedure can be seen as moving of a 630×630 pixel window (red square) along the logo in 10-pixel steps in the vertical and horizontal directions. b) Examples of DMD bitmaps that correspond to three positions of the moving window. The distance between the chosen positions of the centre of the window is 100 pixels.

5.3.5 Acquisition of direct images

To compare HGM with conventional imaging, we acquire intensity images, which are formed by our imaging system of low numerical aperture, of the test set. Each symbol of the alphabet and each pair of lines are acquired via a single shot, unlike the University of Oxford logo and coat of arms for which we simulate the transverse scanning with a stride of 10×10 DMD pixels, as explained in Fig. 5.10 and below.

The procedure for direct imaging of the University of Oxford logo and coat of arms is as follows. To generate DMD bitmaps, we split these objects into 10×10 pixel areas and display each 10×10 pixel frame at the central part of the DMD together with a context of 630 pixels, as shown in Fig. 5.10(b). The size of the context is significantly higher than the waist of the PSF (120 DMD pixels) so that an image of the central area is fairly affected by diffraction from the context. We acquire direct images for each DMD bitmap and separate 10×10 pixel frames from the context by using a mapping between the DMD and camera planes. Then, we join all images of 10×10 pixel frames into a single image.

For comparison, the stride for HGM is 210×210 pixels (and no context). The reason for the finer stride for direct imaging is that, unlike HGM, direct imaging is sensitive to aberrations of the optical system, and border effects occur when

joining directs images taken with a high stride. Scanning procedure with a fine stride benefits direct imaging when it is compared with HGM. However, our purpose is to show that HGM overcomes direct imaging even when the effort is taken to reduce the effect of aberrations on direct images.

5.3.6 Acquisition of photocurrents

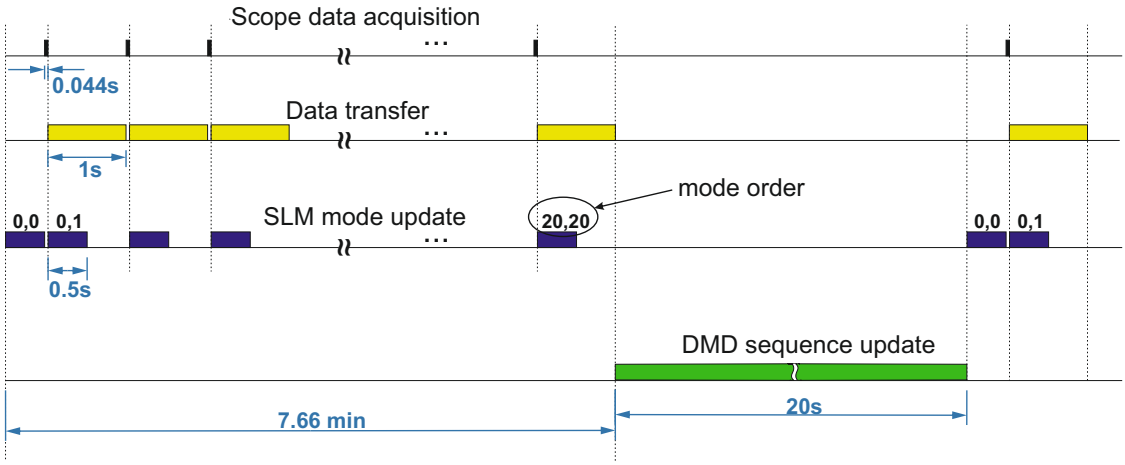


Figure 5.11: Data acquisition timing diagram that shows the duration of operations, such as the acquisition of a trace by the oscilloscope (scope data acquisition), data transfer, update of a hologram on the SLM (together with the liquid crustal stabilization time), and loading new sequence into the DMD internal memory, and the execution order. The length of each coloured block corresponds to the operation duration.

We split the optical power between the signal and LO such that before the heterodyne detector we have ∼200 μW in $HG_{0,0}$ and ∼20 μW in the signal produced by the alignment square (see Fig. 5.9(e)).

To acquire photocurrents, we distribute 26803 DMD samples into 68 DMD sequences. Each sequence (except the last one) contains 398 original bitmaps and 2 special ones (phase reference and alignment square, which are the first and the last elements of the sequence, respectively). The last sequence contains 137 original bitmaps.

Fig. 5.11 shows the timing diagram of data acquisition. We set the acquisition order to minimize the overall measurement time taking into account a low refresh rate of the SLM (we have found that ≈ 0.5 s is needed for liquid crystal molecules to stabilize after the hologram update), a high DMD refresh rate (105 μs) and the

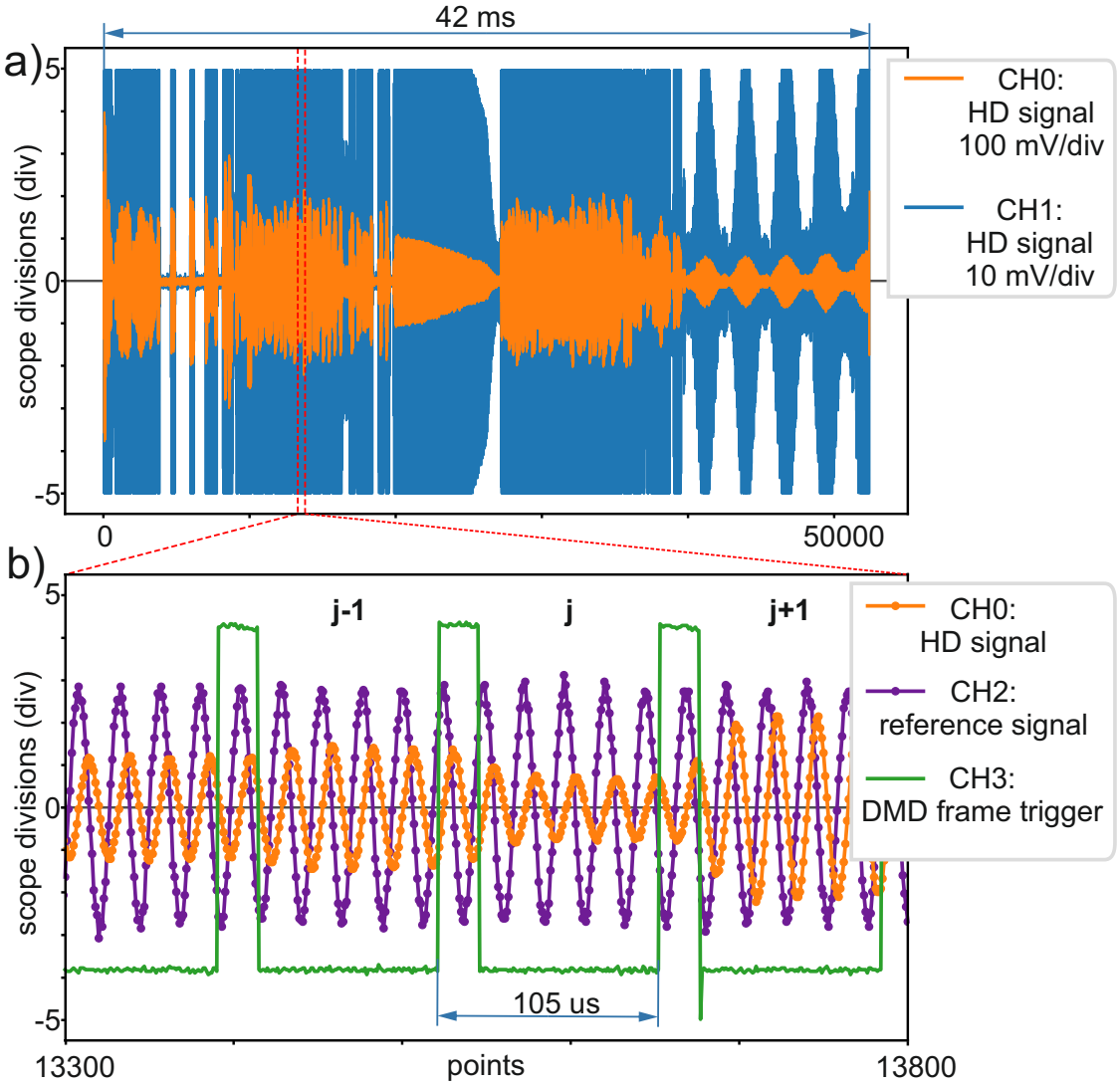


Figure 5.12: a) HD photocurrents measured by CH0 and CH1 of the oscilloscope for some DMD sequence (400 samples) and $HG_{0,0}$ mode. The duration of each DMD sequence is 42 ms, but the oscilloscope acquires data for a slightly longer time (44 ms), which is the duration of one oscilloscope data acquisition operation in Fig. 5.11. b) Zoom of CH0 and the remaining channels (CH2 and CH3). Each green trigger pulse indicates the update of a DMD bitmap (sample) within the sequence. Using the sampling rate of the oscilloscope, which is 1.25 Ms/s, we can evaluate the duration of the zoomed region: $(13800 - 13300)/(1.25 \cdot 10^6) \approx 400$ μs.

time required to load a sequence into the DMD internal memory (≈ 20 s). We first generate a Hermite-Gaussian mode starting with $HG_{0,0}$ and acquire photocurrents for the given sequence, and then update a hologram on the SLM to produce the next mode. After photocurrents for all modes are acquired for this sequence, we load the next one to the DMD and repeat the procedure.

As was mentioned in subsection 5.2.7, we equally split down-converted HD photocurrents between two oscilloscope channels: CH0 and CH1. The reason is that the power of photocurrents varies by several orders of magnitude as a function of the LO mode and DMD sample (from ∼10 pW to ∼1 mW). To measure all photocurrents with high sensitivity, we set the resolution of CH0 to 100 mV/division for high-power photocurrents and the resolution of CH1 to 10 mV/division for low-power photocurrents. Fig. 5.12(a) shows photocurrents in both channels acquired for some DMD sequence with the LO in $HG_{0,0}$ mode. During data processing, we read data from CH0 whenever CH1 saturates.

We acquire the entire sequence of 400 DMD samples within one oscilloscope trace. The DMD sequence trigger synchronizes the oscilloscope and DMD. The sampling rate of the oscilloscope is 1.25 Ms/s. Each trace contains 55000 points in each channel. Thus, the duration of a trace is 44 ms, which is slightly higher than the duration of a sequence ($400 \cdot 105$ μs = 42 ms).

Simultaneously with the HD signal, we acquire the reference signal and DMD frame trigger. Fig. 5.12(b) shows these curves together with photocurrents in CH0 for several samples indicated by j counter within the sequence.

While the trace is transferred from the oscilloscope to a computer in binary format via USB (it takes approximately 1s to transfer 55000 data points and 4 oscilloscope channels), the SLM hologram is updated to produce the next Hermite-Gaussian mode (see Fig. 5.11). Approximately every 3 hours an automatic alignment procedure is executed to assure that LO and signal beams are aligned during the whole experiment. The acquisition of the training set (26501 samples) takes about 9 hours, while the test set (302 samples) takes less than 10 minutes.

5.3.7 Noise and setup misalignment

This subsection provides indicative information about the signal-to-noise ratio of photocurrents, which is limited by technical noise in our experiment.

Due to various sources of temporal noise, the HD photocurrent for a given DMD object and LO mode does not remain constant over time. Photocurrents experience

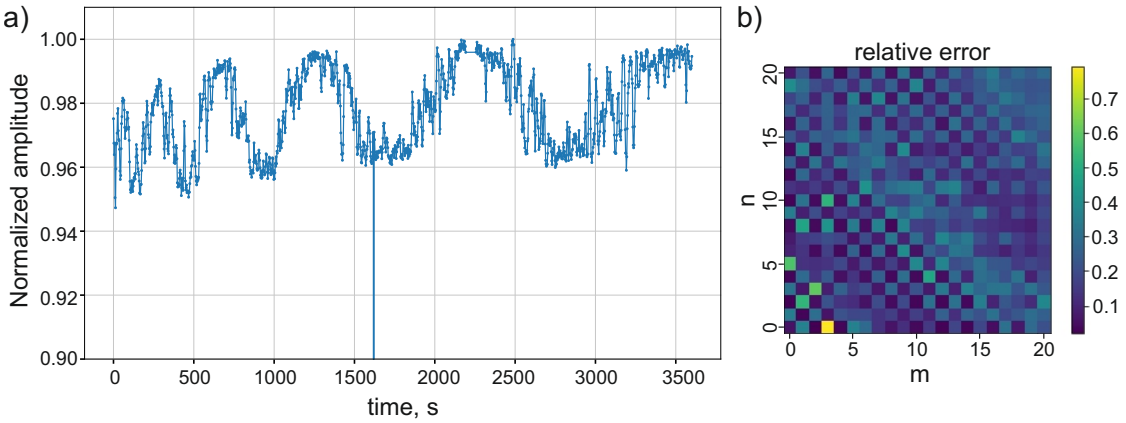


Figure 5.13: a) Alignment square plus $HG_{0,0}$ photocurrent amplitude noise. The curve is normalized by its maximum. The oscilloscope took acquisition every 2.5 s. b) Relative error of the photocurrent amplitude for the alignment square and each $HG_{m,n}$.

random noise, drift and noise at different frequencies. Instabilities can originate from the drift and shaking of kinematic parts of optics, laser power fluctuations, temperature fluctuations, air turbulence, etc. The DMD and SLM also contribute to the temporal noise. For example, the SLM produces noise due to flipping the pixels' voltage. Besides, we have noticed that control cables attached to the LC screen of the SLM are sensitive to vibrations and that the cables transfer vibrations to the LO.

Fig. 5.13(a) shows an example of changes in the amplitude of the photocurrent produced by the alignment square and $HG_{0,0}$ during one hour. The relative error (standard deviation over mean value) of 3-4% is typical for this pair of DMD object and LO mode. However, temporal fluctuations differ for different combinations of the signal and LO, which means that instabilities leading to the mode mismatch affect photocurrents the most. For example, the interference signal of the alignment square, whose image is close to a Gaussian intensity profile, and odd Hermite-Gaussian modes fluctuates more significantly on a relative scale because, theoretically, any mode mismatch gives rise to photocurrents from zero to a finite amplitude. Fig. 5.13(b) shows the experimental relative error of the photocurrent amplitude for the alignment square and each Hermite-Gaussian mode during a data acquisition run (each DMD sequence contains the alignment square, 68 sequences in total).

According to the log files of the automatic alignment procedure for this data acquisition run, LO modes required compensating shifts of up to 3 and 4 SLM pixels

over the acquisition time in horizontal and vertical directions, respectively. The total deviation of the compensating phase gradient slopes was $\frac{0.3}{1280}\frac{\text{rad}}{\text{px}}$ in both directions.

5.4 Data processing

5.4.1 Pre-processing

After converting the saved oscilloscope data for each DMD sequence from binary to integer format, we process the DMD frame trigger to identify the beginning and end of the time period that correspond to each DMD object. To image the jth object of the sequence, we have to extract amplitudes $\mathcal{A}^j_{m,n}$ and phases $\varphi^j_{m,n}$ of photocurrents for all $HG_{m,n}$ modes from the data.

The difficulty in measuring the relative phases between photocurrents in different LO modes is that they fluctuate due to air turbulence, which is significant at the time scale of 0.5 s needed to switch the LO mode. In our experiment, we took the following steps to reduce airflows. We enclosed the optical table with hardboard panels and moved electronic equipment outside the enclosure (to avoid airflows caused by fans and heating of the electronics). Although these measures allowed us to reduce airflows, we could not eliminate them because the optical table occupies a large area, and our enclosure is not hermetic. We take into account residual airflows when collecting and pre-processing data.

