

Quantum Super-resolution

by Siobhan Maeve Tobin



Australian
National
University

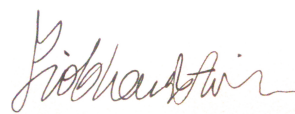
A thesis submitted in partial fulfilment of the
requirements for the Degree of
Bachelor of Philosophy (Science Honours) at
The Australian National University

October 2016

Declaration

This thesis is an account of research undertaken between February 2016 and October 2016 at the Department of Quantum Science, College of Physical and Mathematical Sciences, The Australian National University, Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

A handwritten signature in brown ink, reading "Siobhan Maeve Tobin", written over a horizontal line.

Siobhan Maeve Tobin

October 2016

Acknowledgements

A massive thanks to Fouad and Naeem of the ANFF for giving their time and expertise so generously. We got there in the end and enjoyed some great conversations while we were at it!

Assad, without a doubt the hero of this project. You know so much about so many things and always found time for my questions. Plus you kept things moving forwards in the lab. Your usual response to me running off to the great outdoors - "Sounds fun. Can I come too?" - reassured me that I wasn't going crazy. You are a super physicist and a great mate too.

Thanks to Amanda, Lynne, Zhao Jie and Ruvi for making me laugh! Thanks to Ben and PK "What's a hashtag?" Lam, for keeping me planning on my feet and having the opportunity to start an experiment from scratch. Such a learning experience.

Paula for being an oracle these past four years.

Jodie - I don't know how I would have gotten to this point without your support. Whether it has been channelling my rage, providing a backbone for my dreams or a shoulder to cry on, buying me smoothies or just being happy to chat about home, nothing has been too much, or moreover, too trivial for you. You are my firmest ally and friend at RSPE, thank you. #buninyong 🙌🎓

These people never hesitated to offer a hug even when I didn't know I needed one. Thanks to Jacqueline for keeping the bastards honest. Chris for being my best friend and fighting like a girl. Mark and Hugh for the stats, the maths and the Suds Wars. Matt for being equal parts distraction and inspiration. Isla and Aithne for being two of the bravest and coolest women ever. Parents for teaching me to read and to love.

Room 45: your care, your insight, your persistence this year has been extraordinary. Alix, Bonnie, Jake, Matt and Ruth: the Physics Olympiad has been the best thing I've ever done largely because of your fantastic friendship and pedagogy. Adrian, Ashleigh, Jacob, Peter, Scarlett: motivations to find things out don't come much greater.

I am proud to present my honours thesis ... in three emojis or less:



(Incoherent point sources of light, interferometry, Fisher information.)

- S

Abstract

One can locate incoherent sources of light even when they are very close together within an error bound. This thesis contributes to knowledge by presenting progress towards experimentally realising the super-resolution of two incoherent point sources of light, using Fisher information gleaned through image-inversion interferometry measurements. The diffraction limit has been a limiting construct in optical physics for centuries without it being a fundamental physical law. This study is important because it highlights what rich knowledge is within reach while still working with optical wavelengths.

Techniques in fluorescence microscopy previously bounded by Rayleigh's criterion can be enhanced using interferometry, enabling the non-destructive study of smaller and smaller biological samples. At the other end of the length scale, astrophysical applications include resolving binary star systems and discovering exoplanets. What is common in all these cases is that localisation is possible using quantum estimation theory even if the point spread function for each incoherent source is unable to be obtained.

Here we tackle a experimental super-resolution technique, superlocalisation by image inversion interferometry, alongside advances in construction of optical devices that can act as light sources with sub-Rayleigh separations. Beyond the successful operation of an image-inversion interferometer, future work in this sphere may add in a directional element to this separation through interferometry in multiple dimensions.

Contents

Declaration	i
Acknowledgements	ii
Abstract	iii
List of figures	vii
Table of symbols used	viii
Table of acronyms used	ix
1 Introduction	1
2 Background theory	6
2.1 Estimation Theory	6
2.1.1 Minimum Variance Unbiased Estimation	7
2.1.2 Fisher information	8
2.1.3 Cramér-Rao Lower Bound	10
2.1.4 Maximum likelihood estimation	11
2.2 Quantum Fisher information	11
2.3 Statistical optics	13
2.3.1 Coherence	14
2.3.2 Transverse modes	15
2.4 Photodetection	16
2.4.1 Poisson processes	17
2.4.2 Semiclassical and quantum theories of detection	18
2.5 Resolution	18
2.5.1 Point spread functions	18
2.5.2 Quantifying resolution	19
2.5.3 Resolution for incoherent point sources	23
3 Super-resolution	25
3.1 Recent progress	25
3.2 Interferometry	27
3.2.1 Image-inversion interferometry	29

3.3	Superlocalisation via image inversion interferometry (SLIVER) . . .	29
3.3.1	Statistical optical modelling	30
3.3.2	Cramér-Rao bounds and Fisher information	31
3.3.3	Maximum likelihood estimation	32
3.3.4	The effect of misalignment on SLIVER	33
3.3.5	Pixelated SLIVER	35
3.3.6	Applications	36
3.3.7	Previous experimental progress with image inversion interferometry	37
4	Fabrication	40
4.1	Electron beam evaporation	41
4.2	Electron beam lithography	42
4.3	Inductively coupled plasma etching	43
4.4	Scanning electron microscopy	45
4.5	Discussion of fabrication	49
5	The quest for image inversion interferometry	53
5.1	The image inversion interferometer	53
5.2	Interference with coherent light	54
5.3	Interference with incoherent light	56
5.4	Image processing	58
6	Conclusions and future aims	62
	Bibliography	64
	Appendix	68

List of Figures

1.1	Demonstrating what is meant by resolving point sources	2
1.2	Images of a lysosomal membrane	2
1.3	A binary star system imaged with WISE and GMOS	4
2.1	Intensity profiles of Hermite-Gaussian modes	17
2.2	Gaussian approximations of point spread functions	19
2.3	Different measures of resolution	20
2.4	The integrand of the $\Gamma_0(d)$ function	22
2.5	Average error probability for two binary hypothesis tests of resolution for incoherent point sources	24
3.1	Fisher information versus separation of sources for binary SPADE, SPLICE and IPC	26
3.2	Michelson and Mach-Zehnder interferometers	28
3.3	An image inversion interferometer	29
3.4	Fisher information for average $\epsilon = 10^{-3}$	32
3.5	MSE for average $\epsilon = 5$ and bucket detection at the antisymmetric port	33
3.6	Fisher information for average $\epsilon = 0.5$	34
3.7	MSE for average $\epsilon = 0.2$ and on-off detection at the antisymmetric port	34
3.8	MSE for average $\epsilon = 0.2$ and on-off detection at the antisymmetric port	35
3.9	Fisher information versus separation for pixelated SLIVER	36
3.10	Interferometric resolution improvement	38
4.1	Schematic of optical device design showing pairs of slits and grating	40
4.2	Overview of fabrication process	42
4.3	Opacity of different coating options	42
4.4	Inductively coupled plasma etching set-up	43
4.5	Secondary electron image of slits separated by $\approx 3\mu\text{m}$	46
4.6	Backscattered electron image of slits separated by $\approx 3\mu\text{m}$	46
4.7	Intensity profiles from SEM images of pairs of slits of the same separation but different samples.	47
4.8	Intensity profiles from SEM images of pairs of slits on sample B of different separations.	48
4.9	Intensity profiles from a SEM image of a pair of slits on the same sample B.	48

4.10	Measurements of slit width and separation from device B	49
4.11	Backscattered electron image of slits with separation $\approx 2\mu\text{m}$	49
4.12	Backscattered electron image of slits separated by $\approx 5\mu\text{m}$	50
4.13	Backscattered electron image of grating	51
4.14	Part of an intensity cross-section of the grating.	52
5.1	Experimental set-up for image inversion interferometry	54
5.2	Interference fringes obtained using a coherent source and the piezo- electric mirror scanning frequency set to 1Hz	56
5.3	Sodium lamp fringes	57
5.4	Interference fringes obtained using a sodium lamp source	58
5.5	LED fringes	59
5.6	Interference fringes obtained using an LED and the piezoelectric mirror scanning frequency set to 100mHz	59
5.7	Interference fringes obtained using an LED and the piezoelectric mirror scanning frequency set to 200mHz	60
5.8	Interference fringes obtained using an LED and the piezoelectric mirror scanning frequency set to 1Hz	60
5.9	Preliminary results showing enhancement of fringe visibility through processing image data in <i>Mathematica</i>	61
5.10	Interference fringes obtained using an LED (select area of CCD)	61
5.11	Post-processing enhanced interference fringes obtained using an LED	61

Table of symbols used

Symbol	Meaning	Symbol	Meaning
x	a data set	ψ	opening angle of largest light cone
θ	a parameter	D	diameter of iris
$\hat{\theta}$	an estimator	Θ	limit of resolution (angle)
$p(x \theta)$	probability density function	L	total number of photons counted
$L(\theta x)$	likelihood	\mathfrak{V}	visibility
\mathbb{E}	expectation	Ω	ratio of residual power output to total power
mse	mean squared error	I_{\max}	maximum fringe intensity
var	variance	I_{\min}	minimum fringe intensity
b	bias	ψ	distribution of light
l	log-likelihood	d	separation of sources
X	a random variable	A_k	amplitude
\mathcal{I}	Fisher information	E_{in}	input electric field
\mathcal{J}	quantum Fisher information	E_{out}	output electric field
cov	covariance	\bar{N}_k	average photocount at a port
r	an arbitrary estimator	ϵ	average photocount without interference
$U(t)$	amplitude of a light field	$\delta(d)$	interference photocount
$\Gamma(t)$	time autocorrelation function	$\zeta(d)$	$\delta'(d)$
$\mathcal{G}(\nu)$	power spectral density	M	number of measurements
$\gamma(\tau)$	temporal coherence function		
τ_c	coherence time		
l_c	coherence length		
$S(\nu)$	spectral density		
I	intensity		
$\Delta\nu_c$	spectral linewidth		
$g(r_1, r_2, \tau)$	mutual coherence function		
β	Poisson process arrival rate		
ω	angular frequency		
\hbar	reduced Planck constant		
$j(t)$	photocurrent		
r	limit of resolution (distance)		
λ	wavelength		
NA	numerical aperture		
n	refractive index		

Table of acronyms used

Acronym	Meaning
GFP	green fluorescent protein
UV	ultraviolet
CRLB	Cramér-Rao lower bound
WISE	Wide-field Infrared Survey Explorer
GMOS	Gemini Multi-Object Spectrograph
SLIVER	superlocalisation via image inversion interferometry
PDF	probability density function
MSE	mean squared error
MVU	minimum variance unbiased
MLE	maximum likelihood estimator
qCRLB	quantum Cramér-Rao lower bound
POVM	positive operator valued measure
LUE	locally unbiased estimator
SLD	symmetrised logarithmic derivative
TEM	transverse electromagnetic mode
NA	numerical aperture
FREM	fundamental resolution measure
LED	light-emitting diode
SPADE	spatial mode demultiplexing
SPLICE	superresolved position localisation by inversion of coherence along an edge
IPC	image plane photocounting
RIE	reactive ion etching
ICP	inductively coupled plasma
EBL	electron beam lithography
SEM	scanning electron microscopy
SE	secondary electron
BSE	backscattered electron
PBS	polarising beam splitter
CCD	charged coupled device

Chapter 1

Introduction

*You want a revelation / You want to get it right /
But it's a conversation / I just can't have tonight /
You want a revelation / Some kind of resolution /
Tell me what you want me to say.*

No Light, No Light, Florence + The Machine

Can you see this? Can you see that? The individual leaves on a tree? The road name on that sign? The tail lights of a car speeding off into the night? Two tail lights? Are you sure – don't they now appear as just one bright red blur?

What it means to *see* something has been pondered for centuries. We can see stars light years away with the naked eye, but we are unable to recognise a friend until they are much much closer than that, only a hundred metres away. Quantifying these differences in vision was the aim of scientists Ernst Abbe, and later, Lord Rayleigh. Humans are quite good at spotting single points (acuity) but there is a limit to our ability to distinguish things that are closer together (resolution).[1] Figure 1.1 demonstrates the difference between clearly resolved, just resolved, and unresolved points. Our vision proved to be the motivation for Abbe to develop a relationship that would quantify this limit – the diffraction limit – for optical instruments such as microscopes. For the visible light that we are all so familiar with, this limit is about 200 nm, the size of a big virus or quite a small bacteria. Practically, this means that when a specimen is viewed through a conventional microscope, which uses optical wavelengths to illuminate a sample, you might be able to see that there was *something* there, but you would not be able to make out the features that identify it as a certain type of virus or any other fine structure.

The diffraction limit is known to be dependent on wavelength. Electrons with energies of only a few thousand electron volts have a much shorter wavelength, and so can reveal greater detail. This is exploited in scanning electron microscopes. However, electron microscopy requires samples to be conducting to allow the charge to escape. This means biological samples have to be coated in a conductive material such as carbon or gold. This may obscure fine details, or worse, destroy a

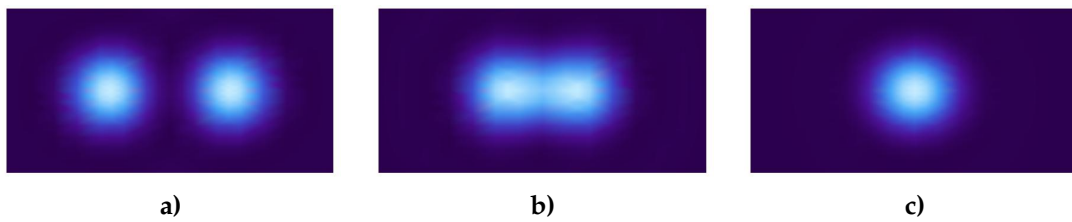


Figure 1.1: As two point sources of light are brought closer and closer together, it becomes increasingly difficult to resolve them individually: **a)** shows two well resolved points, **b)** we are just able to resolve the points by eye, and **c)** the overlap of the profiles of the two sources is too great to resolve them.

delicate sample. The limit to the resolving power of electron microscopes is around two nanometres, depending on the microscope.[2] The challenge is to develop a method of examining the smallest of things without damaging them.

The research groups of Eric Betzig, Stefan Hell, and W.E. Moerner independently answered this challenge, shattering the belief that the diffraction limit could not be overcome. The two techniques in **superresolved fluorescence microscopy** – or more properly, nanoscopy – developed by these scientists saw them awarded the 2014 Nobel Prize in Chemistry.[3] Creatively illuminated biological samples enabled resolutions far better than 200 nm to be reached.[3]

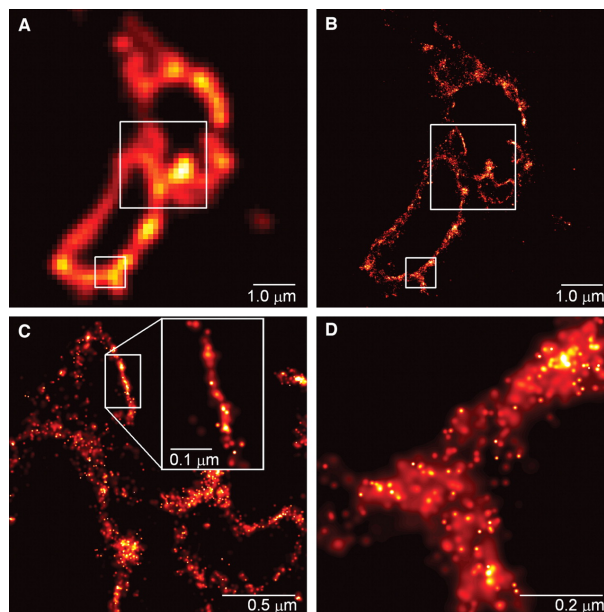


Figure 1.2: Images of a lysosomal membrane. This is from one of Betzig's key publications that saw him share in the 2014 Nobel Prize in Chemistry. **A** shows an image of the membrane with conventional optical microscopy techniques. In **B**, **C** and **D** the power of fluorescence microscopy is exhibited in stunning fashion. [4]

Figure 1.2 shows the power of Eric Betzig's technique in imaging a lysosomal membrane. This particular method worked by connecting fluorescent proteins,

such as the ubiquitous green fluorescent protein (GFP), to samples of bacteria.[4] These proteins then became the light sources for the image. It is hard to determine the exact intensity profile of a light source; it is much easier to pinpoint the centre of a point source of light when there are not any other sources nearby. By shining an ultraviolet (UV) light onto the bacteria, some of these GFPs will emit green light, but only some of them, and it is random which ones do emit the light each time they are exposed to UV. The optical microscope images the positions of these proteins, indicated by the brightest spots of green light. The UV light is turned off, the fluorescence dies, the light is turned on again, and a *different* ensemble of green spots is activated. By repeating this process and layering all the images, a 'join-the-dots' picture of the bacteria is produced at much higher resolution than a single optical photo of the naked bacteria. This is **single-molecule microscopy**, pioneered by Betzig and Moerner.[4] Stefan Hell's experiment in **stimulated emission depletion microscopy**[5] also relied extensively upon knowing where the fluorescence from the sample was originating.

In this way, single-molecule microscopy and stimulated emission depletion microscopy are both techniques that achieve super-resolution but do have constraints. For the former, it is that every time a 'layer' image is taken, the active proteins still have to be more than 200 nm apart in order to be resolved. For the latter, it is that a laser can be focused to a spot size of diameter equal to about half the laser's wavelength, and no smaller: a spot size of about 200nm. The light from active GFPs is also incoherent, meaning that there are many fluctuations in the amplitude and phase of the light, so it is stochastic. Multiple points fluorescing at the same time complicates matters: photon counting statistics can be used to localise these points, but the error blows up the closer the point sources become. This is known as **Rayleigh's curse**, a term coined by Mankei Tsang's Quantum Measurement Group at the National University of Singapore.[6]

The theoretical and simulation work of Tsang and company in the area of quantum super-resolution provides the impetus for this current project. They have studied the quantum **Cramér-Rao lower bound** (CRLB) for the case of two incoherent point sources.[6] This is a bound on the mean square error of an estimator.¹ The mean square error is inversely proportional to the **Fisher information** of an estimator. A small variance indicates a good estimator and that is why Fisher information is exactly that: information. Here the data being analysed are the detected photons from two incoherent point sources and the parameter of interest is the distance separating the point sources. Tsang *et al.* found the CRLB unwavering even as the point sources are brought closer and closer, meaning that an estimate of the separation can always be produced within a finite error.[6] In short, experiments which maximise the Fisher information enable the locations of the point sources of light to be estimated with accuracy far greater than the diffraction limit would suggest.

Another important example of point sources of light being too close together to

¹An estimator is simply a function for generating an estimate of a parameter from a data set.

resolve² is found at the other end of the length scale to microscopy: astronomy. Take the case of two stars, or a star and an exoplanet. When observed through a telescope, these objects may appear as one faint spot. The astronomer asks “Is that one bright star or two less bright stars close together? How far apart are the stars? How can we estimate the separation between the stars if we do not even know if there are stars plural?” Such questions motivated the development of theories in statistical optics: photodetection in the case of dim sources, **binary hypothesis testing** (how do we decide whether there is one source of light or two, and with what certainty?), and the resolution of incoherent point sources.

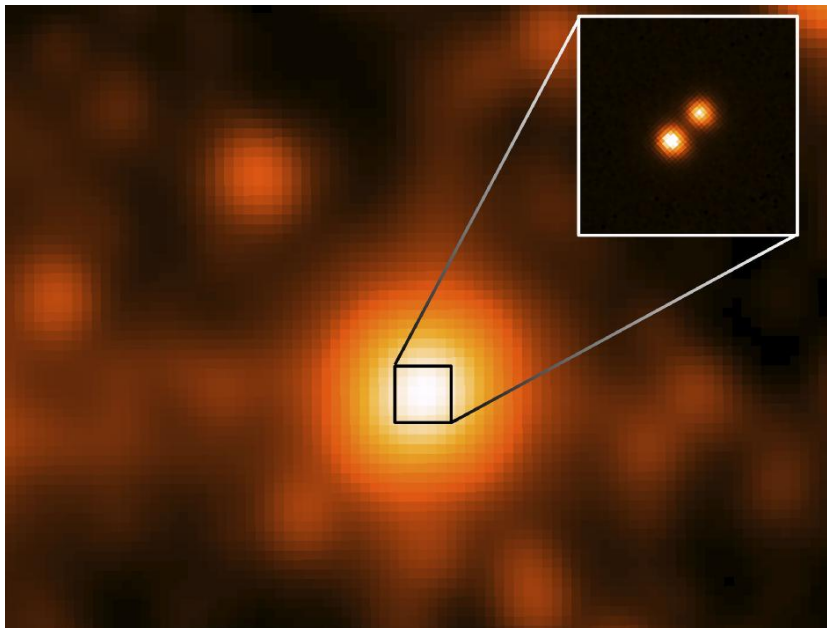


Figure 1.3: *Luhman 16, a binary star system consisting of two brown dwarf stars. The large image is from NASA's WISE. The inset is from GMOS. This image is at a much higher resolution to the main picture, with the two stars clearly distinguishable.[7]*

A further idea in astronomy is that of combining the data from two or more telescopes that are a distance apart in a way that gives equivalent resolving power to a telescope the size of the longest distance between any of the individual telescopes. Such techniques in interferometry have been employed at radio, and more recently, optical wavelengths, demonstrating the power of interferometry to yield additional information. Figure 1.3 shows a binary star system imaged through a single telescope aperture, the Wide-field Infrared Survey Explorer (WISE) space telescope, and the same system pictured through the Gemini Multi-Object Spectrograph (GMOS). GMOS combines data from the twin Gemini telescopes: a much higher resolution image is the result, allowing the two stars to be seen. Future imaging techniques in astronomy could make use of a nulling interferometer so that much of the brighter star's light can be cancelled out or “nulled”.[8] The construction of a nulling interferometer is very similar to a Mach-Zehnder interferometer, but

²From where we are, anyway.

with a π phase shift in one arm. This would involve very similar statistics to those outlined by Tsang et al. with the possibility of uncovering numerous exoplanets.

An experimental proposal that builds on these ideas by combining the delicacy of interferometry with the ambit of mathematical statistics is **superlocalisation via image inversion interferometry** (SLIVER).[6] Image inversion interferometry is along a similar line to nulling interferometry: it destructively interferes any symmetric components of the fields from the sources but preserves light travelling in the off-axis direction (asymmetric), which then appears as interference fringes. Analysing these fringes enables localisation of the point sources. In this work I will address the construction of SLIVER to enable the super-resolution of points of incoherent green light. I also present work resulting in the fabrication of small optical devices, used to create point sources with separations on the order of microns.

From here, the structure of the thesis is as follows:

- Chapter 2 introduces a range of background theory, spread across several different areas, namely mathematical statistics, statistical optics, and quantum information theory.
- Chapter 3 examines interferometry and the theory behind the experimental proposal in greater detail, while also reviewing progress in image inversion interferometry and other super-resolution techniques driven by estimation theory.
- Chapter 4 details the fabrication of the small optical devices utilised in the experiment.
- Chapter 5 reports progress towards realising SLIVER.
- Chapter 6 describes future optical adventures made possible through the SLIVER technique. Conclusions from the image inversion interferometer experiment and the study of the fabrication process are presented here.

The diffraction limit is a state of mind.

- W.E. Moerner [9]

Chapter 2

Background theory

This chapter covers elementary material from several areas of statistics and optics. It aims to give a thorough background to which later chapters can refer, whilst also ensuring the reader feels equipped for tackling a problem that lies at the intersection of optics, statistics and quantum information theory: superresolution. We start with estimation theory in 2.1 before moving to quantum applications of estimation theory in 2.2. An initial dose of statistical optics is provided in 2.3. 2.4 deals with photodetection, the measurement process so crucial to the experimental work conducted for this project. In 2.5 different resolution limits are examined and the specific problem of incoherent point sources is introduced.

2.1 Estimation Theory

Estimation theory, a branch of mathematical statistics, is applied to many disciplines that depend on signal processing. Its primary concern is extracting information in the presence of random noise.[10] As the legendary Carl Helstrom puts it,

Information about the world is acquired by observation and measurement, the results of which are subject to error. One would like to think it could be eliminated if only one built elaborate enough instruments and took sufficient pains. All efforts to be free of error will, however, finally reach a bound set by nature's underlying chaos ... Acquiring information about a physical entity or system involves either decision or estimation ... hypotheses and parameters exist and signify in the context of some theory about the system. That theory also describes the sources of the uncertainty or error that corrupts the observations. The irreducibly minimum component of error in decisions and estimates is discovered by analysing the decision or estimation procedures that minimize some convenient measure of the average amount of error.[11]

A **probability density function** (PDF) may be described in terms of one or more

parameters. At the heart of estimation theory is the problem of inferring an unknown parameter θ in terms of the data $x = x[0], x[1], \dots, x[N - 1]$ of a random variable X .¹ An **estimator** $\hat{\theta}$ is an inversion that returns θ and is performed by a function g :²

$$\hat{\theta} = g(x[0], x[1], \dots, x[N - 1]) \quad (2.1)$$

The inversion is not necessarily unique, accurate or precise; hence, estimation theory also seeks to calculate uncertainties of estimates.

This introduction to estimation theory is supported by many course materials from different institutions available online ([12, 13, 14]). In particular, I would like to acknowledge Associate Professor Songfeng Zheng of Missouri State for generously sharing his excellent notes ([14]), which inspired the approach taken in the following subsections.

