Applications of Scattering-Type Scanning Near-Field Optical Microscopy in the Infrared

Edward Yoxall

October 2013

Experimental Solid State Group
Department of Physics
Imperial College London
United Kingdom

Thesis submitted for the degree of Doctor of Philosophy of Imperial College London and for the Diploma of the Imperial College

Imperial College London

This page has intentionally been left blank.

Abstract

This thesis is split into two broad sections. These are defined by the various applications of scattering-type near-field optical microscopy (s-SNOM) in different parts of the electromagnetic spectrum; the near-infrared ($700-1000$nm) and the mid-infrared ($6-10\mu$m). S-SNOM is a means of imaging surfaces at resolutions well below the diffraction limit - the level of recorded detail does not depend on the wavelength of light (as it does with traditional optical microscopy), but instead on the sharpness of a probe (usually ~ 10nm), meaning an image resolution approaching $\lambda/1000$ in the mid-infrared.

For the work presented in the near-infrared, the focus lies with the modelling and mapping of various plasmonic resonances supported by metallic nanostructures. These resonances have the ability to "squeeze" light into substantially sub-wavelength volumes which is useful for a variety of applications ranging from cancer treatments to molecular sensing.

The mid-infrared section starts with the implementation of a pulsed quantum cascade laser (QCL) as the system's light source. This presents some instrumentation challenges as all s-SNOM imaging to date has been conducted with continuous-wave (CW) lasers. Using a pulsed laser also raises some significant signal-to-noise implications which are quantified and discussed. In terms of the experimental applications of such a setup, the first steps towards ultra-high resolution infrared chemical spectroscopy are made by studying the epithelial cells of an oesophageal biopsy.

The thesis concludes with an examination of the major noise sources faced by s-SNOM, and makes a number of recommendations on how their effects can be mitigated.

This page has intentionally been left blank.

Contents

Declarations **8**

Acknowledgements **9**

Acronyms **10**

1 Introduction **11**
1.1 A Brief History of Near-Field Microscopy 13
1.2 Introducing s-SNOM 14
1.3 Other Super-Resolution Techniques 15
1.4 Thesis Outline 16

2 The s-SNOM Setup **18**
2.1 Introduction 18
2.2 Atomic Force Microscopy 18
2.3 Experimental Setup for s-SNOM 20
2.4 Field Enhancement 21
2.5 Scattering and Suppression: Theoretical Background 21
2.5.1 The Scattering Coefficient 21
2.5.2 Background Suppression 22
2.5.3 Approach Curves 23
2.6 Pseudo-Heterodyne Detection 24
2.6.1 Modelling Pseudo-Heterodyne Detection 25
2.7 Alternative s-SNOM Setups 26

3 The s-SNOM Signal **29**
3.1 Introduction 29
3.2 Fourier Optics 29
3.2.1 The Origin of Evanescence 30
3.2.2 Evanescent Waves and Resolution 31
3.2.3 Detection of Evanescent Waves 31
3.3 Signal Contrast 32
3.3.1 Plasmonic Samples 33

3.3.2 Chemical Mapping . . . 35
3.4 Modelling the Probe-Sample Interaction . . . 35
3.4.1 The Point-Dipole Model . . . 35
3.4.2 The Monopole Model . . . 37
3.4.3 Including the Far-Field Contribution . . . 39
3.4.4 Different Materials, Different Signals . . . 40
3.4.5 Near-Field Spectroscopy . . . 44
3.5 Noise in Measurements . . . 44
3.5.1 Signal, Background and Noise . . . 44
3.5.2 Defining the SNR for s-SNOM Measurements . . . 48
3.5.3 Increasing the SNR . . . 48

4 Plasmonics in the Near-Infrared **50**
4.1 Introduction . . . 50
4.2 Experimental Setup for Plasmonic Samples . . . 50
4.3 Gold Colloids . . . 53
4.3.1 Modelling and Far-Field Spectra . . . 53
4.3.2 Near-Field Images of Colloidal Spheres . . . 54
4.3.3 Aside: Interpretation of Plasmonic s-SNOM Images . . . 56
4.4 Retardation Effects . . . 59
4.4.1 Dolmen Structures: A Cautionary Tale . . . 60
4.5 Pentamer Structures . . . 61
4.5.1 Theory: Oligomer Responses . . . 63
4.5.2 Experimental Images and Discussion . . . 65
4.6 Conclusion . . . 69

5 The Use of Pulsed QCL Sources **71**
5.1 Introduction . . . 71
5.2 Quantum Cascade Lasers . . . 72
5.2.1 Lasertune . . . 73
5.3 Pulsed Sources and s-SNOM . . . 73
5.3.1 Synchronisation of Sampling and Laser Pulse . . . 74
5.3.2 Inherent Pulse-To-Pulse Variation . . . 77
5.4 Broadband Antennas . . . 80
5.4.1 Results: Simulation and Experiment . . . 80
5.5 Conclusion . . . 81

6 Cellular Imaging with s-SNOM **83**
6.1 Introduction . . . 83
6.2 Biological Imaging . . . 83
6.2.1 Oesophageal Cancer Sample . . . 85
6.2.2 Far-Field Spectroscopy . . . 85

6.2.3 Calculating the Dielectric Constant . . . 85
6.2.4 Imaging the Sample . . . 88
6.2.5 Discussion of Results . . . 92
6.3 Conclusion . . . 93

7 Towards Low Noise s-SNOM 95
7.1 Introduction . . . 95
7.2 Electronic Sources of Noise . . . 95
7.2.1 Noise Power Spectral Density . . . 96
7.2.2 Equivalent Noise Circuit for Photodetection . . . 97
7.2.3 Shot Noise . . . 98
7.2.4 Thermal Noise . . . 99
7.2.5 Other Noise Sources . . . 99
7.2.6 Summary of Electronic Noise Sources . . . 100
7.3 Laser Noise . . . 101
7.3.1 Relating the Time and Frequency Domains . . . 101
7.3.2 Comparison of Laser Sources . . . 103
7.3.3 Improving the SNR: Attenuating the Reference Beam . . . 104
7.3.4 Additive vs Multiplicative Noise . . . 105
7.4 Adding a Reference Channel . . . 105
7.4.1 Subtraction vs Division . . . 106
7.4.2 Circuit Design for Voltage Division . . . 107
7.4.3 Testing the Divider . . . 108
7.4.4 Using the Voltage Divider . . . 110
7.5 Conclusion . . . 111

8 Conclusion and Further Work 112
8.1 Conclusion . . . 112
8.2 Further Work . . . 113

Declarations

Declaration of Originality

The work contained within this thesis is my own, and all else has been appropriately referenced.

Copyright Declaration

Acknowledgements

First and foremost, I must thank my supervisor **Chris Phillips** for his guidance; he hit the "Goldilocks zone" with his management style - not too much and not too little. Always ready for a chat when needed, I owe him an enormous debt of gratitude for setting me off in my academic life so ably. Secondly, thanks to **Stefan Maier** for his steady hand on the tiller of all things plasmonic, and **Lesley Cohen** for her pastoral help as head of the experimental solid state group.

Also to the rest of Chris's group; particularly **Hemmel Amrania** for making my first few months such a breeze, but also **Frank Murphy, Andy McCrow, Megumi Yoshida, Laurence Drummond, Alex Bak, Mary Gosal and Oliver Curtin**. Thanks for the breakfasts, the help around the lab and for genuinely making it a pleasure to be at work.

For those postdocs with whom I've worked and published; **Yannick Sonnefraud** has gone above and beyond more times than I care to count - it's no exaggeration to say that plasmonics would still be a fuzzy mess for me without him. Also to **Vincenzo Giannini, Mohsen Rahmani and Miguel Navarro-Cia** - thanks for sharing your wisdom, samples and time with me.

I enormously appreciate the input from those people working at Neaspec GmbH in Munich and at CIC NanoGUNE in San Sebastian - without your help in person, by phone and by email this thesis would never have happened; especially **Nenad Ocelic, Rainer Hillenbrand, Antonija Cvitkovic, Andy Huber, Marcus Diem, Florian Huth and Stefan Schiefer**.

Thanks also to the support staff who have made my time at Imperial so smooth - **Carolyn Dale, Bhavna Patel and Juraci Didone** as group administrators, the mechanical maestros **Steve Cussell and Martin Pettifer**, and **Lee Parker**, whose electronics knowledge (and, furthermore, his ability to explain it) is second to none.

Procrastination is something best practised over coffee, and I've had many willing partners in crime - **Yan, Antonio, Kostas, Tariq, Bing, Rob, Roberto, Will, Ben, Emma** and anyone else who has indulged my proclivity for caffeine - life is never so bad after a biscuit with you.