The phase fluctuations within one DMD sequence for a given LO mode can be neglected (as the duration of a sequence is 42 ms). For this reason, we use the phase reference bitmap (see Fig. 5.9(f)), the phase of which we denote as $\widetilde{\varphi}_{m,n}$, at the beginning of each sequence to measure phases of all other objects in the sequence with respect to it. Thus, relative phases of photocurrents in different LO modes in the absence of fluctuations are equivalent to

$$\varphi^j_{m,n} - \varphi^j_{0,0} \equiv (\varphi^j_{m,n} - \widetilde{\varphi}_{m,n}) + (\widetilde{\varphi}_{0,0} - \varphi^j_{0,0}) + \langle \widetilde{\varphi}_{m,n}(t) - \widetilde{\varphi}_{0,0}(t+\delta t)\rangle_t \,. \quad (5.5)$$

Here, the phase difference in the first two terms on the right-hand side of the equation can be considered unaffected by phase fluctuations because the phase difference in each term is obtained from a single DMD sequence. In the last term, where $\delta t \approx 0.5$

s is the SLM refresh time, phase fluctuations are averaged over time. Note that the last term is the same for all DMD objects and can be learnt by the neural network. We have found that this approach gives better results than a direct measurement. Considering that only relative phases matter, we define $\varphi^j_{m,n} - \varphi^j_{0,0} \equiv \varphi^j_{m,n}$.

To find the phase difference between the jth object and phase reference bitmap, i.e. to evaluate the first two terms in Eq. (5.5), we first find the phases $\varphi^j_{m,n}$ and $\widetilde{\varphi}_{m,n}$ individually with respect to the electronic reference signal. As was described in subsection 5.2.7, the electronic reference signal is the difference frequency signal of the two DDS channels. This signal is shown by the purple curve in Fig. 5.12(b) and has a constant phase and amplitude. To find $\varphi^j_{m,n}$ and $\widetilde{\varphi}_{m,n}$, we adhere to the following procedure.

We take the HD signal for the jth object from sensitive CH1 unless it is saturated, otherwise, we use CH0. Then, we fit the HD and electronic reference signal within the borders defined by the DMD frame trigger (green curve in Fig. 5.12(b)) for the jth object with a cosine and subtract the fitted phases, which gives us $\varphi^j_{m,n}$. We repeat the same procedure for the phase reference bitmap, which is the first object of every DMD sequence, and obtain $\widetilde{\varphi}_{m,n}$. The step of measuring phases with respect to the electronic reference signal is necessary to avoid errors in fitted relative phases between DMD objects due to the fact that the period of the DMD frame trigger is not multiple to the HD signal period.

When fitting the HD signal for any DMD object, we skip the first 30 points from the beginning of the object's frame (i.e. from the falling edge of the DMD frame trigger) because they are affected by the DMD switch. Besides the phase, we extract the HD signal amplitude $\mathcal{A}^j_{m,n}$ from the fit for each DMD object.

The amplitude and phase of jth object form the corresponding complex photocurrent $\mathcal{I}^j_{m,n} = \mathcal{A}^j_{m,n} e^{i\varphi^j_{m,n}}$. For each j, we flatten complex photocurrents acquired for each $HG_{m,n}$ into a vector and normalize this vector by its total power $\mathcal{P}^j = \sum_{m,n} |\mathcal{I}^j_{m,n}|^2$. Then, we split the real and imaginary parts, which is a more convenient format of the input data for the neural network. At that stage, complex photocurrents of all samples are represented by 26803 vectors of 882 elements each

(21×21 HG modes, real and imaginary parts). We sort them according to the corresponding DMD object type (i.e. random matrix, University of Oxford logo, etc) and save them as tensors, ready to be fed to the neural network. The processing of all samples and 441 Hermite-Gaussian modes takes about 4.5h on a Desktop PC (i7-6700 CPU @3.4 GHz, 16GB RAM) using a multi-thread parallel processing.

5.4.2 Neural network

A supervised neural network analyzes training examples with the corresponding labels (the desired outputs) and learns probability-weighted associations between the two. After the training, this neural network can be used to make predictions on previously unseen examples (test set). A tutorial on neural networks can be found in [115].

In imaging, neural networks are vastly used in image recognition, restoration of images distorted by the measurement system, image enhancement to amplify specific features, and segmentation of an image into coherent parts (review [52]). Besides, neural networks have been applied to obtain high-resolution images from low-resolution ones [169]. However, this approach requires high-resolution images for training, and up-sampling of images remains an ill-posed problem, which means that there are always multiple high-resolution images corresponding to a single low-resolution image. Neural networks are capable to increase the resolution when trained with images of chosen categories by guessing inter-pixel values.

In our experiment, a neural network does not guess the features of objects beyond the resolution of HGM. It filters systematic errors from the experimental data and performs post-processing to access sub-Rayleigh information that the data contain due to the physics of HGM. Besides filtering the systematic errors, the neural network is a crucial part of HGM in the presence of noise because even the slightest relative fluctuations of photocurrents, especially for higher-order Hermite-Gaussian modes, make the reconstruction method using Eq. (4.55) impractical. The reason for such sensitivity to noise is that the impact of higher-order modes is amplified via the factorial terms in Eq. (4.56). The neural network is significantly

more robust to noise than the direct reconstruction method since it learns to extract information from noisy data.

Architecture of the neural network

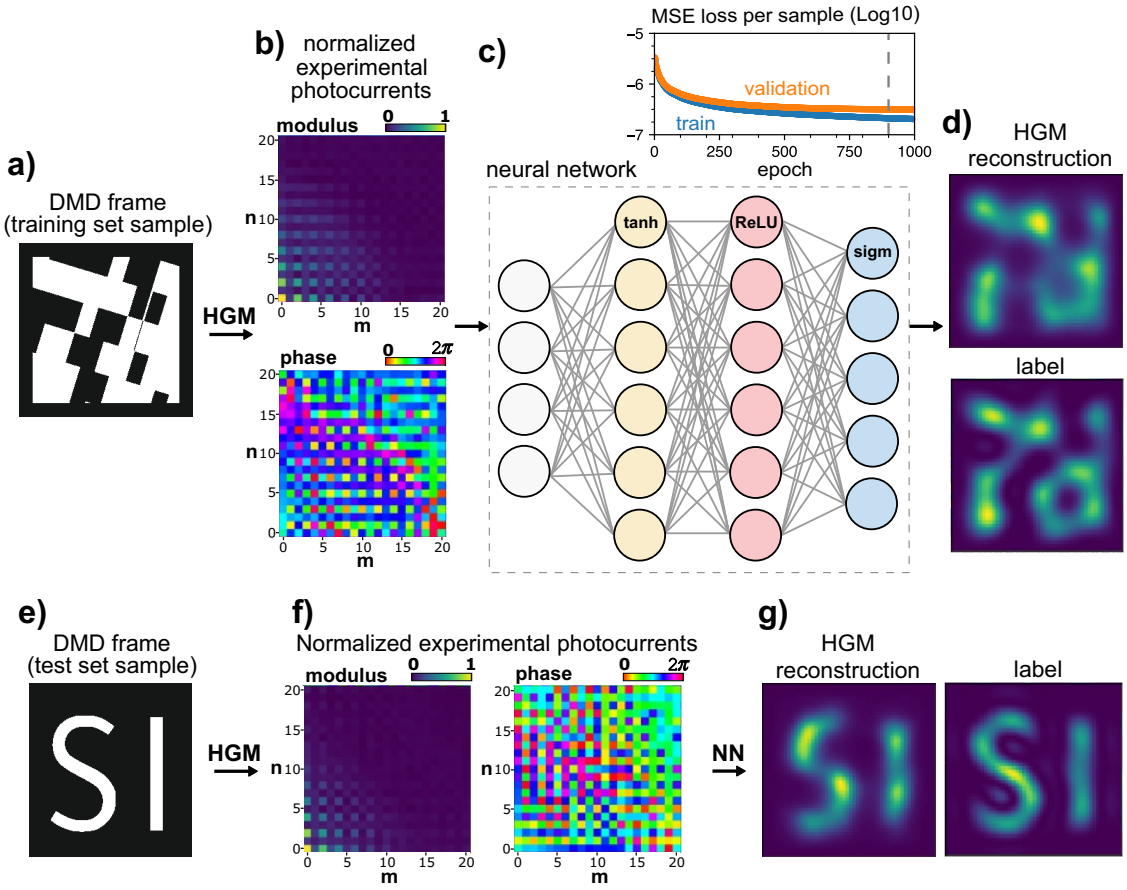


Figure 5.14: Sketch of the data acquisition and processing. A training set sample (a) interfered with $HG_{m,n}$ modes produces a set of complex photocurrents (b) which, after being flattened and split into real and imaginary parts, are fed to the neural network (c)(the inset shows the learning curve, i.e. the training and cross-validation loss versus the training epoch). The neural network (NN) is trained to predict the idealized HGM reconstruction of the sample image (d). After the training, the neural network reconstructs previously unseen test samples (e)-(g).

Fig. 5.14 sketches the data processing using our neural network (see our GitHub repository [104] for the code), where Fig. 5.14(a)-(d) depict the training process and Fig. 5.14(e)-(g) illustrate imaging of the test set. To train the neural network, we send vectors, which contain real and imaginary parts of normalized complex photocurrents of the DMD train set, to the input of the neural network. The

values of the input vectors' elements range in $[-1, 1]$. The labels are simulated ideal HGM reconstructions, i.e. simulated images of DMD objects obtained using the idealized theory of HGM that assumes no systematic errors and noise (see paragraph 'Simulation of labels' below), normalized by their total power. Each label is a 50×50 matrix flattened into a 2500-dimensional vector. The elements of the labels range in $[0, 1]$. The neural network learns the mapping between the inputs and labels.

Using the idealized HGM reconstructions as labels is advantageous over the actual object matrices since HGM reconstructions are smoother, which facilitates the training of the neural network, and mark the "ground truth" achievable by HGM. With this approach, we train the neural network to approximate the underlying HGM model and filter the experimental noise, and not to guess the samples features beyond the resolution capabilities of our imaging technique. On the contrary, using as labels the actual or even a slightly smoothed objects, whose exact reconstruction is not physically achievable by the experiment, leads to overfitting issues that actually degrade the image quality for the test set.

The neural network consists of two hidden layers with 6000 weights each. The activation function of the first hidden layer is the hyperbolic tangent, $\tanh(z) = (e^z - e^{-z})/(e^z + e^{-z})$, to account for $[-1, 1]$ range of inputs. The activation function of the second hidden layer is ReLU, $\mathrm{ReLU}(z) = \max(0, z)$. Finally, we choose the sigmoid function, $\sigma(z) = 1/(1 + e^{-z})$, for the output layer since labels range from 0 to 1.

Simulation of labels

According to the theory provided in section 4.4, we start simulating labels with the calculation of theoretical coefficients of decomposition of the electric field in the object plane into the Hermite-Gaussian basis:

$$\beta_{m,n} = \sum_{x,y} E(x, y) H_m\left(\frac{x}{2\sigma}\right) H_n\left(\frac{y}{2\sigma}\right) e^{-\frac{x^2+y^2}{8\sigma^2}}. \tag{5.6}$$

Here, the electric field $E(x, y)$ of an object is a DMD bitmap from the train set, integration is replaced by the sum over discretized x and y positions of pixels, and σ is the width of a Gaussian PSF, which corresponds to the measured numerical

aperture (NA $= 7.1 \times 10^{-4}$, $\sigma = 0.21\lambda/\text{NA}$). To resolve fine features of Hermite-Gaussian modes, we oversample 210×210 DMD objects and, hence, the simulation space to a higher resolution. The HGM reconstruction (label) of an object is

$$\mathcal{E}(x,y) = \sum_{m,n=0}^{21} \frac{\beta_{m,n} H_m\left(\frac{x}{2\sigma}\right) H_n\left(\frac{y}{2\sigma}\right) e^{-\frac{x^2+y^2}{8\sigma^2}}}{2^{m+n+2} m!n!\pi\sigma^2}. \tag{5.7}$$

After the labels are calculated, we undersample them to 50×50 matrices before converting them into vectors for the neural network to reduce the computational cost of the training.

Training

The loss function is the mean squared error (MSE) loss, which makes the neural network effectively behave as a nonlinear regressor. The neural network is trained for 900 epochs, achieving training and validation $\log_{10}$-loss per sample of -6.674 and -6.498, respectively, as illustrated in the inset of Fig. 5.14(c).

The optimization method is Adaptive Moment Estimation (Adam), with a learning rate of 10^{-4}, exponential decay rate moving parameters for the first and second moment estimates $(0.9, 0.999)$ and weight decay 0. We set the batch size to 512. We notice that a large batch size, apart from reducing the training time, allows us to achieve lower training and validation losses. This may be due to the fact that the acquired experimental data have a stochastic noise component. If we train the neural network over batches of 512 samples, we reduce the impact of this stochastic noise on the trained model.

HGM reconstruction

After the neural network is trained, we use it to produce HGM reconstructions of the test set (Fig. 5.14(e)-(g)). As before, we measure photocurrents for each test set sample and rewrite them in a form of normalized vectors caching normalization factors. For a given input, the neural network produces an image, which we multiply by the normalization factor to preserve the relative power ratio between different frames of the test set (for example, between different frames that compose the University of Oxford logo and coat of arms).

5.5 Results

5.5.1 Comparison with direct imaging

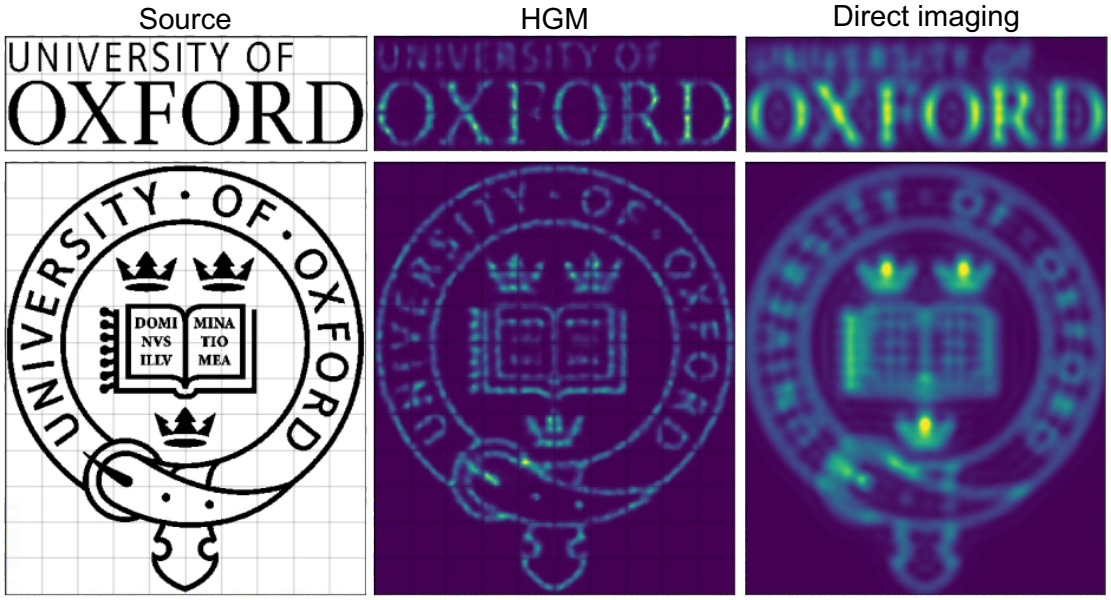


Figure 5.15: Original, HGM-reconstructed and camera images of the University of Oxford logo and coat-of-arms test sets. The grid shows the division of the images into 210×210 DMD pixel sub-images that are fed to HGM sequentially.