2.1.1 Minimum Variance Unbiased Estimation

From the familiar standpoint of conditional probability,³ $p(x|\theta)$ is read as ‘the probability of the data taking the values x given the parameter has a value θ ’. In estimation theory, we turn this around and read it as ‘the likelihood of the parameter having the value θ given the data x ’. $L(\theta|x_i)$ is the **likelihood function**, which is a family of probability distributions over x_i . Each member of the family is a function of a specific parameter θ with the data held constant $x_i = x$, so the numerical value of $L(\theta_0|x) = p(x|\theta_0)$, which goes some way towards explaining why the words “probability” and “likelihood” are interchangeable colloquially. The integral of the likelihood over all possible θ does not necessarily equal one, whereas the definition of a PDF is a nonnegative function that when integrated across all x equals one.

An **unbiased** estimator is one which satisfies $\mathbb{E}[\hat{\theta}] = \theta$, where the expectation (and all others in this section unless specified otherwise) would be performed with respect to the true distribution $p(x|\theta_0)$. As explained above, this is interchangeable with likelihood function of θ_0 : $L(\theta_0|x)$. In more physical terms, an unbiased estimator is an accurate measurement.[12, 13, 14]

The **mean squared error** (MSE) defines the optimality of an estimator

$$\text{mse}[\hat{\theta}] = \mathbb{E}[(\hat{\theta} - \theta)^2] = \text{var}(\hat{\theta}) + b(\theta)^2 \quad (2.2)$$

where the **bias** $b(\theta) = \mathbb{E}[\hat{\theta}] - \theta$ measures the precision of the measurement. Theoretically, an accurate and precise measurement would be represented by an estimator with zero bias and minimal MSE. This would be the **minimum variance unbiased** (MVU) estimator.[12, 13, 14]

¹Here we use uppercase for variables and lowercase for the data corresponding to that variable.

² θ_0 is the true value of the constant parameter θ . $\hat{\theta}$ is a function.

³There are several warring tribes of statisticians and not all of them agree with what follows.

2.1.2 Fisher information

How much information can x provide about θ ?

Intuitively, if something is unlikely to happen, a specific set of conditions must have been in place for it to happen: unlikely events contain more information than likely ones. For a random variable X , which is described by a PDF dependent on a parameter θ with true value θ_0 , the likelihood function would be a maximum at $\theta = \theta_0$, or equivalently, the derivative of the log-likelihood function would be zero. The logarithm is monotonically increasing; therefore, the maximum of the log-likelihood occurs at the same point as the maximum of the likelihood. This is the basic principle of **maximum likelihood estimation**. [12, 13, 14]

The **log-likelihood function** is defined as

$$l(\theta|x) = \log L(\theta|x) \quad (2.3)$$

$$l'(\theta|x) = \frac{\partial}{\partial \theta} (\log L(\theta|x)) = \frac{L'(\theta|x)}{L(\theta|x)}. \quad (2.4)$$

Choosing l over L often makes for simpler calculations.

A flat likelihood function says that any value of θ is equally likely for the outcome x . We interpret a steep gradient in the likelihood function to mean that one value is much more likely than the others; hence, if $[l'(\theta|X)]^2$ is large, the random variable provides much information about θ . $[l'(\theta|x)]^2$ is a measure of the amount of information provided by X ; since X is a random variable, we find the average of $[l'(\theta|x)]^2$. This is the **Fisher information** for θ contained in the random variable X : [12, 13, 14]

$$\mathcal{I}(\theta) = \mathbb{E} \left[[l'(\theta|x)]^2 \right] = \int [l'(\theta|x)]^2 L(\theta|x) dx. \quad (2.5)$$

Although we may have replaced the PDF with the equivalent function $L(\theta|x)$ (we do not know the true value θ_0) to make the next few derivations simpler, we have still integrated over x . The Fisher information is the maximal statistical information regarding the parameter θ which can be extracted from the data set. We now find another two equivalent definitions of the Fisher information. To do this, we need the following results:

$$\begin{aligned} \int L'(\theta|x) dx &= \frac{\partial}{\partial \theta} \int L(\theta|x) dx = 0 \\ \int L''(\theta|x) dx &= \frac{\partial^2}{\partial \theta^2} \int L(\theta|x) dx = 0 \\ \therefore \mathbb{E} [l'(\theta|x)] &= \int l'(\theta|x) L(\theta|x) dx \\ &= \int \frac{L'(\theta|x)}{L(\theta|x)} L(\theta|x) dx \\ &= 0 \end{aligned} \quad (2.6)$$

We combine (2.5) with (2.6) to obtain a second definition of the Fisher information:[14]

$$\mathcal{I}(\theta) = \mathbb{E} \left[[l'(\theta|x)]^2 \right] - (\mathbb{E} [l'(\theta|x)])^2 = \text{var}[l'(\theta|x)]. \quad (2.7)$$

Some tactful rearranging of the second derivative of the log-likelihood provides a third definition of the Fisher information:[14]

$$\begin{aligned} l''(\theta|x) &= \frac{\partial}{\partial \theta} \left(\frac{L'(\theta|x)}{L(\theta|x)} \right) \\ &= \frac{L''(\theta|x)L(\theta|x) - (L'(\theta|x))^2}{(L(\theta|x))^2} \\ &= \frac{L''(\theta|x)}{L(\theta|x)} - (l'(\theta|x))^2 \end{aligned} \quad (2.8)$$

$$\begin{aligned} \therefore \mathbb{E} [l''(\theta|x)] &= \int \left(\frac{L''(\theta|x)}{L(\theta|x)} - (l'(\theta|x))^2 \right) L(\theta|x) dx \\ &= \int L''(\theta|x) dx - \mathbb{E} [l'(\theta|x)]^2 \\ &= -\mathcal{I}(\theta). \end{aligned}$$

The Fisher information for a sample of size n , X_1, X_2, \dots, X_n , can be determined from the joint PDF, which is

$$p_n(\tilde{x}|\theta) = \prod_{i=1}^n p(x_i|\theta). \quad (2.9)$$

A similar definition exists for the joint likelihood $L(\theta|\tilde{x})$, which leads to the joint log-likelihood:

$$\begin{aligned} l_n(\theta|\tilde{x}) &= \log L_n(\theta|\tilde{x}) \\ &= \sum_{i=1}^n \log L(\theta|x_i) \\ &= \sum_{i=1}^n l(\theta|x_i). \end{aligned} \quad (2.10)$$

It can be shown easily that

$$l'_n(\theta|\tilde{x}) = \frac{L'_n(\theta|\tilde{x})}{L_n(\theta|\tilde{x})} \quad (2.11)$$

The Fisher information is then

$$\mathcal{I}_n(\theta) = \mathbb{E} [l'_n(\theta|\tilde{X})^2] = \int \dots \int [l'_n(\theta|\tilde{x})]^2 L_n(\theta|\tilde{x}) dx_1 \dots dx_n \quad (2.12)$$

Adopting a similar strategy to that used to find the Fisher information for one observation (see (2.6)), we have

$$\mathcal{I}_n(\theta) = \text{var} [l'_n(\theta|\tilde{X})] \quad (2.13)$$

$$\mathcal{I}_n(\theta) = -\mathbb{E} [l_n''(\theta|\tilde{X})] \quad (2.14)$$

From the definition of $l_n(\theta|\tilde{x})$ it follows that

$$\begin{aligned} l_n''(\theta|\tilde{x}) &= \sum_{i=1}^n l''(\theta|x_i) \\ \therefore \mathcal{I}_n(\theta) &= -\mathbb{E} \left[\sum_{i=1}^n l''(\theta|X_i) \right] \\ &= -\sum_{i=1}^n \mathbb{E} [l''(\theta|X_i)] \\ &= n \mathcal{I}(\theta) \end{aligned} \quad (2.15)$$

The Fisher information in a random sample of size n is simply n times the Fisher information in a single observation.[12, 13, 14]

2.1.3 Cramér-Rao Lower Bound

How do we use Fisher information to determine the lower bound for the variance of an estimator of the parameter θ ?

Let $\hat{\theta} = g(X_1, \dots, X_n) = g(\tilde{X})$ be an arbitrary estimator of θ , with $\mathbb{E}[\hat{\theta}] = m(\theta)$ and the assumption that the variance of $\hat{\theta}$ is finite. Let us consider the random variable $l_n'(\theta|\tilde{X})$ and recall $\mathbb{E} [l_n'(\theta|\tilde{X})] = 0$ (regularity condition). Therefore the covariance between $\hat{\theta}$ and $l_n'(\theta|\tilde{X})$ is [14]

$$\begin{aligned} \text{cov}[\hat{\theta}, l_n'(\theta|\tilde{X})] &= \mathbb{E} [(\hat{\theta} - \mathbb{E}[\hat{\theta}]) (l_n'(\theta|\tilde{X}) - \mathbb{E} [l_n'(\theta|\tilde{X})])] \\ &= \mathbb{E} [(\hat{\theta} - \mathbb{E}[\hat{\theta}]) l_n'(\theta|\tilde{X})] \\ &= \mathbb{E} [(g(\tilde{X}) - m(\theta)) l_n'(\theta|\tilde{X})] \\ &= \mathbb{E} [g(\tilde{X}) l_n'(\theta|\tilde{X})] - m(\theta) \mathbb{E} [l_n'(\theta|\tilde{X})] \\ &= \mathbb{E} [g(\tilde{X}) l_n'(\theta|\tilde{X})] \\ &= \int \cdots \int g(\tilde{x}) l_n'(\theta|\tilde{x}) L_n(\theta|\tilde{x}) dx_1 \cdots dx_n \\ &= \int \cdots \int g(\tilde{x}) L_n'(\theta|\tilde{x}) dx_1 \cdots dx_n \\ &= \frac{\partial}{\partial \theta} \int \cdots \int g(\tilde{x}) L_n(\theta|\tilde{x}) dx_1 \cdots dx_n \\ &= \frac{\partial}{\partial \theta} (\mathbb{E}[\hat{\theta}]) \\ &= m'(\theta) \end{aligned} \quad (2.16)$$

By the Cauchy-Schwarz inequality and the definition of $\mathcal{I}_n(\theta)$:[12, 13, 14]

$$(\text{cov}[\hat{\theta}, l_n'(\theta|\tilde{X})])^2 \leq \text{var}[\hat{\theta}] \text{var} [l_n'(\theta|\tilde{X})] = \text{var}[\hat{\theta}] \mathcal{I}_n(\theta) \quad (2.17)$$

that is

$$(m'(\theta))^2 \leq \text{var}[\hat{\theta}] \mathcal{I}_n(\theta) = n \mathcal{I}(\theta) \text{var}[\hat{\theta}] \quad (2.18)$$

Finally we get the lower bound of variance of an arbitrary estimator $\hat{\theta}$ as [14]

$$\text{var}[\hat{\theta}] \geq \frac{(m'(\theta))^2}{n\mathcal{I}(\theta)} \quad (2.19)$$

This inequality is called the **information inequality**, also known as the **Cramér-Rao inequality** (Cramér and Rao both developed it independently). The information inequality shows that as $\mathcal{I}(\theta)$ increases, the variance of the estimator decreases; therefore, the quality of the estimator increases. That's why $\mathcal{I}(\theta)$ is called information![14]

If $\hat{\theta}$ is an unbiased estimator, then $m(\theta) = \mathbb{E}[\hat{\theta}] = \theta$, therefore $m'(\theta) = 1$. So for an unbiased estimator:

$$\text{var}[\hat{\theta}] \geq \frac{1}{n\mathcal{I}(\theta)} \quad (2.20)$$

This is the **Cramér-Rao lower bound** (CRLB): under certain conditions, no other unbiased estimator of the parameter θ based on a sample of size n can have a variance smaller than the CRLB.[12, 13, 14]

2.1.4 Maximum likelihood estimation

There might not be a MVU estimator. It is still possible to find an asymptotic approximation for which the error decreases as the number of data sets increases. The **maximum likelihood estimator** (MLE) is defined as the value of θ as a function of the data x which maximises the likelihood function $L(\theta|x)$.[12, 13, 14]

If we assume that the true value of θ is θ_0 and the MLE of θ is $\hat{\theta}$, then the probability distribution of $\sqrt{n\mathcal{I}(\theta_0)}(\hat{\theta} - \theta_0)$ tends towards a standard normal distribution. This is a key result: $\sqrt{n\mathcal{I}(\theta_0)}(\hat{\theta} - \theta_0) \sim N(0, 1)$ asymptotically.[14]

This theorem indicates the asymptotic optimality of the MLE since the asymptotic variance of MLE can achieve the CRLB. For this reason, the MLE is frequently used with large samples. The beauty of MLE is that it will return the MVU estimator if it exists.

2.2 Quantum Fisher information

In this section, we focus on one of the tenets of quantum information theory, **quantum Fisher information**. To introduce quantum Fisher information comprehensively requires a thorough background in quantum mechanics. In the interest of maintaining enthusiasm for the subject, this section will be kept light on rigour and short on space. The reader is encouraged to immerse themselves in *Quantum Detection and Estimation Theory* by Carl Helstrom[11] if they find their appetite for quantum information theory suitably piqued. We adopt Helstrom's approach to introducing quantum Fisher information, and subsequently the **quantum Cramér-Rao lower bound** (qCRLB).

In the classical case, a PDF $p(x)$ can be utilised to obtain moments and the expectation value of any observable f .

$$\langle f \rangle = \int f(x)p(x)dx \quad (2.21)$$

In quantum mechanics, a PDF doesn't quite cut it. Instead, we use the **density operator**, ρ . Physical systems may be described as statistical mixtures of many wavefunctions, each essentially acting like a PDF. A density operator encapsulates such ensembles of states which have classical uncertainty as well as the probabilistic weirdness of quantum mechanics. Properties of ρ , and details of its calculation and use are given in [15]. Here we aim at the quantum Cramér-Rao lower bound, and so we need ρ to replace the PDF. Any classical regime is simply a special case of this generalised quantum approach. The expectation value of an observable \hat{F} in quantum mechanics is

$$\langle \hat{F} \rangle = \text{Tr} [\hat{F}\rho]. \quad (2.22)$$

Measurement in quantum mechanics is a can of worms best opened elsewhere. Here we will stick to an intuitive explanation of what is meant by a **positive operator valued measure** (POVM), or a “quantum measurement”. Most generally, a POVM is a map that associates a positive operator with every subset of measurement outcomes.[16] Here, we take \mathcal{M} to be a POVM that when acting on a quantum system generates measurement outcomes, which are then the input for an estimator $\hat{\theta}$ of parameters θ of $\rho(\theta)$.⁴ [11, 17] The probability of the measurement outcome $\mathcal{M}(x)$ being observed given the state $\rho(\theta)$ is [18]

$$p(\mathcal{M}(x)|\theta) = \text{Tr} [\rho(\theta)\mathcal{M}(x)]. \quad (2.23)$$

From this PDF we can construct a function that gives the likelihood of the parameters having value θ given outcome $\mathcal{M}(x)$. Let us define some statistics using our density operator and POVM. Firstly, the expectation value:

$$\mathbb{E}[\hat{\theta}_j|\mathcal{M}(x)] = \int_X \hat{\theta}_j \text{Tr} [\rho(\theta)\mathcal{M}(x)]. \quad (2.24)$$

When $\mathbb{E}[\hat{\theta}_j|\mathcal{M}(x)] = \theta_j$, the estimator $\hat{\theta}_j$ is unbiased. We will only contend with unbiased estimators here. In the quantum case, it is also useful to introduce the concept of a **locally unbiased estimator** (LUE). This is when at a certain point θ_0 , $\mathbb{E}[\hat{\theta}_0|\mathcal{M}(x)] = \theta_0$ up to the first order of the Taylor expansion about θ_0 . [17]

Secondly, the covariance matrix \mathcal{B} :

$$\mathcal{B}_{ij} = \mathbb{E} [(\hat{\theta}_i - \bar{\theta}_i)(\hat{\theta}_j - \bar{\theta}_j)|\mathcal{M}(x)] = \int_X (\hat{\theta}_i - \bar{\theta}_i)(\hat{\theta}_j - \bar{\theta}_j)\text{Tr}[\rho(\theta)\mathcal{M}(x)]. \quad (2.25)$$

To obtain the classical Fisher information, we differentiated the log-likelihood function. How do we differentiate the logarithm of ρ ? The derivative of ρ can be

⁴ $\mathcal{M} = \{\mathcal{M}(x); x \in X\}$ and $\hat{\theta} : X \rightarrow \Theta$ [17]

defined in several different ways, and so there are several answers to this question. Let us adopt operators $\mathcal{L}_{\theta,k}$ of $\rho(\theta)$ for this purpose, defined by:

$$\frac{\partial \rho}{\partial \theta_k} = \frac{1}{2} (\mathcal{L}_{\theta,k} \rho(\theta) + \rho(\theta) \mathcal{L}_{\theta,k}). \quad (2.26)$$

These $\mathcal{L}_{\theta,k}$ are the **symmetrised logarithmic derivatives** (SLDs). The **quantum Fisher information matrix** \mathcal{J} has elements

$$\mathcal{J}_{ij} = \frac{1}{2} \text{Tr} [\rho(\theta) (\mathcal{L}_{\theta,i} \mathcal{L}_{\theta,j} + \mathcal{L}_{\theta,j} \mathcal{L}_{\theta,i})] = \text{Tr} \left(\frac{\partial \rho}{\partial \theta_i} \mathcal{L}_j \right). \quad (2.27)$$

In the case where we are after an unbiased estimate of one parameter, the covariance matrix becomes a variance matrix, and (2.27) becomes

$$\mathcal{J} = \text{Tr}(\rho(\theta) \mathcal{L}^2) = \text{Tr} \left(\frac{\partial \rho}{\partial \theta} \mathcal{L} \right). \quad (2.28)$$

Analogous to the classical case, the quantum Cramér-Rao Lower Bound is

$$\mathcal{B} \geq (\mathcal{J})^{-1}. \quad (2.29)$$

In the case of single parameter estimation, the qCRLB as defined by the SLD can always be achieved.[19] Now for multiple parameter estimation, it should be pointed out that not often is there a LUE which satisfies the equality of the qCRLB. This is because the SLDs may not commute. Indeed the ‘best’ LUE intrinsically depends on the value of unknown parameter.[17] Earlier we presented the asymptotic optimality of the MLE for the classical CRLB, and it turns out that MLE can be used in the quantum context to deliver the best LUE. This is through an adaptive quantum estimation scheme: a small sample \sqrt{n} is approximately measured to find a rough estimate of θ , then local estimation is done using $n - \sqrt{n}$ samples with an optimised measurement based on the rough estimate of θ . [17] The estimator used is the MLE, and the iteration in [17] is described by

$$L_n(\theta | \mathcal{M}(x)) = \prod_{i=1}^n p(\mathcal{M}(x_i) | \theta, \mathcal{M}(x_n, \hat{\theta}_{i-1})) \quad (2.30)$$

where $L_n(\theta | \mathcal{M}(x))$ is the likelihood function to be maximised.

The amazing thing about the qCRLB is that it does not depend directly on the measurement, unlike the classical CRLB. It tells us the best we can do given the state $\rho(\theta)$. If we find a POVM saturating \mathcal{B} , we know we cannot do better than that.

2.3 Statistical optics

A statistical approach to optics, rather than a deterministic one, is required when the physical situation encountered has one or more aspects of randomness to it.

Randomness is an emergent property of quantum mechanics, which describes light and matter at the smallest scales. It follows that to most accurately describe optics we need statistical analysis. Culprits of randomness in the case of light could include a thermal source (e.g. an incandescent lamp) or irregularities in the medium in which the light propagates (e.g. random changes in refractive index along an optical path). There are no well-defined physical transformations providing an analytic solution for light entering a photodetector where there is shot noise, and none to account for the fluctuations in phase and amplitude of a real laser beam. Here a brief summary of some of the most important concepts in statistical optics is given. For further details, the reader is encouraged to peruse [20] or a similar tome.

2.3.1 Coherence

In statistical optics, we require that the amplitude of the light field $U(t)$ is a random variable. Simply put, we cannot predict the amplitude of the light field at a given position and time in advance, but $U(t)$ encompasses both the amplitude values and their probabilities and so is a complete statistical model.

The time **autocorrelation function** $\Gamma(\tau)$ associated with $U(t)$ measures the amount of random fluctuations relative to the propagation of the field.

$$\Gamma(\tau) = \langle U^*(t)U(t + \tau) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} U^*(t)U(t + \tau) dt \quad (2.31)$$

The autocorrelation function is also the Fourier transform of the **power spectral density**, $\mathcal{G}(\nu)$:

$$\Gamma(\tau) = \int_{-\infty}^{\infty} \mathcal{G}(\nu) \exp(-i2\pi\nu\tau) d\nu \quad (2.32)$$

The function $\gamma(\tau) = \Gamma(\tau)/\Gamma(0)$ is normalised, and is referred to as the **temporal coherence function**. Temporal coherence is the property of a beam to interfere with an identical beam that has been delayed in time, and as such, $|\gamma(\tau)|$ measures the correlation between $U(t)$ and $U(t + \tau)$. The **coherence time** is defined as the **power-equivalent width**, which is basically the width of $|\gamma(\tau)|$:

$$\tau_c = \int_{-\infty}^{\infty} |\gamma(\tau)|^2 d\tau \quad (2.33)$$

and could be scaled by the light intensity if not normalised. The **coherence length** is defined as $l_c = c\tau_c$. Both quantities are measures which determine whether different parts of the field are strongly correlated or not.

For **quasimonochromatic light** - light with a coherence length greater than the maximum optical path difference - the greater the spectral linewidth the greater degree of decoherence. For a given spectral density $S(\nu)$, one definition of the

spectral linewidth is

$$\Delta\nu_c = \left(\int_0^\infty S(\nu) d\nu \right)^2 \left(\int_0^\infty S(\nu)^2 d\nu \right)^{-2} \quad (2.34)$$

which was chosen to relate to the coherence time by simply $\Delta\nu_c = 1/\tau_c$.

For example, the linewidth of the laser used in the course of the experimental work is approximately $\Delta\nu_c = 1 \times 10^6$ Hz. This gives a coherence length of 300 m.

The other type of coherence is **spatial coherence**, the property of a beam to interfere with an identical beam that has been shifted in space without being delayed. Spatial coherence is a measure of the distortions in wavefronts. This spatial dependence could be incorporated by considering the **mutual coherence function**, which allows us to deal with shifts of the beam's twin in space and time relative to the beam:

$$g(r_1, r_2, \tau) = \frac{\langle U^*(r_1, t) U(r_2, t + \tau) \rangle}{\sqrt{I(r_1) I(r_2)}} \quad (2.35)$$

The autocorrelation for spatial coherence would be recovered when the delay τ is set to zero.

The **coherence area** can be defined in a manner similar to the coherence time, and gives an indication of the degree of coherence over the wavefront. For an incoherent source of radius a , the coherence area is

$$A_c = \frac{\lambda^2 z^2}{\pi a^2} \quad (2.36)$$

where z is the distance from the source.[20] Pinholes on the order of micrometers are required for effective spatial filtering.

The interference of two partially coherent light sources could be characterised by considering the correlation between the two fields U_1 and U_2 , as given by the mutual coherence function $g_{12} = g(r_1, r_2, 0)$. The mean intensity would be given by

$$\langle I \rangle = \langle |U_1 + U_2|^2 \rangle = I_1 + I_2 + 2\sqrt{I_1 I_2} \Re\{g_{12}\} \quad (2.37)$$

For perfectly correlated light, $g_{12} = \exp(i\phi)$ which gives the relative phase shift. The interference varies between constructive and destructive, depending on the phase angle ϕ . For fields with no correlation, g_{12} vanishes so that there is no interference, and the intensities add directly.

2.3.2 Transverse modes

The distribution of the electric field may be described by a **transverse mode**. Here light travels in the positive z direction. The profile of the electric field in the xy plane will satisfy the wave equation, and also any boundary conditions of the system. The wave equation is

$$\nabla^2 U - \frac{1}{c^2} \frac{\partial^2 U}{\partial t^2} = 0 \quad (2.38)$$

where U_0 is the complex amplitude of the wave. The wave equation admits well-known plane wave solutions, and in the case of a beam with a varying amplitude perpendicular to the z direction, solutions of the form

$$U(x, y, z) = \psi(x, y, z) \exp(-ikz)T(t) \quad (2.39)$$

can be found. In the case of a narrow beam of light, the paraxial approximation can be made and the wave equation reduces to [21]

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} - 2ik \frac{\partial \psi}{\partial z} = 0 \quad (2.40)$$

Solutions of this equation are of the form

$$\psi(x, y, z) = E_0 \frac{w_0}{w} H_m \left(\frac{x\sqrt{2}}{w} \right) H_n \left(\frac{y\sqrt{2}}{w} \right) \cdot \exp \left(-(x^2 + y^2) \left(\frac{1}{w^2} + \frac{ik}{2R} \right) - i(m + n + 1) \arctan \left(\frac{z}{z_0} \right) \right)$$

where $R = z \left(1 + \frac{\pi^2 w_0^4}{\lambda^2 z^2} \right)$ is the radius of curvature of the phase front,

$$w = w_0 \sqrt{1 + \frac{\lambda^2 z^2}{\pi^2 w_0^4}} \text{ is the waist of the beam, with minimum waist } w_0,$$

$$k = \frac{2\pi}{\lambda},$$

(2.41)

and H_m, H_n are the **Hermite polynomials** of order m and n respectively.[21, 22] These equations describe the **Hermite-Gaussian modes** with mode numbers mn , also known as transverse electromagnetic modes (TEM_{mn}). The output of many lasers is TEM_{00} . There are other beam profiles such as Laguerre-Gaussian modes and Bessel modes.[22]

2.4 Photodetection

Detecting the presence of light is a subtle experimental task. Photodetectors absorb optical energy and output a corresponding electrical signal. The quantum nature of light cannot be ignored and so the theory of photodetection can be dealt with either semiclassically or by full quantisation. The former can be thought of in terms of the wave-particle duality, and suffices for most practical purposes (e.g. the photoelectric effect, coherent states). The latter stipulates that the electric field be quantised and is only necessary in the case of non-classical states of light (e.g. squeezed states, Fock states). A thorough introduction to Poisson processes is given here. This is necessitated by their use in semiclassical photodetection theory. Lastly, we compare the semiclassical and quantum descriptions of photodetection.