Penultimately, to **my parents** for getting me here in the first place - I couldn't possibly have dreamed of a better start in life - and also to those people who have had to live with me over the past few years - **Dan, Cath, Roz and Dave**. You had an uncanny knack of knowing whether I needed stew, beer or gin.

Finally, to my girlfriend **Nic** who has had to bear the brunt of my frustrations (and successes) over the past three years. You've had my back, kept me frosty and filled my basket more times than I care to remember - the time I've spent working towards this thesis has been infinitely better for the fact that I've shared it with you.

Acronyms Used in This Thesis

Page number is where each is introduced.

Page	Acronym	Standing For:
11	EM	Electron Microscopy
11	SPM	Scanning Probe Microscopy
11	STED	Stimulated Emission Depletion microscopy
11	PALM	Photo Activated Localisation Microscopy
12	STORM	Stochastic Optical Reconstruction Microscopy
12	SNOM	Scanning Near-field Optical Microscopy (aperture SNOM)
13	s-SNOM	Scattering-type SNOM (also apertureless SNOM)
13	SEM	Scanning Electron Microscopy
14	AFM	Atomic Force Microscopy
14	STM	Scanning Tunnelling Microscopy
14	TERS	Tip Enhanced Raman Spectroscopy
14	SERS	Surface Enhanced Raman Spectroscopy
16	QCL	Quantum Cascade Laser
25	nano-FTIR	Nanoscale Fourier Transform Infrared spectroscopy
32	SPP	Surface Plasmon Polariton
33	LSPR	Localised Surface Plasmon Resonance
43	NFBDR	Near-Field to Background and Disturbance Ratio
52	FDTD	Finite-Difference Time-Domain modelling method
63	CDA	Coupled Dipole Approximation
73	ADC	Analogue to Digital Conversion
79	TFSF	Total Field Scattered Field plane wave source
100	RIN	Relative Intensity Noise

Chapter 1

Introduction

"Pick the low-hanging fruit." **Prof. Chris Phillips**

Light has long been our primary means of gathering information about the world around us. Even in the scientific age, optical imaging remains one of the most important techniques in any researcher's toolkit. Its benefits hardly need stating - light is a reliable, cheap and fast method for studying all kinds of samples. One area in which optical microscopy has been usurped, however, is resolution. Feynmann's prophetic lecture of 1959, "There's Plenty of Room Left at the Bottom", provides a useful marker for the start of an explosion in the activity and applications of the nanosciences. Such a growth has been matched by a requirement for truly nanoscale imaging, but the wave nature of light - and its subsequent diffraction - limits the resolution of traditional "eyepiece" microscopy to just a few hundred nanometres in the visible part of the spectrum [1].

The effects of the "diffraction limit" are neatly demonstrated by the example of two point-like sources imaged by an optical microscope (figure 1.1). The image of a point source is not a well defined spot, but rather a series of bright and dark rings known as an Airy disk. When two point sources are located sufficiently far apart (image (a)), the points are easily resolved. As they move closer together, however, they eventually become impossible to tell apart from a single, brighter point source (image (c)). The smallest resolvable distance between two point sources is usually defined by the Rayleigh criterion. For a microscope, this states that the minimum separation distance R for two point sources to be distinguishable is given by:

$$R = 0.61\frac{\lambda}{\mathrm{NA}} \tag{1.1}$$

where λ is the wavelength and NA is the numerical aperture of the system. The level of detail that can be observed, therefore, is dependent on the imaging wavelength.

Although a resolution limit of approximately $\lambda/2$ may not seem particularly limiting for visible light, it is much more problematic when the wavelength grows and one enters the infrared regime. Here, the level of observable detail may only be 10μm - about the same size as a biological cell - and many orders of magnitude larger than the 22nm CMOS integrated circuits used in modern electronic devices. It is useful to be able to image structures with infrared light for its ability to

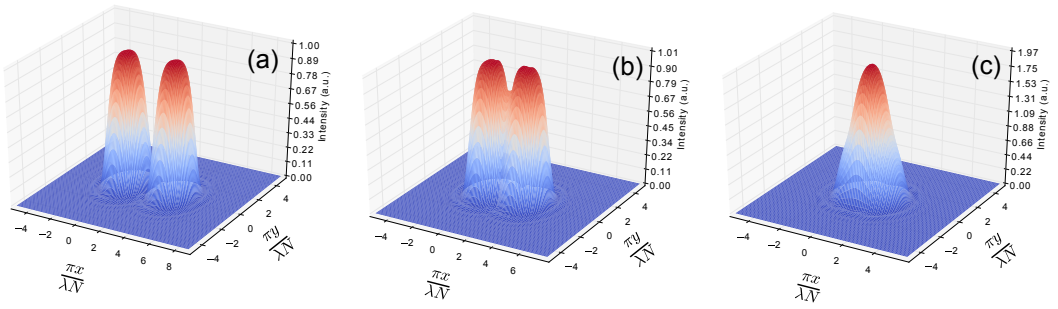

Figure 1.1: Intensity patterns (known as Airy disks) recorded by an imaging system for two point sources when (a) they are well enough separated to be easily resolved, (b) they are just resolvable according to the Rayleigh criterion (see equation 1.1) and (c) they are insufficiently separated and appear as a single, brighter source (note the z-axis scale difference). λ represents the wavelength and N is the f-number of the imaging system.

perform chemical spectroscopy - identifying the constituent molecules of a sample from the way its absorption changes at different wavelengths.

The microscopy gap left by the diffraction limit has, historically, been filled by two broad technological families; electron microscopy (EM) and scanning probe microscopy (SPM). The former uses electrons (with their shorter De Broglie wavelength and thus much smaller diffraction limit) in place of photons as the information carriers, whereas the latter scans the sample's surface with a very sharp probe to gain topographical and other details.

Both have their benefits and drawbacks; with EM, samples must be metallised to prevent overheating due to the resulting electron current and measurements are often required to be carried out in a vacuum, but the technique can reach resolutions of less than 50pm [2]. With SPM, however, the resolution is limited to the size of the probe's apex - typically 10nm (although they can provide atomic scale resolutions [3]) - and images can take a considerable amount of time to gather as the sample must be raster scanned pixel by pixel, line by line. On the other hand, a sample can be non-destructively investigated as the technique need not require anything other than ambient conditions, and a host of other properties - electrical, mechanical or magnetic, for example - can be examined at the same time as the topography [4, 5].

More recently, fluorescence imaging has also grown in stature as a family of techniques capable of resolving sub-diffraction limited details. As with traditional optical microscopy, it is completely non-destructive, but it does require the sample itself to be fluorescent. This is particularly easily achieved in the biological sciences, where samples can be labelled with a fluorescent stain, or be made to express a fluorescent protein. Some methods, such as stimulated emission depletion microscopy (STED), work by structuring the shapes of a series of incident light pulses to selectively excite and then deplete the sample's flurophores, resulting in a sub-diffraction limited area of fluorescence [6, 7]. Another method, known as photoactivated localisation microscopy (PALM), achieves the same effect by activating only a small portion of the sample's flurophores at a given time [8]. This allows the position of each to be precisely located via digital post-processing as the centre of the observed Airy disk. A further closely related technique to PALM is stochastic optical

reconstruction microscopy (STORM), where the random "blinking" on and off of fluorophores is measured [9]. For all of these fluroescence techniques, the resolution is generally around 30nm.

1.1 A Brief History of Near-Field Microscopy

Scanning near-field optical microscopy, or SNOM, falls into the SPM category of sub-diffraction limit imaging techniques. Despite its relatively recent emergence, the basic idea behind SNOM was first suggested by an Irish physicist, Edward Synge, in 1928 [10]. He conceived of a system in which an opaque screen would be placed extremely close to a sample surface. If a hole with dimensions less than the wavelength was placed in the screen, then by raster scanning the hole around the sample, the amount of light which penetrated the hole could be mapped pixel by pixel - this would be a measure of the sample's absorption. Synge's envisioned setup is shown in figure 1.2.