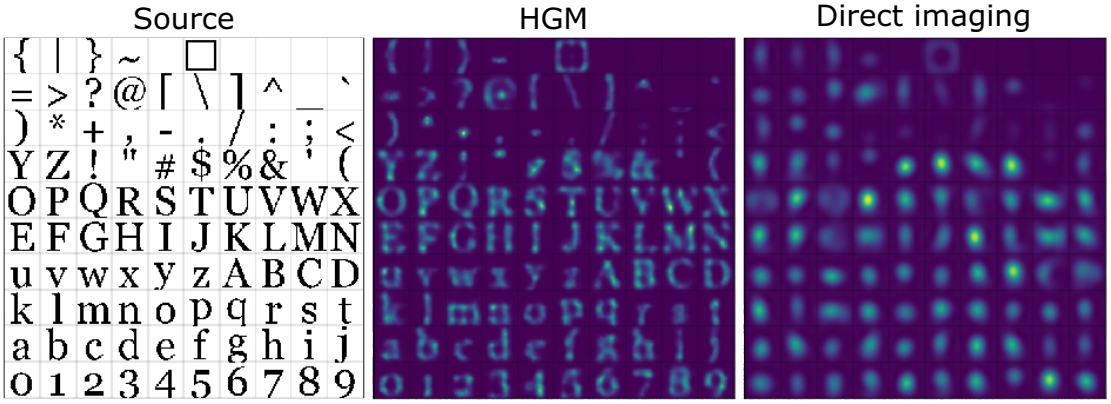


Figure 5.16: Original, HGM-reconstructed and camera images of the alphabet test set. Each symbol corresponds to one DMD frame, i.e. every symbol is displayed as a whole (no division into sub-images).

Fig. 5.15 and Fig. 5.16 show our main results. We compare the samples of the test set, HGM reconstructions, and direct images. As described in subsection 5.3.5, direct images of the University of Oxford logo and coat of arms are obtained using

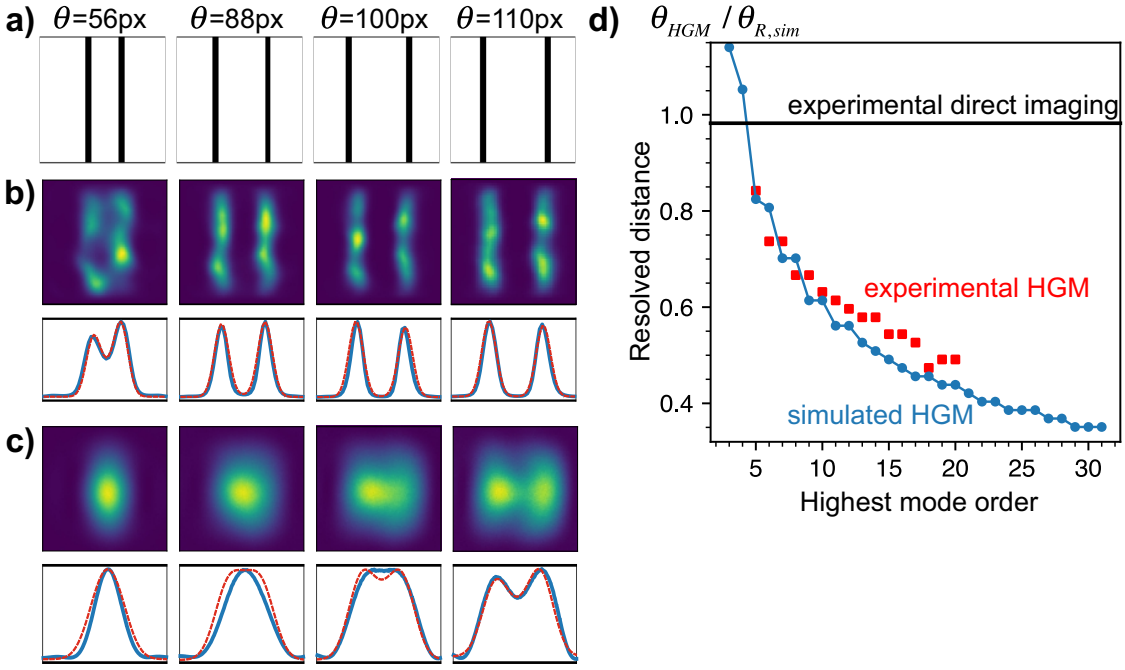


Figure 5.17: Quantitative assessment of the HGM resolution enhancement with respect to direct imaging. Original (a), HGM reconstructed (b) and camera images (c) of lines pairs of increasing spacing θ (in DMD pixels, px). A plot below each image shows the image's integral along the vertical axis (blue curve). The red curve represents a theoretically expected one-dimensional image of two coherent sources assuming a Gaussian PSF, $\psi(x) = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \exp\left(-\frac{x^2}{4\sigma^2}\right)$, i.e. the curve is the square of the sum of two Gaussian functions. For camera images, the waist of the Gaussian functions corresponds to $\sigma = 0.21\lambda/\mathrm{NA}$. For HGM reconstructions, the waist is 0.47 of that for camera images. In both (b) and (c), the separation between Gaussian functions is assumed to be known and equal to the separation between lines, and amplitudes are fitted to the blue curve. d) Experimental (red squares) and theoretical (blue connected dots) resolution enhancement of HGM with respect to direct imaging vs number of used modes in each dimension. Along the vertical axis is the ratio between the minimum resolvable distance for HGM (θ_{HGM}) and simulated direct imaging ($\theta_{R,sim}$). The continuous black line indicates the level of the experimental direct imaging resolution.

scanning of a fine stride, which potentially benefits direct imaging. Nevertheless, we can see that HGM reconstructions in Fig. 5.15 are significantly sharper than direct images and we distinguish fine features and details. The experiment with alphabet symbols (see Fig. 5.16) compares HGM with direct imaging with no scanning involved. Here, the supremacy of HGM is even more notable.

To evaluate quantitatively the resolution enhancement, we use HGM reconstructions and direct images of pairs of lines. Examples are shown in Fig. 5.17(a)-(c). The resolution enhancement is calculated using the Rayleigh criterion. For both

methods (HGM and direct imaging), we compare the distance between the two lines where the midpoint gap intensity of the image integrated along the vertical axis (blue curves on the plots of Fig. 5.17(b) and (c)) is about 75% of the maximum intensity. We have found that this distance is about 53 DMD pixels for the HGM method when $HG_{0\leq m,n\leq 20}$ basis is used, while, for direct imaging, it is approximately 108 pixels. Thus, the minimum resolvable distance for HGM is one-half of that for direct imaging, and, hence, the resolution of HGM is twice the resolution of direct imaging. In other words, HGM can resolve two sources at approximately one-half of the diffraction limit.

In addition to comparing the minimum resolvable distance for HGM and direct imaging, we demonstrate that the twofold resolution enhancement is consistent with our data by comparing the experimental results (blue curves on the plots in Fig. 5.17(b) and (c)) with theoretical predictions (red curves). The red curves for direct imaging are obtained assuming a Gaussian PSF of waist $\sigma_{DI} = 0.21\lambda/\text{NA}$. The HGM reconstructed images are best described by an effective PSF of waist $\sigma_{HGM} = 0.47\sigma_{DI}$. While blue and red curves are in agreement for HGM reconstructions, confirming that the HGM resolution is approximately twice the resolution defined by the Rayleigh limit, the resolution of direct images for $\theta = 88$ px and $\theta = 100$ px (see Fig. 5.17(a) and (c)) is worse than that expected theoretically due to aberrations. For this reason, the HGM resolution seems to be more than twice the resolution of direct imaging for the aforementioned separations. However, for an ideal diffraction-limited imaging system (without aberrations), the resolution enhancement would be approximately twofold for all separations, which follows from the ratio $\sigma_{HGM}/\sigma_{DI} = 0.47$. Therefore, we report a twofold resolution enhancement as a result of our experiment.

Fig. 5.17(d) shows how the resolution scales with the highest mode order in the Hermite-Gaussian basis in each dimension. Theoretically, we expect the resolution to scale approximately as the inverse square root of the highest mode order. As seen, the experimental results are close to the simulated theoretical curve. However, the discrepancy may occur for the further increase of the highest mode order because

the experimental resolution is fundamentally limited by the shot noise of heterodyne detection. The power of HD photocurrents rapidly drops with the mode order until the HD signal is covered by the noise. Besides, it is increasingly challenging to generate higher-order modes with high fidelity due to the limited SLM resolution.

5.5.2 Comparison with deconvolution of direct images

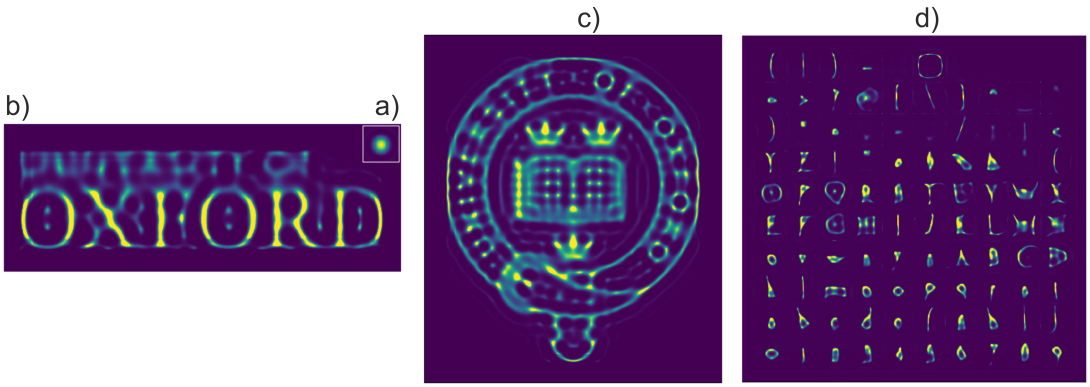


Figure 5.18: Richardson-Lucy deconvolution applied to direct images of University of Oxford logo (b) and coat of arms (c) and alphabet (d) datasets using the Airy PSF (a).

To further evaluate the performance of HGM, we post-process the experimental camera images applying one of the commonly used deconvolution algorithms, Richardson-Lucy (RL) deconvolution. RL is an iterative algorithm that aims to estimate the original object $O(x, y)$, given the point spread function $\psi(x, y)$ and the experimental camera image $I(x, y)$, via

$$\hat{O}_{t+1} = \hat{O}_t \cdot \left(\frac{I}{\hat{O}_t * \psi} * \psi \right) \tag{5.8}$$

with $\hat{O}_t$ being the estimate of the object at the iteration t and asterisks denote convolution.

Fig. 5.18(b)-(d) shows the result of RL applied to the test datasets, after 1000 iterations and using the Airy PSF reported in Fig. 5.18(a). The size of the Airy PSF is chosen to match the experimental PSF. The program used to perform RL can be found in [104].

By qualitatively comparing the HGM predictions (Fig. 5.15 and Fig. 5.16) with the results of RL (Fig. 5.18), we conclude that HGM allows for better reconstruction of the original object, with a higher resolution and fewer artifacts. The quantitative comparison of the resolution using pairs of lines is not possible because the deconvolved images contain artifacts that prevent their correct interpretation (i.e. it is hard to tell from the image if the object is a pair of lines). Due to the artifacts, the alternative method to estimate the resolution improvement described below is also not applicable.

5.5.3 Alternative resolution improvement benchmark

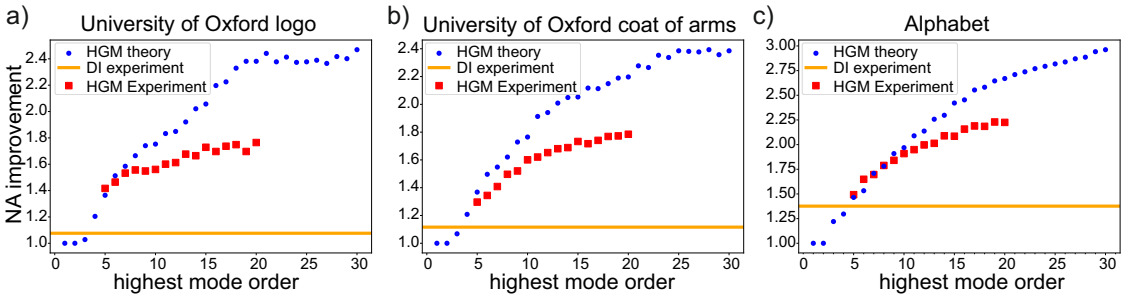


Figure 5.19: Resolution improvement evaluated through the MSE loss as described in the text.

An alternative way of estimating the resolution improvement is by evaluating the mean squared error (MSE) between the reconstructed images and the original objects. We calculate the MSE $M_{\mathrm{DI}}(\mathrm{NA})$ between theoretically calculated direct camera images with varying numerical apertures and the original objects, and then compare these values with the MSE $M_{\mathrm{HGM}}(\mathrm{NA}_0)$ for the HGM reconstruction with the numerical aperture NA_0. Because HGM is generally superior to direct imaging, $M_{\mathrm{DI}}(\mathrm{NA}_0) > M_{\mathrm{HGM}}(\mathrm{NA}_0)$. However, the function $M_{\mathrm{DI}}(\mathrm{NA})$ monotonically decreases with NA and becomes equal to $M_{\mathrm{HGM}}(\mathrm{NA}_0)$ at some $\mathrm{NA}_1 > \mathrm{NA}_0$: in other words, the image quality obtained via direct imaging with the numerical aperture NA_1 is similar to that with HGM and numerical aperture NA_0. The quantity $\mathrm{NA}_1/\mathrm{NA}_0$ corresponds to the resolution enhancement [13] and is plotted, as a function of the number of modes in each dimension used in HGM, in Fig. 5.19.

Using $HG_{0\leq m,n\leq 20}$ basis, we achieved resolution improvements of 1.76, 1.78 and 2.22 for University of Oxford logo, University of Oxford coat of arms and alphabet symbols, respectively.

6 | Conclusion

6.1 Summary

This thesis has presented the first experimental implementation of Hermite-Gaussian microscopy, which is a far-field and linear-optical super-resolution imaging technique. Hermite-Gaussian microscopy originates from the spatial-mode demultiplexing (SPADE) technique for separation estimation between two point sources with quantum optimal precision. Unlike prior experiments based on SPADE, we have applied Hermite-Gaussian microscopy to full imaging.

Using our technique, we successfully imaged two-dimensional objects generated by a coherently illuminated digital micromirror device. Instead of technically complex mode-sorting, we performed heterodyne detection with the local oscillator prepared in each Hermite-Gaussian mode up to $HG_{20,20}$. We achieved approximately two-fold resolution enhancement beyond the Rayleigh limit, established by the objective lens of a small numerical aperture (NA = 7.1×10^{-4}). This result is in agreement with the theoretical prediction.

Besides direct imaging, we compared our results with the commonly used Richardson-Lucy deconvolution algorithm. Hermite-Gaussian microscopy overcomes deconvolution in resolution and shows fewer undesirable artifacts in reconstructed images.

To achieve remarkable results, we had to solve issues related to systematic errors, including optical aberrations, and noise in our experiment. The solution we chose was to filter systematic errors and noise with a supervised neural network, which learns the mapping between experimental heterodyne detector photocurrents and theoretical Hermite-Gaussian microscopy reconstructions. With this approach, the

neural network cleans up experimental noise and does not guess features of objects beyond the resolution capabilities of the imaging technique.

To conclude, our experiment on Hermite-Gaussian microscopy paves the path towards a universally applicable super-resolution imaging technique. Indeed, unlike other far-field super-resolution imaging techniques, HGM does not rely on non-linear properties of samples and does not require a specific type of illumination. Hence, HGM can be applied to various imaging scenarios from microscopy to astronomy after a few adaptations to the current setup, which we discuss in section 6.3. HGM is compatible with existing optical microscopes and telescopes and can enhance their resolution [13].