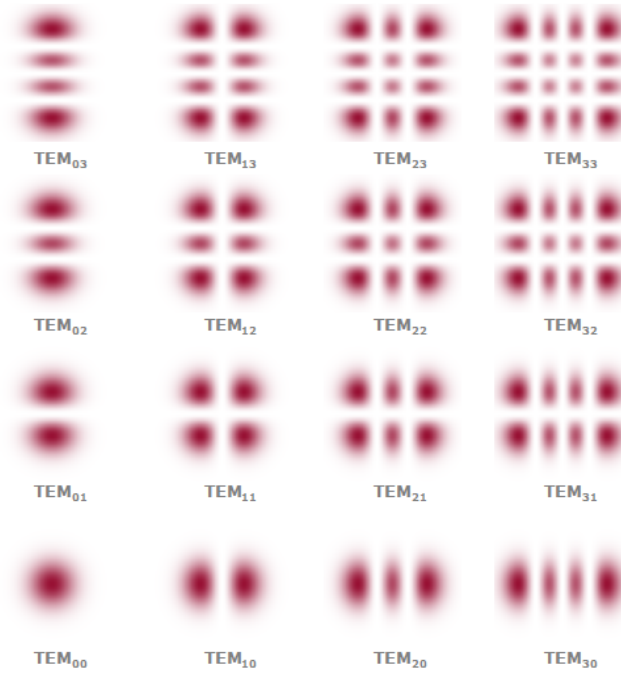


Figure 2.1: Some intensity profiles of Hermite-Gaussian modes.[22]

2.4.1 Poisson processes

A **Poisson process** is an example of a stochastic process used for arrival events, or simply, arrivals. The probability of an arrival during a time period Δt_1 is independent of the probability of an arrival during a different time period Δt_2 . As an arrival process, the Poisson process starts at time zero. The intervals between arrivals form an exponential distribution function. Denoting β as the rate such that the number of arrival events in an interval of length t is βt , each random variable X has a density $\beta \exp(-\beta x)$.

Poisson processes are what is called memoryless. This means that if an arrival event has not occurred yet at a time u then the distribution of waiting time for an arrival event on either side of u is the same. So the remaining waiting time does not know that there has been waiting in the past.[23]

An **inhomogeneous Poisson process** is one where the arrival rate is a function of time. $\beta(t)$ is positive. The distribution of the number of arrivals in a time interval (t, τ) is given by

$$\Pr [N(t, \tau) = n] = \frac{(\int_t^\tau \beta(u) du)^n \exp(-\int_t^\tau \beta(u) du)}{n!} \quad (2.42)$$

If the arrival events in a Poisson process are independently directed to different places, and there is a fixed probability associated with being directed to a particular place, then the arrivals at each place form individual, independent Poisson

processes.[23] This property allows us to consider the noise due to the pixellation of photodetectors as a Poisson process, as the photodetector can be considered as many arrival places.

2.4.2 Semiclassical and quantum theories of detection

For semiclassical applications, monochromatic light incident on a detector can be described by the field $E(t) \exp(-i\omega t)$. The average power of the light illuminating the detector will then be $\hbar\omega|E(t)|^2$. The photon count $N(t)$ is given by an inhomogeneous Poisson process with rate $\beta(t)$, dictated by: statistically independent increments, and Poisson distributed increments according to (2.42).[24] The rate function is given by $\beta(t) = \eta|E(t)|^2$. The mean photocount and the mean photocurrent are thus

$$\langle N(t) \rangle = \int_0^t \eta |E(\tau)|^2 d\tau \quad (2.43)$$

$$\langle j(t) \rangle = q\eta |E(t)|^2 \quad (2.44)$$

The covariances of the uncertainties of both these quantities give shot noise and excess noise terms.

In fully quantised photodetection, the light field is also quantised $\hat{E}(t) \exp(-i\omega t)$. The photocurrent is $\hat{j}(t) = q\hat{E}^\dagger \hat{E}$ and the photocount is $\hat{N}(t) = \frac{1}{q} \int_0^t \hat{j}(\tau) d\tau$.

This project uses classical sources of light. Is quantum mechanics overdoing it slightly? Quantum mechanics is the most powerful theory of light available: we discover what is possible in optics and detection through quantum mechanics. Any unbiased estimation using measurement, classical or quantum, will be bounded and optimised by the qCRLB. Through quantum optics, semiclassical approximations can be shown to be rigorous on a case-by-case basis, while respecting fundamental quantum principles governing noise.

2.5 Resolution

What does it mean to resolve something? Are we getting closer to resolving this issue? I don't know. Read this section and find out!

2.5.1 Point spread functions

Before we discuss resolution, it is important to set out what is meant by a **point spread function** (PSF). As the name suggests, this is a function that describes the spreading out of a point source or object as it is seen through an imaging system. Any object can be expressed as the superposition of many point-like objects, which is why the PSF is so useful.

An example of a point source is light shone through a very small circular aperture. Fraunhofer (far-field) diffraction around the aperture will produce an intensity pattern

$$I(\Theta) \propto \left(\frac{J_1(\pi D \Theta / \lambda)}{(\pi D \Theta / \lambda)} \right)^2 \quad (2.45)$$

where $J_1(r)$ is a Bessel function of the first kind, D the diameter of the aperture, and Θ is the angle subtended from the centroid. The pattern, with its central bright disk and concentric light and dark rings, is known as an Airy pattern.[25]

The Fraunhofer diffraction pattern is given by the two-dimensional Fourier transform of the aperture. A rectangular aperture of infinite length and width W , produces an intensity pattern

$$I(\Theta) \propto \left(\frac{\sin(\pi W \Theta / \lambda)}{\pi W \Theta / \lambda} \right)^2 \quad (2.46)$$

In practise, the finite length of the slit or the spot size of the beam incident upon it limits the vertical extent of the diffraction pattern (when the slit is oriented vertically). The PSFs (2.45) and (2.46) may be nicely approximated by Gaussian curves. Examples of this approximation are shown in Figure 2.2.

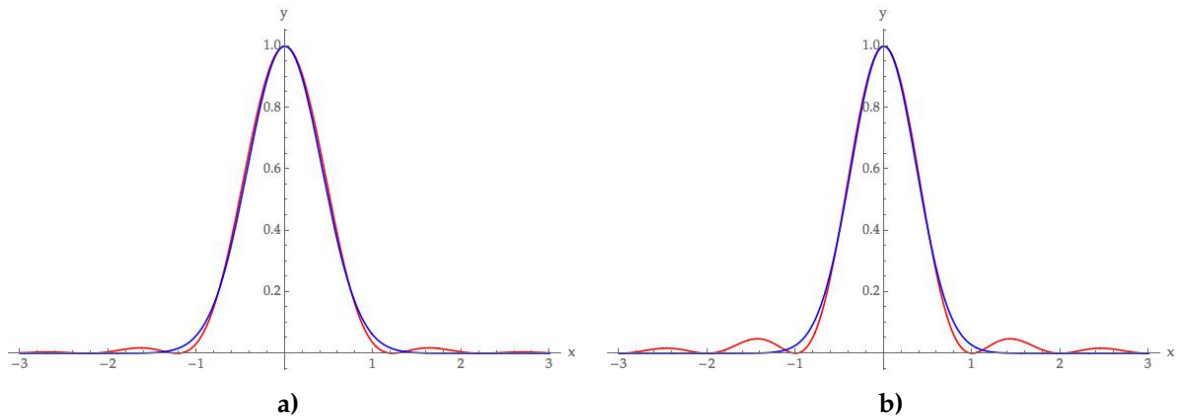


Figure 2.2: Gaussian approximations of point spread functions: **a)** in red $y = 4 \left(\frac{J_1(\pi x)}{\pi x} \right)^2$, in blue $y = \exp(-2.83x^2)$; **b)** in red $y = \left(\frac{\sin(\pi x)}{\pi x} \right)^2$, in blue $y = \exp(-3.65x^2)$.

2.5.2 Quantifying resolution

There exist several different measures of resolution. Here $2r$ is the smallest distance you can place two sources apart while still being able to distinguish that they are separate sources through conventional optical means.

Ernst Abbe was a pioneering microscope manufacturer and is generally credited as the first to attempt to quantify resolution. Abbe's **diffraction limit** is [26]

$$r_{\text{Abbe}} = \frac{0.50\lambda}{\text{NA}} \quad (2.47)$$

where λ is the wavelength used and NA is the **numerical aperture**. The numerical aperture is $n \sin \theta$ for an optical system involving a lens: n is the refractive index of the medium between the objective lens and the image plane, and θ is the half the opening angle of the largest cone of light on either side of the lens. **Rayleigh's criterion** is another resolution measure, attributed to Lord Rayleigh. Two sources are said to be resolved by this criterion if the first minimum of the diffraction pattern of one source lies at the maximum of the other. If the sources are any closer together than this, they are unable to be resolved.[27] The expression below is derived from the Fourier transform of a circular aperture, which is a Bessel function of the first kind.

$$r_{\text{Rayleigh}} = \frac{0.61\lambda}{\text{NA}} \quad (2.48)$$

For angular resolution of an iris of diameter D , the Rayleigh criterion becomes

$$\Theta_{\text{Rayleigh}} = 1.22 \frac{\lambda}{D} \quad (2.49)$$

Sparrow's limit is attained when there is a constant brightness between the two sources of light in an image featuring both sources. Sparrow, an astrophysicist, reasoned that if you can detect a dip in intensity between two sources, then you can resolve them.[26] Sparrow's limit is when it is no longer possible to resolve sources and is mathematically presented as

$$r_{\text{Sparrow}} = \frac{0.47\lambda}{\text{NA}} \quad (2.50)$$

Other diffraction measures exist, but the Rayleigh's criterion and Abbe's limit are by far the most common. Figure 2.3 shows the differences between (2.47), (2.48) and (2.50) in practise.

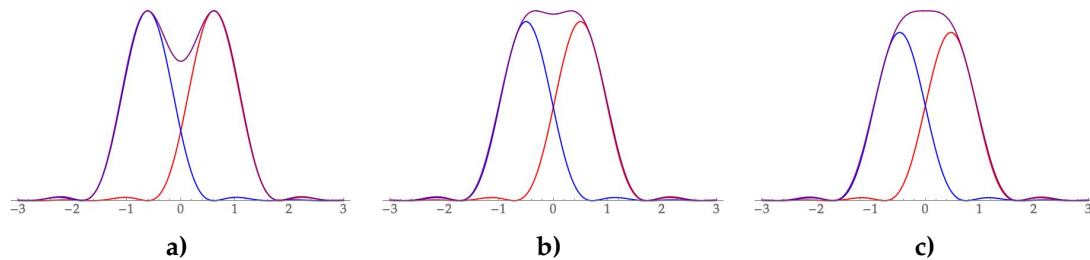


Figure 2.3: Different measures of resolution for two point sources (red and blue), with the combined intensity profile is shown in purple: **a)** sources separated by $2r_{\text{Rayleigh}}$, **b)** sources separated by $2r_{\text{Abbe}}$, and **c)** sources separated by $2r_{\text{Sparrow}}$.

Although the diffraction limit was believed to be a boundary for experimental work for decades, it has never been - or claimed to be - a fundamental physical law. There has been increasing dissatisfaction with these kinds of resolution measures, as they do not account for the intensity of the light or the duration of the measurement. Indeed Rayleigh and Abbe's work was motivated in part by the human eye, which is not a state-of-the-art optical instrument. For visible wavelengths ($\approx 500\text{nm}$) and

a NA of 1, Abbe, Rayleigh, and Sparrow suggest that we can resolve point sources down to 250 nm, 305 nm and 235 nm respectively.

Photon counting statistics now play a part in many diverse measurements of light. Beginning from a stochastic viewpoint, it is possible to define other resolution measures such as the **fundamental resolution measure** or FREM.[28] This is a measure, rather than a limit, as it provides a metric for the estimation of separation between sources instead of dictating a separation at which resolution becomes impossible.

$$\text{FREM} = \frac{1}{\sqrt{4\pi\Lambda_0(t-t_0)\Gamma_0(d)}} \frac{\lambda}{\text{NA}} \quad (2.51)$$

This retains the some form of earlier resolution limits as far as dependence on wavelength and numerical aperture are concerned; however, it also includes factors due to the intensity of each point source (Λ_0), the acquisition time interval ($[t, t_0]$), and the point spread function ($\Gamma_0(d)$) for sources separated by distance d . The point spread function described in [28] is involves a complicated integral across the image plane, where the integrand consists of first- and second- order Bessel functions of the first kind (familiar from Airy disk diffraction patterns). The argument of the Bessel functions is the product of the radial distance from a point source and the ratio λ/NA .

Plots of the integrand of $\Gamma_0(d)$ are shown in Figure 2.4. The integrand is basically zero everywhere except for a region about the origin, and becomes much flatter for greater λ/NA . If λ is less than d , then the integrand dramatically changes shape: it approaches zero at the origin, with peaks becoming taller, sharper, and tending towards two distinct locations. The separation of the pairs of peaks hints at increased resolving power at shorter wavelengths. Indeed the FREM exhibits a non-linear dependence on λ/NA . The FREM indicates that with an increasing number of detected photons, sources at separations approaching 50 nm can be resolved. As the separation tends to zero, the FREM blows up due to $\Gamma_0(d)$ tending to zero.

As emphatically stated in [20], statistical methods can – and should – be applied at every stage of the analysis of an optics experiment. The FREM is a gauge of how accurate the resolution is. In [28], the FREM treats the photodetector as though it has pixels of finite size and so accounts for shot noise. The authors then include additional sources of Poissonian (dark current, autofluorescence of the GFP) and Gaussian (detector readout) noise in their treatment to create the “practical” resolution measure using Fisher information and parameter estimation. With this more realistic analysis and experimental verification using GFP and cyanine-5 molecules, it was concluded that increasing photocount does allow for precise estimation of the parameter d beyond Rayleigh’s criterion. For example, the FREM for $d = 235\text{nm}$, a detection rate of 3000 photons per second, a detection time of 1 second and a NA of 1.4, is $\approx 2\text{nm}$ for visible light. Unlike the limits discussed earlier, the FREM encapsulates extra information including the intensity of the

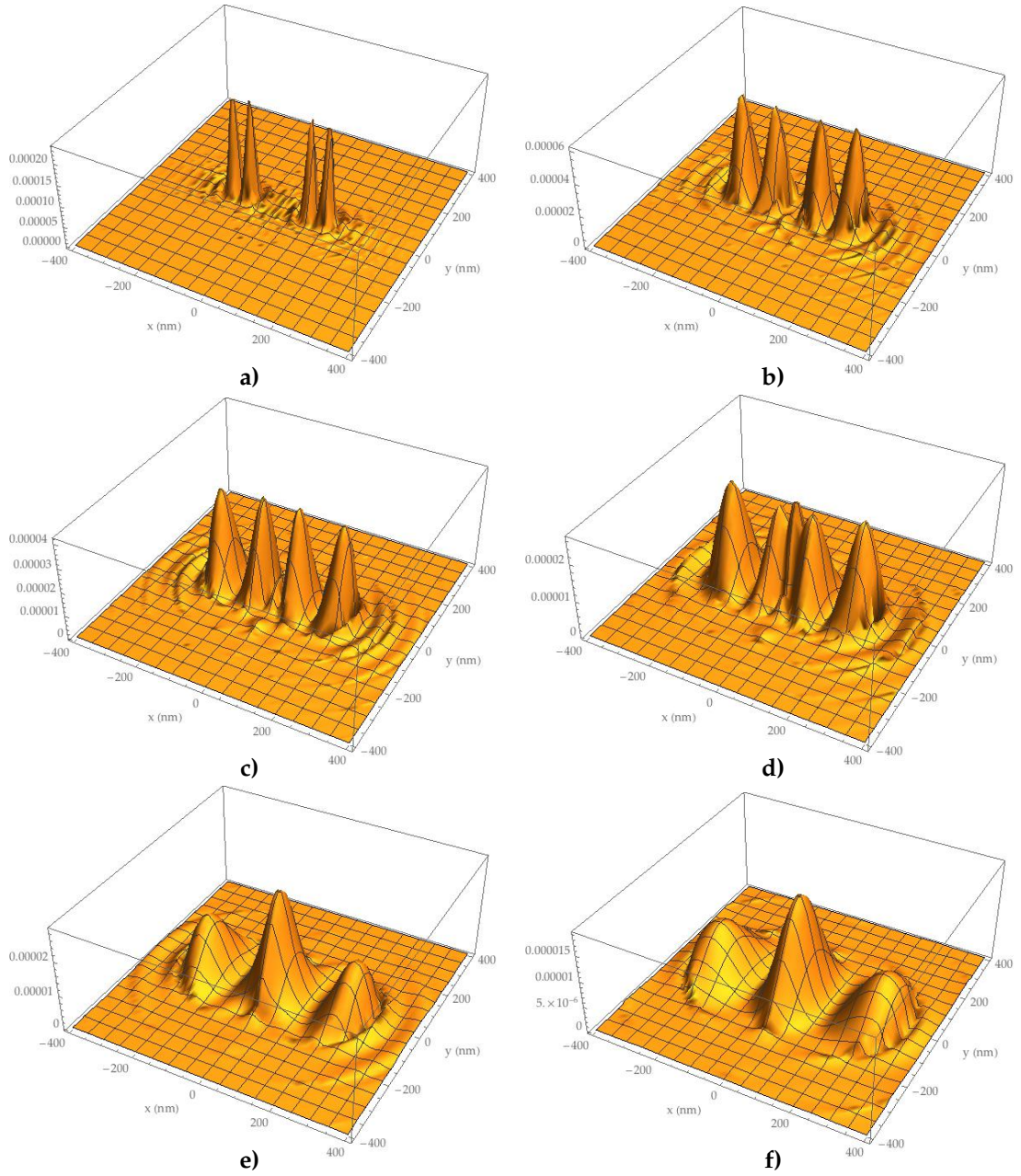


Figure 2.4: The integrand of the $\Gamma_0(d)$ function for $d = 235$ nm, $NA = 1.4$ and different λ : a) 100 nm, b) 200 nm, c) 250 nm, d) 300 nm, e) 400 nm, and f) 500 nm.

sources and the duration of the measurement. This allows different experimental conditions to be compared for objects consisting of point sources separated by the same distance. This work has two main caveats: higher photocounts are not always possible in fluorescence microscopy due to bleaching of the fluorophores, and both the FREM and the practical resolution measure asymptote to infinity at $d = 0$. Nevertheless, [28] removes the blinkers of a hard diffraction limit from the eyes of physicists and shows the agility of the estimation theory approach to resolution.

The problem with using more photons alone to resolve as d tends to zero shown in [28] is titled **Rayleigh's curse** by Tsang et al.[6]. This is responsible for the divergent behaviour of the FREM. Rayleigh's curse is as follows:

Any (unbiased) estimate of the separation between the sources from the image-plane photocounts must suffer a divergent mean squared error for a given mean photon number as the separation tends to zero.[6]

2.5.3 Resolution for incoherent point sources

Pioneering work on the resolution of incoherent point sources of light via estimation theory was done by Carl Helstrom in the 1970s.[11, 29] He considered two problems, each with profound implications for signal processing and imaging. Each problem concerns a choice between two hypotheses. The first: there are two point sources of incoherent light separated by a known distance σ and digits 0 and 1 are signalled by the illumination of one or other of the sources. Both sources are never on simultaneously.[11] From a distance away that is much larger than the separation between the sources x , have we received a 0 (hypothesis H_0) or a 1 (hypothesis H_1)? What is the average probability of an incorrect decision? This error probability indicates the resolvability of the sources.

Helstrom treats the binary hypothesis test as a state selection problem. Each hypothesis is assigned a density matrix ρ_0, ρ_1 , which represents the electric field at the receiving aperture under H_0 or H_1 respectively. The operator equation

$$(\rho_1 - \Lambda\rho_0)|\eta_k\rangle = \eta_k|\eta_k\rangle \quad (2.52)$$

gives the eigenvectors $|\eta_k\rangle$, which are used to construct the optimum detection operator Π . The constant Λ is equal to one in the case where the prior probabilities of the two hypotheses are equal (that is, one source is no more likely than the other to be turned on without knowing anything else). The resulting optimum detection measurement has two outcomes 0 or 1, corresponding to decisions accepting either H_0 or H_1 . In the case of no photons entering the aperture, H_0 and H_1 are equally likely according to this operator. The error probability is calculated and is shown in Figure 2.5.

The second estimation theory problem considered by Helstrom can be unpacked in much the same way, but with different density operators (and more algebra again!). Are there two incoherent sources of equal intensity located at $(x, 0, 0)$ and

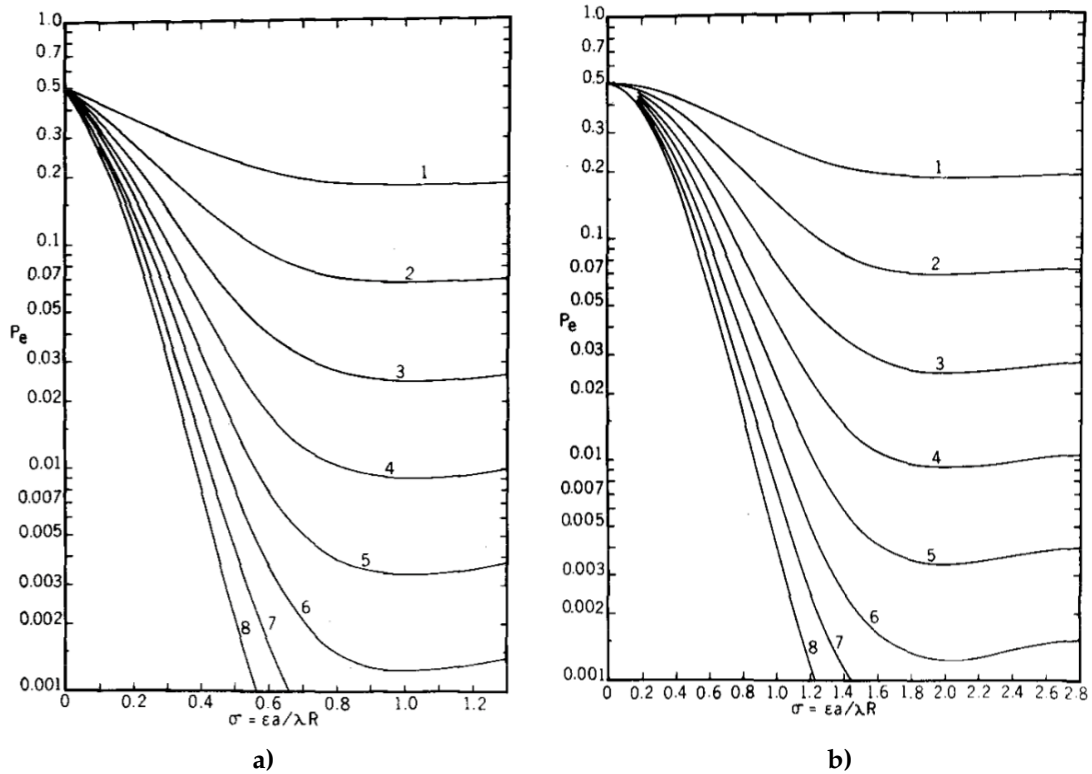


Figure 2.5: Average error probability P_e in terms of the separation parameter σ for different values of mean received photon number (indices of curves). The separation parameter is expressed in terms of the wavelength of light λ , the extent/diameter of each source in the object plane a , the separation of the two sources (if there are two sources!) ϵ , and the distance between the object and image planes R . **a)** is a plot of P_e in deciding which of two sources is radiating, **b)** is a plot of P_e in deciding whether there is one source present or two.[11, 29]

$(-x, 0, 0)$ or is there one source of twice the intensity at $(0, 0, 0)$?[29] The average probability that an incorrect hypothesis will be chosen is plotted in Figure 2.5, and is remarkably similar in form to that of the first problem.

For decades, Helstrom's work has been peerless in the area of quantum resolution. However, the point sources he considers have a known separation. In a sense, this begs the question: what's the point of resolving sources when you already know how far apart they are? We must now move forwards using key results from this chapter into the new world of super-resolution.

Chapter 3

Super-resolution

We have arrived at the hot topic of this thesis: super-resolution. In the first section of this chapter, we look at new techniques utilising Fisher information for imaging point sources separated by distances less than Rayleigh's criterion. We then move on to interferometry and its specific applications in the field of super-resolution through image inversion interferometry. Finally, superlocalisation via image inversion interferometry is introduced.

3.1 Recent progress

The literature of this section draws upon quantum metrology, statistical optics and estimation theory. Here several methods of achieving super-resolution are reviewed. All have been uploaded to the arXiv in the past year and there is not a GFP in sight. These techniques actively violate the diffraction limit by exploiting information in the transverse modes of the image plane. While none of the works presented in this section rely on interferometry, they serve to highlight the power of quantum estimation theory in the realm of super-resolution.