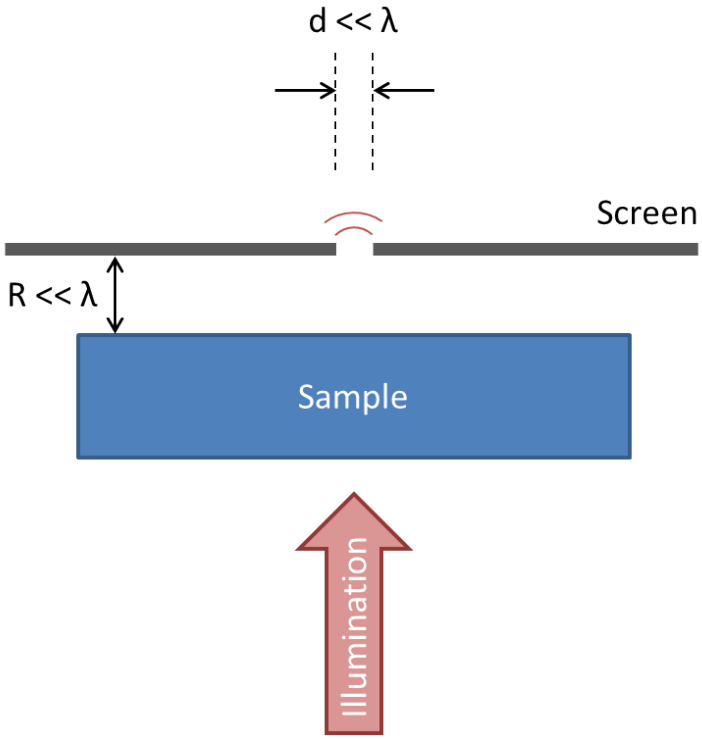

Figure 1.2: Diagram of Synge's setup for sub-diffraction limited optical imaging. A sub-wavelength hole is placed in a screen which is raster scanned very close to a sample surface. The light intensity measured on the other side of the hole (to the illuminating light) is related to the sample's absorption.

Synge's idea relies on the presence of the "near field" (or evanescent waves). This is in contrast to the "far field" (or propagating waves) and is central to sub-diffraction limited optical imaging. Simply put, when light is scattered from a surface or is emitted from a source, those waves which carry the very fine details of the sample or source decay away very rapidly - for visible light this is usually within a few tens of nanometres. Only those waves which encode the rougher details on

length scales $\sim \lambda$ can propagate into the far field. The concept is rooted in Fourier optics and is explained in more detail in section 3.2. The key for sub-diffraction limited imaging, therefore, is to measure the near-field waves before they decay away, and thus to preserve the higher resolution information. In Synge's case (figure 1.2) this was achieved by placing the aperture in the screen very close to the sample surface - the resolution becomes dependent on the size of the hole rather than the light's wavelength.

Interestingly, Synge wrote a letter to Einstein about his design on 22$^{\text{nd}}$ April 1928 in which he states:

> *By means of the method [presented,] the present theoretical limitation of the resolving power in microscopy seems to be completely removed and everything comes to depend on technical perfection.*

In the 1920s, engineers did not possess the technical prowess to put his ideas into practise, and it wasn't until the 1970s that they were first experimentally demonstrated in the microwave regime [11]. A working near-field microscope at optical frequencies had to wait until 1984 at which point it was given the rather prosaic name of "optical stethoscopy" [12]. From these early beginnings grew the field of SNOM (or more specifically "aperture SNOM"), which used a tapered optical fibre in place of Synge's hole, and it has since become a common technique for examining the optical properties of sample surfaces.

1.2 Introducing s-SNOM

Although aperture SNOM allows for significantly sub-diffraction limited imaging, it is still - to some extent - limited by the imaging light's wavelength. As the fibre must be tapered to provide the high resolution, a significant amount of light is backscattered by the taper before it can exit through the aperture. The smaller the hole, the greater the backscatter, so this effectively limits the size of the aperture (and, therefore, the resolution) to 50nm for visible light. For the same reason, the resolution is worse in the infrared part of the spectrum where the backscatter is more pronounced due to the longer wavelength. On top of this, the fibre technology required by aperture SNOM becomes much more challenging in the infrared. Figure 1.3 shows a scanning electron microscope (SEM) image of a typical aperture SNOM probe.

For a truly wavelength independent system, therefore, a new approach was required. Throughout the 1980s, scientists were becoming aware of potential applications for the large electric field enhancements that occur in the vicinity of point-like structures (particularly those scientists within the Raman spectroscopy community [14–17]). It became clear that a sharp probe could act like a nanoscale light source when an external beam was focused at its apex [18, 19], and that this could form the basis of a new type of SNOM, known as apertureless, or scattering-type SNOM (a or s-SNOM). The first demonstration of such a system based on the scattering of light from a sharp tip, rather than the amount of light transmitted through an optical fibre, was made in 1994 [20, 21]. The s-SNOM system itself was not invented from scratch, however: much of its technology is based on the atomic force microscope (AFM - developed in 1986) which uses a probe

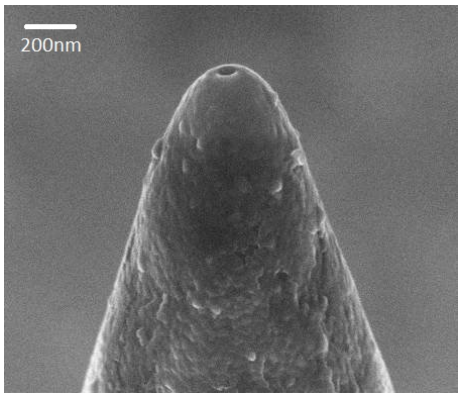

Figure 1.3: An SEM image of a typical SNOM probe where the tapered fibre has been coated with metal. The aperture is clearly visible at the apex of the probe. Taken from [13].

to measure a surface's topography (without recording any optical information) [22]. AFMs, in turn, inherited much from the scanning tunnelling microscope (STM), which was introduced in 1981 (garnering its inventors the 1986 Nobel Prize for Physics) [23, 24].

With the advent of s-SNOM, near-field microscopy became completely independent of the wavelength. Resolution was now dependent on the size of the probe (which are nowadays commercially widely available with 10nm diameters) with no restriction on the light's delivery as it only needed to be focused to the probe apex. As such, terahertz radiation has been used to image samples with a $\lambda/100,000$ resolution [25–27]. This access to longer wavelength spectral regimes becomes particularly useful when the chemical constituents of surfaces need to be identified, as spectroscopic measurements can be performed at nanoscale resolutions by scanning through a range of illuminating wavelengths [28–31].

Since the turn of the millennium, s-SNOM has found applications in a huge number of fields [32–34]. For surface analysis measurements, simple optical contrast has been demonstrated [35], as well as strain visualisation in semiconductors [36], charge carrier concentration mapping [37, 38], and the imaging of subsurface structures [39]. Photonic and optoelectronic devices have also been examined [40], as have anisotropic systems [41, 42].

One of the more widely used applications is the mapping of surface electric fields such as those created by plasmonic resonances (see chapter 4). Here s-SNOM has been put to use mapping plasmon and phonon-polariton propagation [43,44] as well as localised surface plasmons [45]. Even the near-field optical properties of graphene based devices have recently been probed [46, 47]. Cryogenic versions of s-SNOM have also been developed to allow for low temperature measurements [48, 49], and, intriguingly, it may prove possible to adapt the technique for nanoscale tomography [50].

1.3 Other Super-Resolution Techniques

While s-SNOM has developed over the last decade, a number of other techniques that allow for sub-diffraction limited optical (or infrared) imaging have also sprung up [51]. Some of these are

also based on sharp probes [52], while others rely on special substrate preparation using plasmonics or other, post-processing techniques to extract the sub-diffraction limit information [53–55].

Another set of techniques that has proved itself capable of high-resolution chemical mapping is Raman spectroscopy. Its sub-diffraction limited forms come in two major versions; tip-enhanced (TERS) [56] and surface-enhanced Raman spectroscopy (SERS) [57]. Rather than looking at elastically scattered light (as is the case with s-SNOM), inelastically scattered light from the samples is analysed (i.e. different colours to the incident beam are collected), and, depending on the intensity peaks and troughs at different wavelengths, this allows for material identification.

While all of the super-resolution techniques listed so far bear little resemblance to traditional "eyepiece" microscopy, a paper published in 2000 by John Pendry [58] sparked an explosion of interest in metamaterials and "perfect lenses" (or "superlenses") that might, in the words of Nicolay Zheludev [51] make it possible that:

> *One day, thanks to nanotechnology, a schoolboy will be able to screw a nanoarray lens to his science class microscope and see a DNA molecule.*

These superlenses work by amplifying the evanescent waves close to a surface instead of allowing them to decay. In this way, all of the fine details are retained and, therefore, any reconstruction of the object in an image plane is no longer diffraction limited [59–64]. One of the major difficulties, however, is that as soon as the light exits the superlens itself, the evanescent components start decaying again [65]. As such, many of the practical demonstrations of superlenses have been performed with s-SNOM, as a sub-diffraction limited imaging system is required to verify that the image formed by the lens is itself sub-diffraction limited [66, 67].

The field of metamaterials is involved with the manipulation of light, and so the physics of superlenses is closely linked with that of optical cloaking [68–70] and negative refraction [71–76]. Attempts have been made to convert the evanescent components within the superlens to propagating waves which would fulfil the dream of the schoolboy looking directly at DNA (metamaterial lenses which do this have been dubbed "far-field superlenses" or "hyperlenses") but as yet they have been plagued by loss and a narrow spectral operating range [53].