Although the resolution enhancement demonstrated in our experiment is inferior to that of such techniques as STED, STORM, and PALM, we believe that it is a remarkable achievement considering the versatility of HGM. An important note is that the resolution of our technique can be further increased by extending the Hermite-Gaussian basis and is only limited by the shot noise of heterodyne detection.

In our experiment, the frame rate of HGM is limited by the refresh rate of the SLM. It takes at least 3.7 minutes to obtain data for a single frame because there is a waiting time of 0.5 seconds after each update of a hologram on the SLM to let LC molecules stabilize, and the total number of updates is 441 (size of our Hermite-Gaussian basis). However, HGM can be much faster. The refresh rate of modern SLMs achieves 422 Hz (for example, Meadowlark HSP1920), which brings the frame rate of HGM to 1 Hz. By utilizing mode-sorting instead of heterodyne detection, the data acquisition can be performed almost instantaneously. As was mentioned before, some of the mode-sorting techniques, such as MPLC-based techniques (see the corresponding paragraph and the second part of "Conclusion" in subsection 4.3.1), allow one to project signal field onto different Hermite-Gaussian modes simultaneously.

The speed of an imaging technique can be limited by scanning. To obtain HGM images of the University of Oxford logo and coat of arms, we scanned these objects because their size exceeds the working area of the DMD and the field of

view of HGM, as well as to reduce aberrations. The scanning, however, is not a part of the working principle of HGM, which we demonstrated by imaging alphabet symbols (each symbol is imaged without scanning). The field of view of HGM can be increased along with the resolution by extending the Hermite-Gaussian basis since higher-order modes cover a larger area. To filter systematic errors, the training set for the neural network should be updated according to the field of view. As seen from this and the previous paragraph, there are no fundamental limitations to the speed of HGM, which makes HGM a promising technique for imaging of moving objects in real time.

Besides the main project on super-resolution imaging, this thesis has covered accompanying work on performance optimization of phase-only spatial light modulators. We proposed a comprehensive model that describes the joint effect of three major SLM imperfections: back panel curvature, inner Fabry-Perot cavity, and pixel crosstalk. Based on the model, we developed a novel method of quantitative characterization of the imperfections and worked out the compensation algorithm, which can be applied to the generation of arbitrary beam profiles. For the validation, we produced $HG_{12,12}$ with the highest fidelity among other compensation methods.

6.2 Discussion.

Our experiment is quantum-inspired, however, it was performed in the classical regime, i.e. we did not count photons. Thus, the following questions remain: whether HGM gives an advantage in terms of Fisher information and, if it does, whether the resolution of HGM approaches the quantum limit. The former question refers to the following controversy. The resolution of direct images can be improved by applying post-processing algorithms such as deconvolution, which is impractical in the presence of noise, however, we discuss the theoretical possibility of improving the resolution here. Despite the resolution enhancement, none of the post-processing algorithms give growth in the Fisher information. Thus, the question of whether HGM extracts additional information from the image field is equivalent to the question of whether the resolution enhancement that HGM provides is a fundamental

or practical improvement. If HGM provides an advantage in Fisher information, the next question to be asked is whether this advantage is maximum possible, i.e. whether HGM is a quantum optimal measurement.

As HGM is a technique for full imaging, which can be seen as the estimation of a large number of parameters (positions of all point sources composing the object and their relative intensities), answering these questions requires an extended theoretical investigation. A numerical study carried out in our group (DPhil work of Jose Da Costa Filho) gives preliminary answers. In this study, the resolution enhancement (with respect to direct imaging) achievable by HGM based on mode-sorting and heterodyne detection was compared with the quantum optimal resolution enhancement for imaging of five incoherent point sources. The crucial point is that the calculation of the resolution enhancement was based on the Fisher information for multiparameter estimation rather than on comparison of direct and HGM images.

The results are as follows. In both settings, with heterodyne detection and mode-sorting, HGM overcomes direct imaging in terms of Fisher information for small separations between the sources, and HGM with mode-sorting approaches the quantum limit for small separations. HGM with both heterodyne detection and mode-sorting was simulated projecting the image field not only into Hermite-Gaussian modes but also into their superpositions since the sources are incoherent (see the discussion of HGM for incoherent light in subsection 4.4.2). The limitation of this study is that the number of sources was assumed to be known and fixed. For completing the theoretical research on the nature of HGM resolution enhancement and on whether HGM is a quantum optimal measurement, this limitation should be eliminated.

6.3 Future outlook

We propose to test HGM in various settings to move towards practical microscopy.

Incoherent light. In practice, most objects in imaging are incoherent. As was discussed in subsection 4.4.2, heterodyne detection still can be applied for quasi-monochromatic incoherent objects. However, phases of photocurrents are random in this case, and powers of photocurrents contain only even moments of the object intensity. This issue can be solved by placing an object to a single quadrant of the reference frame or using superpositions of Hermite-Gaussian modes.

Our setup can be easily adapted to simulate incoherent objects as follows. We display DMD bitmaps pixel by pixel and acquire a set of photocurrents for Hermite-Gaussian modes and superpositions of them for each pixel. By discarding phases and summing amplitudes of photocurrents from each pixel, we simulate projections of incoherent objects into LO. Using the adapted setup, we obtained preliminary experimental results demonstrating that HGM performs super-resolution imaging of incoherent objects with approximately the same resolution enhancement as demonstrated in the experiment with coherent light.

Mode-sorting. For broadband objects, usage of heterodyne detection is limited due to the limited bandwidth of electronics. Therefore, one of the mode-sorting methods reviewed in subsection 4.3.1 should be used instead. Besides being the solution for broadband objects, mode-sorting is not limited by the shot noise.

Imaging of microscopic objects. The final challenge is to perform actual microscopy using microscopic samples and a microscope objective followed by a mode-sorter. For an imaging system with a high-NA objective, we expect to achieve comparable results to that demonstrated in the current work, i.e. scaling of the resolution enhancement as the inverse square root of the highest mode order in the Hermite-Gaussian basis and a two-fold resolution enhancement beyond the Rayleigh limit for $HG_{20,20}$ as the highest-order mode.

To demonstrate HGM in microscopic settings, we will use fluorescent patterns produced by Argolight company (ARGO-HM slide) and an objective of NA ≥ 0.75. The Argolight slide contains 3D microscopic and nanoscopic samples, which can be exited by a wavelength ranging from 300 to 550 nm. For future experiments, we

chose the excitation wavelength of 405 nm. The broadband emission of samples is a continuum from slightly above the excitation wavelength to 800 nm. We are going to exploit confocal imaging setup due to the 3D shape of objects.

For the neural network training in microscopic settings, one could rely on off-the-shelf calibration slides and microplates for optical microscopes [37, 5]. A calibration slide of a few tens of μm size, containing several thousand training objects of size 0.5–1 μm, can be fabricated with a resolution of a few tens of nanometers by way of lithography or laser writing. This slide can be scanned in front of the microscope objective in various orientations to increase the training set size. If the variety and amount of the training data are still not sufficient, one could implement data augmentation techniques [145], adapting them to the HGM neural network.

Appendices

A | Derivation of the quantum Fisher information for separation estimation

Here we derive the right-hand side of Eq. (4.14). Let us represent the symmetric logarithmic derivative in equation

$$\mathcal{K}_Q \approx Tr[\hat{\rho}_1 \mathcal{L}_\theta^2] \tag{A.1}$$

as

$$\mathcal{L}_\theta = \frac{1}{2}(\mathcal{L}_{X_2} - \mathcal{L}_{X_1}), \tag{A.2}$$

where X_1 and X_2 are coordinates of the point sources, so that $\theta = X_1 - X_2$.

We introduce the following notation:

$$\begin{aligned}
|\psi_{11}\rangle &\equiv \frac{1}{\Delta k}\int_{-\infty}^{\infty} \frac{\partial\psi(x-X_1)}{\partial X_1}\,|x\rangle\,dx,\\
|\psi_{22}\rangle &\equiv \frac{1}{\Delta k}\int_{-\infty}^{\infty} \frac{\partial\psi(x-X_2)}{\partial X_2}\,|x\rangle\,dx,\\
\Delta k^2 &\equiv \int_{-\infty}^{\infty}\left[\frac{\partial\psi(x)}{\partial x}\right]^2 dx,\\
c_3 &\equiv \sqrt{\Delta k^2 + b^2 - \frac{\gamma^2}{1-\delta}},\\
c_4 &\equiv \sqrt{\Delta k^2 - b^2 - \frac{\gamma^2}{1+\delta}},\\
b^2 &\equiv \int_{-\infty}^{\infty} \frac{\partial\psi(x-X_1)}{\partial X_1}\frac{\partial\psi(x-X_2)}{\partial X_2}dx,\\
\delta &\equiv \int_{-\infty}^{\infty} \psi(x-X_1)\psi(x-X_2)dx,\\
\gamma &\equiv \int_{-\infty}^{\infty} \frac{\partial\psi(x)}{\partial x}\psi(x-\theta).
\end{aligned} \tag{A.3}$$

Assuming the PSF to be real, we choose the following set of orthonormal eigenvectors to span the support of $\hat{\rho}_1$ and $\partial\hat{\rho}_1/\partial\theta$:

$$\begin{aligned}
|e_1\rangle &= \frac{1}{\sqrt{2(1-\delta)}}(|\psi_1\rangle - |\psi_2\rangle),\\
|e_2\rangle &= \frac{1}{\sqrt{2(1+\delta)}}(|\psi_1\rangle + |\psi_2\rangle),\\
|e_3\rangle &= \frac{1}{c_3}\left[\frac{\Delta k}{\sqrt{2}}(|\psi_{11}\rangle + |\psi_{22}\rangle) - \frac{\gamma}{\sqrt{1-\delta}}|e_1\rangle\right],\\
|e_4\rangle &= \frac{1}{c_4}\left[\frac{\Delta k}{\sqrt{2}}(|\psi_{11}\rangle - |\psi_{22}\rangle) + \frac{\gamma}{\sqrt{1+\delta}}|e_2\rangle\right].
\end{aligned} \tag{A.4}$$

In this eigenbasis, $\hat{\rho}_1 = \sum_i p_i |e_i\rangle\langle e_i|$, where

$$p_1 = \frac{1-\delta}{2},\ p_2 = \frac{1+\delta}{2},\ p_3 = 0,\ p_4 = 0, \tag{A.5}$$

which follows from Eq. (4.12) and Eq. (A.4).

According to Eq. (4.6), elements of the symmetric logarithmic derivative $\mathcal{L}_{X_{\mu=1,2}}$ are

$$\mathcal{L}^{ij}_{X_\mu} = \frac{2}{p_i + p_j}\langle e_i|\frac{\partial\hat{\rho}_1}{\partial X_\mu}|e_j\rangle. \tag{A.6}$$

The partial derivative of $\hat{\rho}_1$ with respect to X_μ is

$$\frac{\partial\hat{\rho}_1}{\partial X_\mu} = \frac{\partial p_1}{\partial X_\mu}|e_1\rangle\langle e_1| + \frac{\partial p_2}{\partial X_\mu}|e_2\rangle\langle e_2| + \left[p_1\frac{\partial|e_1\rangle}{\partial X_\mu}\langle e_1| + p_2\frac{\partial|e_2\rangle}{\partial X_\mu}\langle e_2| + H.c.\right], \tag{A.7}$$

where $H.c.$ denotes the Hermitian conjugate.

Substituting Eq. (A.7) to Eq. (A.6) and taking into account Eq. (A.4) and Eq. (A.5), we get the unique and nonzero elements of real and symmetric matrix $\mathcal{L}^{ij}_{X_\mu}$:

$$\begin{aligned}
\mathcal{L}^{11}_{X_1} &= -\mathcal{L}^{11}_{X_2} = \frac{\gamma}{1-\delta},\\
\mathcal{L}^{12}_{X_1} &= \mathcal{L}^{12}_{X_2} = \frac{\gamma\delta}{\sqrt{1-\delta^2}},\\
\mathcal{L}^{13}_{X_1} &= -\mathcal{L}^{13}_{X_2} = \frac{c_3}{\sqrt{1-\delta}},\\
\mathcal{L}^{14}_{X_1} &= \mathcal{L}^{14}_{X_2} = \frac{c_4}{\sqrt{1-\delta}},\\
\mathcal{L}^{22}_{X_1} &= -\mathcal{L}^{22}_{X_2} = \frac{-\gamma}{1+\delta},\\
\mathcal{L}^{23}_{X_1} &= \mathcal{L}^{23}_{X_2} = \frac{c_3}{\sqrt{1+\delta}},\\
\mathcal{L}^{24}_{X_1} &= -\mathcal{L}^{24}_{X_2} = \frac{c_4}{\sqrt{1+\delta}}.
\end{aligned} \tag{A.8}$$

We use Eq. (A.8) to calculate $\mathcal{L}_\theta$ according to Eq. (A.2). Finally, substituting $\mathcal{L}_\theta$ to Eq. (A.1), we obtain that

$$\mathcal{K}_Q \approx \Delta k^2. \tag{A.9}$$

B | Derivation of the maximum-likelihood estimator for the separation

Maximum likelihood estimation is a method that will find the value of parameter $Q \equiv \theta^2/(16\sigma^2)$, such that the observed data (i.e. the set of photon counts $\{m_q\}$) is the most probable under the assumed model. The probability to detect a photon in qth mode is given by Eq. (4.22). Thus, the probability to detect m_q photons in qth mode is $[P(q)]^{m_q}$, and the probability of the observed data is

$$P\left(\{m_q\}\right) = \prod_q [P(q)]^{m_q} = \prod_q \left[\exp(-Q)\frac{Q^q}{q!}\right]^{m_q}. \tag{B.1}$$

The *maximum likelihood estimate* of Q is the value that maximizes $P\left(\{m_q\}\right)$. It can be found by differentiating the expression given by Eq. (B.1). However, it is computationally simpler to differentiate the natural logarithm of the expression, which gives the same result for the maximum likelihood estimate since the logarithm is a monotonically increasing function:

$$\begin{aligned} &\frac{\partial}{\partial Q}\log P\left(\{m_q\}\right) = \frac{\partial}{\partial Q}\sum_q m_q \log P(q) = \frac{\partial}{\partial Q}\sum_q m_q(-Q + q\log Q - \log q!) \\ &= \sum_q \left(-m_q + \frac{qm_q}{Q}\right) = -L + \sum_q \frac{qm_q}{Q}. \end{aligned} \tag{B.2}$$

Here, we used that $\sum_q m_q = L$, where L is the total amount of photons. Setting Eq. (B.2) to zero, we find that the maximum likelihood estimates for Q and θ can be evaluated applying the following estimators:

$$\check{Q} = \frac{1}{L}\sum_q q m_q, \; \check{\theta} = 4\sigma\sqrt{\check{Q}}. \tag{B.3}$$

C | An optical Eratosthenes' sieve for large prime numbers

This appendix reports our work [98] on the first experimental demonstration of prime number sieve via linear optics.

C.1 Motivation

The study of prime numbers, although millenia old, is still a trending topic in mathematics. On the theoretical side, there are relevant fundamental problems which remain open, such as the Riemann hypothesis and the primes' distribution [82]. On a more applied note, prime factorization, which is the cornerstone of modern cryptography, motivates a search for more efficent algorithms.

In the past decades, physics emerged as a fruitful venue for progress on prime numbers research. Most efforts were focused on factorization, culminating in the celebrated Shor's algorithm [144] and the current quest for practical quantum computers. However, the application of analogue physical tools for more fundamental mathematical problems is also being investigated, for example with respect to the Riemann hypothesis [138, 16].

A new and relatively unexplored testbed for the connection between physics and prime numbers is classical optics. On the factorization front, multi-wavelength interferometry [35] and optoelectronically assisted algorithms [141] schemes have been proposed, and experimental demonstrations were carried out based on the Talbot effect [120] and polychromatic interference in a multipath interferometer [153]. On the fundamental side, Berry [18, 19] proposed the reading of Riemann function's

zeros from radiation patterns using the property of the far field propagation to represent the Fourier transform.