Splitting the image plane into its constituent Hermite-Gaussian modes (2.3.2) and measuring the amplitude of each is a process proposed by Tsang, Nair and Lu in [19] to superresolve incoherent point sources. This method is titled **spatial mode demultiplexing** (SPADE). For point sources with zero separation, there is effectively one point source that is completely spatially symmetric, and hence all of the light detected in a paraxial scheme is coupled into the Gaussian TEM_{00} mode. As the separation between sources increases, the amount of detected light in TEM_{00} decreases. In short, transverse modes are sensitive to any separation but do not indicate which source is on or off within a given (very short) coherence time. The Hermite-Gaussian basis is ideal for estimating the separation between point sources. Indeed the Fisher information calculated in this basis ($\mathcal{J}^{(HG)}$) does

not depend on the separation parameter:

$$\mathcal{J}^{(\text{HG})} = \frac{N}{4\sigma^2} \quad (3.1)$$

where N is the average photon number and σ the width of the PSF from each source. If the total number of photons counted over all measurements is L , with m_q being the number of photons in each mode TEM_{q0} , then the MLE estimator for the separation is

$$\hat{d} = 4\sigma \sqrt{\frac{1}{L} \sum_q q m_q} \quad (3.2)$$

With an increasing number of measurements, maximum likelihood estimation approaches the QCRLB. Monte-Carlo simulation in [19] confirms this. Experimentally, it is proposed that highly multimode waveguides would be a natural choice for measuring the photocounts in the Hermite-Gaussian basis, for these transverse modes are the waveguide modes of a quadratic-index waveguide. Simpler set-ups have been suggested (**binary SPADE**, [19] and **superresolved position localisation by inversion of coherence along an edge** (SPLICE) [30]). Common to both experiments is measuring the amount of light in the TEM_{00} mode, and the amount of light not in this mode, either as is ([19]) or projected into a single mode orthogonal to TEM_{00} ([30]). The beauty of this simplification is as the separation decreases, far less light is found in the modes ‘further’ from TEM_{00} , so little information is lost for small separations. At larger separations, the sources can of course be resolved through image plane photocounting.

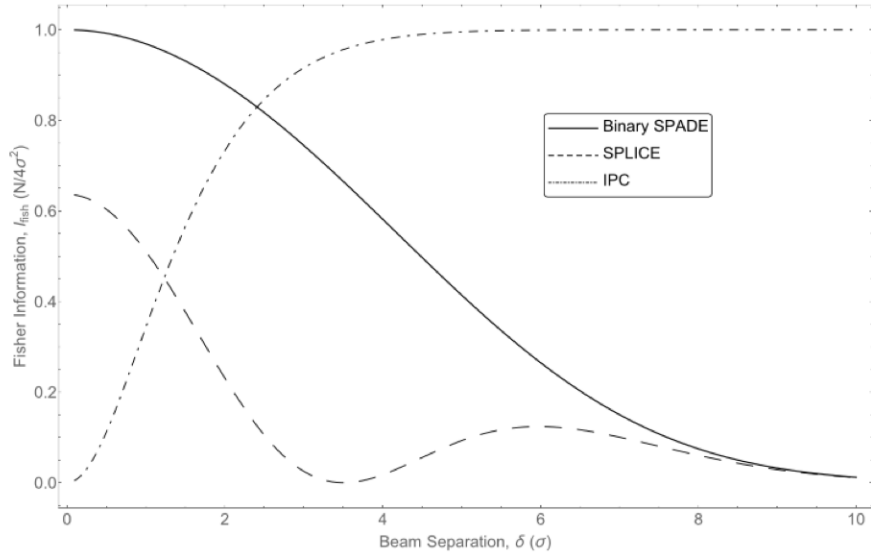


Figure 3.1: Fisher information: binary SPADE, SPLICE and image plane photocounting (IPC). The Fisher information is given in units of $N/4\sigma^2$ and the separation of the sources in units of σ . [30]

For binary SPADE, the Fisher information is given by

$$\mathcal{I}^{(Bi)} = \frac{N}{4\sigma^2} \frac{Q \exp(-Q)}{1 - \exp(-Q)}, \quad Q = \frac{d^2}{16\sigma^2} \quad (3.3)$$

and the MLE is simplified to

$$\hat{d}^{(Bi)} = 4\sigma \sqrt{-\ln \frac{m_0}{L}}. \quad (3.4)$$

Recent alternative uses of higher order transverse modes to achieve super-resolution include optical heterodyne detection[31] and digital holography.[32] The former treats the light from the two incoherent point as the signal beam. By preparing the local oscillator beam in TEM₀₁, the output power of the detector samples only the TEM₀₁ mode of the signal beam. In this way, information about the separation between these sources can be gleaned. This technique has delivered more precise resolution for sources separated by 0.18mm,[31] but mode-matching difficulties in practise and the limitations caused by shot noise have meant smaller separations have not been imaged. The latter method projects onto different modes by way of an amplitude spatial light modulator that creates a hologram, and measures the photocounts in the TEM₀₀ and TEM₀₁ modes.[32] The initial demonstration resolved sources with separations down to $\approx 8\mu\text{m}$.

The above works draw similar conclusions through a variety of means. Transverse modes beyond TEM₀₀ are dense in information about the separation of point sources. With an increase in the number of modes measured comes an increased range of separations for which the measurement saturates the qCRLB.[33] As these techniques are not yet a year old, we can expect sub-Rayleigh separations to be resolved with improvements in the near future. Already the errors accompanying the estimates of separations in [30, 31, 32] significantly better the precisions established by Rayleigh's criterion through conventional image plane methods. It should be noted that none of the demonstrations in this section used thermal sources, only pseudo-thermal sources that were mutually incoherent. While it is claimed that the performance of each of the methods [30, 31, 32] would not suffer if the illumination came from classically incoherent light, no one has actually bothered to do this yet. Using incoherent light is an important test of super-resolution methods prior to their implementation in optical fields such as fluorescence microscopy.

3.2 Interferometry

Interferometry is the science of combining waves – usually electromagnetic, so henceforth light will be used as the example – together to generate some new data that will tell you something about where or what the waves came from. Interferometers come in many designs depending on their purpose. For example, there is the legendary Michelson-Morley experiment using a two-arm interferometer

with a single beam splitter at the intersection of the two arms and a single detection port. This experiment demonstrated that light waves do not require a medium in which to propagate, and disproved the purported existence of the luminiferous aether.

A more subtle interferometric experimental set-up is the Mach-Zehnder interferometer, which features again two arms but this time two beam splitters and two detection ports. Light comes in and is split along two paths. Each path is normally depicted as being two sides of a square. The light is then recombined at the final beam splitter. Half the light from each path goes to each detector. The square has been completed in a sense. One of the detection ports is a so-called ‘dark’ port, where destructive interference is seen. The other is the ‘bright’ port, where constructive interference is observed. By changing the optical path length difference between the two arms, you can scroll through interference fringes. Which fringes you get at either port is determined by the relationship between the optical path and the wavelength. A whole number of wavelengths as a path length difference gives constructive interference - the beams from each arm have arrived at the detector in phase. In this way, interferometers can detect very small changes in distance or optical path length. This is well known and these kinds of interferometers feature in the basic experimental study of physics.

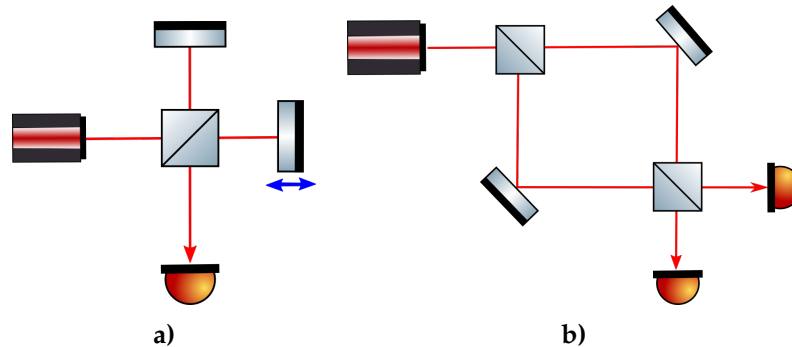


Figure 3.2: Schematic diagrams of **a)** a Michelson interferometer with one mirror translating (blue arrow) and **b)** a Mach-Zehnder interferometer

The **visibility** of an interferometer is defined as

$$\mathfrak{V} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (3.5)$$

where I_{\max} and I_{\min} are the maximum and minimum intensities of the fringe. As the relative pathlength difference grows between the arms of the interferometer, the visibility will drop off, owing to the finite coherence length of the light. This harks back to the principles outlined in 2.3.1: when the time delay in one arm approaches or exceeds the coherence time, visibility will go to zero. When the path length difference is zero, there is still an overarching envelope seen in the fringe pattern, that is, the central fringe is the brightest. This indicates spatial coherence.[20]

It is useful to define the ratio of residual power output at a port of an interferometer,

Ω :

$$\Omega = \frac{\text{Total power output at port}}{\text{Power in first arm plus the power in the second arm}}. \quad (3.6)$$

3.2.1 Image-inversion interferometry

An image-inversion interferometer is basically a Mach-Zehnder interferometer with an arrangement of mirrors, or another inversion device, in one arm to flip the beam about one axis. Any light that is on axis remains so, and there is still destructive and constructive output ports for this light. Any antisymmetric, or off-axis, components of the field entering the interferometer will be preserved at the dark port.

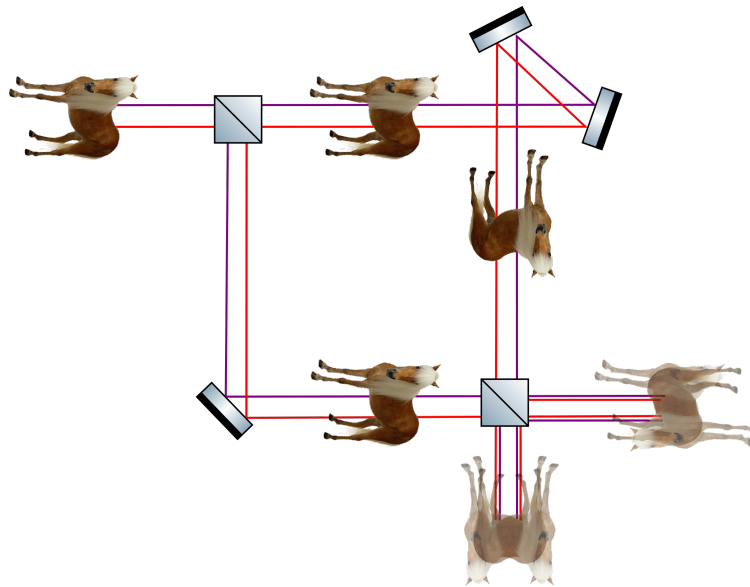


Figure 3.3: A schematic diagram of an image inversion interferometer. Horses are for illustration only.

3.3 Superlocalisation via image inversion interferometry (SLIVER)

Superlocalisation via image inversion interferometry, or SLIVER, is a super-resolution technique put forwards by Ranjith Nair and Mankei Tsang in late 2015.[6] In essence, information about the separation of two point sources centred about the optical axis is contained in the output of the dark port of an image inversion interferometer. As stated above, this is because the dark port is sensitive to off-axis components of the electric field; thus, it enables an estimator for the separation to be formulated, as this explicitly depends on how off-axis the points of light are. Below the SLIVER method is studied from the perspective of estimation

theory, closely following the work of Nair and Tsang in [6]. As such, all equations and results in this section can be assumed to be sourced from [6] unless otherwise referenced. There are five aims of this section (relevant subsections in parentheses):

1. to detail how SLIVER achieves super-resolution through the calculation of key statistical quantities (3.3.2,3.3.3);
2. to familiarise the reader with the experiment ahead of the specific experimental discussion in Chapter 5;
3. to present a possible extension of the method, **pixelated SLIVER** (3.3.5);
4. to highlight the applicability of SLIVER in several scenarios, including the binary hypothesis problems of 2.5.3 (3.3.6);
5. to inform the reader of empirical progress made with SLIVER independently of this project during the course of the year (3.3.7).

3.3.1 Statistical optical modelling

The interferometer requires incoherent thermal sources.¹The incoherence is necessary for each source to interfere only with itself in the interferometer. Examples are discharge lamps and light emitting diodes, and the optical states of such sources could be modelled as thermal states. The incoherence is introduced by using random variables for the amplitudes. The input field entering the interferometer is given as

$$E(\rho) = A_1\psi(\rho + d/2) + A_2\psi(\rho - d/2) \quad (3.7)$$

where d denotes source separation, A_1 and A_2 the strengths of the fields for each source, and the distribution ψ of the light is a Gaussian

$$\psi(\rho) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-|\rho|^2/(4\sigma^2)). \quad (3.8)$$

The input field can be decomposed mathematically into the symmetric and anti-symmetric parts $E = E_a + E_s$:

$$E_k(\rho) = \frac{E(\rho) \pm E(-\rho)}{2} = \frac{A_k}{2} (\psi(\rho + d/2) \pm \psi(\rho - d/2)) \quad (3.9)$$

where the subscript $k = s, a$ corresponds to quantities at the symmetric and antisymmetric ports respectively. Similarly, the addition and subtraction in the \pm sign correspond to $k = s, a$ respectively. Symmetry gives $A_k = A_1 \pm A_2$.

The inversion arm of the interferometer would act on the input field as $E_{out}(\rho) = E_{in}(-\rho)$, thus the recombination at the output beam-splitter would be such that the symmetric and antisymmetric parts exit in different ports. In short, the image inversion interferometer functions to separate the symmetric and antisymmetric parts of the input field. Due to the symmetry of ψ , it suffices to invert in just

¹For coherent sources, there are easier methods to estimate the separation precisely.

one coordinate $E_{out}(x, y) = E_{in}(-x, y)$ through some mirror plane. This allows for easier realisations of the inversion in practice.

The integrated photocount at each port would be given by the modulus squared of the amplitude integrated over the image plane. As usual, there will be an interference term. Direct computation gives the average photocount at each port

$$\bar{N}_k = \epsilon(1 \pm \delta(d)), \quad (3.10)$$

a function of the average photocount without interference $\epsilon = \mathbb{E}[|A_s^2|]/2 = \mathbb{E}[|A_a^2|]/2$ and the interference term

$$\delta(d) = \Re \int \psi(\rho)\psi(\rho - d)d\rho. \quad (3.11)$$

For an incoherent thermal source, the photon statistics could be modelled by the Bose-Einstein distribution

$$P_k(n) = Pr [N_k = n] = \frac{1}{N_k} \left(\frac{\bar{N}_k}{\bar{N}_k + 1} \right)^n, \quad (3.12)$$

which would also serve as the likelihood function. In particular, the source separation would affect the degree of interference, which would in turn affect the photocount. The likelihood function quantitatively expresses this dependence.

3.3.2 Cramér-Rao bounds and Fisher information

Armed with the likelihood function, one can now proceed to calculate key quantities of estimation theory. The classical Fisher information \mathcal{I} is

$$\mathcal{I}_k = \mathbb{E} \left[\frac{\partial}{\partial d} (\log p_k) \right]^2. \quad (3.13)$$

The expectation would be performed over the likelihood function as dictated by the theory, which in this case is the Bose-Einstein distribution. Computing the expression explicitly leads to

$$\mathcal{I}_k = \frac{\epsilon\gamma(d)^2}{1 \pm \delta(d)} \frac{1}{1 + \epsilon(1 \pm \delta(d))}. \quad (3.14)$$

A reminder that addition and subtraction in the \pm sign corresponds to the symmetric and antisymmetric ports respectively. The function $\gamma(d)$ is the derivative of δ , and as δ is an even function of the separation d , γ is an odd function and thus vanishes for zero separation.

Inversion of the Fisher information gives the Cramér-Rao bound

$$\text{var}[\hat{d}] = \mathbb{E} \left[\hat{d}(N_k) - d \right]^2 \geq \mathcal{I}_k^{-1} = \frac{1 \pm \delta(d)}{\epsilon\gamma(d)^2} + \left(\frac{1 \pm \delta(d)}{\gamma(d)} \right)^2. \quad (3.15)$$

In the limit of zero separation, $\delta(d) \rightarrow 1$ and $\gamma(d) \rightarrow 0$. Hence $\mathcal{I}_s^{-1} \rightarrow \infty$ pushing the MSE to infinity as well. This was expected: the symmetric port suffers from Rayleigh's curse.

The indeterminate form for \mathcal{I}_a^{-1} turns out to be finite in the limit and is equal to $1/2M\epsilon(\Delta k)^2$, where Δk is the spectral width of the Gaussian ψ given by

$$\Delta k = \sqrt{\int \left| \frac{\partial \psi(x, y)}{\partial x} \right|^2 dx dy}. \quad (3.16)$$

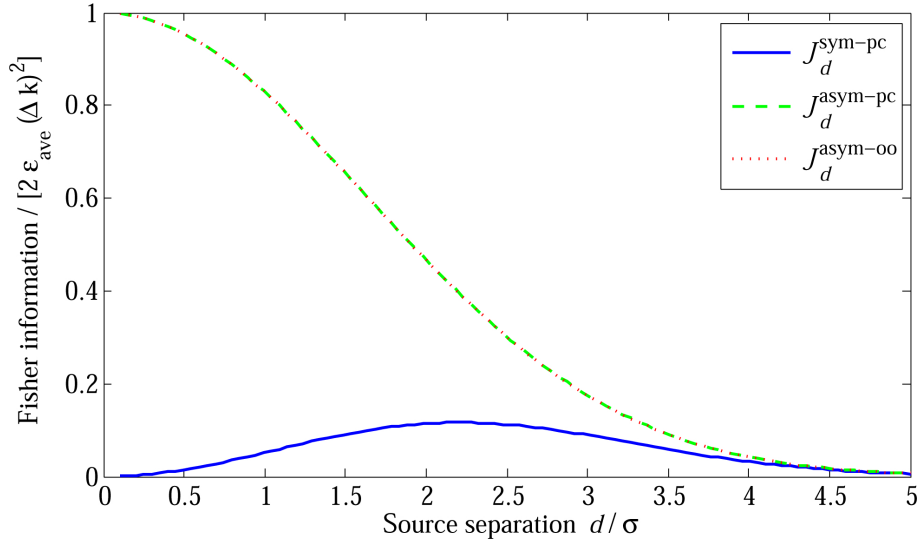


Figure 3.4: Fisher information as a function of source separation for average $\epsilon = 10^{-3}$. The blue, green, and red curves represent the information obtained through photoncounting at the symmetric port, photoncounting at the antisymmetric port (number-resolved detection), and photoncounting at the antisymmetric port (on-off detection).[6]

3.3.3 Maximum likelihood estimation

If $P = \sum_{j=1}^M p_j$ denotes the photocount over M measurements, direct computation shows that the maximum likelihood estimator is given by

$$\hat{d} = 2\sigma \sqrt{-2 \log \left(1 - \frac{P}{M\epsilon} \right)} \quad (3.17)$$

under the assumption $P < M\epsilon$. Note that this assumption is physical since P is the measured photocount with interference, while $M\epsilon$ is that without. At the antisymmetric port where the interference is destructive, the measured value would have to be smaller. With increasing M , the estimator asymptotically converges to the Cramér-Rao bound, and the MSE remains finite for small separations. The finite MSE allows for the measurement of the separation in the antisymmetric port. As the measurement depends on the interference within the interferometer, increased

separation weakens the interference and hence decreases the Fisher information. The MSE can be simulated for different separation values using Monte Carlo methods. For this purpose, the estimator returns a value of 2σ for cases where $P > M\epsilon$.

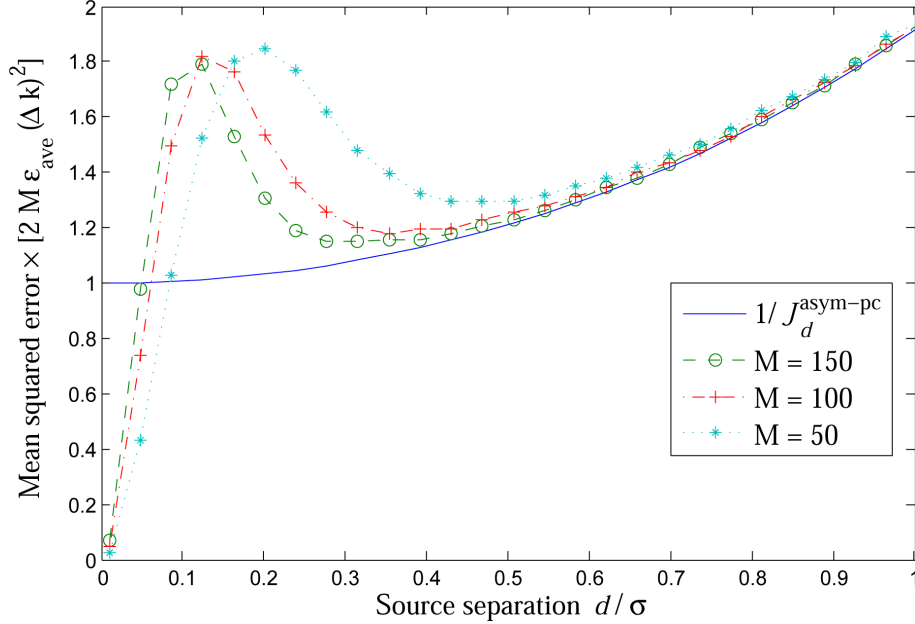


Figure 3.5: Simulated MSE for bucket detection at the antisymmetric port as a function of source separation for average $\epsilon = 5$. The blue curve is the normalised CRLB for photoncounting at the antisymmetric port. The remaining curves present the simulation results for different numbers of measurements ($M = 50, 100, 150$).[6]

Throughout the analysis in [6], number-resolved semi-classical photodetection has been employed. This is not a necessary condition for producing a non-divergent Cramér-Rao bound at the antisymmetric port. On-off detection also admit super-resolution in a similar way. In terms of Fisher information, on-off agrees with number-resolved detection for weak sources ($\epsilon \ll 1$ photon). For stronger sources, the Fisher information for on-off detection decays off much faster compared to number-resolved detection and hence has less resolving power.

3.3.4 The effect of misalignment on SLIVER

Up until this point, all formulations of statistical quantities have been based on the centroid of the two sources being known, and located on the optical axis at that. The sensitivity of the SLIVER method to displacements of the centroid has not yet been examined in detail in any preprints or publications; however, Tsang's group has simulated the MSE for a centroid located $\xi\sigma$ off the optical axis in the plane of the two point sources. The result, which indicates the resilience of SLIVER to misalignment in the plane of the sources, is shown in Figure 3.8. For small misalignments ($\xi = 0.1$) the MSE surpasses the CRLB for aligned on-off detection.