1.4 Thesis Outline

This thesis starts by looking at the experimental setup of the s-SNOM system (chapter 2), with a particular focus on how the near-field information is recovered from the light scattered by the probe using the so-called pseudo-heterodyne detection technique. This is followed by a chapter on the theoretical background of s-SNOM measurements (chapter 3) in which the origin of evanescent waves is explained, as is the origin of signal contrast in both cases of measurements of plasmonic structures, and measurements of chemical differences across the sample. Two analytical models are also introduced to help with the latter; the point-dipole model and the monopole model.

Chapter 4 marks the start of the experimental chapters with an examination of three different plasmonic structures in the near-infrared. The first is a simple gold colloidal sphere which supports a dipole mode - the simplest of plasmonic resonances - as a proof-of-principle demonstration of our

setup. The second is a dolmen structure (a Stonehenge-esque arrangement of three blocks as viewed from the side), which supports a Fano resonance. Actually imaging the effects of the resonance, however, proves to be impossible due to the reflection mode in which our system operates. The final structure studied here is a cruciform pentamer, which has the property that its far-field response is unchanged by the incident radiation's polarisation, whereas its near-field distribution is highly dependent upon it. Theoretical, far-field and near-field results are presented to provide evidence for this behaviour.

Chapter 5 is the first of the experimental chapters where mid-infrared light is used, as the first ever implementation of a pulsed quantum cascade laser (QCL) with an s-SNOM is presented. These new infrared sources are widely tuneable (6 - 10μm in our case) and allow for benchtop, turn-key operation; their tuneability in particular marks them out as promising candidates for performing nanoscale spectroscopy in conjunction with s-SNOM. Another plasmonic sample - a broadband log-periodic antenna - is examined at different wavelengths as evidence for the setup's tuneability; the positions of maximum field enhancement shift across the stucture as the illuminating light's frequency changes.

The following chapter (chapter 6) takes this combination of pulsed the QCL and the s-SNOM and attempts to apply it to spectroscopic imaging of biological material. High resolution images of oesophagial epithelial cells are collected at a number of wavelengths, but truly spectroscopic measurements are stymied by the inherent pulse-to-pulse variation of our QCL at all but a select few wavelengths. Far-field, bulk tissue spectra are also presented to provide a comparison between near and far-field measurements of tissue samples.

The final experimental chapter (chapter 7) is a detailed examination of the major noise sources present in s-SNOM measurements. It is seen that, for most of our experimental parameters, laser intensity fluctuations prevent the shot noise limit from being reached, not electronic noise sources. An attempt is made to remedy these issues by introducing a reference channel in conjunction with an analogue voltage divider to ratio out the fluctuations.

Chapter 2

The s-SNOM Setup

"A man's got to do what a man's got to do." **Prof. Chris Phillips**

2.1 Introduction

As a form of microscopy, s-SNOM is unique in its ability to optically and non-destructively image samples with nanometric precision, regardless of the wavelength of light being used. As such, it has found applications in all sorts of fields, from plasmonics to biology [31, 36, 38, 77–80]. Broadly speaking, it is based on a near-field interaction between a very sharp probe and the sample surface, where the near-field can be defined as the non-propagating solutions to Maxwell's equations that only exist within a few nanometres of a surface. These non-propagating - or evanescent - solutions encode the fine details of a surface, and so the diffraction limit is circumvented. If the far-field, propagating solutions are used - as in conventional microscopy or telescopy - diffraction limits the achievable resolution to approximately $\lambda/2$.

The purpose of the s-SNOM probe, therefore, is to scatter light from the near-field into the far-field where it can be detected and analysed at each pixel so that information regarding this scattering - its magnitude and phase - can be displayed as varying contrast in the images.

2.2 Atomic Force Microscopy

The atomic force microscope (AFM) forms the basis of the s-SNOM system. Having first been demonstrated in 1986 [22], atomic force microscopy is a mature and widespread technology [81]. The well-developed AFM industry brings an important benefit to s-SNOM; many of the parts used are interchangeable between the two techniques meaning that s-SNOM spares are cheap and readily available.

An AFM consists of a probing tip which is held very close to a sample attached to a piezoelectric scanner. This scanning stage typically allows for nanometrically precise placement of the tip in the horizontal directions and Angstrom precision in the vertical. The probe itself is attached to a cantilever, which, when vibrated, exhibits a resonant frequency (much like a tuning fork).

During measurements it is driven at this frequency causing the probe to oscillate vertically with an amplitude of around 30nm [82, 83].

In this "tapping mode", the probe is just brought into contact with the sample when at the lowest point in its oscillation cycle. This is measured by a laser beam which is brought to a focus at the top side of the cantilever, whereupon it is reflected. It then falls upon a knife-edge before hitting a detector (see figure 2.1). The presence of the knife edge means that the measured voltage is sensitive to the deflection of the cantilever as, depending on the angle of reflection, more or less light reaches the detector. The signal from the detector is therefore used in a feedback loop to regulate the probe-sample separation; the sample is lowered if a maximum cantilever deflection is measured. Not only does this prevent damage to the probe, but, by recording the vertical position of the sample required to keep the probe's vibration at a constant amplitude, the topography of the surface can be measured pixel by pixel with a resolution equal to the size of the probe.

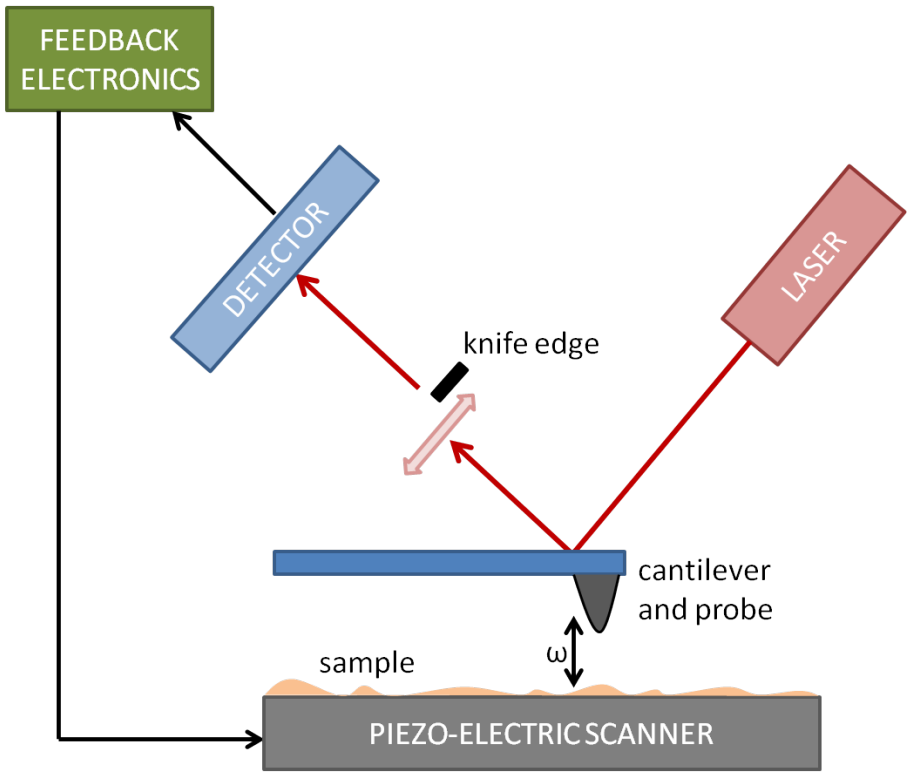

Figure 2.1: Schematic representation of an AFM setup. As the cantilever deflects due to the oscillation of the probe, the amount of light reaching the detector varies due to the presence of the knife edge (light red, doubled-headed arrow). This signal is then used in a feedback loop to regulate the position of the sample (keeping the oscillation amplitude constant), and thus the topography of the surface is measured. The focusing optics have been excluded for clarity.

2.3 Experimental Setup for s-SNOM

An s-SNOM differs from an AFM in one key respect; in an AFM measurement one laser is used to measure the cantilever deflection. With s-SNOM, however, a second laser is focused onto the apex of the probe. The light that is consequently scattered from the probe is affected - via a near-field interaction between the surface and the tip - by the sample beneath, and so this scattered light varies in both magnitude and phase depending on the dielectric constant of the surface (see section 2.5.1). Differences in the magnitude and phase form the basis of the contrast within an image.