Inspired by Berry's proposal, Petersen et al. [122] have recently proposed a physical prime numbers sieve based on the superposition of identical waves. The scheme is based on the sieve of Eratosthenes [38], one of the oldest algorithms for finding prime numbers. Given the sequence of integers n from 2 to M, Eratosthenes' sieve iteratively discards their multiples, such that all remaining integers lower than M^2 are necessarily primes. Ref. [122] proposes to realize the sieving by gratings of equally spaced n point-like sources: in the far field profile, each grating interferes constructively at locations corresponding to the multiples of n, and destructively at other integers. The pattern resulting from all these gratings has the prime numbers encoded in its intensity zeros, and locating them permits the realization of the sieve.

In our work [98], we implement and further develop the scheme devised in [122] by means of a phase-only spatial light modulator (SLM), where the apertures are realized by phase masks illuminated by a monochromatic laser. The far field intensity profile is measured by a CMOS camera, and the optical sieve corresponds to intensity minima of that profile. By making use of additional diffraction gratings and sequential camera acquisitions we can extend the sieving range beyond the pixel resolution of the SLM and the camera. The only remaining limitation associated with the SLM pixel resolution is the number of gratings that can be fitted in the phase mask. Capped by this limitation, we demonstrate an optical sieve capable of identifying all prime numbers below $149^2 = 22201$.

C.2 Theory

C.2.1 Sieving mask.

To illustrate the optical sieve proposed in [122], consider an amplitude grating of n equally spaced point-like sources, aligned along the horizontal direction x and with spatial frequency $1/n$:

$$g_n(x) = \sum_{i=0}^{n-1} \delta(x - i/n), \tag{C.1}$$

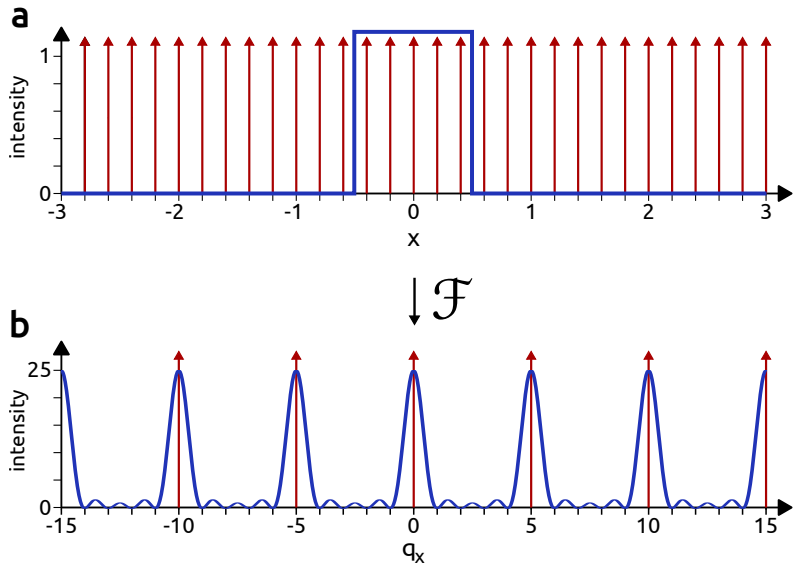


Figure C.1: Light intensity profiles at near (a) and far (b) fields, for $n = 5$ pinholes per unit length. Red arrows represent an infinite array of pinholes and the corresponding effect on the far field. When this array is restricted to five pinholes [which corresponds to multiplication by the blue envelope in (a)], the far field pattern (which is the result of a Fourier transformation $\mathcal{F}$ of the near field) becomes convolved with a sinc function, but the intensity at all integers except multiples of 5 remains at zero.

where $\delta(x)$ is the Dirac delta function. This can equivalently be seen as an infinite grating of pinholes illuminated by a plane monochromatic wave over a unit length window, as shown in Fig. C.1. Since far field propagation corresponds to taking a Fourier transform

$$\tilde{g}_n(q_x) = \int_{-\infty}^{\infty} g_n(x) e^{-2\pi i x q_x} dx, \tag{C.2}$$

where q_x is the horizontal coordinate in the far field plane, this configuration generates a field

$$\tilde{g}_n(q_x) = \left(n \sum_{j=-\infty}^{\infty} \delta(q_x - nj) \right) * \mathrm{sinc}(q_x), \tag{C.3}$$

where $\mathrm{sinc}(t) = \sin(\pi t)/\pi t$, of equally spaced peaks (due to the pinholes) convolved with a sinc envelope (due to the illumination window). The resulting intensity pattern has zeros at all integers except the multiples of n.

Superposing, in the object plane, these gratings g_n for $n \in \{2, 3, 5, .., M\}$ leads to a far field interference pattern where, akin to the sieve of Erastothenes, the multiples of the n's are "discarded" (by becoming intensity peaks) and the remaining

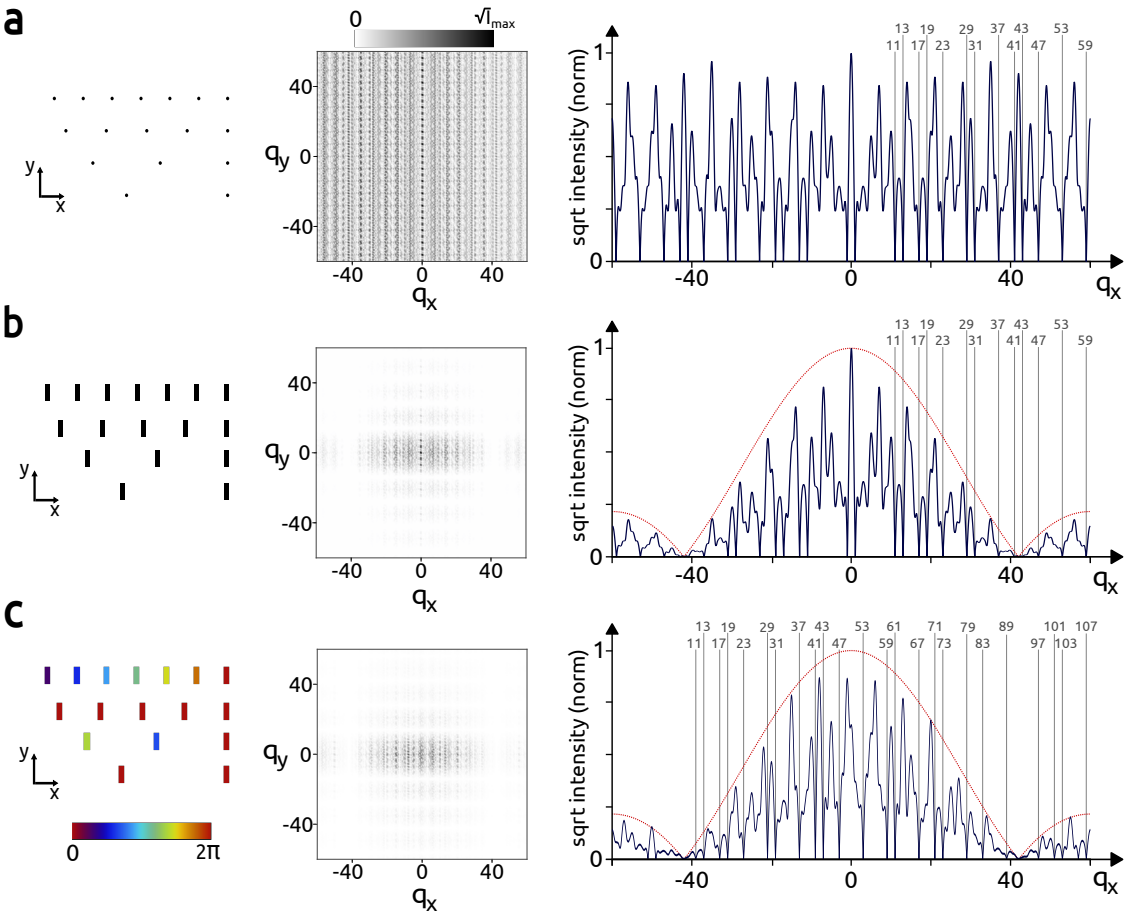


Figure C.2: Optical sieve diffraction mask (left) in the case of $\{n\} = \{2, 3, 5, 7\}$. The square root of the corresponding far field intensity profile (centre) and its horizontal marginal (integral over q_y, right), which reveals the prime numbers in the intensity zeros, highlighted by vertical grey lines. (a) Sieving mask proposed in [122], composed of point-like apertures. (b) Sieving mask with finite-size rectangular apertures of width $w = 1/42$. (c) Sieving mask superposed with displacement grating, which shifts the primes distribution by $d_x = 50$.

zeros correspond to q_x being a prime number. Therefore, via the mapping of the far field zeros, the scheme sieves the prime numbers in the range $M' \leq q_x < M'^2$, where M' is the next prime after M.

The superposition of gratings (which we refer to as the *sieving mask*) can be implemented by displaying each of them at a different vertical position, as shown in Fig. C.2 (left column). In the far field plane, Fig. C.2 (centre column), the waves from different gratings interfere, but the intensity at the primes' q_x positions remains zero for all q_y. Therefore, a horizontal profile of the far field [Fig. C.2 (right column)], obtained by integration along q_y, allows sieving of the

prime numbers as described before.

C.2.2 Displacement grating.

The above treatment assumes ideal point-like pinholes. A more realistic model would consider finite-size rectangular apertures, as depicted in Fig. C.2(b), which are mathematically represented by the convolution

$$E_n(x) = g_n(x) * T_w(x), \tag{C.4}$$

where $T_w(x)$ is the top-hat function and w is the aperture width. The resulting far field pattern will be modulated by a sinc envelope $\mathrm{sinc}(wq_x)$. This envelope affects the pattern's visibility, which progressively degrades for higher values of $|q_x|$, and also introduces additional zeros, which do not correspond to prime numbers. In practice, it limits the sieve's working range to the central lobe of the sinc.

This issue is addressed by adding a position-dependent linear function, dubbed here *displacement grating*, to the phase to each aperture (see Fig. C.2(c)):

$$g'_n(x) = g_n(x)e^{i2\pi x d_x}. \tag{C.5}$$

It results in a spatial shift of the intensity zeros by d_x in the q_x coordinate of the far field,

$$\tilde{g}'_n(q_x) = \tilde{g}_n(q_x - d_x) \tag{C.6}$$

so that

$$\tilde{E}'_n(x) = \tilde{g}'_n(q_x) \cdot w\mathrm{sinc}(wq_x). \tag{C.7}$$

Because we keep the phase inside each aperture constant, the modulation envelope $\mathrm{sinc}(wq_x)$ remains undisplaced. By progressively increasing the slope d_x of the displacement grating we are able to scan different sections of the Eratosthenes' sieve over the central lobe of the sinc envelope, therefore covering the full sieving range.

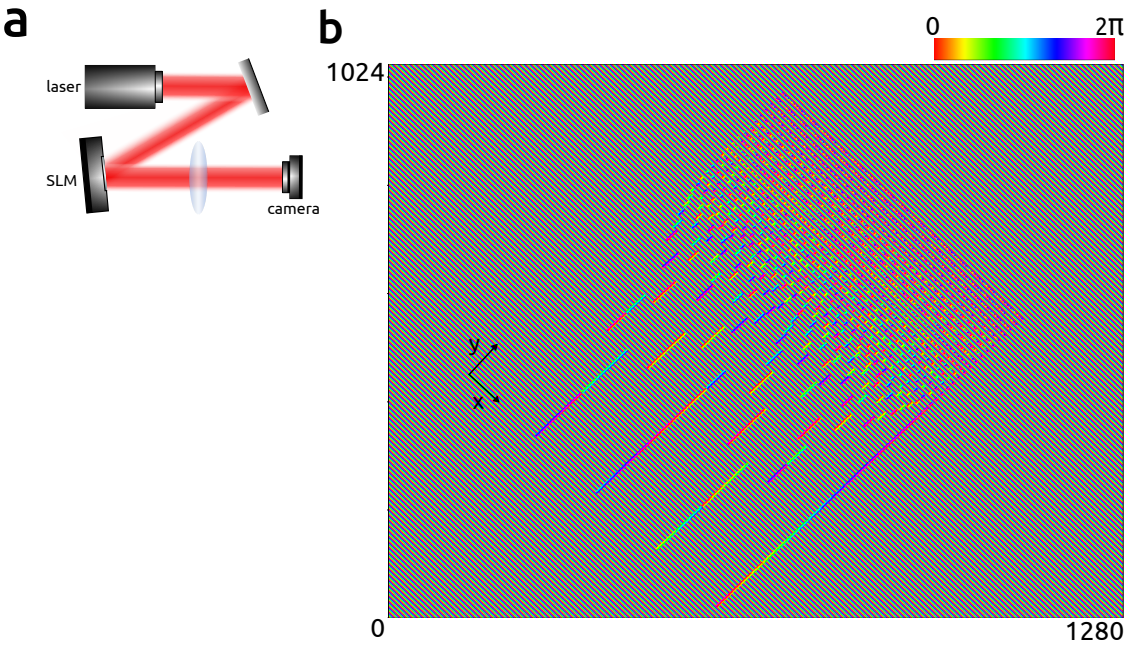


Figure C.3: Experimental implementation of the optical prime sieve. a) Scheme of the experiment. b) Phase mask hologram as displayed on the SLM screen.

C.3 Experiment

As illustrated in the scheme of the experimental setup in Fig. C.3(a), the optical prime sieve consists of a phase-only liquid crystal on silicon (LCoS) SLM illuminated almost uniformly by a continuous-wave laser (Eagleyard EYP-DFB-0785 with $\lambda = 785$ nm) beam, along with a camera placed in its far field.

The spatial profile of the emitted laser beam is filtered by a single mode fibre and magnified by a telescope before illuminating the SLM screen. At the SLM, the laser power is approximately 0.1mW. The phase-only LCoS-SLM (Hamamatsu X13138-02) has a spatial resolution of 1280×1024 pixels, 12.5um pixel pitch, 60Hz refresh rate and 256 grey levels.