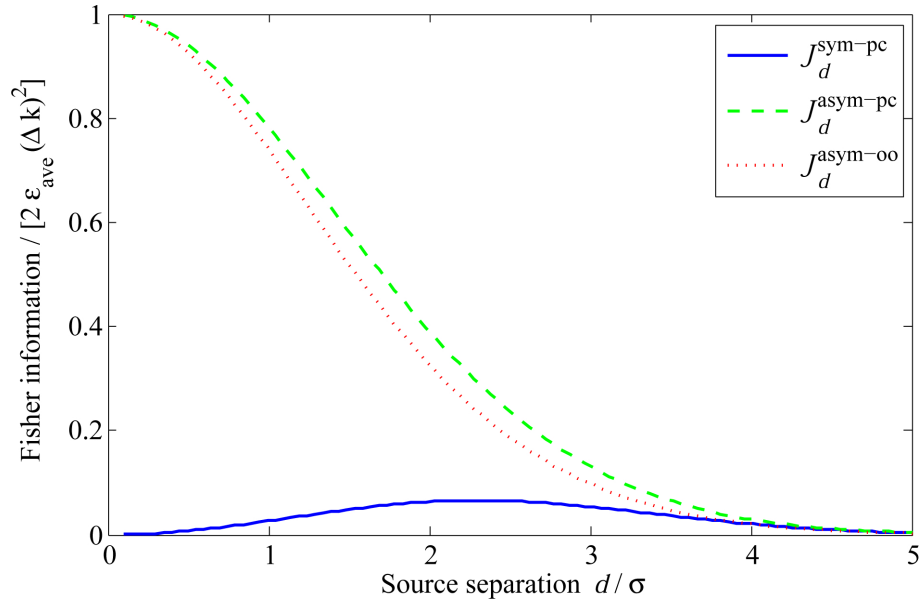


Figure 3.6: Fisher information as a function of source separation for $\epsilon = 0.5$. The blue, green, and red curves represent the information obtained through photoncounting at the symmetric port, photoncounting at the antisymmetric port (number-resolved detection), and photoncounting at the antisymmetric port (on-off detection).[6]

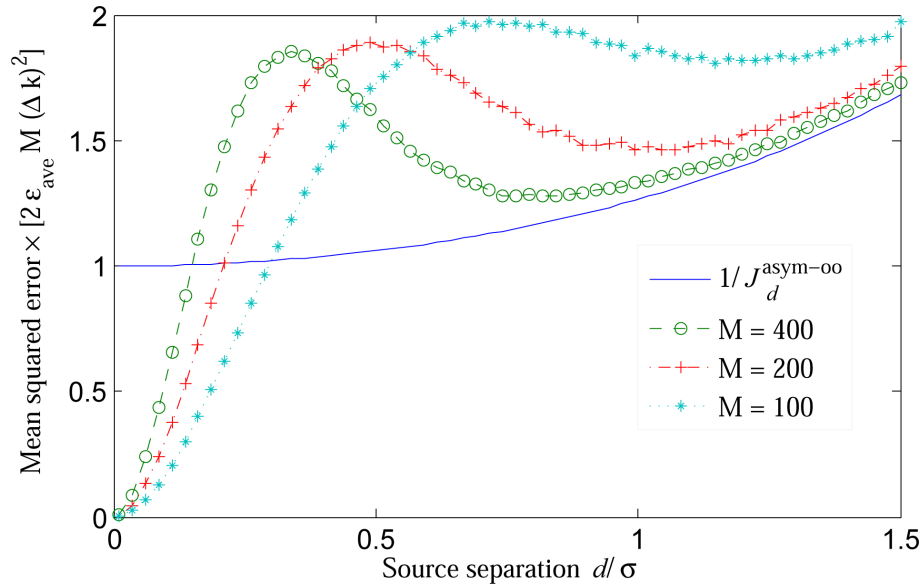


Figure 3.7: Simulated MSE for on-off detection at the antisymmetric port as a function of source separation for average $\epsilon = 0.2$. The blue curve is the normalised CRLB for photoncounting at the antisymmetric port. The remaining curves present the simulation results for different numbers of measurements ($M = 100, 200, 400$).[6]

Increasing misalignment does not hamper the performance of the estimator at larger separations; while as the separation goes to zero, the MSE increases but importantly does not diverge. Obtaining the quantum Fisher information in the case

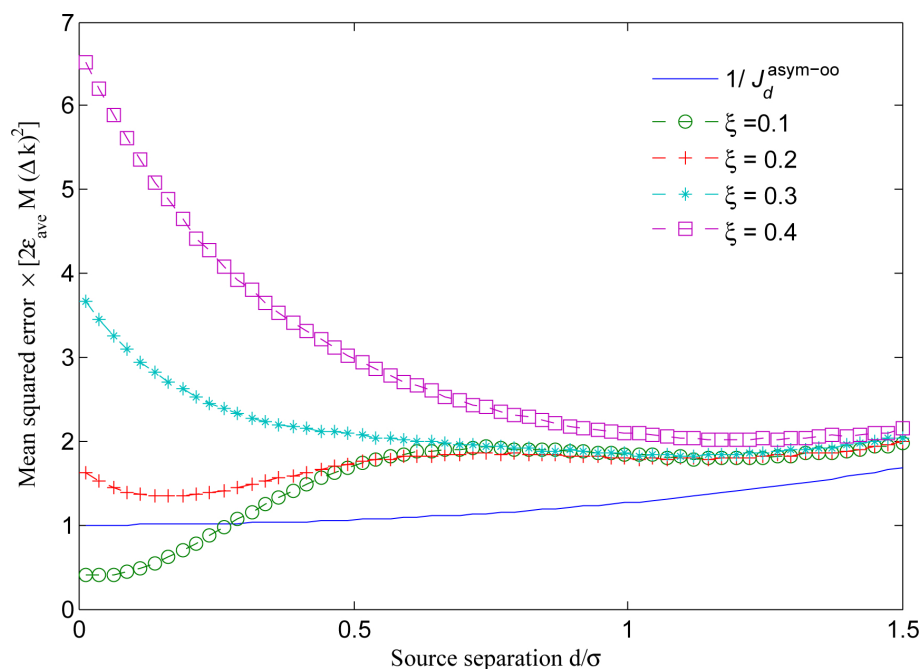


Figure 3.8: Simulated MSE for on-off detection at the antisymmetric port as a function of normalised source separation for average $\epsilon = 0.2$. The blue curve is the normalised CRLB for on-off detection at the antisymmetric port. The remaining curves present the simulation results for different numbers of measurements ($M = 100, 200, 400$).[6]

when both the separation and the centroid are unknown forms a two-parameter estimation theory problem. The form of this matrix has not been deduced. Direct imaging of the centroid is not subject to a divergent error, so it can be located using conventional techniques and aligned accordingly.

3.3.5 Pixelated SLIVER

While SLIVER delivers super-resolution for separations close to zero, the Fisher information via this method decreases with increasing separation. Eventually, the separation may be large enough to resolve the point sources through established image plane photocounting, but an approach that yielded high Fisher information at any separation would be preferred. Tsang and Nair suggest collecting spatial information from both ports of an image inversion interferometer that is otherwise lost through photodetection alone by using spatially-resolved detectors consisting of many pixels.[33] This is called **pixelated SLIVER** (pix-SLIVER). For two sources separated by a distance d along the x -axis and then interfered with an inversion in the same axis for one arm, the P pixels of the array of width W would be vertical strips of equal x width. Instead of one on-off detection measurement consisting of a 1 (photon detected) or a 0 (no photon detected), an on-off measurement would be a vector with a 1 or a 0 component for each pixel on both detectors.

Parameter estimation from such a vector quantity is much more complicated

to describe, and indeed the Fisher information depends on the covariances of detections at different pixels on the two different arrays, different pixels on the same array, and the same pixels on different arrays. For this reason, Tsang and Nair place a lower bound on the Fisher information available via pix-SLIVER using on-off detection in [33], rather than providing an explicit calculation for number- and spatially-resolved detection at both ports. Their results are shown in Figure 3.9.

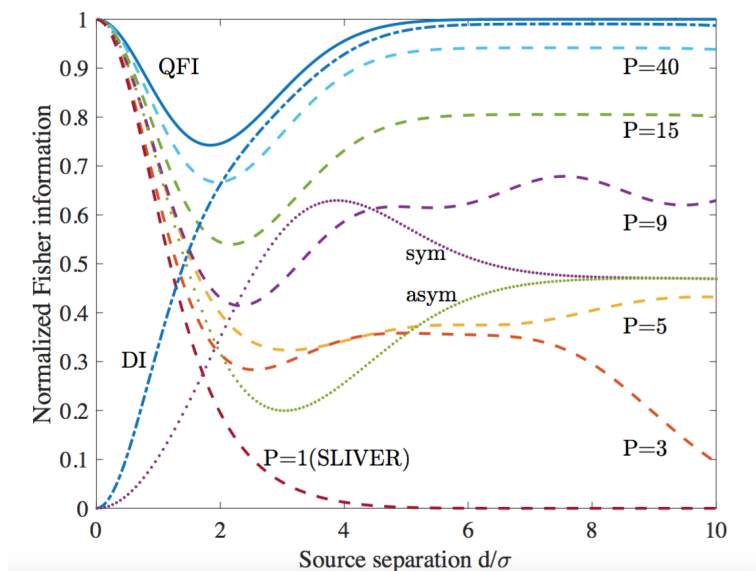


Figure 3.9: Fisher information versus separation for pix-SLIVER. The quantum Fisher information for two sources of arbitrary strength is shown in solid blue (expression (3.18)). The lower bound on the Fisher information for pix-SLIVER with $P = 1, 3, 5, 9, 25,$ and 40 are the dashed lines. The lower bound for $P = 40$ is also deconstructed into contributions from the antisymmetric and symmetric ports, shown respectively by the green and pink dotted lines. Finally the blue dashed-dotted line represents the Fisher information obtained through direct imaging. In all cases, the arrays are taken to have $W = 17\sigma$ and so each pixel is $17\sigma/P$ wide. [33]

Intuitively, the superior performance of pix-SLIVER compared to SLIVER across a range of separations can be explained quickly by considering the spread of the light from the sources as d increases. How spread out the light is when it arrives at an array is sensitive to the separation. Hence, a measurement that accounts for this spread, rather than merely adding up the number of photons, will return more information at larger separations.

3.3.6 Applications

The super-resolution method SLIVER has the potential to launch new developments in imaging previously imprisoned by Rayleigh's criterion. This is discussed in the final section of this thesis. The SLIVER framework may be applied directly to the binary hypothesis problems considered first by Helstrom some forty years ago (see

2.5.3) to great effect.[34]

As an extension of the estimation theory presented above, Nair and Tsang next considered sources of arbitrary strength, broadening the applicability of their approach to include incoherent sources in the microwave and infrared, and high-temperature stars.[33] Strong sources introduce the possibility of multiphoton coincidences at the detector. For two sources of arbitrary strength separated by d , the quantum Fisher information is

$$\mathcal{J} = -2 \left(\frac{\partial \gamma(d)}{\partial d} \Big|_{d=0} \right) N_s - 2\gamma^2(d) \left(\frac{(1 + N_s)N_s^2}{(1 + N_s)^2 - \delta^2(d)N_s^2} \right) \quad (3.18)$$

where first term is identified as twice the mean-squared spatial bandwidth of the PSF in the x -direction.[33]

Adding more dimensions presents another challenge for both theory (more parameters to estimate) and experiment (in the case of SLIVER, a second interferometer in the vertical direction). In the preprint [35], the quantum Fisher information matrix for the four parameters needed to describe sources separated in two directions is presented. Since the quantum Fisher information for multiple parameters was glossed over in 2.2, here we merely alert the reader that a proposal exists for a 2D SLIVER.[35]

The work of Tsang and co. has provided a leading example of quantum estimation theory forging ahead in problems of localisation and resolution, which has in turn motivated refinements in existing super-resolution techniques that have merely skirted around Rayleigh's criterion rather than outright defying it. The estimation of the position of a fluorescing molecule is revisited in [36], no doubt anticipating the application of SLIVER, SPADE or similar to the field of fluorescence microscopy.

3.3.7 Previous experimental progress with image inversion interferometry

One of the first appearances of image inversion interferometry in the literature as a technique for resolution enhancement is the 2007 work of Kai Wicker and Rainer Heintzmann.[37] They proposed imaging through a standard confocal microscope, or via an extended focus technique, followed by an image inversion interferometer. Their work does not consider the parameter estimation problem of localising point sources but rather introduces image inversion as a means to underscore off-axis components of the object to be imaged, and proceeds by analysing the optical transfer function for a single point source a distance away from the optical axis. The difference in intensity between the images at the two ports of the interferometer is used effectively to narrow the point spread function of the imaged points, improving the lateral resolution. This is demonstrated in Figure 3.10. Combining Wicker and Heintzmann's basic instrumentation proposal with recent works from the past twelve months, such as [36], which apply quantum estimation theory in the context of fluorescence microscopy, offers increased potential to expand resolution

capabilities of microscopes. Extensions of [37] include performing the inversion across several image rotation interferometers rather than in one image inversion interferometer. This would reduce the constant offset in the intensity at the dark port, which contains no information about the off-axis position of the fluorescing protein.

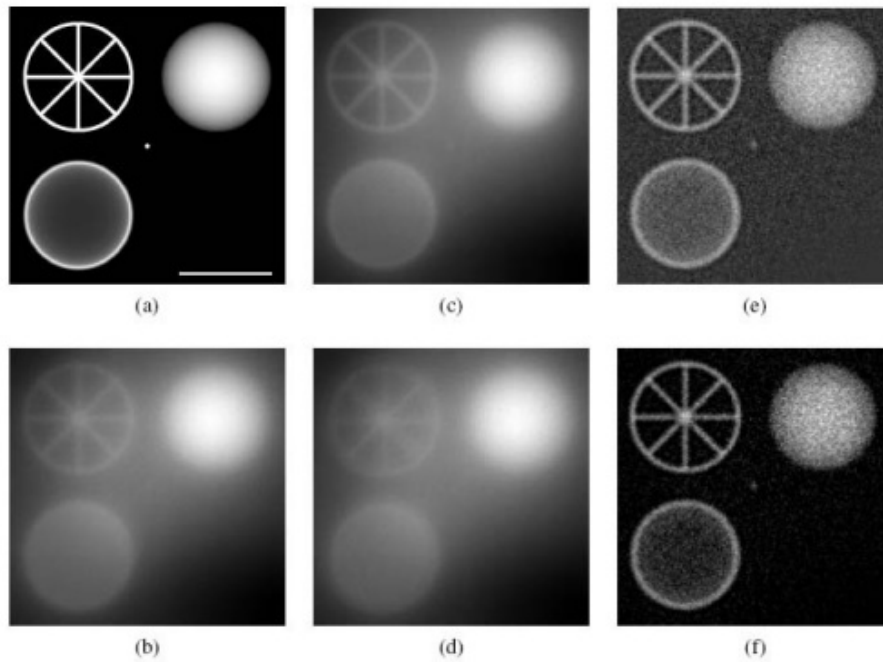


Figure 3.10: *Simulated images via Bessel beam excitation of a synthetic sample (a). b) simulation of Bessel beam excitation of sample with no interferometry. c) Image plane of the symmetric port after simulated image inversion interferometry. d) Image plane of the antisymmetric port. e) The difference of the two output ports. f) Contrast enhanced version of e). Image from [37], where further specifics can be found.*

In mid-2016, a demonstration of super-resolution through image inversion interferometry for mutually incoherent sources was completed by a group at the Centre for Quantum Technologies² at the National University of Singapore.[38] This was the first instance of SLIVER being tested experimentally. However, there were three key differences between the work in [38] and the original proposal of Tsang et al. in [6]. The two point sources in question in [38] originated in fact from the same HeNe laser. The laser light was split by a polarising beam splitter, giving two beams of orthogonal polarisation, which could then be recombined at a 50:50 beam splitter. The two beams were separated spatially before this second beam splitter. A configuration of two lenses, rather than an additional reflection in an arm of the interferometer, performed the inversion, but this was seen to introduce additional uncertainty in measurements due to the dispersive nature of the lenses.[38] Furthermore, the estimator implemented in [38] was a function of the residual power

²Not affiliated with Tsang's Quantum Measurement Group at the same institution.

output of the interferometer rather than the photocount across all measurements. For comparison, the estimator in [38] is

$$\hat{d} = \sigma \sqrt{2 \log \left(\frac{1}{1 - \Omega} \right)} \quad (3.19)$$

where $\Omega \approx 1 - \exp \left(\frac{-d^2}{2\sigma^2} \right)$.

Of these three differences (mutually incoherent sources that were not thermal, utilising lenses rather than mirror geometry for the inversion, and a different measured quantity used in estimation of the separation), the first is particularly noted by the authors of [38]: completely incoherent sources would present an additional challenge in construction of an image inversion interferometer, owing to the far shorter coherence length of the light.

Chapter 4

Fabrication

In order to create two incoherent point sources, optical devices with pairs of small apertures were fabricated. Each aperture then acted as a point source emitter when illuminated from behind by an incoherent source. Each pair was to be made up of two apertures of the same shape and size, ensuring that the point sources were of equal strength. Pairs differed in the separation between individual apertures, which is the primary independent parameter of interest to this project. To demonstrate resolution of green sources separated by less than the diffraction limit, the separations¹ would have to be on the order of a couple of hundred nanometers; hence, custom fabrication was required. This chapter documents the fabrication process.

The apertures considered were pairs of pinholes and slits. For such small devices, it would be far easier to align long rectangular slits than pinholes, and so the fabrication of pairs of slits was prioritised over pinholes. Dimensions of $100\mu\text{m}$ by $1\mu\text{m}$ were decided upon for the slits, with separations ranging from $2\mu\text{m}$ to $5\mu\text{m}$. Smaller slits with separations as tiny as 100nm were also planned.



Figure 4.1: Schematic of optical device design showing pairs of slits and grating (white) in an opaque material (blue).

The aim of the fabrication process was to produce macroscopic optical devices

¹As in previous chapters, the separation is defined as the distance between the two midpoints of the apertures.

each with sets of microscopic slits of different separations. A grating, consisting of slits spaced at progressively smaller intervals, was also included in the design (Figure 4.1) to test the process for smaller separations. It was clear from the outset that the main challenge would be obtaining the desired dimensions for the slits, as fortunately rectangles are a simple shape. Several sequences of techniques were discussed. An overview of the method described in this chapter is shown in Figure 4.2. The end result was five devices made of titanium-coated glass, with the slits etched into the titanium. Sub-micron separations were not achieved but a discussion of suggested modifications to the process appears in the final section (4.5) of this chapter. Here is a summary of the main steps of the fabrication process, with relevant sections in parentheses:

1. The glass slide was coated with a thin layer of titanium using electron beam evaporation (4.1);
2. A layer of electron beam resist was applied on top of the titanium (4.2);
3. Electron beam lithography drew the pattern of slits into the resist (4.2);
4. Development of the resist removed the areas drawn on by the electron beam (4.2);
5. Inductively coupled plasma etching transferred the pattern into the titanium (4.3);
6. The resist was removed and the fabrication was complete. (4.3)

Finally, scanning electron microscopy of the devices allowed the dimensions of the slits to be measured. This is detailed in section 4.4.

4.1 Electron beam evaporation

Electron beam evaporation (e-beam evaporation) is a physical vapour deposition (PVD) technique. Under high vacuum, a current is applied to a tungsten filament, which gives off electrons. The electrons are then formed into a beam and steered by a magnetic field towards a sample of the material that is to form the coating. Attacked by the electron beam, the material vaporises. This vapour then precipitates and coats everything in sight, including the substrate of interest, and a quartz crystal monitor also located in the vacuum chamber.[39] The oscillation frequency of the crystal changes with the thickness of the coating applied to it, and so can be used in real time to measure indirectly the amount of material deposited on the substrate.

We chose to investigate Ti and W as possible materials for the slits as these two metals can be etched in fluorine-based plasmas. Glass slides underwent e-beam evaporation PVD with either titanium or tungsten to a thickness of 50nm or 100nm. These slides were then tested for opacity with 532nm laser light. Results are displayed in Figure 4.3.

The 100nm Ti coating was chosen as it offered superior opacity; however, the thicker the coating, the more difficult it would be to etch.

4.2 Electron beam lithography

Electron beam lithography (EBL) is a process that uses an electron beam to draw tiny patterns on a layer of resist, which changes its chemical composition upon being exposed to the beam.[40] The e-beam resist may be positive (exposed areas removed when developed) or negative (unexposed areas removed when developed). Due to the small size of the slits compared to the surrounding area, a positive resist was chosen: ZEP520 resist.[41] This resist offers high resolution and good resistance to plasma etching, which is discussed in the next section. The resist was dropped onto the glass slide with the titanium side up. After spinning for

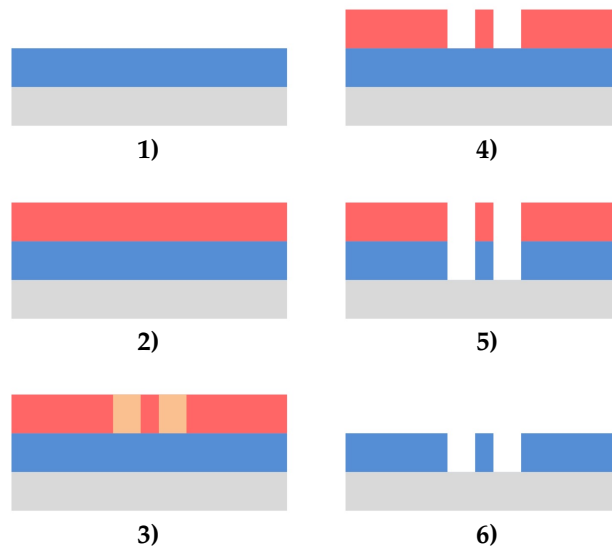


Figure 4.2: Schematic overview of fabrication process (microscope slide viewed side on): **1)** one side of a glass microscope slide (grey) was coated with titanium (blue), **2)** the titanium was topped with a layer of resist (red), **3)** electron beam lithography (area targeted by electron beam in orange), **4)** the resist was developed, **5)** the titanium was etched using inductively coupled plasma etching, **6)** the remaining resist was removed and the sample was ready for use.

Material	Thickness (nm)	Light transmitted
W	50	4.41%
W	100	0.08%
Ti	50	1.79%
Ti	100	0.05%

Figure 4.3: Opacity of different coating options

30 seconds at 3000rpm, the slide was then baked at 180°C for three minutes. This process of adding resist, spinning, and baking was repeated twice more until the desired resist thickness of >1000nm was reached. The rectangular geometry of the slide meant that the resist layer had uneven thickness after spinning, which gave rainbow thin film interference patterns when the slide was viewed from above.

The pattern used for the EBL was five repeats of the scheme of apertures shown in Figure 4.1. Each repetition of this pattern formed a column of five writing fields on the slide.² A gap of 5mm was left between writing fields in the same column. Each column was separated from the next by at least 15mm. The electron energy used was 20keV. After EBL, the slide was placed in ZEP developer,[41] which dissolved the resist that had been exposed to the beam. The slide was then ready for dry etching.

4.3 Inductively coupled plasma etching

Reactive ion etching (RIE) is a dry etching method that uses plasma ions to etch materials. The chemical reactions between the many species of ions and radicals in the plasma and the substrate create volatile compounds that fly away from the surface of the substrate.[42] **Inductively-coupled plasma** (ICP) etching is a specific type of RIE capable of etching very anisotropic patterns (patterns dominated by deep trenches into substrate with narrow openings).[43] A diagram of an ICP etching set-up is shown in Figure 4.4.

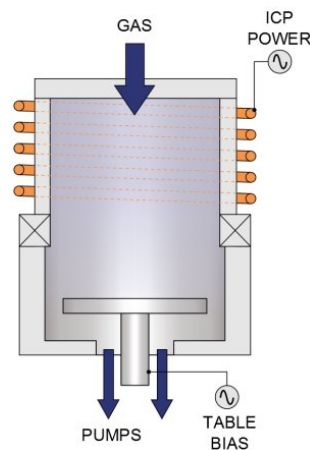


Figure 4.4: Inductively coupled plasma etching set-up showing coils in orange [44]

In an ICP etching system, a coil carrying a radio frequency current envelops the chamber, creating a magnetic field. The wall of the chamber is a dielectric material, often quartz, ensuring good coupling of the magnetic field to the inside of the chamber.[42] By Faraday's law, charged particles inside the reactor will experience

²A writing field is the maximum area the EBL can 'draw' on without moving the sample.

an electric field. The resulting current nulls the magnetic field. ICP etching is a low pressure (<10mTorr) process with particles mostly vertically incident on the sample. Applying a bias voltage (DC or low frequency AC) to the sample table increases the energy of ions in the plasma, which in turn increases the energy of radicals bombarding the substrate and the rate of the etching process.[43] Plasma density may be increased by adjusting parameters such as the ICP power and/or the magnetic field with multipole magnets to further confine electrons.[42, 45] The combination of high ICP power (high density), low pressure, and a moderate radio frequency table bias ensures highly directional or anisotropic etching can be achieved.[43] The formation of compounds on the surface of the substrate is generally to be avoided as this is the opposite of etching, but controlled deposition of polymers may enhance etching in the case of anisotropic patterns.[45] The sample table is helium-cooled to slow the degradation of the resist, and also to minimise sample damage by ion bombardment; however, longer etching times are always accompanied by greater sample damage.

The aim of the process scientist is to find a combination of initial resist thickness, chemistry, etch time, ICP power and table bias that maximises the **selectivity** and anisotropy of the etch. The pattern used (Figure 4.1) is isotropic as far as RIE is concerned, as its narrowest openings are $1\mu\text{m}$ wide with a depth of 100nm; therefore, the focus was on selectivity. Selectivity refers to preferential etching of one material over another – the resist should obviously resist the etch, while the substrate material should be eaten away more quickly. At this point in time, there is no way of predicting which process will succeed other than prior experience; hence, extensive tests of ICP etching comparing the etch rate of various thicknesses of ZEP resist with titanium were undertaken. ZEP resist-coated glass is clearly distinguishable from plain glass by eye; similarly, ZEP resist-coated titanium can be easily contrasted with titanium, due to the dramatic decrease in reflectivity introduced by the ZEP resist.

Ellipsometry was used to determine quantitatively the amount of ZEP left on the slides. This technique measures the change in amplitude and phase for p- and s-components of polarisation after a reference beam is reflected off the surface.[46] Various parameters can be estimated through modelling of the data, including the thickness of thin films on the surface. The resist was considered to have successfully ‘resisted’ the etching process if there was >100nm of resist left on the surface after etching. For two glass slides (one coated with 100nm of Ti, the other with three layers of resist [total resist thickness >700nm]) the process found to completely remove the Ti – while leaving sufficient resist – was:

- four etches of 100s duration, then one etch of 20s duration, separated by 30s “rests” to allow cooling of the samples
- with 10 mTorr pressure,
- 400 W ICP power,
- 25 W radio frequency bias,

- 183 V DC bias,
- and 40 sccm (standard cubic centimetres per minute) SF₆.

The ICP system used was an Oxford Instruments Plasmalab 100. After EBL and ICP etching, the final devices were cleaned by soaking in ZEP remover[41] followed by a final O₂ plasma clean in a barrel etcher. The devices were then cut to size and identified with a label A - E.

4.4 Scanning electron microscopy

Scanning electron microscopy (SEM) uses a focused beam of electrons to scan a sample. Electrons are produced by an electron gun with energies ranging from hundreds to tens of thousand electron volts, then focused with magnetic lenses and directed at the sample with scanning coils. The formation of an electron beam for SEM is the same as with EBL, but SEM directs the beam in a raster scan rather than a specific pattern. Those electrons that have interacted with the sample, carry information that can be used to form an image. The two primary types of electrons used to create images in SEM are the **secondary electrons** (SE) and the **backscattered electrons** (BSE). The former electrons are produced through inelastic scattering and are of low energy. This means only SEs near to the surface end up at the SE detector, hence SEs are sensitive to edges, roughness and general topography of the sample. As there are more secondary electrons escaping from edges, edges appear brighter in an SE image. For an example, see Figure 4.5. As the name suggests, the latter electrons are backscattered from the original electron beam hitting the sample. These electrons have high energy, of the same order as the original beam. The amount of backscattering depends on the average atomic number of the material the electrons are interacting with, and so BSEs yield information about the composition of the sample. The greater the atomic number, the more backscattering, and the brighter the area will appear on an image. A BSE image has less information about the surface of the sample than an SE image, for the BSEs come from a greater variety of depths within interaction volume of the sample owing to their higher energy.[47] An example of a BSE image is given in Figure 4.6. This is the same field of view as in Figure 4.5.