For the experiments in this thesis, either a helium-neon laser (fixed at $\lambda = 633$nm), a titanium-sapphire laser (tuneable from $\lambda \sim 700 - 1000$nm) or a pulsed quantum cascade laser (tuneable from $\lambda \sim 6 - 10\mu$m) were used in conjunction with a commercially available s-SNOM (*NeaSNOM, Neaspec, Germany*). The precise experimental setup changes slightly depending on the type of measurement being performed (plasmonics or chemical mapping, for example) - each individual case will be fully described in the relevant chapters. The general scheme on which all setups were based, however, is shown in figure 2.2.

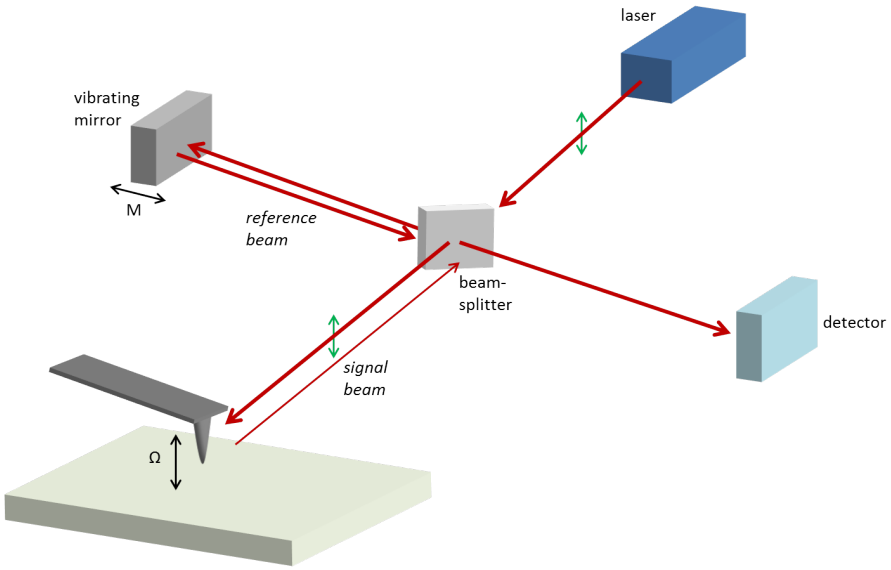

Figure 2.2: Schematic representation of the s-SNOM setup. The AFM laser and detector have been excluded for clarity, as have the optics which bring the beam to a focus at the probe's apex. The green arrows represent the light's polarisation.

In terms of the benchtop optical setup, the light from the laser is first passed through a beam-splitter. One of the resulting two beams is focused onto the apex of the vibrating probe (with oscillating frequency $\Omega \sim 75 - 250$kHz) by a reflective parabolic mirror (not shown in figure 2.2). Light scatters from the probe in all directions, but only the backscattered light which retraces the illumination pathway is collected. This is the *signal beam*. The second beam - the *reference beam* - is directed towards a vibrating mirror which introduces a phase modulation ($M \sim 300$Hz). This modulation later allows for a reconstruction of the magnitude and phase of the near-field scattering (see 2.5) and is known as *pseudo-heterodyne detection* (see 2.6). The two beams are recombined

at the beamsplitter and pass on to the detector.

Practically speaking, reflective optical components are used as far as possible to allow for easy interchanges between measurements in different spectral regions. In fact, the only component which cannot be reflective is the beamsplitter, and so this, along with the detector, is the only part of the setup which needs changing. In terms of system alignment, the most difficult task is positioning the laser's focal spot at the probe's apex. This is particularly challenging in the visible and near-infrared spectral regions where the diffraction limited focal spot is just a few hundred nanometres in diameter - indeed, it is only possible because the focusing parabolic mirror is mounted on a three axis piezoelectric positioning stage.

2.4 Field Enhancement

The reason that s-SNOM is capable of such sensitive measurements is the high level of field enhancement at the apex of the probe. This means that the near-field interaction between the probe and the sample - which then affects the scattering - is sufficiently strong to be detectable in the far-field [84, 85]. The large field enhancement is primarily caused by the geometry of the probe - the "lightning rod" effect [86]. Incident light polarised along the axis of the probe forces the free electrons to oscillate along its length. The curvature of the probe at its tip then creates a high charge density as many electrons are forced together at the very apex (see figure 2.3). This creates a large, highly confined electric field, with enhancement factors typically in the hundreds [87]. When the polarisation of the incident light is perpendicular to the axis of the probe, the lightning rod effect is much reduced and near-field measurement becomes much less sensitive (although not impossible) [32]. Other effects - notably surface plasmon-polaritons (see chapter 4) - can also play a role in the field enhancement around the probe apex, and probes have been designed specifically to exploit these effects [88, 89].

The radius of curvature of the s-SNOM probe defines the achievable resolution, as it is in the area immediately underneath the probe that the field enhancement is strong enough to affect the scattering into the far-field. The radius of curvature of commercially available probes is typically 10-20nm, allowing for optical resolution far below the traditional diffraction limit.

2.5 Scattering and Suppression: Theoretical Background

2.5.1 The Scattering Coefficient

In simple terms, the entire interaction between the probe and the sample can be described by the following equation:

$$E_{out} = \sigma E_{in}. \tag{2.1}$$

More prosaically, the light scattered from the probe is related to the incident light by some complex scattering coefficient σ. Fundamentally, it is changes in the magnitude and phase of σ across a sample that are measured by s-SNOM. The complexity of the interaction, however, and its

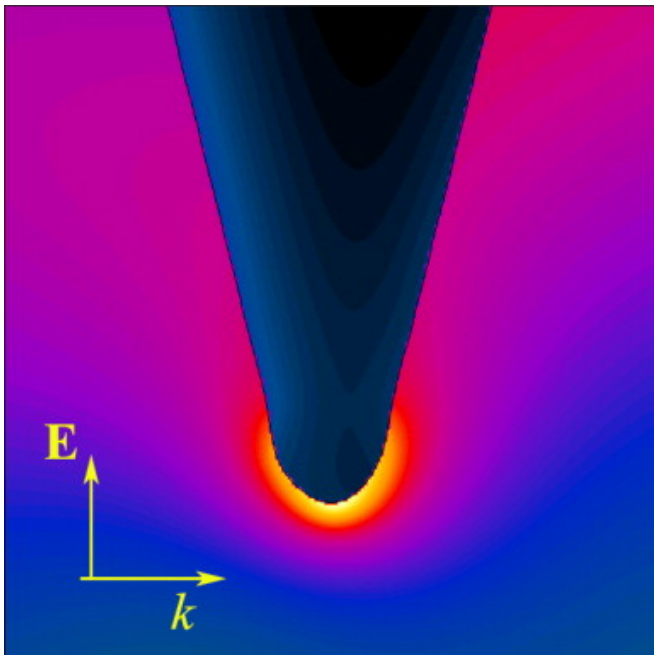

Figure 2.3: When incoming light is polarised along the tip axis, the resulting electric field at the apex is greatly enhanced and confined (enhancement factors are typically in the hundreds). This is not the case for polarisation normal to the axis for which charges just align themselves along the lateral sides of the probe (taken from [87]). E and k represent the incident light's electric field and wavevector respectively.

dependence on many variables mean that - in modelling terms - solutions for σ can only be found with complete accuracy by numerical simulations. There are, however, several analytical models of varying intricacy which can provide good approximations and an intuitive understanding of the probe-sample interaction (see section 3.4).

The scattering coefficient can be split into two components,

$$\sigma = \sigma_N + \sigma_B \tag{2.2}$$

where σ_N is the near-field component of the overall scattering coefficient - the *near-field scattering signal* - or in other words the scattered light which has actually been affected by the near-field interaction between the probe and the sample, and σ_B is the background component - the *background scattering signal* - which would exist regardless of whether the probe was close to a sample or just hanging in free space. In absolute terms, σ_B is much larger than σ_N and so the challenge is to extract purely the near-field scattering signal from the light that arrives at the detector.

2.5.2 Background Suppression

One of the many parameters on which σ depends is the probe height z above the sample. Fortunately, the two components of σ, σ_B and σ_N, are affected very differently by changes in z. Changes in σ_B are largely due to interference between incoming and reflected beams, and so variations are on length scales comparable with the wavelength which is typically much larger than z. The near-

field signal σ_N, however, is much more sensitive to z as the near-field only exists within a few nanometres of the surface [90].

Changes in the probe height, therefore, can be used to distinguish between the rapidly decaying σ_N and the slowly varying σ_B. These constituent parts of the scattering coefficient σ, however, must be extracted from the quantity that is actually measured - the voltage at the detector, $u(t)$.