A set of optical lenses ($f_1 = 250$mm, $f_2 = 75$mm, $f_3 = 150$mm) produces and magnifies the far field pattern. A monochrome CMOS sensor (uEye UI-3270 CP-M-NO-R2), of pixel pitch 3.45um, aligned with the SLM diagonal and anti-diagonal axis, acquires the far field intensity profile. The camera spatial resolution (2056×1542 pixels) and the chosen optics allow a proper spatial sampling rate, with approximately 10 camera pixels corresponding to $\Delta q_x = 1$. The camera's

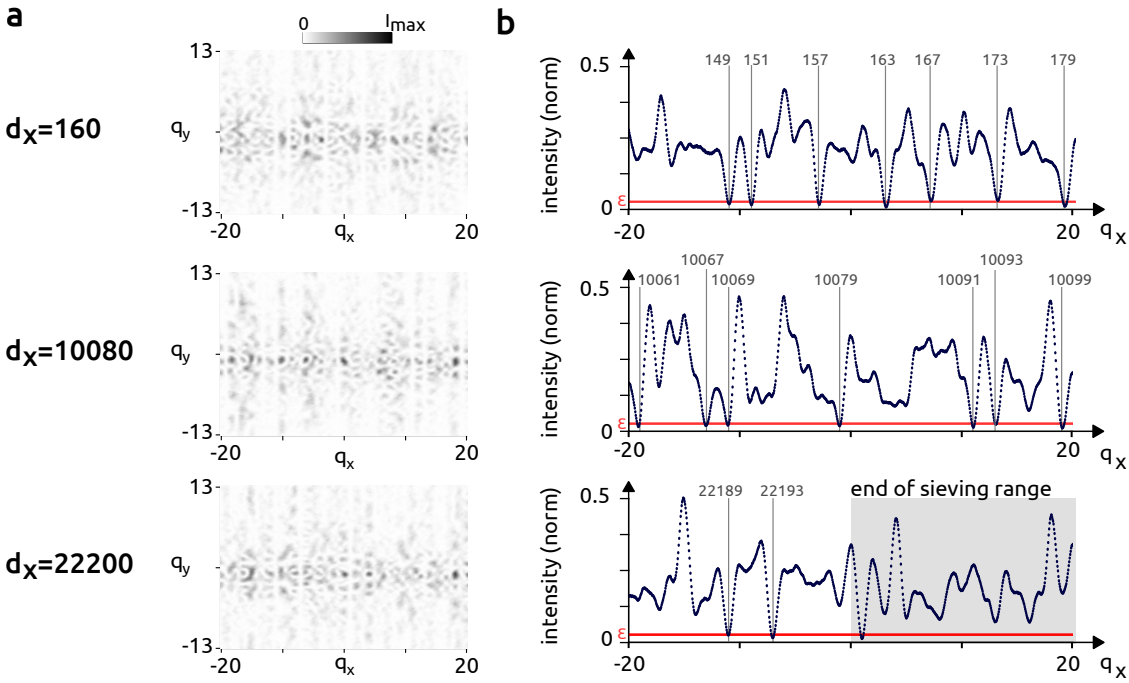


Figure C.4: Experimental results. a) Far field intensity images as acquired by the camera for different displacement grating slopes d_x, rescaled into the units of q_x, q_y. b) Integrals of the acquired intensity images along q_y. All minima below the intensity threshold ϵ (horizontal red line) are classified as prime numbers (vertical grey lines).

digitiser is set to 12-bit depth and the exposure time to 3.15 ms, to maximize intensity resolution and avoid saturation.

The sieving mask is implemented by phase holograms in the SLM screen, as exemplified in Fig. C.3(b). The background outside the "apertures" is filled with a blazed grating whose lines are oriented along the x axis. The function of this grating is to diffract the light away from the sieving area, so only the light reflected from the apertures is captured by the camera. The mask is arranged diagonally at 45°, in order to prevent the light scattered from both the SLM screen borders and the space between pixels, in the horizontal and vertical directions, from contaminating the image (in the laboratory reference frame, the SLM is mounted upright and the camera is physically tilted). Additionally, we apply a linear phase displacement to the apertures, i.e. we multiply $E'_n(x)$ by $e^{i2\pi x d_0}$, which displaces the entire far field pattern by d_0, thereby separating it from the SLM screen's specular reflection.

The patterns printed on the SLM extend over 420 pixels along x, accommodating apertures which are 3 pixels wide, corresponding to the effective width $w = 1/140$.

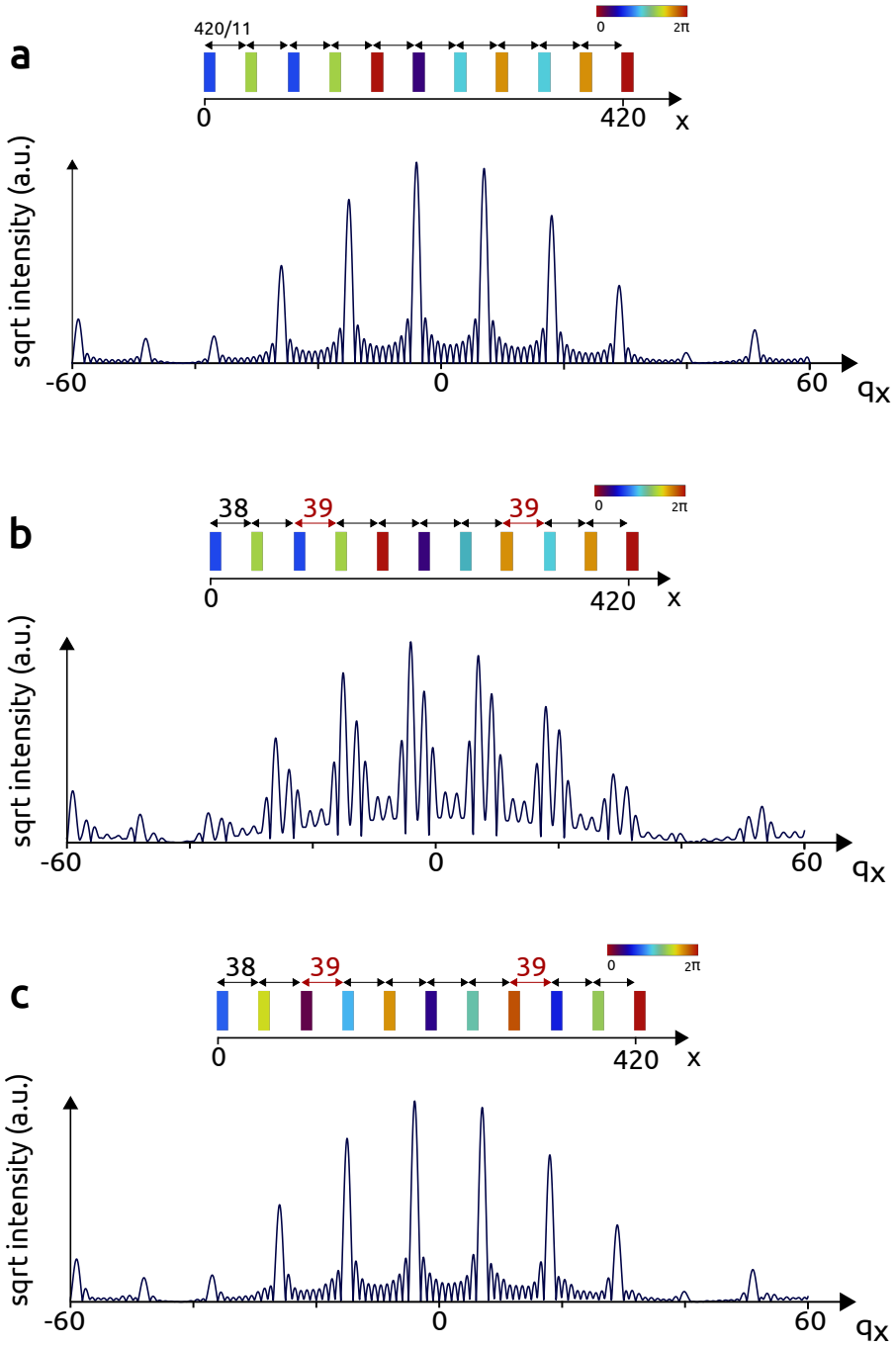


Figure C.5: Effect of the SLM's spatial resolution on the produced far field intensity profile for $n = 11$ and a displacement grating of slope $d_x = 180$ (theoretical model). (a) Evenly spaced apertures (infinite resolution); (b) Discrete pixels with a total pattern width of $W = 420$ pixels: the apertures are spaced by either 38 or 39 pixels. The integers not multiples of 11 are no longer zeros in the far field intensity profile. (c) $W = 420$ with a displacement grating of slope $d_x^{(11)} = 180 \mod 11 = 4$: the desired structure of the far field is closely restored.

These patterns include gratings corresponding to prime numbers up to $M = 139$, allowing a sieving range of $139 < q_x < 149^2 = 22201$, 149 being the next prime number after 139. The displacement gratings' slopes d_x are varied from zero to 22180 in increments of $R = 40$, covering the whole sieving range. While the size of the central lobe of the sinc function is $2/w = 280$, we chose a smaller increment value for higher contrast.

To improve the contrast, we set the slits' length along y to round$(200/n)$ pixels, so that each grating g_n contributes equally to the far field intensity pattern. We substitute g_2 and g_3 by pairs of gratings with $n = 4$ and $n = 6$, respectively, in order to reduce the space required by the grating with the longest slits. Additional displacement gratings are applied to one element of each pair so the collective action of the pair of elements is equivalent to that of g_2 and g_3. Further, we extend the total length of g_4 by a factor of 1.6, since it is located close to the SLM screen border and the laser power distribution, which is not perfectly uniform, is slightly lower there.

A key requirement on the optical prime sieve is the even spacing of its apertures, combined with their precise positioning. This is difficult to achieve in practice given a large number of gratings that need to be printed, as well as the finite SLM pixel resolution. If the apertures' spacing is rounded to the closest integer, the far field pattern becomes severely distorted and can no longer be used for sieving (see Fig. C.5(a,b)).

This distortion is significant when the function $\tilde{g}_n(\cdot)$ has a large argument, i.e., according to Eq. (C.6), for large displacement grating slopes d_x. We address this issue by recalling that each function $\tilde{g}_n(\cdot)$ is periodic with zeros at the integers not multiples of n. Hence displacing it by d_x is equivalent to displacing it by the remainder of the division of d_x by n. Therefore the slope of each displacement grating can be modified according to $d_x^{(n)} = d_x \mod n$ without any effect on the sieve, albeit at the cost of a different displacement slope for each line g_n. The effect of this correction is illustrated in Fig. C.5(c). This compensation allows us to keep the argument of $\tilde{g}_n(\cdot)$ low for all g_n, thereby achieving correct sieving.

The aberrations of the setup, due to both the lenses along the optical path and the intrinsic SLM curvature, distort the wavefront of the diffracted beam, reducing its spatial coherence and thus the interference visibility. This is corrected by adding a global phase profile to the SLM mask, which is measured via the calibration method proposed in [34].

Fig. C.4(a) shows the recorded far field intensity image for the SLM holograms of Fig. C.3(b). To maximize the contrast between primes zeros and composites minima, each recorded frame is integrated over 600 pixels along q_y, generating an average intensity plot $I(q_x)$ (see Fig. C.4(b)). During this operation, we also subtract the background noise, recorded after applying to the SLM a hologram without the apertures. To cover the whole sieving range, we need to acquire $22000/40 = 550$ frames. When the intensity value $I(q_x)$ is lower than the threshold ϵ, q_x is labeled as a prime.

We find that an intensity threshold that is in between 4.8% and 5.9% of the global maximum allows us to correctly sieve all the primes (2436 numbers) and discard all the composites (19626 numbers) within the sieving range (see Fig. C.6). This range quantifies the robustness of our scheme.

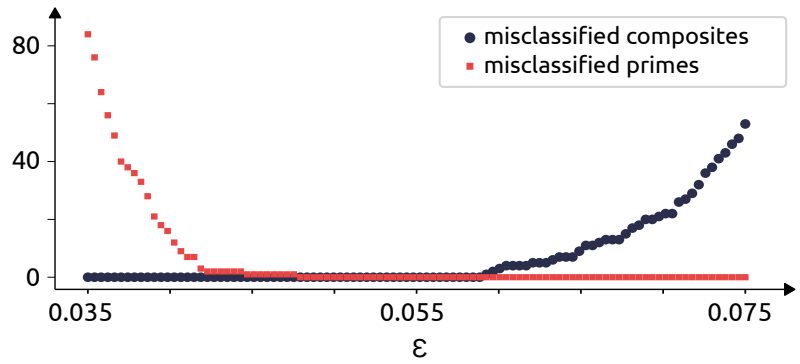


Figure C.6: Number of misclassified primes (red squares) and composites (blue dots) within the measured sieving range $139 < q_x < 22201$ as a function of the intensity threshold ϵ. By choosing $0.048 < \epsilon < 0.059$, the optical sieve correctly sieves all primes and discards composites.

C.4 Discussion

This work is the first experimental realization of an optical prime number sieve. By generating high resolution and on-demand phase masks via an SLM, we prove the effectiveness of the sieve by correctly classifying integers up to 22201 as either primes or composites.

The scalability of the sieving range, dictated by the number of apertures that can be fitted in a phase mask, is ultimately limited by the SLM spatial resolution. It is also limited by the SLM refresh rate, since a sequence of images has to be displayed in order to cover the whole sieving range. As SLM technology improves with respect to this two features, we expect better optical sieves to be implemented. The principally achievable range of the sieve can be estimated as the square of the SLM pixel resolution.

Note, however, that calculating the corrected displacement grating for each displacement d_x involves multiple modulo operations, which is equivalent, in terms of computational complexity, to checking the primality of d_x. Therefore our technique does not completely eliminate this type of calculations, but only reduces it by a factor equal to the increment value R. This value is limited by the width $2/w$ of the central lobe of the function $\mathrm{sinc}(wq_x)$, and hence cannot exceed the doubled SLM resolution.

Our method can be straightforwardly adapted to other kinds of sieves, such as Gaussian or twin primes sieves [122]. More generally, it demonstrates the potential of free-space linear optics in tasks that transcend imaging or communication, such as computation and solving mathematical problems. We can leverage the spatial coherence and superposition properties of light to achieve parallelism in computing.

Bibliography

[1] Ernst Abbe. "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung". In: *Archiv für mikroskopische Anatomie* 9.1 (1873), pp. 413–468.

[2] George Biddell Airy. "On the diffraction of an object-glass with circular aperture". In: *Transactions of the Cambridge Philosophical Society* 5 (1835), p. 283.

[3] Vahid Ansari et al. "Achieving the ultimate quantum timing resolution". In: *PRX Quantum* 2.1 (2021), p. 010301.

[4] D Appelt and R Heintzmann. "Structured illumination microscopy (SIM)". In: *Encyclopedia of Biophysics, G. Roberts, ed.(Springer, 2013)* (2013).

[5] *ArgoLight*. `http://argolight.com/technology/`. 2021.

[6] Victor Arrizón et al. "Pixelated phase computer holograms for the accurate encoding of scalar complex fields". In: *JOSA A* 24.11 (2007), pp. 3500–3507.

[7] Mehdi Asghari-Khiavi et al. "Exploring the origin of tip-enhanced Raman scattering; preparation of efficient TERS probes with high yield". In: *Journal of Raman Spectroscopy* 43.2 (2012), pp. 173–180.

[8] Andreas Assion et al. "Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses". In: *Science* 282.5390 (1998), pp. 919–922.

[9] Satyanarayana Bade et al. "Fabrication and characterization of a mode-selective 45-mode spatial multiplexer based on multi-plane light conversion". In: *2018 Optical Fiber Communications Conference and Exposition (OFC)*. IEEE. 2018, pp. 1–3.

[10] Brent Bailey et al. "Enhancement of axial resolution in fluorescence microscopy by standing-wave excitation". In: *Nature* 366.6450 (1993), pp. 44–48.

[11] Christopher J Bardeen et al. "Effect of pulse shape on the efficiency of multiphoton processes: implications for biological microscopy". In: *Journal of Biomedical Optics* 4.3 (1999), pp. 362–367.

[12] Thomas Bauer et al. "Observation of optical polarization Möbius strips". In: *Science* 347.6225 (2015), pp. 964–966.

[13] Katherine KM Bearne et al. "Confocal super-resolution microscopy based on a spatial mode sorter". In: *Optics Express* 29.8 (2021), pp. 11784–11792.

[14] Saraswati Behera and Joby Joseph. "Design and realization of functional metamaterial basis structures through optical phase manipulation based interference lithography". In: *Journal of Optics* 19.10 (2017), p. 105103.

[15] Marco W Beijersbergen et al. "Astigmatic laser mode converters and transfer of orbital angular momentum". In: *Optics Communications* 96.1-3 (1993), pp. 123–132.

[16] C.M. Bender, D.C. Brody, and M.P. Müller. "Hamiltonian for the zeros of the Riemann zeta function". In: *Phys. Rev. Lett.* 118.13 (2017), p. 130201.

[17] Gregorius CG Berkhout et al. "Efficient sorting of orbital angular momentum states of light". In: *Physical review letters* 105.15 (2010), p. 153601.