The purpose of conducting SEM on the fabricated devices was to accurately measure the dimensions of the slits etched. Whilst the slits have length as well as width and separation, accurate measurements of only the last two parameters are of chief interest to the project. The choice was made to use BSE images to measure the width and the separation of the slits. Titanium, the material surrounding the slits, is sufficiently different in average atomic number to the material of the slits, silica, to provide adequate contrast in a BSE image. If the slits are positioned so that they are vertical in the image, horizontal intensity profiles provide cross-sections, enabling the distances between titanium boundaries to be measured.

Three devices were imaged in the SEM, and labelled samples A, B, and C. Pairs

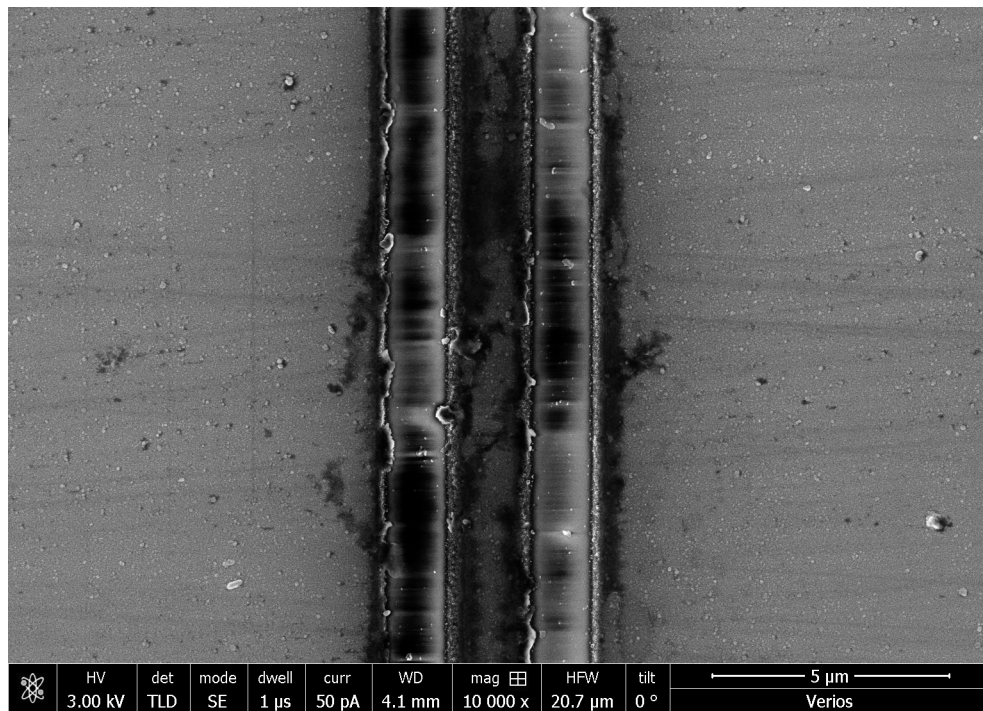


Figure 4.5: Secondary electron image of slits on sample B separated by $\approx 3\mu\text{m}$

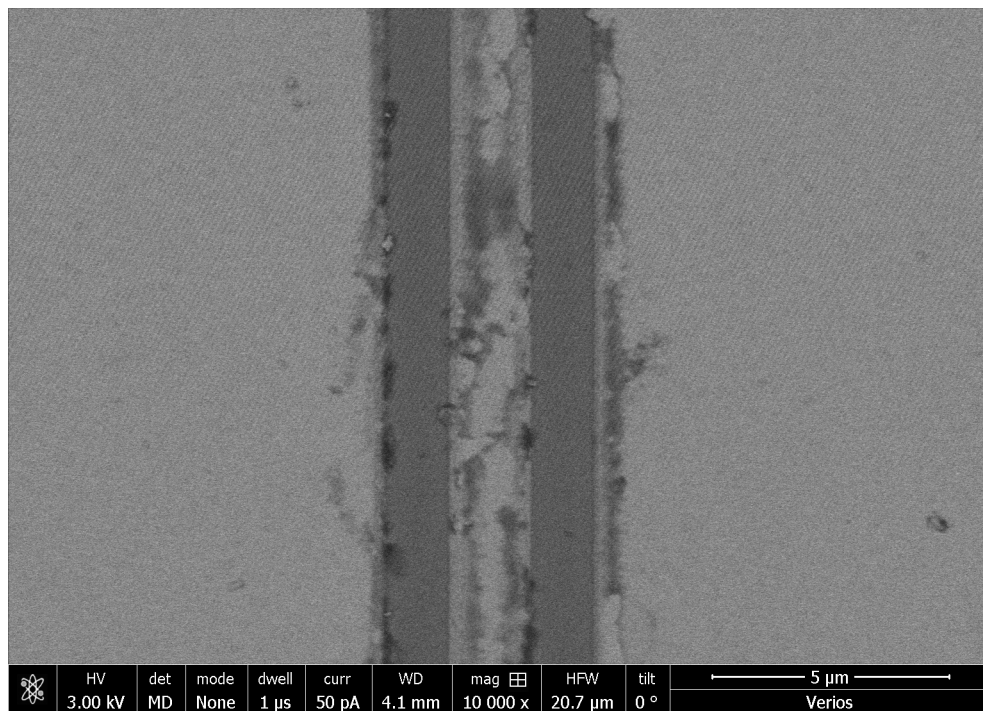


Figure 4.6: Backscattered electron image of slits on sample B separated by $\approx 3\mu\text{m}$

of slits created using an EBL pattern describing a separation of $2\mu\text{m}$ were then indexed with a 2, and so on for other separation values ('A3' denotes a pair of

slits on device A that has separation $\approx 3\mu\text{m}$). Between these samples there was considerable variation in the appearance of the slits under the microscope even if they had resulted from the same EBL pattern. Figure 4.7 shows some of this variation for three pairs of slits, one from each of samples A, B, and C, but all with the same separation according to the EBL pattern used.

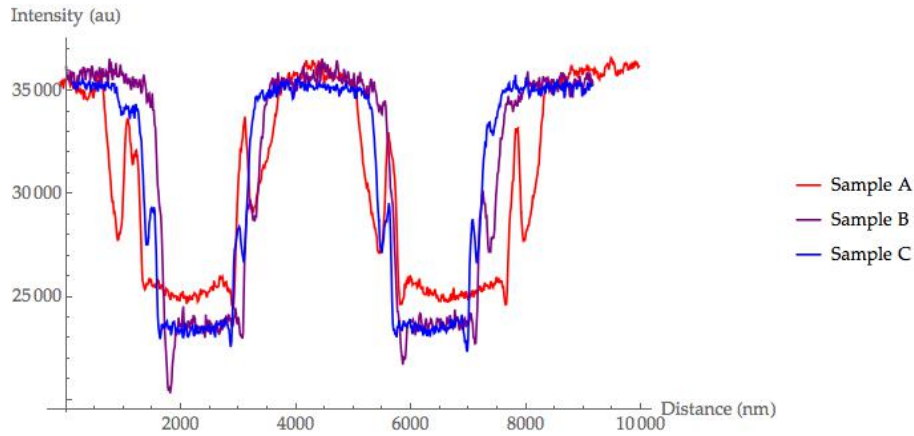


Figure 4.7: Intensity profiles from SEM images of pairs of slits of the same separation but different samples (A, B, and C). The EBL pattern was for a separation of $4\mu\text{m}$.

Pairs of slits of different separations on device B are compared in Figure 4.8. Clear steps in the separation can be seen from $\approx 2\mu\text{m}$ to $\approx 5\mu\text{m}$, but there are differences in the width of each slit between pairs of slits, and even within a pair of slits. It should be noted that the intensity profiles presented in Figure 4.8 are a selection illustrating the presence of these differences, and many more cross-sections from different locations were analysed for a given image in order to measure the dimensions of the slits. Each cross-section featured here consists of integrated intensity over a $3\mu\text{m}$ length of the slits. High resolution BSE images that featured the greater part of an entire pair of slits were used for this purpose (a couple of examples are seen in Figures 4.11, 4.12). For one particular pair of slits (B5), three of the ten profiles, which were collected across much of the length of the slits, are shown in Figure 4.9. Little variation in either separation or width of the slits is seen. In general, the width of slits is regular along the length of a pair, but irregularities were observed including ragged edges, partially-etched stripes next to slits, and islands of titanium in the middle of slits. A combination of by-eye qualitative inspection of BSE images and quantitative results of intensity profile measurements were used to rank the samples in terms of average slit quality. Only one pair of slits was written off, as it was badly damaged by a scratch invisible to the naked eye.

The first step in measuring the width and separation of slits from a set of intensity data was to find the average intensity for titanium I_{Ti} (high intensity) and exposed glass I_{glass} (low intensity) areas. The edge of a slit could then be defined by an intensity I_{edge}

$$I_{\text{edge}} = I_{\text{Ti}} - 0.85 (I_{\text{Ti}} - I_{\text{glass}}) \quad (4.1)$$

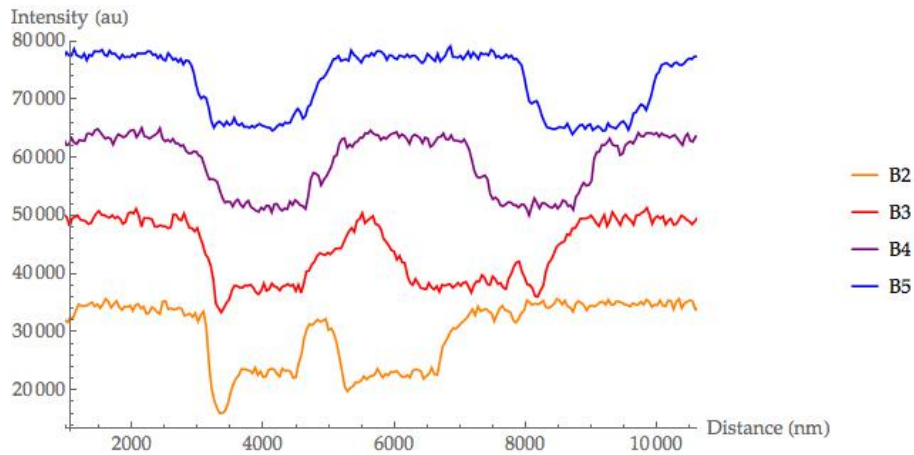


Figure 4.8: Intensity profiles from SEM images of pairs of slits on sample B of different separations. The EBL patterns were for separations of $2\mu\text{m}$ (orange), $3\mu\text{m}$ (red), $4\mu\text{m}$ (purple), and $5\mu\text{m}$ (blue). Intensity profiles have been translated vertically to allow for easier comparison of pairs of slits and do not necessarily reflect the true contrast attained.

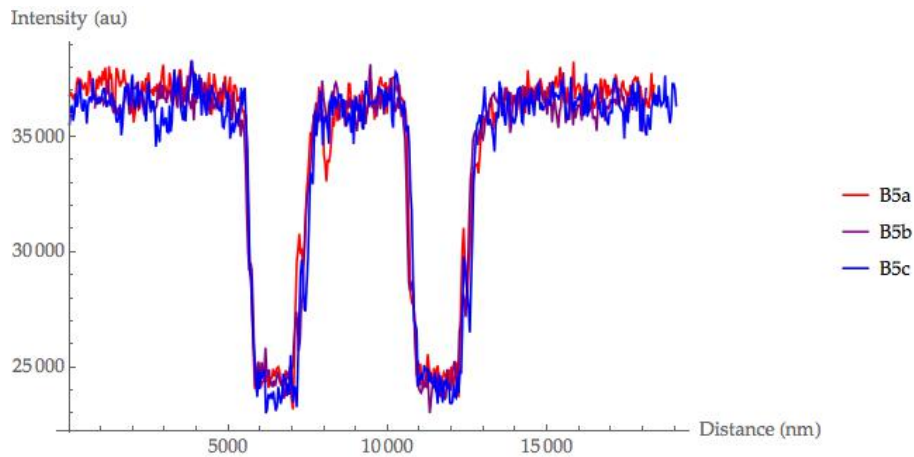


Figure 4.9: Intensity profiles from a SEM image of a pair of slits on the same sample B. The EBL pattern was for a separation of $5\mu\text{m}$.

The position of an edge was interpolated from the intensity data. Labelling the left and right edges of slits 1 and 2 by $L_1, R_1, L_2,$ and R_2 , the width of each slit w_1, w_2 , and the separation d are easily found:

$$\begin{aligned}
 w_1 &= R_1 - L_1 \\
 w_2 &= R_2 - L_2 \\
 d &= \left(\frac{L_2 + R_2}{2} \right) - \left(\frac{L_1 + R_1}{2} \right)
 \end{aligned} \tag{4.2}$$

For each pair of slits, this process was conducted ten times using ten intensity cross-sections from different locations along the slits. The means of w_1, w_2 , and d were found and the uncertainty taken as the standard deviation. Device B was found to have the most precise widths and separations overall, which was reinforced by the

nice clean look of the slits in the SEM images. The results for device B are tabulated in Figure 4.10.

Pair of slits	Width slit 1 (μm)	Width slit 2 (μm)	Separation (μm)
B2	1.38 ± 0.08	1.49 ± 0.05	2.09 ± 0.02
B3	1.45 ± 0.02	1.74 ± 0.09	3.2 ± 0.2
B4	1.38 ± 0.06	1.28 ± 0.06	4.11 ± 0.06
B5	1.36 ± 0.09	1.42 ± 0.06	5.14 ± 0.09

Figure 4.10: Measurements of slit width and separation from device B

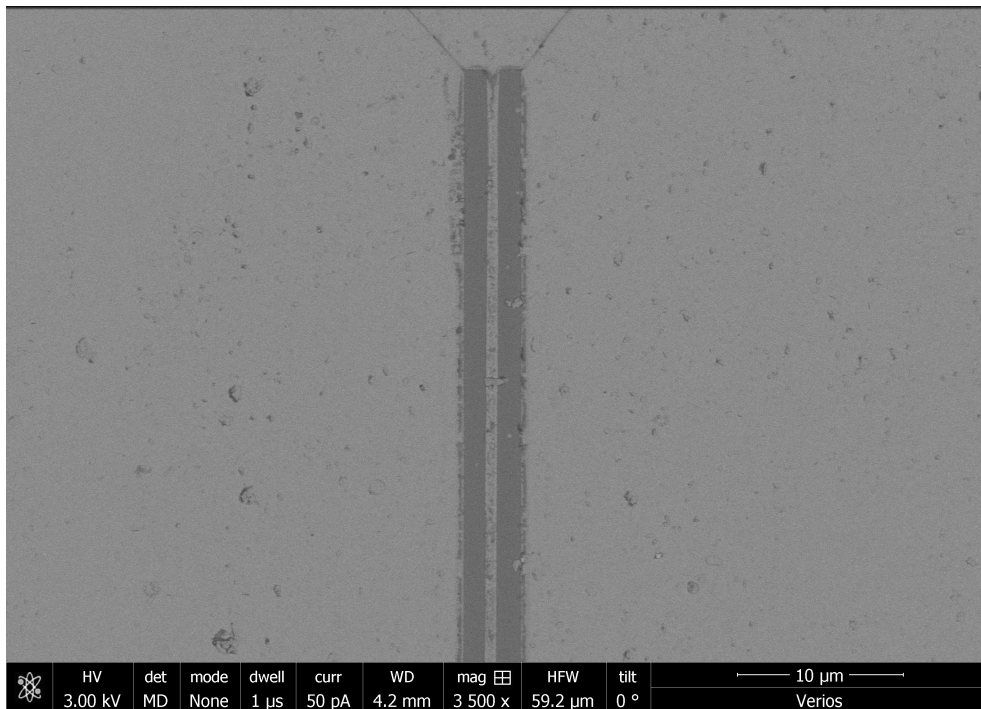


Figure 4.11: Backscattered electron image of slits on sample B with separation $\approx 2\mu\text{m}$.

4.5 Discussion of fabrication

The choice of resist and fabrication process were dictated by the preliminary study we conducted to determine the exact conditions that would lead to successful fabrication. Optical lithography was considered in the early brainstorming stages of discussion, but the desired slit dimensions are at the very limit of resolution for this technique ($1\mu\text{m}$). Extensive testing of dry etching of Ti along with ZEP resist established the required thickness for the resist ($1\mu\text{m}$) and the ICP etch conditions (duration, RF and DC biases, ICP power).

The fabrication process chronicled above successfully resulted in the production of five devices made of titanium and glass, each with four pairs of slits with

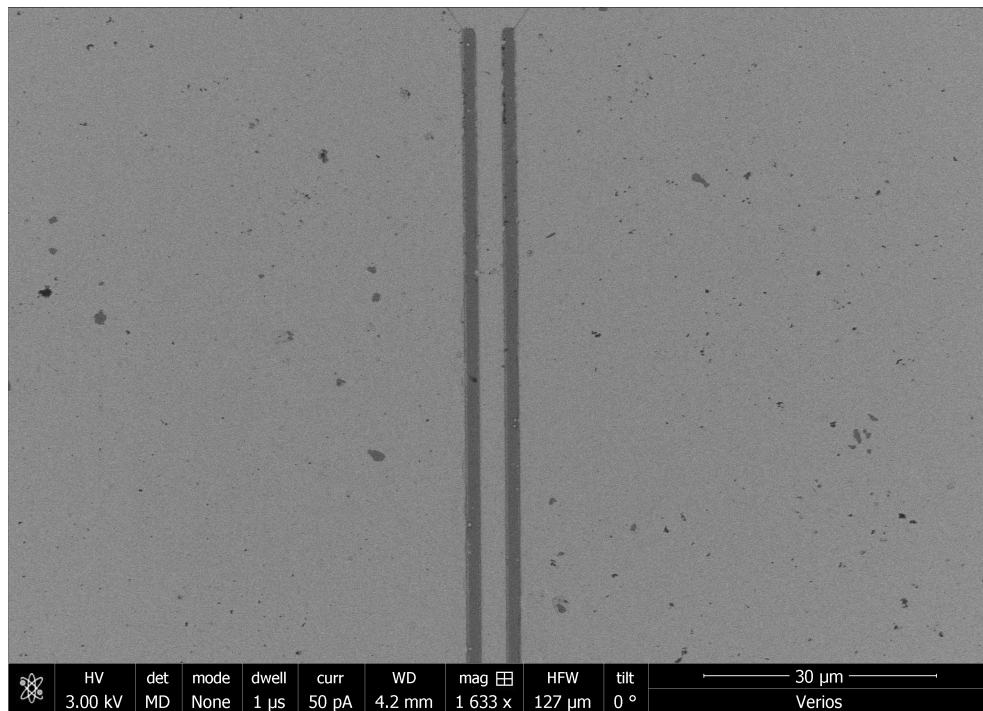


Figure 4.12: Backscattered electron image of slits on sample B separated by $\approx 5\mu\text{m}$

separations $2\text{-}5\mu\text{m}$ and one grating. The grating provided a test of the process for smaller spacings of slits. The grating consisted of a sequence of slits repeated six times, with a gap of $3\mu\text{m}$ between each repetition. The sequence was seven slits of $1\mu\text{m}$ width and various separations of $3, 2, 1.9, 1.8, 1.7, 1.6,$ and $1.5\mu\text{m}$. Inspection of the BSE image presented in Figure 4.13 and the associated intensity data, some of which is shown in Figure 4.14, reveals that at best only the first two slits in each sequence were etched with some accuracy. The remaining slits were etched but the titanium material separating was also removed. Here we discuss possible explanations for why this happened, and suggest future issues and modifications of the process for smaller slits and separations.

While sufficient selectivity was achieved, superior selectivity would likely result if there was a reduction in the thickness of the titanium to be etched. On the other hand, decreasing the thickness of the resist will enable the etching of higher resolution patterns. ICP etching can accomplish the transferral of EBL patterns into substrate with dimensions of 20nm or less, but with just tens of nanometres of resist. An alternative technology is to use an intermediate layer or mask between the titanium and the resist. Possible materials that could be trialled include SiO and TiO_2 . [48] The pattern would be etched first into the intermediate material, and then in a second process, etched into the substrate. By suitably adjusting parameters of the ICP system, selectivity would be optimised for each process, transferring finer patterns into the substrate while only using one layer of ZEP resist.

A key consideration when performing EBL to imprint patterns of small scale is the

proximity effect. As the electron beam is aimed at specific areas of resist, some electrons will travel through the resist and interact directly with the substrate. If backscattering occurs at the boundary between the substrate and the resist, electrons may interact with areas of the resist that would otherwise remain unexposed.[49] In the case of two slits with small separation, the much narrower strip of resist between the slits would especially suffer this effect due to the material either side being exposed. The significance of the proximity effect would be studied through dose testing to find the optimal exposure time for each pixel of the pattern. In addition to adjusting the exposure, further corrections may be made including ghost background exposure, multiple layers of resist and pattern modification.[49]

Acknowledgements

I worked with several accomplished scientists over several months on the fabrication side of this project. Dr Fouad Karouta oversaw the construction of the microslits from brainstorming to final inspection, and taught me all I know about clean rooms and etching. Dr Naeem Shahid conducted the e-beam evaporation, the ellipsometry, and assisted greatly with the EBL. I would also like to thank Dr Mark Lockrey for his SEM skills and tutelage.

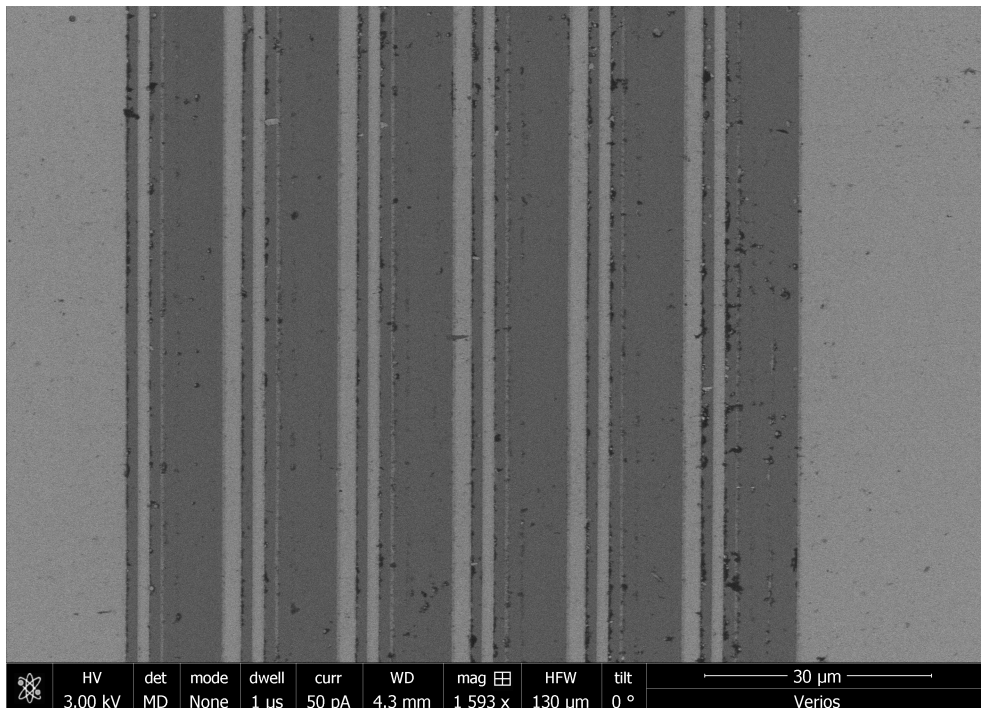


Figure 4.13: Backscattered electron image of grating

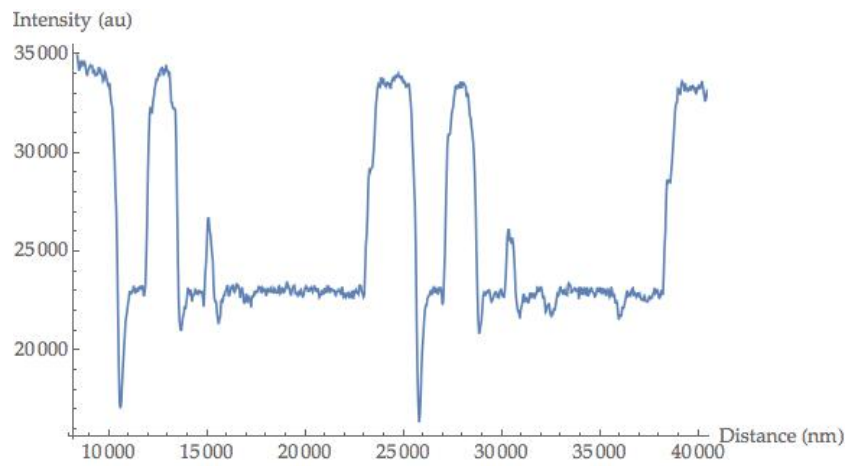


Figure 4.14: Part of an intensity cross-section of the grating.