Given that an s-SNOM is based on an AFM operating in tapping mode (section 2.2), any light scattered by the probe is modulated at the probe oscillation frequency Ω. Because of this periodic modulation we can express the voltage registered by the detector $u(t)$ as a Fourier series:

$$u(t) = \sum_{n=-\infty}^{\infty} u_n e^{in\Omega t} \tag{2.3}$$

where the Fourier coefficients u_n (n being the probe frequency harmonic) are obtained by:

$$u_n = \frac{1}{T}\int_0^T u(t)e^{-in\Omega t} \tag{2.4}$$

Practically speaking, the values of u_n can be found with a lock-in amplifier, or by performing a fast fourier transform (FFT) on a number of samples of $u(t)$. The reason for measuring the higher harmonics u_n of the voltage signal from the detector $u(t)$ is straightforward: the higher harmonics contain more of the rapidly changing near-field scattering signal σ_N than the slowly changing background scattering signal σ_B. At a sufficiently high harmonic, the contribution from σ_B at the harmonic n (denoted $\sigma_{B,n}$) to u_n becomes negligibly small and so $\sigma_{N,n}$ - the useful near-field scattering signal - dominates the measurement. Further details on how $\sigma_{N,n}$ is extracted from measurements of u_n are presented in section 2.6. The precise harmonic at which $\sigma_{N,n}$ comes to dominate $\sigma_{B,n}$ depends on the wavelength; typically n=3 is required for visible and near-infrared experiments, while n=2 suffices for the mid-infrared [91].

It should be noted that although finding u_n at higher harmonics results in a *relative* increase in the measurement of $\sigma_{N,n}$ compared to $\sigma_{B,n}$, the *absolute* values of $\sigma_{N,n}$ and $\sigma_{B,n}$ decrease. This means that, in practise, a balance must be struck between suppressing the background signal $\sigma_{B,n}$ and having a measurably large near-field signal $\sigma_{N,n}$ (see section 3.5).

2.5.3 Approach Curves

As evidence that the measurement of u_n at a sufficiently high value of n effectively suppresses the background signal $\sigma_{B,n}$, the height of the probe can be varied and the overall scattering coefficient σ_n measured at different harmonics. Doing so results in traces known as "approach curves", as shown in figure 2.4 [92]. These plot the magnitude of σ_n against the probe height above the sample. The approach curves clearly show a rapid signal decay at the third harmonic - to be expected of a near-field scattering $\sigma_{N,n}$ dominated measurement of σ_n as the near-field decays away within a few nanometres of the surface. A much slower decay is observed at the first harmonic, typical of a signal dominated by the background scattering $\sigma_{B,n}$ which varies over length scales $\sim \lambda$.

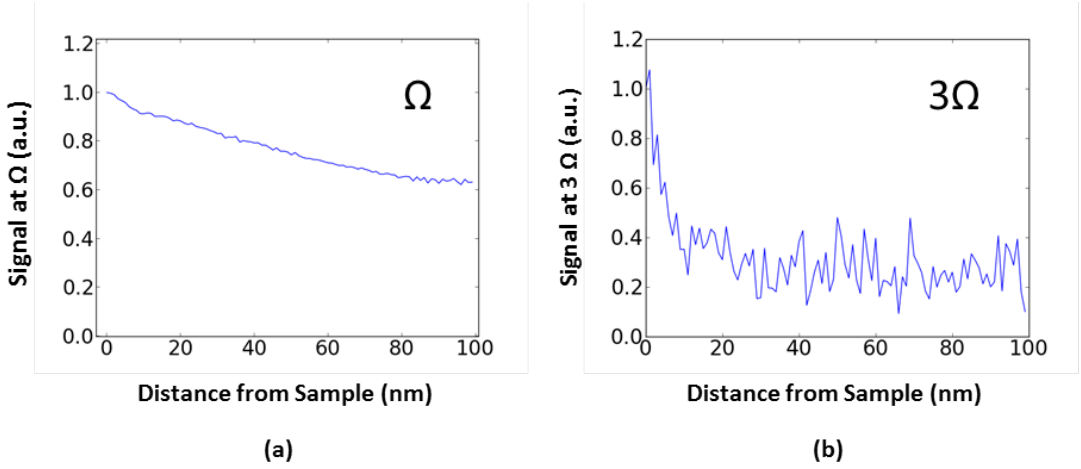

Figure 2.4: Signal values at different probe heights at the first (a) and third (b) harmonics of the probe oscillation frequency measured at $\lambda = 750$nm (titanium-sapphire laser) on a gold sample. At the first harmonic the background scattering signal $\sigma_{B,n}$ is still a significant portion of the overall measured scattering σ_n and so the signal decays slowly. At the third harmonic, however, the background has been effectively suppressed and the near-field signal $\sigma_{N,n}$ is the dominant contribution. Now the signal decays extremely rapidly to the noise floor. The rate of decay is related to the radius of the curvature of the probe's tip, which is typically around 10nm.

2.6 Pseudo-Heterodyne Detection

Historically, three different detection schemes have been used to take advantage of the higher harmonic suppression of the background scattering signal $\sigma_{B,n}$; homodyne, heterodyne and pseudo-heterodyne detection [93, 94]. The pseudo-heterodyne method is our preferred technique because it requires the simplest optical components (all of which can be reflective except a beamsplitter and the detector allowing for simple switching of the setup between different spectral regions). It is also the only technique capable of measuring the magnitude s and phase ϕ of the scattering coefficient σ simultaneously. In essence it works like a Michelson interferometer (see figure 2.2) where one of the arms ends at the probe - the *signal beam* - and the other at a mirror. This mirror is sinusoidally vibrated at a frequency M (where $M << \Omega$) to provide the phase modulated *reference beam* (see figure 2.2).

The reference beam and the signal beam are recombined at the beamsplitter. Their coherence mean that they interfere, causing a time-varying voltage trace at the detector $u(t)$. If this trace is Fourier transformed, there are two types of Fourier components visible:

- Peaks at the harmonics of the probe vibration frequency, u_n.
- Sidebands of these initial peaks separated by the frequency of the vibrating mirror $C_{n\Omega,mM}$.

A diagram of a typical pseudo-heterodyne generated spectrum is shown in figure 2.5. Note that the figure is not to scale - in reality the probe oscillation frequency is usually several hundred kilohertz and the reference mirror vibration frequency is around three orders of magnitude smaller at several hundred hertz.

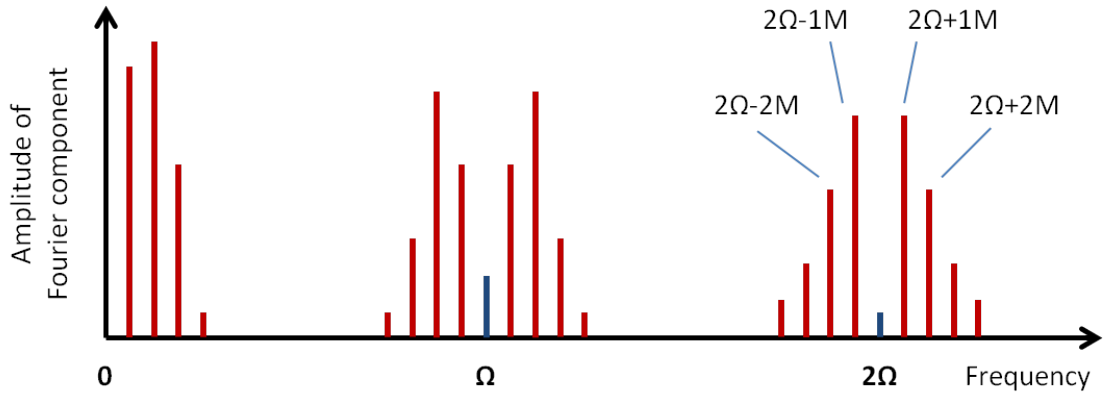

Figure 2.5: The effect of the phase modulated reference beam is to split the harmonics of the signal from the probe (blue) into sidebands (red). The relationship between odd and even order sidebands gives the near-field amplitude and phase. Note that the figure is not to scale; Ω is typically between 75 and 250kHz while M is just 300Hz.

Extracting the amplitude s_n and phase ϕ_n from the overall scattering coefficient at a given harmonic σ_n from the values of the sidebands $C_{n\Omega,mM}$ is rather mathematically involved and so the full derivation is left to reference [95]. The relationship between the odd and even number sidebands provide the final values, given by:

$$s_n = \sqrt{|C_{n\Omega+M}|^2 + |C_{n\Omega+2M}|^2} \tag{2.5}$$

and

$$\phi_n = tan^{-1} \frac{|C_{n\Omega+M}|}{|C_{n\Omega+2M}|} \tag{2.6}$$

where $C_{n\Omega+M}$ represents the value of the sideband at frequency $n\Omega + M$.