[18] M.V. Berry. "Riemann zeros in radiation patterns". In: *J. Phys. A* 45.30 (2012), p. 302001.

[19] M.V. Berry. "Riemann zeros in radiation patterns: II. Fourier transforms of zeta". In: *J. Phys. A* 48.38 (2015), p. 385203.

[20] Eric Betzig et al. "Imaging intracellular fluorescent proteins at nanometer resolution". In: *Science* 313.5793 (2006), pp. 1642–1645.

[21] Eric Betzig et al. "Near field scanning optical microscopy (NSOM): development and biophysical applications". In: *Biophysical journal* 49.1 (1986), pp. 269–279.

[22] Carolin Blum et al. "Tip-enhanced Raman spectroscopy–an interlaboratory reproducibility and comparison study". In: *Journal of Raman Spectroscopy* 45.1 (2014), pp. 22–31.

[23] Eliot Bolduc et al. "Exact solution to simultaneous intensity and phase encryption with a single phase-only hologram". In: *Optics letters* 38.18 (2013), pp. 3546–3549.

[24] Pauline Boucher et al. "Spatial optical mode demultiplexing as a practical tool for optimal transverse distance estimation". In: *Optica* 7.11 (2020), pp. 1621–1626.

[25] Alexandre Bouhelier et al. "Near-field second-harmonic generation induced by local field enhancement". In: *Physical review letters* 90.1 (2003), p. 013903.

[26] Markus Brehm et al. "Infrared spectroscopic mapping of single nanoparticles and viruses at nanoscale resolution". In: *Nano Letters* 6.7 (2006), pp. 1307–1310.

[27] Przemysław R Brejna and Peter R Griffiths. "Electroless deposition of silver onto silicon as a method of preparation of reproducible surface-enhanced Raman spectroscopy substrates and tip-enhanced Raman spectroscopy tips". In: *Applied spectroscopy* 64.5 (2010), pp. 493–499.

[28] Ward Brullot et al. "Resolving enantiomers using the optical angular momentum of twisted light". In: *Science advances* 2.3 (2016), e1501349.

[29] Miguel Carbonell-Leal and Omel Mendoza-Yero. "Encoding complex fields by using a phase-only optical element: mitigation of pixel crosstalk effects". In: *arXiv preprint arXiv:1903.06046* (2019).

[30] Nicolas J Cerf et al. "Security of quantum key distribution using d-level systems". In: *Physical review letters* 88.12 (2002), p. 127902.

[31] Chi Chen, Norihiko Hayazawa, and Satoshi Kawata. "A 1.7 nm resolution chemical analysis of carbon nanotubes by tip-enhanced Raman imaging in the ambient". In: *Nature communications* 5.1 (2014), pp. 1–5.

[32] Jianing Chen et al. "Optical nano-imaging of gate-tunable graphene plasmons". In: *Nature* 487.7405 (2012), pp. 77–81.

[33] Doo-Han Chung et al. "Alignment mechanism of nematic liquid crystal on rubbed polymer surface studied by subsequent processes of rubbing and photoalignment". In: *Japanese Journal of Applied Physics* 39.3A (2000), p. L185.

[34] Tomáš Čižmár, Michael Mazilu, and Kishan Dholakia. "In situ wavefront correction and its application to micromanipulation". In: *Nature Photonics* 4.6 (2010), p. 388.

[35] J.F. Clauser and J.P. Dowling. "Factoring integers with Young's N-slit interferometer". In: *Phys. Rev. A* 53.6 (1996), p. 4587.

[36] *Confocal microscopy*. Available at `https://ibidi.com/content/216-confocal-microscopy`.

[37] Alexander D Corbett et al. "Microscope calibration using laser written fluorescence". In: *Optics express* 26.17 (2018), pp. 21887–21899.

[38] R. Crandall and C.B. Pomerance. *Prime numbers: a computational perspective*. Vol. 182. Springer Science & Business Media, 2006.

[39] Jennifer E Curtis, Brian A Koss, and David G Grier. "Dynamic holographic optical tweezers". In: *Optics communications* 207.1-6 (2002), pp. 169–175.

[40] Neda Cvijetic et al. "Detecting lateral motion using light's orbital angular momentum". In: *Scientific reports* 5 (2015), p. 15422.

[41] A Cvitkovic, N Ocelic, and R Hillenbrand. "Analytical model for quantitative prediction of material contrasts in scattering-type near-field optical microscopy". In: *Optics express* 15.14 (2007), pp. 8550–8565.

[42] Matthias Danckwerts and Lukas Novotny. "Optical frequency mixing at coupled gold nanoparticles". In: *Physical Review Letters* 98.2 (2007), p. 026104.

[43] Jeffrey A Davis et al. "Encoding amplitude information onto phase-only filters". In: *Applied optics* 38.23 (1999), pp. 5004–5013.

[44] Herbert De Smet et al. "Design, fabrication and evaluation of a high-performance XGA VAN-LCOS microdisplay". In: *Displays* 23.3 (2002), pp. 89–98.

[45] Patrick Degenaar and Eiichi Tamiya. "Near-Field Optics in Biology". In: *Micromachines as Tools for Nanotechnology*. Springer, 2003, pp. 83–119.

[46] AL Demming, F Festy, and D Richards. "Plasmon resonances on metal tips: Understanding tip-enhanced Raman scattering". In: *The Journal of chemical physics* 122.18 (2005), p. 184716.

[47] Dong-Sheng Ding et al. "Quantum storage of orbital angular momentum entanglement in an atomic ensemble". In: *Physical review letters* 114.5 (2015), p. 050502.

[48] Jiangli Dong, Kin Seng Chiang, and Wei Jin. "Compact three-dimensional polymer waveguide mode multiplexer". In: *Journal of Lightwave Technology* 33.22 (2015), pp. 4580–4588.

[49] John M Donohue et al. "Quantum-limited time-frequency estimation through mode-selective photon measurement". In: *Physical review letters* 121.9 (2018), p. 090501.

[50] Ralf Dorn, Susanne Quabis, and Gerd Leuchs. "Sharper focus for a radially polarized light beam". In: *Physical review letters* 91.23 (2003), p. 233901.

[51] Robert C Dunn. "Near-field scanning optical microscopy". In: *Chemical reviews* 99.10 (1999), pp. 2891–2928.

[52] Michael Egmont-Petersen, Dick de Ridder, and Heinz Handels. "Image processing with neural networks—a review". In: *Pattern recognition* 35.10 (2002), pp. 2279–2301.

[53] David Engström et al. "Calibration of spatial light modulators suffering from spatially varying phase response". In: *Optics express* 21.13 (2013), pp. 16086–16103.

[54] Nicolas K Fontaine et al. "Hermite-Gaussian mode multiplexer supporting 1035 modes". In: *Optical Fiber Communication Conference*. Optical Society of America. 2021, pp. M3D–4.

[55] Nicolas K Fontaine et al. "Laguerre-Gaussian mode sorter". In: *Nature communications* 10.1 (2019), pp. 1–7.

[56] Nicolas K Fontaine et al. "Scalable mode sorter supporting 210 Hermite-Gaussian modes". In: *Optical Fiber Communication Conference*. Optical Society of America. 2018, Th4B–4.

[57] Andrew Forbes, Angela Dudley, and Melanie McLaren. "Creation and detection of optical modes with spatial light modulators". In: *Advances in Optics and Photonics* 8.2 (2016), pp. 200–227.

[58] Heinrich G Frey et al. "Enhancing the resolution of scanning near-field optical microscopy by a metal tip grown on an aperture probe". In: *Applied Physics Letters* 81.26 (2002), pp. 5030–5032.

[59] Shiyao Fu and Chunqing Gao. "Influences of atmospheric turbulence effects on the orbital angular momentum spectra of vortex beams". In: *Photonics Research* 4.5 (2016), B1–B4.

[60] MF Garcia-Parajo et al. "Near-field optical microscopy for DNA studies at the single molecular level". In: *Bioimaging* 6.1 (1998), pp. 43–53.

[61] Philippe Genevaux et al. “6-mode spatial multiplexer with low loss and high selectivity for transmission over few mode fiber”. In: *2015 Optical Fiber Communications Conference and Exhibition (OFC)*. IEEE. 2015, pp. 1–3.

[62] Faouzi Ghorbel et al. “Reconstructing with geometric moments”. In: *Proc. int. conf. on machine intelligence: Acidca-icmi*. Citeseer. 2005, pp. 5–7.

[63] Alexander A Govyadinov et al. “Quantitative measurement of local infrared absorption and dielectric function with tip-enhanced near-field microscopy”. In: *The journal of physical chemistry letters* 4.9 (2013), pp. 1526–1531.

[64] David G Grier and Yael Roichman. “Holographic optical trapping”. In: *Applied optics* 45.5 (2006), pp. 880–887.

[65] Mats GL Gustafsson. “Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution”. In: *Proceedings of the National Academy of Sciences* 102.37 (2005), pp. 13081–13086.

[66] Mats GL Gustafsson. “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy”. In: *Journal of microscopy* 198.2 (2000), pp. 82–87.

[67] Mats GL Gustafsson, David A Agard, and John W Sedat. “Sevenfold improvement of axial resolution in 3D wide-field microscopy using two objective-lenses”. In: *Three-Dimensional Microscopy: Image Acquisition and Processing II*. Vol. 2412. International Society for Optics and Photonics. 1995, pp. 147–156.

[68] Benjamin Harke et al. “Resolution scaling in STED microscopy”. In: *Optics express* 16.6 (2008), pp. 4154–4162.

[69] Norihiko Hayazawa et al. “Near-field Raman scattering enhanced by a metallized tip”. In: *Chemical Physics Letters* 335.5-6 (2001), pp. 369–374.

[70] Bert Hecht et al. “Scanning near-field optical microscopy with aperture probes: Fundamentals and applications”. In: *The Journal of Chemical Physics* 112.18 (2000), pp. 7761–7774.

[71] Stefan Hell and Ernst HK Stelzer. “Fundamental improvement of resolution with a 4Pi-confocal fluorescence microscope using two-photon excitation”. In: *Optics Communications* 93.5-6 (1992), pp. 277–282.

[72] Stefan W Hell and Jan Wichmann. “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy”. In: *Optics letters* 19.11 (1994), pp. 780–782.

[73] Peter Hermann et al. “Imaging and strain analysis of nano-scale SiGe structures by tip-enhanced Raman spectroscopy”. In: *Ultramicroscopy* 111.11 (2011), pp. 1630–1635.

[74] Nathaniel Hermosa et al. “Nanostep height measurement via spatial mode projection”. In: *Optics letters* 39.2 (2014), pp. 299–302.

[75] Samuel T Hess, Thanu PK Girirajan, and Michael D Mason. "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy". In: *Biophysical journal* 91.11 (2006), pp. 4258–4272.

[76] Markus Hiekkamäki, Shashi Prabhakar, and Robert Fickler. "Near-perfect measuring of full-field transverse-spatial modes of light". In: *Optics express* 27.22 (2019), pp. 31456–31464.

[77] Ming-Kuei Hu. "Visual pattern recognition by moment invariants". In: *IRE transactions on information theory* 8.2 (1962), pp. 179–187.

[78] Shu Huazhong, Luo Limin, and JL Coatrieux. "Moment-based approaches in imaging. Part 1. Basic features". In: *IEEE Eng. Med. Biol. Mag* 26 (2007), pp. 70–74.

[79] Florian Huth et al. "Resonant antenna probes for tip-enhanced infrared near-field microscopy". In: *Nano letters* 13.3 (2013), pp. 1065–1072.

[80] Jaroslav Icha et al. "Phototoxicity in live fluorescence microscopy, and how to avoid it". In: *BioEssays* 39.8 (2017), p. 1700003.

[81] Taro Ichimura et al. "Tip-enhanced coherent anti-Stokes Raman scattering for vibrational nanoimaging". In: *Physical review letters* 92.22 (2004), p. 220801.

[82] A.E. Ingham. *The distribution of prime numbers*. 30. Cambridge University Press, 1990.

[83] CL Jahncke, HD Hallen, and MA Paesler. "Nano-Raman Spectroscopy and Imaging with a Near-Field Scanning Optical Microscope". In: *Journal of Raman spectroscopy* 27.8 (1996), pp. 579–586.

[84] Wiebke Jahr, Philipp Velicky, and Johann Georg Danzl. "Strategies to maximize performance in STimulated Emission Depletion (STED) nanoscopy of biological specimens". In: *Methods* 174 (2020), pp. 27–41.

[85] Felix Jünger, Philipp v Olshausen, and Alexander Rohrbach. "Fast, label-free super-resolution live-cell imaging using rotating coherent scattering (ROCS) microscopy". In: *Scientific reports* 6.1 (2016), pp. 1–11.

[86] Fritz Keilmann and Rainer Hillenbrand. "Near-field microscopy by elastic light scattering from a tip". In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 362.1817 (2004), pp. 787–805.

[87] Taekyeong Kim et al. "Multilayered nano-prism vertex tips for tip-enhanced Raman spectroscopy and imaging". In: *Analyst* 138.19 (2013), pp. 5588–5593.

[88] Andrei Kirilyuk. "Nonlinear optics in application to magnetic surfaces and thin films". In: *Journal of Physics D: Applied Physics* 35.21 (2002), R189.

[89] Thomas A Klar et al. "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission". In: *Proceedings of the National Academy of Sciences* 97.15 (2000), pp. 8206–8210.

[90] C Kohler, F Zhang, and Wolfgang Osten. "Characterization of a spatial light modulator and its application in phase retrieval". In: *Applied optics* 48.20 (2009), pp. 4003–4008.

[91] Ranjeet Kumar et al. "Versatile wideband balanced detector for quantum optical homodyne tomography". In: *Optics Communications* 285.24 (2012), pp. 5259–5267.

[92] Guillaume Labroille et al. "Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion". In: *Optics express* 22.13 (2014), pp. 15599–15607.

[93] Hugo Larocque et al. "Optical framed knots as information carriers". In: *Nature communications* 11.1 (2020), pp. 1–8.

[94] Hugo Larocque et al. "Reconstructing the topology of optical polarization knots". In: *Nature Physics* 14.11 (2018), p. 1079.

[95] Martin PJ Lavery et al. "Detection of a spinning object using light's orbital angular momentum". In: *Science* 341.6145 (2013), pp. 537–540.

[96] Grigory Lazarev et al. "Beyond the display: Phase-only liquid crystal on Silicon devices and their applications in photonics". In: *Optics express* 27.11 (2019), pp. 16206–16249.

[97] AL Lereu, Ali Passian, and Ph Dumas. "Near field optical microscopy: a brief review". In: *International Journal of Nanotechnology* 9.3-7 (2012), pp. 488–501.

[98] Bohan Li et al. "Optical Eratosthenes' sieve for large prime numbers". In: *Optics express* 28.8 (2020), pp. 11965–11973.

[99] Ye Li, Lin Yu, and Yixin Zhang. "Influence of anisotropic turbulence on the orbital angular momentum modes of Hermite-Gaussian vortex beam in the ocean". In: *Optics express* 25.11 (2017), pp. 12203–12215.

[100] Christian Lingel, Tobias Haist, and Wolfgang Osten. "Optimizing the diffraction efficiency of SLM-based holography with respect to the fringing field effect". In: *Applied optics* 52.28 (2013), pp. 6877–6883.

[101] Yong Liu and Cuifang Kuang. "Frequency shifting confocal microscopy via azimuthally polarized Bessel-Gaussian beam". In: *Advanced Optical Imaging Technologies II*. Vol. 11186. International Society for Optics and Photonics. 2019, 111861B.

[102] AI Lvovsky. "Iterative maximum-likelihood reconstruction in quantum homodyne tomography". In: *Journal of Optics B: Quantum and Semiclassical Optics* 6.6 (2004), S556.