Chapter 5

The quest for image inversion interferometry

*Now we barely falter / Now the long drive has begun /
North to the ocean / Hotter than the sun /
Do you see what I see? Do you see what I see?
Do you see what I see? Do you see what I see?*

Do You See What I See? Hunters & Collectors

Brandishing my newly fabricated microslits and steeling myself with knowledge of quantum Fisher information, I quixotically entered the quantum optics laboratory, determined to resolve things once and for all. However, we focused on the imaging journey rather than the destination of super-resolution. In this chapter, practical developments in the area of image-inversion interferometry are reported. We move from using a coherent source to incoherent illumination, with tangible gains in visibility being made. Basic image analysis is explored through *Mathematica* software.

5.1 The image inversion interferometer

The optical set-up begins with the light sources used. The experiment is constructed for easy switching between a coherent source (the laser), and an incoherent source (the LED) of the same colour. This is enabled by the positioning of a flip mirror. The intensity of each source can be adjusted: the LED has a knob for this purpose, while the intensity of the laser transmitted through the polarising beam splitter (PBS) depends on the angle of the $\lambda/2$ waveplate. The beam is then steered onto two converging lenses. The first lens has a focal length of 75mm; the second lens has a focal length of 50mm. The choice of focal lengths for the lenses and the distance separating them (125mm) is such that the beam narrows to a waist between the lenses, and emerges collimated. A collimated laser beam is essential for alignment

of the image-inversion interferometer. The beam then encounters a 50-50 beam splitter. The transmitted beam is inverted using two planar mirrors (all mirrors are highly reflective for broadband 532nm).

The reflected beam is bounced between two mirrors on a translation stage and a third mirror, the position of which is controlled by a piezoelectric device. The piezoelectric mount is connected to a signal generator, which provides a 3V peak-to-peak triangular wave at an adjustable frequency, via a high voltage amplifier.

The two beams are combined at a second 50-50 beam splitter and sent to a photodetector or a charged coupled device (CCD) at each port. The photodetector was used for initial alignment of the interferometer as it was easy to gain real-time intensity data that varied with the scanning of the piezoelectric device. The visibility of the fringes could be gauged using an oscilloscope, and positions of various optical devices optimised. Once the first stage of alignment of the interferometer was complete, a pinhole of $10\mu\text{m}$ diameter was placed at the waist between the collimating lenses, i.e. 75mm after the first lens, and 50mm before the second. The pinhole was mounted on a 3D translation stage, offering precise positioning along the beam, as well as in the horizontal and vertical directions. Each translation stage was operated manually via turning a micrometer of $0.005\mu\text{m}$ precision for each direction. The 3D translation stage was also used for positioning of the pairs of slits.

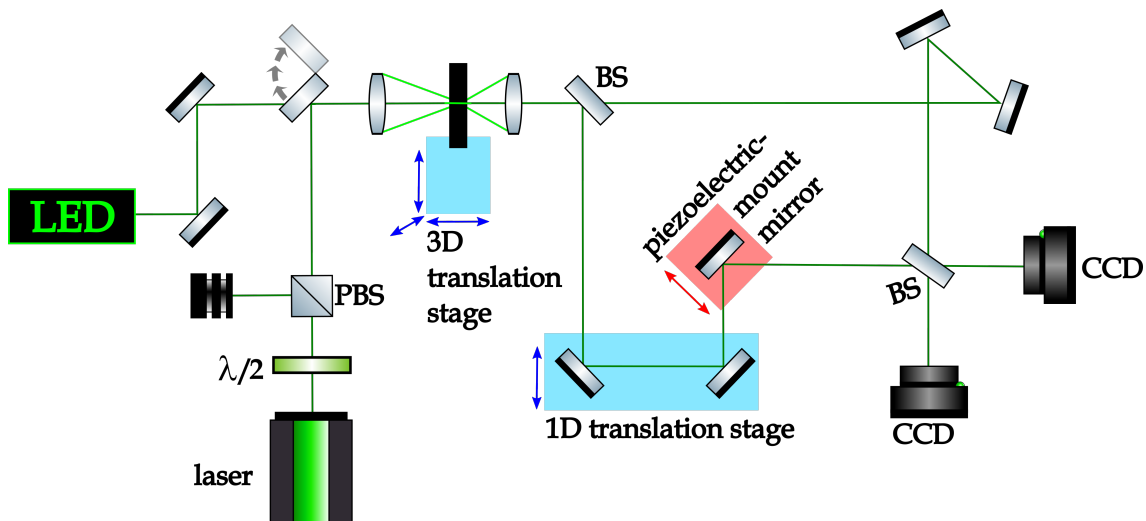


Figure 5.1: Experimental set-up for image inversion interferometry. LED = light-emitting diode, BS = beam splitter, PBS = polarising beam splitter, CCD = charged coupled device.

5.2 Interference with coherent light

The coherent light source used was a Nd:YAG laser (Innolight Mephisto) with a frequency doubling crystal (Innolight Diablo) with maximum power 1W. For quality alignment, a much lower power (2mW) was used to avoid saturating any detectors.

The resulting continuous wave emission was at 532nm with a coherence length of $\approx 300\text{m}$ (estimated from the known linewidth of the laser); hence, interference with coherent light may be observed over a broad range of path length differences between the arms of the interferometer. The fringes are far more sensitive to the positioning of each beam into the final beam splitter. If the beams emerge aligned in the near- and far-fields, then interference is seen. Both beam splitters were checked for even transmission and reflection of 532nm light by placing a power meter in each arm. With no pinhole, the maximum visibility was over 98%. A combination of a photodetector, an oscilloscope edge triggered by the signal generator sending the triangular wave to the piezoelectric device, and a 50Hz scan on the piezoelectric device allowed monitoring of the temporal profile of the fringes. The 50Hz frequency was high enough for the oscilloscope to be triggered constantly. The peak-to-peak voltage across the piezoelectric device was adjusted so that the fringes appeared as close to sinusoidally varying in time without clipping.

The $10\mu\text{m}$ pinhole acted as a spatial filter to ensure the resulting beam had a Gaussian profile, while also reducing the power into the interferometer. Some adjustment of the pinhole vertically or horizontally output higher TEM modes, which were attractive to look at although not very applicable to the project. With the pinhole in place, the output intensity was low enough to use CCDs for detection without any filters. CCDs have the advantage of providing a spatial profile of the output light. When well aligned, the fringes appeared as concentric rings, with those at the centre appearing and disappearing with the scanning of the piezoelectric device. A more sensitive photodetector and amplifier were also used to observe fringes with the laser passing through the pinhole. Fine alignment through the shape of the fringes could be confirmed by incremental increases in the visibility calculated from the amplitude of the fringes on the oscilloscope. Raw interference data collected from the CCD for one coherent point source ($10\mu\text{m}$ pinhole at the waist) and a 1Hz piezoelectric scanning frequency is plotted in Figure 5.2. With the CCD having a maximum frame rate of 15 frames per second, 1Hz was about the fastest scanning frequency that resulted in interference being easily observed in real time. No correction for any ambient light has been made, as the alignment of the interferometer for the laser beam was simply intended to provide a launchpad for subsequent alignment with incoherent light. Autocontrast on the CCD was turned off. However, a discrete Fourier transform was performed on the data to profile the frequency distribution of the interference. The first component of the Fourier series for this data had a frequency of 1.02Hz, which can be attributed confidently to the scanning of the mirror attached to the piezoelectric crystal. There was evidence for other frequencies, but these components had lower amplitudes.

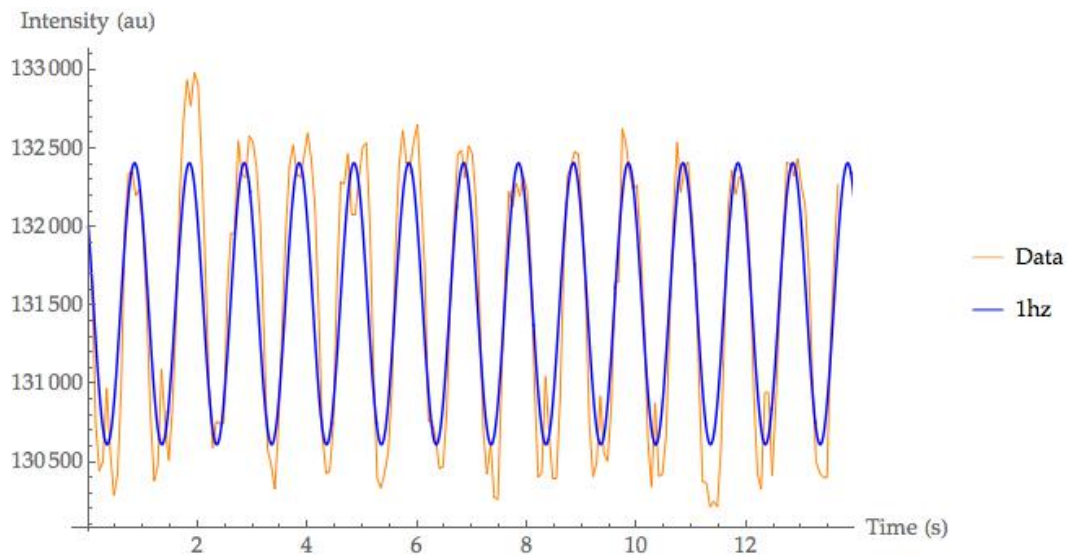


Figure 5.2: Interference fringes obtained using a coherent source and the piezoelectric mirror scanning frequency set to 1Hz. Intensity data from the CCD is shown in orange, while the first Fourier series term is shown in blue ($h=1.02\text{Hz}$).

5.3 Interference with incoherent light

Building an interferometer for use with incoherent light requires orders of magnitude more refinement.¹ The path lengths of the two arms must be identical if interference is to occur, owing to the very short coherence length of the light. Given the geometry of our interferometer deviated from a typical ‘square’ Mach-Zehnder interferometer, this was especially challenging. The placement of optics on the bench was revised so that every component stood on a gridline of the table, including the mirrors constituting the inversion device (minor movements of these caused clipping of the beam). The two mirrors on the one-dimensional translation stage were realigned precisely so that a beam incident on the first mirror was parallel to the beam reflecting off the second mirror. In this way, movement of the translation stage across its maximum range of 2.5cm minimally affected subsequent alignment of the interferometer.

As an intermediate step between laser and the LED, a sodium lamp was used. The emission spectrum of sodium has an orange doublet: a pair at 588.9950nm and 589.5924nm, with the line at 588.9950nm having twice the intensity of the other.[50] The coherence length for the sodium lamp is approximately 0.6mm, which is six (three) orders of magnitude smaller (larger) than that of the laser (LED). Since the sodium lamp is very weak compared to the laser, no pinhole was used, and the lamp positioned immediately prior to the first beam splitter. The extremely low intensity of the sodium lamp motivated the paper-cardboard-stick-tape construction of several ‘houses’, ‘tents’, and ‘tunnels’ for the optics to exclude

¹During the course of this project, I rebuilt the interferometer at least half a dozen times over: to make it smaller, to make it squarer, to make it darker, to make it cleaner...

the majority of stray light in the dark laboratory. Of course, with all the optics on the table being geared for green light, the visibility of any fringes was bound to be limited. Nevertheless some fringes were observed. The interference from the single sodium source formed a pattern of vertical fringes within a larger vertical strip (see Figure 5.3). Varying the path length of the non-inverting arm of the interferometer by more than 0.6mm caused the interference to be lost. The integrated intensity across the CCD is plotted as a function of time in Figure 5.4. The visibility of the raw data is 3.4%. A Fourier transform of the data returned the most prominent five frequency components within the range 4.9 - 5.1Hz, which matches the strong beating seen in Figure 5.4. The limitations of the sodium lamp as a source for this experiment meant we did not look into this further.

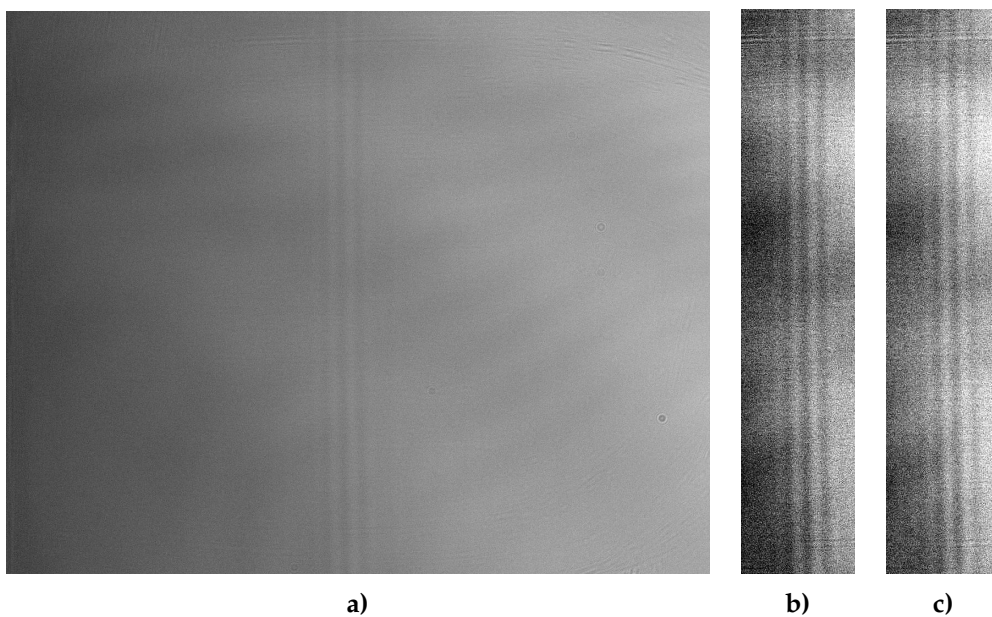


Figure 5.3: Sodium lamp fringes: *a)* example of a raw CCD image, *b)* minimum integrated intensity (cropped, adjusted contrast), *c)* maximum integrated intensity (cropped, adjusted contrast)

The LED obtained for this project was a SCOPELED G-series 250T tabletop microscope illuminator with a maximum radiant power of 2.350W and a coherence length of $10\mu\text{m}$ (it changes colour perceptibly with power).² The LED was always tuned to the lowest power setting while switched on. The light from the LED was steered with a series of mirrors so that it followed the same path as the laser beyond the flipped mirror. The centre of the circular LED beam, which had a diameter of $\approx 0.5\text{cm}$ at the waist of the two lenses, was positioned at the pinhole. By slowly moving the mirrors on the translation stage, and pausing often to observe the CCD live video, faint fringes were seen just 0.04mm away from the calculated position of the translation stage to equalise the path lengths. The interference was so slight

²532nm safety goggles were needed when the LED was uncovered and going at full pelt, and even these only reduced its power by a fraction. The mean green LED developed quite a reputation. Don't mess with it.

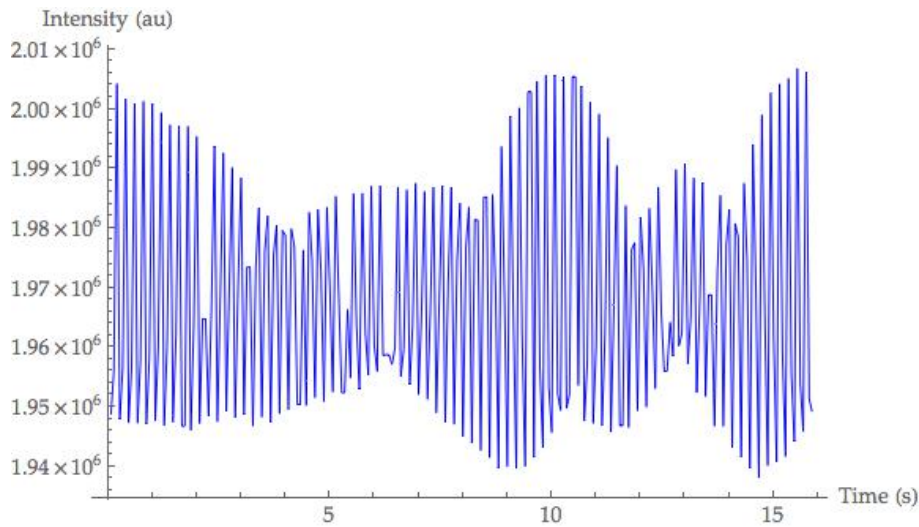


Figure 5.4: Interference fringes obtained using a sodium lamp source (piezoelectric scanning frequency of 1Hz).

that distinguishing whether it was a pattern of rings or vertical strips was nigh impossible by eye. Adjusting the translation stage by $4\mu\text{m}$ (a path length difference of $\approx 8\mu\text{m}$) was sufficient to ruin the interference. Examples of dark and light fringes are depicted in Figure 5.5. Scanning the moving mirror at different frequencies revealed that this positioning of the interferometer was likely outputting a signal of inversion interference: the frequency of the fringes was proportional to the frequency of the scan. The LED data, along with primary Fourier series terms, are plotted in Figures 5.6, 5.7, and 5.8. Assuming the relationship between the frequency of the scan f , and the main frequency of the interference h is of the form $h = \kappa f$, then $\kappa \approx 0.253 \pm 0.002$. Many more measurements are required, but this is the most promising signal found to date. Note that this is for one incoherent source (the pinhole), not two: interference using the fabricated slits as point sources has not yet been observed. Laser light incident upon device B has yielded the relative positions of the pairs of slits B2-B5 in the (x, y, z) co-ordinates of the three-dimensional translation stage. The iconic double slit interference patterns are indeed distinctive and very localised features (see Appendix for pictures).

5.4 Image processing

In addition to honing our fringe-spotting skills in the laboratory, preliminary investigations into image processing using CCD data have been conducted. Two methods relatively easy to implement in *Mathematica* are cropping the images to select the areas where interference is most noticeable, and increasing the contrast of the images. The code used appears in the Appendix. The preliminary results of applying these techniques to data collected for the sodium lamp and LED sources are summarised in Figure 5.9. All of the Na lamp images were cropped to the same

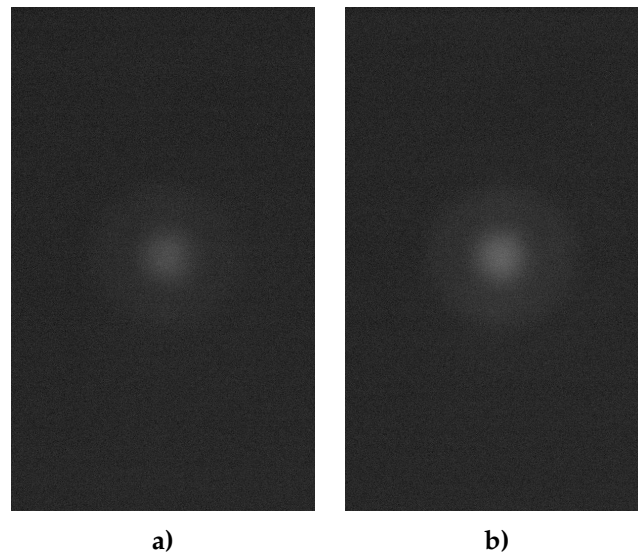


Figure 5.5: LED fringes: *a)* minimum integrated intensity, *b)* maximum integrated intensity

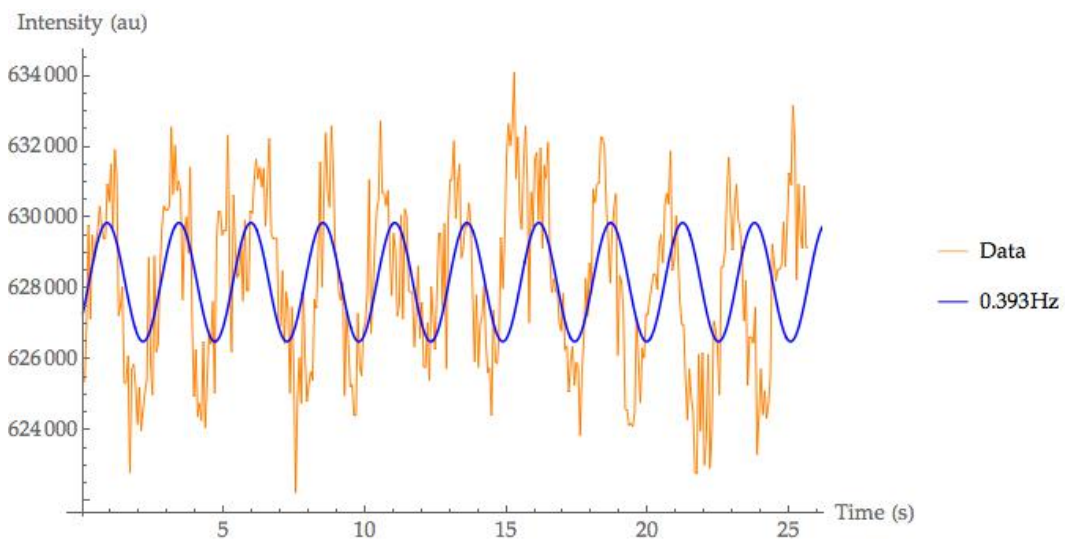


Figure 5.6: Interference fringes obtained using an LED and the piezoelectric mirror scanning frequency set to 100mHz. Intensity data from the CCD is shown in orange, while the first Fourier series term is shown in blue ($h = 0.393\text{Hz}$).

range of pixels corresponding to a set of vertical strips (for examples of cropped and contrast-boostered Na lamp images, see Figure 5.3). The images using LED illumination were cropped to a different size to cover all visible ‘rings’ and the full vertical extent of the CCD (for examples of cropped LED images, see Figure 5.5). The reason for not reducing the size of the LED images in the vertical direction is the uncertainty as to the presence of vertical, stripe-like interference.

The large increase in visibility after increasing the contrast is artificial, particularly in the case of the LED: frames of minimum and maximum intensity are indistinguishable by eye, and comparison of the vertical scales in the plots of the “enhanced”

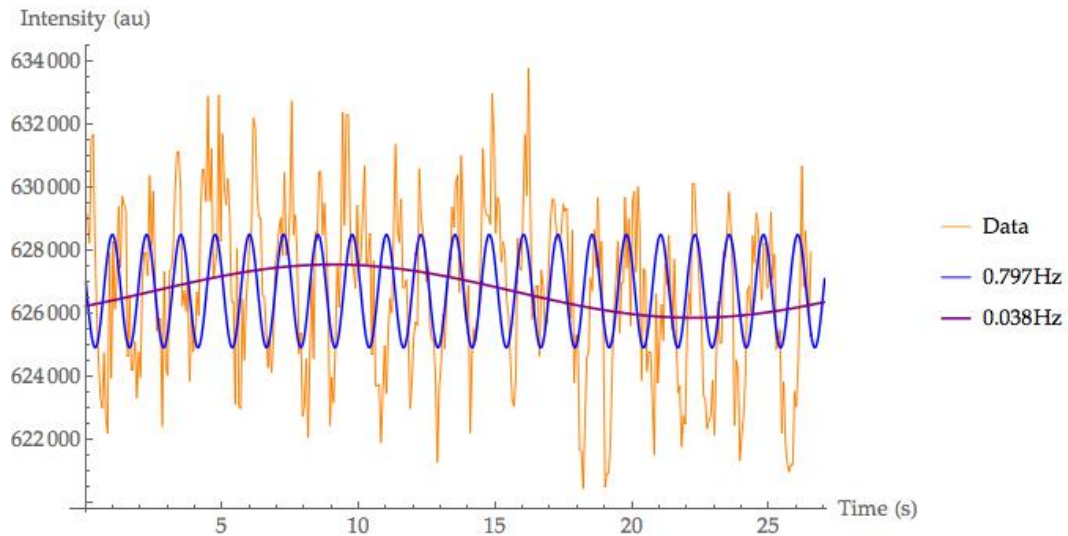


Figure 5.7: Interference fringes obtained using an LED and the piezoelectric mirror scanning frequency set to 200mHz. Intensity data from the CCD is shown in orange, while the first Fourier series term is shown in blue ($h= 0.797\text{Hz}$). A low frequency Fourier term ($h = 0.038\text{Hz}$) is also shown in purple; however, this time series is too short to measure such low frequencies with any accuracy.