2.6.1 Modelling Pseudo-Heterodyne Detection

To gain a more intuitive grasp on the parameters of pseudo-heterodyne detection, it is useful to model the interference between the reference beam and the signal beam. The reference beam can be described as a phase modulated plane wave:

$$E_R = Ae^{i(kz-\omega t+\gamma sin(Mt))} \tag{2.7}$$

where γ represents the *modulation depth*, or the amplitude of the vibrating mirror. The signal beam can similarly be described by:

$$E_S = Bs(t)e^{i(kz-\omega t+\phi(t))} \tag{2.8}$$

where $s(t)$ and $\phi(t)$ represent the time varying magnitude and phase of the scattering coefficient σ as the probe moves up and down. The voltage registered by the detector is therefore proportional to the square of these two beams combined:

$$u(t) = k(E_R + E_S)^2 \tag{2.9}$$

where k is the proportionality constant of the detector. When the full expressions for E_R and E_S are substituted, the following expression arises:

$$u(t) = k[A^2 + B^2 s(t)^2 + 2ABs(t)cos(\gamma sin(Mt) - \phi(t))] \tag{2.10}$$

The first two terms represent the individual contributions from the reference beam and the signal beam. In frequency space, the Fourier components of these appear at DC and the harmonics of the probe vibration frequency u_n respectively. The third term, however, concerns the mixing of the reference and signal beams and, as such, its Fourier components appear as the sidebands of the probe harmonics $C_{n\Omega,mM}$. Figure 2.6 shows the time and frequency domain plots of equation 2.10. For clarity, A and B have been set to 1, M set to 1000Hz and $s(t)$ and $\phi(t)$ calculated using the simplest model of the probe-sample interaction - the point-dipole model (see section 3.4).

Note that the sidebands are larger than the central peaks of the probe harmonics. This is because the reference beam is considerably more powerful than the signal beam, which is just the light that is backscattered along the illumination pathway. It is instructive to rewrite equation 2.10 in terms of the detected photovoltages. Considering that the time average of the cosine is 1/2, it can be re-expressed as:

$$\langle u(t) \rangle = V_{\mathrm{REF}} + V_{\mathrm{SIG}} + \sqrt{V_{\mathrm{REF}} V_{\mathrm{SIG}}} \tag{2.11}$$

where $V_{\mathrm{REF}} = kA^2$ and $V_{\mathrm{SIG}} = kB^2 \langle s(t) \rangle^2$. It is now clear that the size of the sidebands is dictated by $\sqrt{V_{\mathrm{REF}} V_{\mathrm{SIG}}}$. This is the fundamental reason that the pseudo-heterodyne technique is capable of such highly sensitive measurements; the weak near-field scattering signal from the probe, V_{SIG} is amplified by the much stronger reference beam, V_{REF}.

The amplification of the signal beam also has a useful signal-to-noise (SNR) characteristic. Assuming that the measurement is shot noise limited (which is only the case with a very stable laser - see chapter 7), the noise scales with $\sqrt{V_{\mathrm{REF}}}$. Given that the detected quantity scales with $\sqrt{V_{\mathrm{REF}} V_{\mathrm{SIG}}}$, the SNR scales purely with $\sqrt{V_{\mathrm{SIG}}}$. This is a remarkable result - the signal beam can be amplified by an arbitrarily large amount without affecting the amount of noise in the measurement. Note that this is only the case if the detector noise is negligible.

2.7 Alternative s-SNOM Setups

The focus of this chapter so far (and indeed the focus for the rest of this thesis) has been on s-SNOM using single wavelength sources. Recent efforts have been made, however, to use s-SNOM in conjunction with broadband sources, both laser and thermal [96–99]. When used in this way, s-SNOM is usually referred to as nano-FTIR (Fourier transform infrared spectroscopy) for the similarities that it shares with its far-field cousin. Its setup is very similar to that of single wavelength s-SNOM (see figure 2.2). The main differences are the source and the mirror in the reference beam as shown in figure 2.7. In the pseudo-heterodyne system, this vibrates

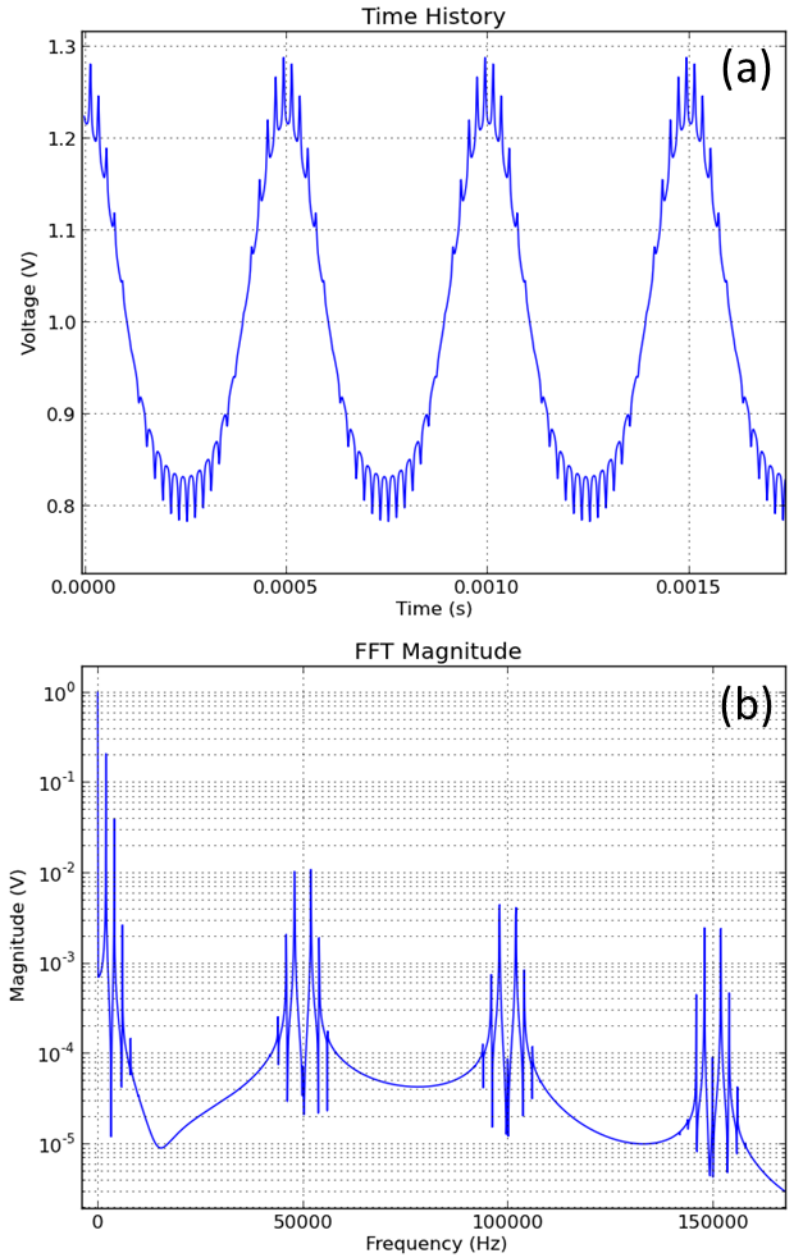

Figure 2.6: **(a)** Example of a modelled voltage waveform in time measured by the detector in the pseudo-heterodyne setup. The large, slow undulations are caused by the vibrating mirror in the reference arm ($M = 1000$Hz) and the faster, smaller pattern imposed on this is representative of the near-field scattering signal which rises and falls as the probe vibrates ($\Omega = 50$kHz). **(b)** The Fourier transform of (a). The sidebands, separated by the mirror vibration frequency $M = 1000$Hz are clearly visible around the central probe vibration harmonics; $\Omega = 50$kHz, $2\Omega = 100$kHz...

to phase modulate the light. In nano-FTIR, however, this mirror is slowly scanned around the white light position of the interferometer (which changes the optical path difference between the two interferometer arms) to create an interferogram. By subsequently Fourier transforming this interferogram, a complete near-field spectrum is obtained at each pixel.