[103] Mhlambululi Mafu et al. "Higher-dimensional orbital-angular-momentum-based quantum key distribution with mutually unbiased bases". In: *Physical Review A* 88.3 (2013), p. 032305.

[104] Giorgio Maltese et al. *Hermite-Gaussian Microscopy Github repository*. `https://github.com/giomalt/HGM`. 2021.

[105] José Luis Martínez et al. "Analysis of multiple internal reflections in a parallel aligned liquid crystal on silicon SLM". In: *Optics express* 22.21 (2014), pp. 25866–25879.

[106] *Materials processing using unconventional ultrafast laser beams*. Available at `https://physics.anu.edu.au/laser/research/processing.php`. 2019.

[107] Christian Maurer et al. "Tailoring of arbitrary optical vector beams". In: *New Journal of Physics* 9.3 (2007), p. 78.

[108] Peiwen Meng, Silvania Pereira, and Paul Urbach. "Confocal microscopy with a radially polarized focused beam". In: *Optics express* 26.23 (2018), pp. 29600–29613.

[109] Yu V Miklyaev et al. "Superresolution microscopy in far-field by near-field optical random mapping nanoscopy". In: *Applied Physics Letters* 105.11 (2014), p. 113103.

[110] Jean-François Morizur et al. "Programmable unitary spatial mode manipulation". In: *JOSA A* 27.11 (2010), pp. 2524–2531.

[111] Simon Moser, Monika Ritsch-Marte, and Gregor Thalhammer. "Model-based compensation of pixel crosstalk in liquid crystal spatial light modulators". In: *Optics express* 27.18 (2019), pp. 25046–25063.

[112] Claus B Müller and Jörg Enderlein. "Image scanning microscopy". In: *Physical review letters* 104.19 (2010), p. 198101.

[113] Mark AA Neil et al. "Method for the generation of arbitrary complex vector wave fronts". In: *Optics letters* 27.21 (2002), pp. 1929–1931.

[114] Ute Neugebauer et al. "Towards a detailed understanding of bacterial metabolism—spectroscopic characterization of Staphylococcus epidermidis". In: *ChemPhysChem* 8.1 (2007), pp. 124–137.

[115] Michael A Nielsen. *Neural networks and deep learning*. Vol. 25. Determination press San Francisco, CA, 2015.

[116] Qifeng Niu, Chunyang Wang, and Hongwei Shi. "Phase modulation characteristics analysis of liquid crystal spatial light modulator under oblique incidence". In: *Review of Scientific Instruments* 90.5 (2019), p. 055001.

[117] Joaquín Otón et al. "Multipoint phase calibration for improved compensation of inherent wavefront distortion in parallel aligned liquid crystal on silicon displays". In: *Applied optics* 46.23 (2007), pp. 5667–5679.

[118] Valentina Parigi et al. "Storage and retrieval of vector beams of light in a multiple-degree-of-freedom quantum memory". In: *Nature communications* 6 (2015), p. 7706.

[119] Martin Paúr et al. "Achieving the ultimate optical resolution". In: *Optica* 3.10 (2016), pp. 1144–1147.

[120] K. Pelka et al. "Prime number decomposition using the Talbot effect". In: *Opt. Express* 26.12 (2018), pp. 15009–15014.

[121] Martin Persson, David Engström, and Mattias Goksör. "Reducing the effect of pixel crosstalk in phase only spatial light modulators". In: *Optics express* 20.20 (2012), pp. 22334–22343.

[122] T.C. Petersen et al. "Simple wave-optical superpositions as prime number sieves". In: *Phys. Rev. Lett.* 122.9 (2019), p. 090201.

[123] Dieter W Pohl, Winfried Denk, and Mark Lanz. "Optical stethoscopy: Image recording with resolution $\lambda/20$". In: *Applied physics letters* 44.7 (1984), pp. 651–653.

[124] A. A. Pushkina et al. "Superresolution Linear Optical Imaging in the Far Field". In: *Phys. Rev. Lett.* 127 (25 2021), p. 253602. DOI: 10.1103/PhysRevLett.127.253602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.253602.

[125] AA Pushkina et al. "Comprehensive model and performance optimization of phase-only spatial light modulators". In: *Measurement Science and Technology* 31.12 (2020), p. 125202.

[126] Anastasiia Pushkina et al. *Python program to generate phase-only SLM holograms*. 2020. URL: https://github.com/giomalt/SLM_hologram_generation (visited on 06/20/2020).

[127] Markus B Raschke and Christoph Lienau. "Apertureless near-field optical microscopy: Tip–sample coupling in elastic light scattering". In: *Applied Physics Letters* 83.24 (2003), pp. 5089–5091.

[128] Lord Rayleigh. "XV. On the theory of optical images, with special reference to the microscope". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 42.255 (1896), pp. 167–195.

[129] J Řeháček et al. "Intensity-based axial localization at the quantum limit". In: *Physical review letters* 123.19 (2019), p. 193601.

[130] Jaroslav Rehacek et al. "Optimal measurements for resolution beyond the Rayleigh limit". In: *Optics letters* 42.2 (2017), pp. 231–234.

[131] Eva Rittweger et al. "STED microscopy reveals crystal colour centres with nanometric resolution". In: *Nature Photonics* 3.3 (2009), p. 144.

[132] Brian Robertson et al. "Demonstration of multi-casting in a 1× 9 LCOS wavelength selective switch". In: *Journal of lightwave technology* 32.3 (2013), pp. 402–410.

[133] Michael J Rust, Mark Bates, and Xiaowei Zhuang. "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". In: *Nature methods* 3.10 (2006), pp. 793–796.

[134] Pablo Salaverria. *Super-Resolution Structured Illumination Microscopy (SR-SIM)*. Available at https://www.cherrybiotech.com/scientific-note/microscopy/super-resolution-structured-illumination-microscopy-sr-sim.

[135] Jean-Sébastien Samson et al. "Setup of a scanning near field infrared microscope (SNIM): Imaging of sub-surface nano-structures in gallium-doped silicon". In: *Physical Chemistry Chemical Physics* 8.6 (2006), pp. 753–758.

[136] Manish Saxena, Gangadhar Eluru, and Sai Siva Gorthi. "Structured illumination microscopy". In: *Advances in Optics and Photonics* 7.2 (2015), pp. 241–275.

[137] Valerio Scarani et al. "The security of practical quantum key distribution". In: *Reviews of modern physics* 81.3 (2009), p. 1301.

[138] D. Schumayer and D.A.W. Hutchinson. "Colloquium: Physics of the Riemann hypothesis". In: *Rev. Mod. Phys.* 83.2 (2011), p. 307.

[139] Fabio Sciarrino. "Complete experimental toolbox for alignment-free quantum communication". In: *Bulletin of the American Physical Society* 58 (2013).

[140] Mohamed S Seghilani et al. "Vortex Laser based on III-V semiconductor metasurface: direct generation of coherent Laguerre-Gauss modes carrying controlled orbital angular momentum". In: *Scientific reports* 6 (2016), p. 38156.

[141] A. Shamir. "Factoring large numbers with the TWINKLE device". In: *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer. 1999, pp. 2–12.

[142] *Shear force microscopy*. Available at `https://www.ntmdt-si.com/resources/spm-principles/snom/shear-force-microscopy`.

[143] Colin JR Sheppard, Shalin B Mehta, and Rainer Heintzmann. "Superresolution by image scanning microscopy using pixel reassignment". In: *Optics letters* 38.15 (2013), pp. 2889–2892.

[144] P. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer". In: *SIAM J. Comput.* 26.5 (1997), pp. 1484–1509.

[145] Connor Shorten and Taghi M Khoshgoftaar. "A survey on image data augmentation for deep learning". In: *Journal of Big Data* 6.1 (2019), pp. 1–48.

[146] Hari Shroff et al. "Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics". In: *Nature methods* 5.5 (2008), pp. 417–423.

[147] EH Singe. "Suggested method for extending microscopic resolution into the ultramicroscopic regime". In: *Phil. Mag* 6 (1928), p. 356.

[148] Alicia Sit et al. "High-dimensional intracity quantum cryptography with structured photons". In: *Optica* 4.9 (2017), pp. 1006–1010.

[149] II Smolyaninov et al. "Near-Field Second Harmonic Microscopy of Thin Ferroelectric Films". In: *MRS Online Proceedings Library Archive* 596 (1999).

[150] Jens Steidtner and Bruno Pettinger. "Tip-enhanced Raman spectroscopy and microscopy on single dye molecules with 15 nm resolution". In: *Physical Review Letters* 100.23 (2008), p. 236101.

[151] Raoul M Stöckle et al. “Nanoscale chemical analysis by tip-enhanced Raman spectroscopy”. In: *Chemical Physics Letters* 318.1-3 (2000), pp. 131–136.

[152] Satoshi Takahashi et al. “High resolution photon scanning tunneling microscope”. In: *Nanotechnology* 8.3A (1997), A54.

[153] V. Tamma et al. “Factoring numbers with a single interferogram”. In: *Phys. Rev. A* 83.2 (2011), p. 020304.

[154] Zong Sheng Tang, Kadir Durak, and Alexander Ling. “Fault-tolerant and finite-error localization for point emitters within the diffraction limit”. In: *Optics express* 24.19 (2016), pp. 22004–22012.

[155] BVR Tata and Baldev Raj. “Confocal laser scanning microscopy: Applications in material science and technology”. In: *Bulletin of Materials Science* 21.4 (1998), pp. 263–278.

[156] Thomas Taubner, R Hillenbrand, et al. “Nanoscale-resolved subsurface imaging by scattering-type near-field optical microscopy”. In: *Optics Express* 13.22 (2005), pp. 8893–8899.

[157] Ron Tenne et al. “Super-resolution enhancement by quantum image scanning microscopy”. In: *Nature Photonics* 13.2 (2019), pp. 116–122.

[158] Weng-Kian Tham, Hugo Ferretti, and Aephraim M Steinberg. “Beating Rayleigh’s curse by imaging using phase information”. In: *Physical review letters* 118.7 (2017), p. 070801.

[159] *The Principles of Operation of Scanning Tunneling Microscope (STM).* Available at `https://www.azom.com/article.aspx?ArticleID=4436`.

[160] Mankei Tsang. “Subdiffraction incoherent optical imaging via spatial-mode demultiplexing”. In: *New Journal of Physics* 19.2 (2017), p. 023054.

[161] Mankei Tsang and Ranjith Nair. “Comment on" Resurgence of Rayleigh’s curse in the presence of partial coherence"”. In: *arXiv preprint arXiv:1810.12166* (2018).

[162] Mankei Tsang, Ranjith Nair, and Xiao-Ming Lu. “Quantum theory of superresolution for two incoherent optical point sources”. In: *Physical Review X* 6.3 (2016), p. 031033.

[163] My-Phung Van et al. “Enhanced lithographic resolution using longitudinal polarization state of light”. In: *Journal of Micro/Nanolithography, MEMS, and MOEMS* 14.4 (2015), p. 043509.

[164] Jeroen Vangindertael et al. “An introduction to optical super-resolution microscopy for the adventurous biologist”. In: *Methods and applications in fluorescence* 6.2 (2018), p. 022003.

[165] SA Wadood et al. “Superresolution of partially coherent light sources using parity sorting”. In: *arXiv preprint arXiv:2102.01603* (2021).

[166] Jian Wang et al. “Terabit free-space data transmission employing orbital angular momentum multiplexing”. In: *Nature photonics* 6.7 (2012), p. 488.

[167] Mi Wang et al. "LCoS SLM study and its application in wavelength selective switch". In: *Photonics*. Vol. 4. 2. Multidisciplinary Digital Publishing Institute. 2017, p. 22.

[168] Xi-Lin Wang et al. "Quantum teleportation of multiple degrees of freedom of a single photon". In: *Nature* 518.7540 (2015), pp. 516–519.

[169] Zhihao Wang, Jian Chen, and Steven CH Hoi. "Deep learning for image super-resolution: A survey". In: *IEEE transactions on pattern analysis and machine intelligence* (2020).

[170] Andrew M Weiner. "Ultrafast optical pulse shaping: A tutorial review". In: *Optics Communications* 284.15 (2011), pp. 3669–3692.

[171] Jolly Xavier and Joby Joseph. "Tunable complex photonic chiral lattices by reconfigurable optical phase engineering". In: *Optics letters* 36.3 (2011), pp. 403–405.

[172] Fajun Xiao et al. "Selective excitation of a three-dimensionally oriented single plasmonic dipole". In: *Photonics Research* 7.6 (2019), pp. 693–698.

[173] Xiaodong Xun and Robert W Cohn. "Phase calibration of spatially nonuniform spatial light modulators". In: *Applied optics* 43.35 (2004), pp. 6400–6406.

[174] Keisaku Yamane et al. "Optical pulse compression to 3.4 fs in the monocycle region by feedback phase compensation". In: *Optics Letters* 28.22 (2003), pp. 2258–2260.

[175] Fan Yang et al. "Far-field linear optical superresolution via heterodyne detection in a higher-order local oscillator mode". In: *Optica* 3.10 (2016), pp. 1148–1152.

[176] Fan Yang et al. "Fisher information for far-field linear optical superresolution via homodyne or heterodyne detection in a higher-order local oscillator mode". In: *Physical Review A* 96.6 (2017), p. 063829.

[177] Alison M Yao and Miles J Padgett. "Orbital angular momentum: origins, behavior and applications". In: *Advances in Optics and Photonics* 3.2 (2011), pp. 161–204.

[178] Anatoly V Zayats and Vahid Sandoghdar. "Apertureless scanning near-field second-harmonic microscopy". In: *Optics communications* 178.1-3 (2000), pp. 245–249.

[179] Anatoly V Zayats and Igor I Smolyaninov. "Near-field photonics: surface plasmon polaritons and localized surface plasmons". In: *Journal of Optics A: Pure and Applied Optics* 5.4 (2003), S16.

[180] Hongxin Zhang, Jian Zhang, and Liying Wu. "Evaluation of phase-only liquid crystal spatial light modulator for phase modulation performance using a Twyman–Green interferometer". In: *Measurement Science and Technology* 18.6 (2007), p. 1724.

[181] Yaoju Zhang and Jianping Bai. "Improving the recording ability of a near-field optical storage system by higher-order radially polarized beams". In: *Optics express* 17.5 (2009), pp. 3698–3706.

[182] Zichen Zhang, Zheng You, and Daping Chu. "Fundamentals of phase-only liquid crystal on silicon (LCOS) devices". In: *Light: Science & Applications* 3.10 (2014), e213.

[183] Haining Zhong. "Photoactivated localization microscopy (PALM): an optical technique for achieving˜ 10-nm resolution". In: *Cold Spring Harbor protocols* 2010.12 (2010), pdb–top91.

[184] Yiyu Zhou et al. "Hermite–Gaussian mode sorter". In: *Optics letters* 43.21 (2018), pp. 5263–5266.

[185] Yiyu Zhou et al. "Quantum-limited estimation of the axial separation of two incoherent point sources". In: *Optica* 6.5 (2019), pp. 534–541.

[186] Zong-Quan Zhou et al. "Quantum storage of three-dimensional orbital-angular-momentum entanglement in a crystal". In: *Physical review letters* 115.7 (2015), p. 070502.

[187] Long Zhu and Jian Wang. "Arbitrary manipulation of spatial amplitude and phase using phase-only spatial light modulators". In: *Scientific reports* 4 (2014), p. 7441.

[188] Xiaowei Zhuang. "Unraveling DNA condensation with optical tweezers". In: *Science* 305.5681 (2004), pp. 188–190.

[189] Lothar Zitzler, Stephan Herminghaus, and Frieder Mugele. "Capillary forces in tapping mode atomic force microscopy". In: *Physical Review B* 66.15 (2002), p. 155436.