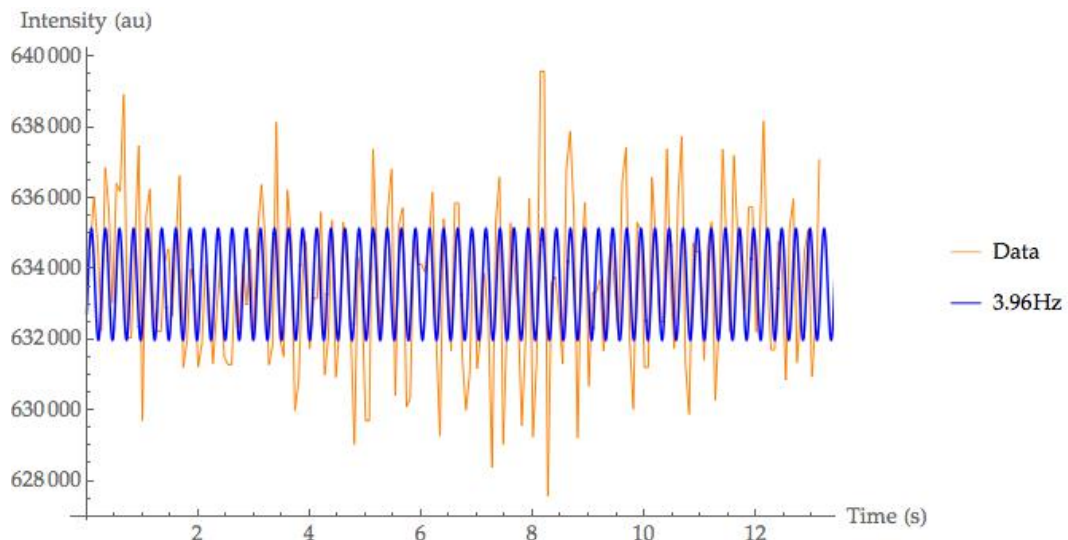


Figure 5.8: Interference fringes obtained using an LED and the piezoelectric mirror scanning frequency set to 1Hz. Intensity data from the CCD is shown in orange, while the first Fourier series term is shown in blue ($h= 3.96\text{Hz}$).

fringes (Figure 5.11) and the “cropped” fringes indicates that much information has been potentially discarded. If we knew which frequencies we were looking for in the data, then tailored contrast enhancement may be an option to extract a better signal from the noise. At this stage we lack the *a priori* understanding to undertake more complicated processing of image data. This is not conclusive analysis by any means. In the results presented here, the background intensity and any noise

Processing technique	Na lamp \mathfrak{V} (%)	LED \mathfrak{V} (%)
None	3.4	2.1
Cropped images	3.6	3.1
Cropping and contrast adjustment	14.9	76.4

Figure 5.9: Preliminary results showing enhancement of fringe visibility through processing image data in Mathematica

arising from the CCD itself is completely ignored. Background measurements of longer duration would provide data suitable for detecting low-frequency noise as well as ascertaining the magnitude of the noise floor.

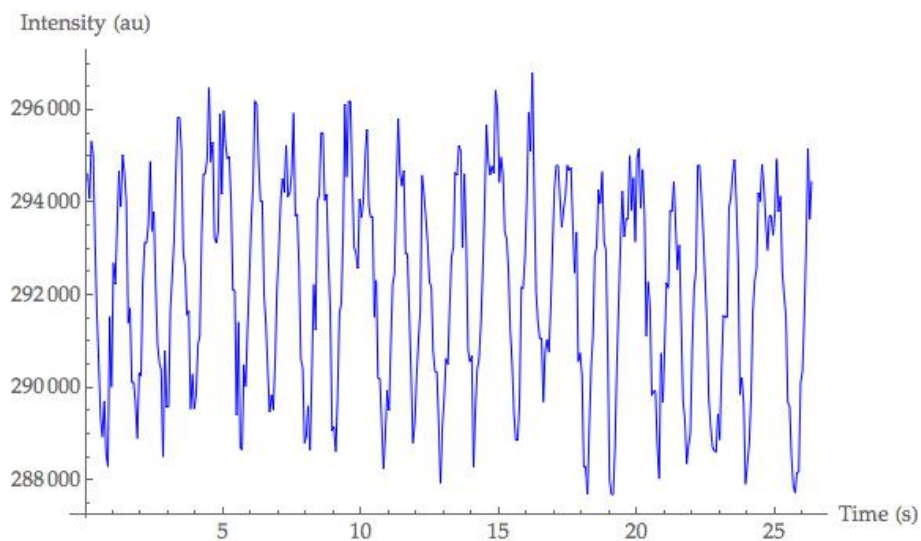


Figure 5.10: Interference fringes obtained using an LED (select area of CCD)

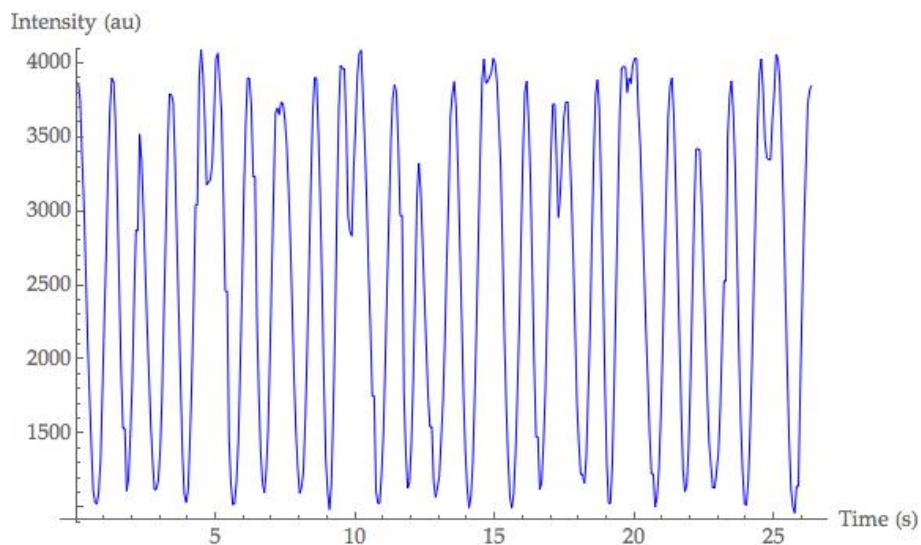


Figure 5.11: Post-processing enhanced interference fringes obtained using an LED

Chapter 6

Conclusions and future aims

Super-resolution through quantum estimation theory is an exciting area of research that is fresh on the scene of quantum metrology. The super-resolution of incoherent thermal sources has not yet been achieved using interferometry. This project took a first step towards realising the construction of a SLIVER.

We now have incoherent point-like sources of light to which the SLIVER technique can be applied. The creation of pairs of slits with dimensions in microns required the perfection of a new recipe for the plasma etching of titanium-coated glass. The development of this empirical process is a useful addition to the body of knowledge in the field of micro- and nano-fabrication. When combined with a powerful LED, the slits will enable a high-quality performance test of SLIVER, given that the dimensions of the slits have been accurately measured through scanning electron microscopy. Investigating alternative fabrication technologies may yield much narrower slits with separations of tens of nanometres or enable the etching of finer structures such as pinholes.

An image inversion interferometry is an instrument of relatively simple design that has much potential to enhance imaging processes at optical wavelengths. However, any path difference between the arms of the interferometer had to be within $8\mu\text{m}$ for interference to be observed using green incoherent light. Building an interferometer of this precision was an extended iterative process of alignment. We found an interference signal for one incoherent source, which will be further optimised in the near future. Additional measurements will no doubt 'shed some light' on the presence of noise in the detectors, the background illumination, and why the frequency of the signal for the incoherent light does not correspond to the frequency of the scan across a small range of path lengths. A more thorough characterisation of the interference will enable the development of appropriate post-processing methods for the image data. A goal for the immediate future is to utilise the fabricated slits with incoherent illumination, to analyse the output of both ports, and to estimate the separation parameter using these measurements.

The power of statistical quantities such as the Cramér-Rao lower bound to overturn historical limits such as Abbe's diffraction limit is being realised, with several theoretical proposals appearing over the past year aiming to improve existing techniques in imaging and fluorescence microscopy. As yet unpublished theory now places a bound on the minimum Fisher information that may be obtained through SLIVER for thermal sources of any power, broadening the potential reach of the technique to astronomical interferometry. Specific extensions of the practical SLIVER set-up include spatially-resolved detection and image inversion in higher dimensions.

The race belongs not to the swift, but to those who keep on running.

Bibliography

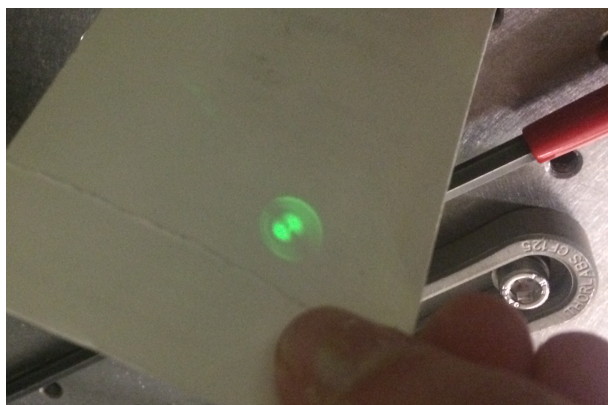
- [1] N. Wolchover. How far can the human eye see? [Online]. Available: <http://www.livescience.com/33895-human-eye.html>
- [2] What is electron microscopy? [Online]. Available: https://www.jic.ac.uk/microscopy/intro_EM.html
- [3] Nobel Prize in Chemistry 2014 press release. [Online]. Available: https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2014/press.html
- [4] E. Betzig *et al.*, “Imaging intracellular fluorescent proteins at nanometer resolution,” *Science*, vol. 313, 2006.
- [5] T. A. Klar *et al.*, “Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission,” *PNAS*, vol. 97, pp. 8206–8210, 2000.
- [6] R. Nair and M. Tsang, “Interferometric superlocalization of two incoherent optical point sources,” *Opt. Express*, vol. 24, no. 4, pp. 3684–3701, 2016.
- [7] NASA/JPL/Gemini Observatory/AURA/NSF. Wide-field Infrared Survey Explorer (WISE) multimedia. [Online]. Available: http://www.nasa.gov/mission_pages/WISE/multimedia/pia16872.html
- [8] R. N. Bracewell, “Detecting nonsolar planets by spinning infrared interferometer,” *Nature*, vol. 274, 1978.
- [9] W. E. Moerner, “The story of single molecules, from early spectroscopy in solids, to superresolution nanoscopy in cells and beyond,” RSPE Director’s Colloquium, 9th February 2016.
- [10] R. Togneri. (2005) Estimation theory for engineers. [Online]. Available: http://people.duke.edu/~hpgavin/SystemID/References/Estimation_Theory.pdf
- [11] C. W. Helstrom, *Quantum Detection and Estimation Theory*. Academic Press, 1976.
- [12] A. Ly *et al.* A Tutorial on Fisher Information. [Online]. Available: <http://www.ejwagenmakers.com/submitted/LyEtAlTutorial.pdf>
- [13] C. J. Geyer. Stat 5102 Notes: Fisher Information and Confidence Intervals Using Maximum Likelihood. [Online]. Available: <http://www.stat.umn.edu/geyer/old03/5102/notes/fish.pdf>

- [14] S. Zheng. Math 541: Statistical Theory II – Fisher Information and Cramér-Rao Bound. [Online]. Available: http://people.missouristate.edu/songfengzheng/Teaching/MTH541/Lecture%20notes/Fisher_info.pdf
- [15] “The Density Matrix , 5.74 Introductory Quantum Mechanics II,” <http://ocw.mit.edu>, 2009.
- [16] Quantum measurement theory. [Online]. Available: http://www.quantum.umb.edu/Jacobs/QMT/QMT_Chapter1.pdf
- [17] A. Fujiwara, “Strong consistency and asymptotic efficiency for adaptive quantum estimation problems,” *J. Phys. A: Math. Gen.*, vol. 39, 2006.
- [18] J. Fiurasek. Maximum-likelihood estimation of quantum measurement. [Online]. Available: <http://cds.cern.ch/record/482663/files/0101027.pdf>
- [19] M. Tsang, R. Nair, and X. Lu, “Quantum theory of superresolution for two incoherent optical point sources,” 2015, arXiv:1511.00552v1.
- [20] J. W. Goodman, *Statistical Optics*. John Wiley & Sons, 1985.
- [21] C. C. Davis, “Optics of Gaussian Beams,” <http://www.ece.umd.edu/davis/chapter16.pdf>.
- [22] Modes. [Online]. Available: <https://www.rp-photonics.com/modes.html>
- [23] R. Gallager, “Poisson Processes, 6.262 Discrete Stochastic Processes,” <http://ocw.mit.edu>, 2011.
- [24] J. H. Shapiro, “Continuous-Time Photodetection, 6.453 Quantum Optical Communication,” <http://ocw.mit.edu>, 2008.
- [25] S. Ferraro. Physical Optics and Diffraction. [Online]. Available: <http://www.astro.princeton.edu/~gk/AST542/simone.pdf>
- [26] Super-resolution microscopy tutorial. [Online]. Available: <http://advanced-microscopy.utah.edu/education/super-res/>
- [27] B. R. Masters. Lord Rayleigh: A scientific life. [Online]. Available: http://www.osa-opn.org/home/articles/volume_20/issue_6/features/lord_rayleigh_a_scientific_life/
- [28] S. Ram, E. S. Ward, and R. J. Ober, “Beyond Rayleigh’s criterion: A resolution measure with application to single-molecule microscopy,” *PNAS*, vol. 103, no. 12, 2006.
- [29] C. W. Helstrom, “Resolution of point sources of light as analyzed by quantum detection theory,” *IEEE Trans. Inf. Theory.*, vol. 19, no. 4, 1973.
- [30] W. Tham, H. Ferretti, and A. M. Steinberg, “Beating Rayleigh’s curse by imaging using phase information,” 2016, arXiv:1606.02666v2.
- [31] F. Yang *et al.*, “Far-field linear optical superresolution via heterodyne detection in a higher-order local oscillator mode,” 2016, arXiv:1606.02662.

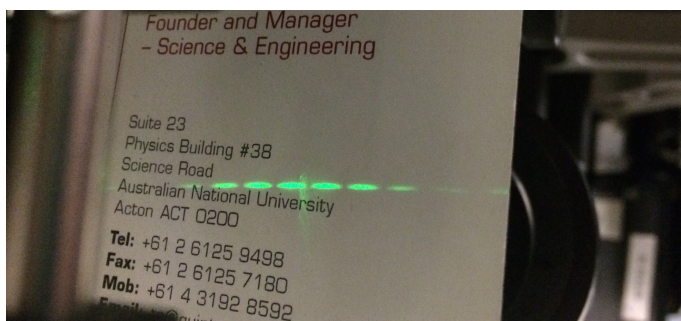
- [32] M. Paur *et al.*, "Achieving the ultimate optical resolution," *Optica*, vol. 3, no. 10, 2016.
- [33] R. Nair and M. Tsang, "Far-field Super-resolution of Thermal Electromagnetic Sources at the Quantum Limit," 2016, arXiv:1604.00937v2.
- [34] X. Lu, R. Nair, and M. Tsang, "Quantum-optimal detection of one-versus-two incoherent sources with arbitrary separation," 2016, arXiv:1609.03025v1.
- [35] S. Z. Ang, R. Nair, and M. Tsang, "Quantum limit for two-dimensional resolution of two incoherent optical point sources," 2016, arXiv:1606.00603v1.
- [36] J. Chao, E. S. Ward, and R. J. Ober, "Fisher information theory for parameter estimation in single molecule microscopy: tutorial," *J. Opt. Soc. Am. A*, vol. 33, no. 7, 2016.
- [37] K. Wicker and R. Heintzmann, "Interferometric resolution improvement for confocal microscopes," *Opt. Express*, vol. 15, no. 19, 2007.
- [38] Z. S. Tang, K. Durak, and A. Ling, "Fault-tolerant and finite-error localization for point emitters within the diffraction limit," 2016, arXiv:1605.07297.
- [39] E. Chen. Characteristics of Thin Film Deposition. [Online]. Available: <http://www.mrsec.harvard.edu/education/ap298r2004/Erli%20chenFabrication%20II%20-%20Deposition-1.pdf>
- [40] M. Altissimo, "Imaging intracellular fluorescent proteins at nanometer resolution," *Biomicrofluids*, vol. 4, 2010.
- [41] ZEP520A Technical Report. [Online]. Available: <https://www.zeonchemicals.com/pdfs/ZEP520A.pdf>
- [42] F. Karouta, "A practical approach to reactive ion etching," *J. Phys. D*, vol. 47, 2014.
- [43] P. Verdonck. Plasma Etching. [Online]. Available: <http://wcam.engr.wisc.edu/Public/Reference/PlasmaEtch/Plasma%20paper.pdf>
- [44] Oxford Instruments Inductively Coupled Plasma (ICP) Etching. [Online]. Available: <https://www.oxford-instruments.com/products/etching-deposition-and-growth/plasma-etch-deposition/icp-etch>
- [45] L. Fuller. Plasma Etching. [Online]. Available: https://people.rit.edu/lffee/Plasma_Etch.pdf
- [46] What is ellipsometry? [Online]. Available: <https://www.jawoollam.com/resources/ellipsometry-tutorial/what-is-ellipsometry>
- [47] AMMRF. MyScope SEM. [Online]. Available: <http://www.ammrf.org.au/myscope/sem/introduction/>
- [48] E. R. Parker *et al.*, "Inductively Coupled Plasma Etching of Bulk Titanium for MEMS Applications," *J. Electrochem. Soc.*, vol. 152, no. 10, 2005.

- [49] A. van de Kraats and R. Murali. Proximity Effect in E-beam Lithography. [Online]. Available: <http://nanolithography.gatech.edu/proximity.htm>
- [50] Sodium spectrum. [Online]. Available: <http://hyperphysics.phy-astr.gsu.edu/hbase/phyopt/na.html>

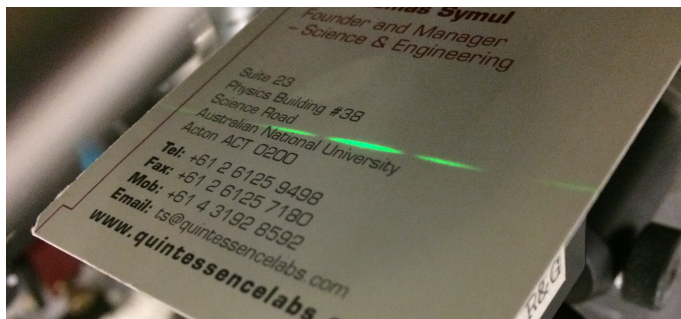
Appendix



Pretty modes through the pinhole!



Double slit interference!



More double slit interference!

(*FOURIER TRANSFORM TEMPLATE NOTEBOOK*)

(*FUNCTIONS FOR DAYSSSSS*)

```
Clear[findPeriod];
findPeriod[data_, threshold_] :=
Module[{fs, s1, s = {}, i, a0f, af, pf, pos, fr, frpos, fdata, fdatac, n, per},
n = Length[data];
fs = Fourier[data];
s1 = Drop[fs, -Floor[Length[fs]/2]];
For[i = 1, i < Length[s1], i++,
If[Abs[fs][[i + 1]] > threshold, AppendTo[s, i + 1]]];
a0f = Abs[fs][[1]]/Sqrt[n];
af = 2/Sqrt[n] Abs[fs][[s]];
pf = Arg[fs][[s]];
{a0f, Transpose[{s, af, pf}]}]

Clear[reconstruct];
reconstruct[data_, fp_] := Module[{n, n = Length[data];
Show[ListLinePlot[data, PlotStyle -> Orange],
Plot[fp[[1]] + Sum[fp[[2, j, 2]] Cos[2 Pi (fp[[2, j, 1]] - 1)/n t - fp[[2, j, 3]]],
{j, 1, Length[fp[[2]]}], {t, 0, n}, PlotStyle -> Purple]]]

specific[data_, fp_, freq_, amp_, ph_] := Module[{n, n = Length[data];
Show[ListLinePlot[data, PlotStyle -> Orange],
Plot[fp[[1]] + amp Cos[2 Pi (freq - 1)/n t - ph], {t, 0, n}, PlotStyle -> Purple]]]

tlb = Import["/Users/siobhantobin/Library/Mobile
Documents/com-apple~CloudDocs/project/siobhanD/led200mhz3vpp.avi",
"ImageList"];

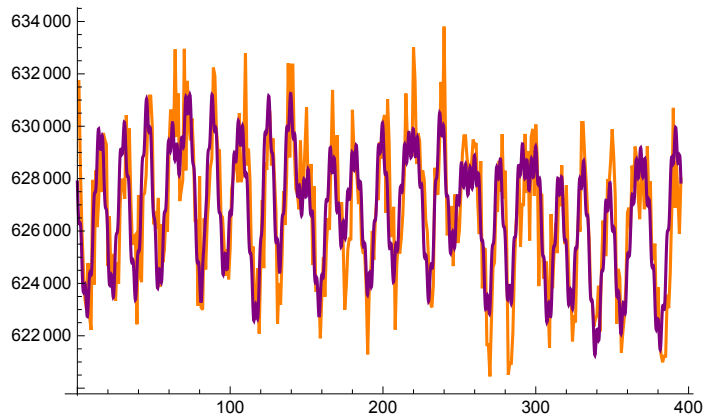
yeahlb = findPeriod[h1b, 3500]
```

```
{626734., {{2, 846.573, 2.16272}, {3, 627.977, 1.79809},
{6, 432.794, -1.1629}, {8, 369.893, 2.19693}, {11, 590.082, 2.80782},
{12, 598.093, 0.0507368}, {13, 452.624, 0.536648}, {17, 638.105, -0.663192},
{18, 417.605, -0.993301}, {19, 741.027, -0.93036}, {22, 1791.87, -2.14066},
{26, 845.406, -1.14466}, {27, 1121.19, -0.289729}, {30, 855.572, 1.90168},
{32, 538.35, -1.73435}, {33, 402.587, 1.59794}, {139, 517.181, -1.35093}}}
```

```
TableForm[Sort[yeahlb[[2]], #1[[2]] > #2[[2]] &],
TableHeadings -> {None, {"Freq", "Amp", "Phase"}}]
```

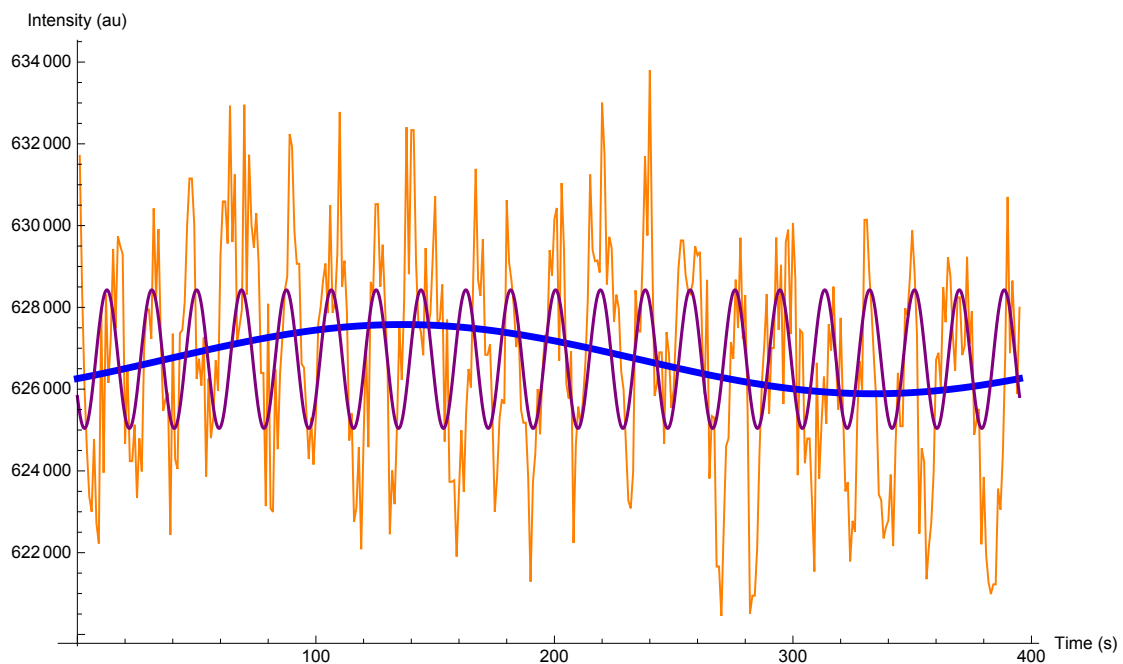
Freq	Amp	Phase
22	1791.87	-2.14066
27	1121.19	-0.289729
30	855.572	1.90168
2	846.573	2.16272
26	845.406	-1.14466
19	741.027	-0.93036
17	638.105	-0.663192
3	627.977	1.79809
12	598.093	0.0507368
11	590.082	2.80782
32	538.35	-1.73435
139	517.181	-1.35093
13	452.624	0.536648
6	432.794	-1.1629
18	417.605	-0.993301
33	402.587	1.59794
8	369.893	2.19693

```
reconstruct[h1b, yeah1b]
```



```
specific32[data_, fp_, freq_, amp_, ph_, freq2_, amp2_, ph2_] :=
  Show[ListLinePlot[data, PlotStyle -> {Orange, Thickness[0.002]},
    AxesLabel -> {"Time (s)", "Intensity (au)"}, ImageSize -> Large],
    Plot[{fp[[1]] + amp Cos[2 Pi (freq - 1) / 395 t - ph],
      fp[[1]] + amp 2 Cos[2 Pi (freq2 - 1) / 395 t - ph2]},
      {t, 0, 395}, PlotStyle -> {{Blue, Thickness[0.007]}, Purple}]]
```

```
specific32[h1b, yeah1b, 2, 846.573, 2.1672,
  22, 1791.8679561099423`, -2.1406561438701837`]
```



```
nilb = Table[ImageAdjust[t1b[[i]], 1], {i, 1, 395}];
```

```
y2b = Table[{i/15, Total[Flatten[
  ImageData[ImageTake[nilb[[i]], {1, 1000}, {400, 1000}]]]}], {i, 1, 395}];
```

```
y3b = Table[{i/15, Total[Flatten[ImageData[nilb[[i]]]]]}], {i, 1, 395}];
```

```
y4b = Table[{i/15, Total[Flatten[
  ImageData[ImageTake[t1b[[i]], {1, 1000}, {400, 1000}]]]}], {i, 1, 395}];
```

```
ListLinePlot[y4b, AxesLabel → {"Time (s)", "Intensity (au)"},  
PlotStyle → {Blue, Thickness[0.002]}, ImageSize → Large,  
LabelStyle → {FontFamily → "Palatino", FontSize → 14}]
```