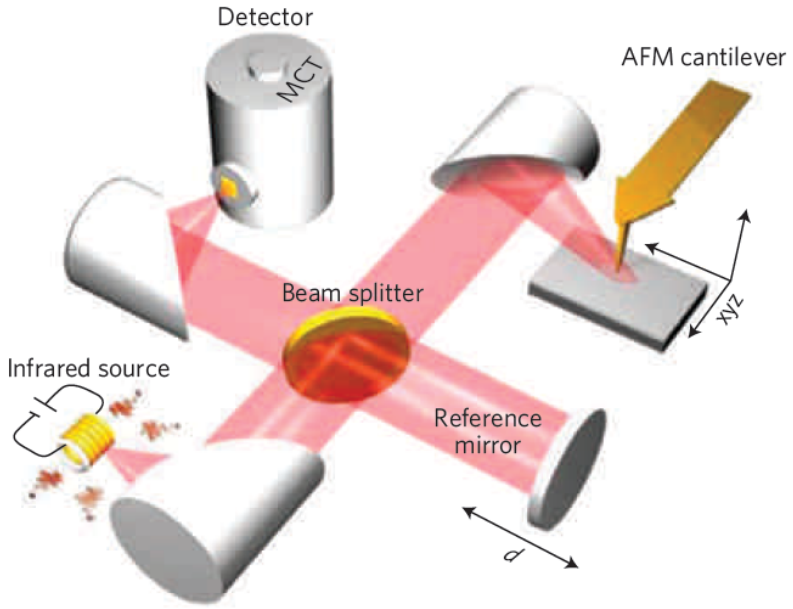

Figure 2.7: The experimental setup for a typical nano-FTIR measurement. Similarly to a single wavelength source, an interferometer is used for detection. In this case, however, the vibrating mirror is replaced with a translating mirror which allows an interferogram to be taken (the same as far-field FTIR). Fourier transforming this interferogram allows a complete near-field spectrum to be obtained at each pixel. Taken from [98].

The benefits of nano-FTIR over single wavelength s-SNOM are obvious - a full spectrum cannot be gained with a laser source unless the laser can be tuned over a large range. There are also, however, drawbacks. For an interferogram to be formed, data must be gathered at each point of the reference mirror's scan range at each pixel. For images of a reasonable size ($\sim$ 100 x 100 pixels), this can lead to extremely long image acquisition times - often several hours. The same scan at a single wavelength would take several minutes. Furthermore, the spectral resolution cannot be increased for broadband sources without increasing the number of scanning steps of the mirror in the reference arm, which further increases the acquisition time. Given these relative strengths and weaknesses of nano-FTIR and single wavelength s-SNOM, it seems likely that the two techniques will continue to be developed in tandem; the choice of which to use will depend on the requirements of the specific application.

Chapter 3

The s-SNOM Signal

"Research: 1% inspiration and 99% perspiration." **Prof. Chris Phillips**

3.1 Introduction

The purpose of this chapter is to provide an overview of the mechanisms by which the s-SNOM scattering signal is produced. This is important in understanding the sources of contrast within an image; it explains what makes one pixel brighter than another. Before this, however, an important concept in Fourier optics - evanescent waves - is covered, to explain how information regarding the fine details of a sample surface is converted into propagating waves that can be detected in the far-field.

3.2 Fourier Optics

To build an intuitive picture of the interaction between an s-SNOM probe and a surface, it is first necessary to understand the nature of evanescent waves [100]. One way of approaching this is to use the *angular spectrum* representation of electric fields. In brief, the angular spectrum describes a complicated electric field distribution in space (for example that of a laser beam) as a superposition of plane waves, each travelling at a different angle, and each with a particular amplitude [101]. This concept is illustrated in figure 3.1.

Once a complex field has been described as a sum of simple plane waves, it is mathematically much easier to determine the new field distribution after some kind of disturbance - an aperture or an obstacle for instance. The power of this technique lies in its ability to predict how a given field distribution will appear after passing the disturbance - which components are still propagating, for example, and which components start decaying away exponentially. It is these latter components which decay away that are defined as evanescent, and the following section looks in more detail at how they are formed.

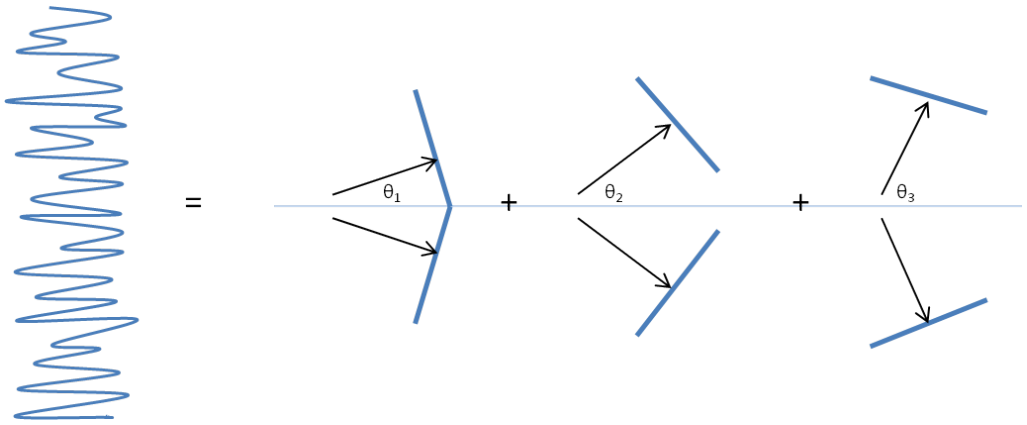

Figure 3.1: In the angular spectrum representation, any field distribution can be expanded as a superposition of plane waves of different amplitudes travelling at different angles.

3.2.1 The Origin of Evanescence

The individual plane waves of the angular spectrum representation can be described using complex notation as:

$$U(x, y, z) = Ae^{(\mathbf{k}.\mathbf{r}-\omega t)} \tag{3.1}$$

where the magnitude of the wavevector k is given by:

$$k^2 = k_x^2 + k_y^2 + k_z^2. \tag{3.2}$$

As a wave, equation 3.1 must satisfy the Helmholtz equation (the time-independent wave equation):

$$(\nabla^2 + k^2)U(x, y, z) = 0 \tag{3.3}$$

Assuming that the overall beam is travelling in the z direction, it can be shown that there are two possible situations with very different physical meanings for which equation 3.3 is satisfied:

1. $k_x^2 + k_y^2 < k^2$: When the transverse components of the wavevector are smaller than its absolute value, the wave propagates as usual. In other words, the wave exists at all values of z and so can be collected by a detector or another imaging system in the far-field.

2. $k_x^2 + k_y^2 > k^2$: In this case the transverse components are larger than the absolute value of k. For equation 3.2 to still be satisfied, k_z must become an imaginary number. When an imaginary k_z is inserted into equation 3.1, the wave no longer propagates freely, but instead decays exponentially in the z direction. This type of wave is *evanescent* and typically exists within a few nanometres of a light source or an illuminated object, hence evanescent waves are those which make up the near-field of an object.

3.2.2 Evanescent Waves and Resolution

By definition, evanescent waves have large transverse wavevector components $k_{x,y}$. These transverse wavevector components are related to the spatial frequency $\nu_{x,y}$ by:

$$\nu_{x,y} = \frac{k_{x,y}}{2\pi} \tag{3.4}$$

The resolution of an imaging system is directly linked to the largest spatial frequency it can record. This principle is demonstrated in figure 3.2; the higher the detectable spatial frequency, the finer the detail that can be measured.

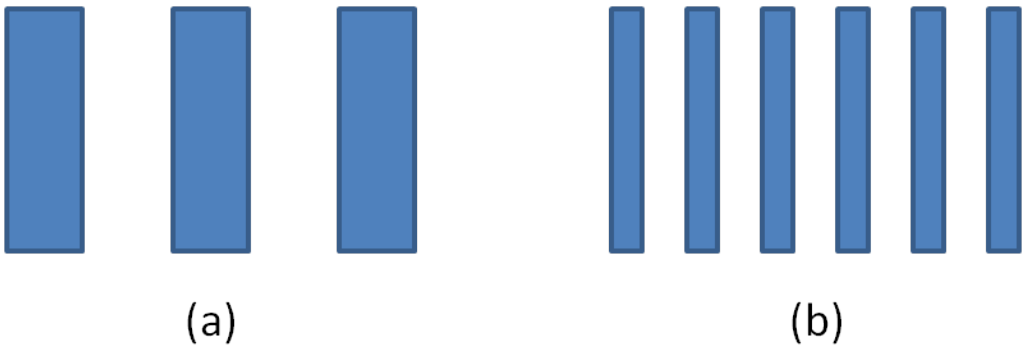

Figure 3.2: Schematic representation of patterns with (a) low spatial frequency and (b) high spatial frequency. The resolution of the high spatial frequency pattern is better than that of with a lower spatial frequency.

Evanescent waves can have arbitrarily large transverse wavevector components and hence correspond to arbitrarily high spatial frequencies. The limiting factor in traditional, far-field imaging systems - cameras, telescopes and "eyepiece" microscopes, for instance - is that these large transverse wavevector plane waves decay before they reach the detector. The high levels of detail they carry are therefore lost, and when the far-field imaging system tries to focus the remaining light into an image, the result is an imperfect reconstruction of the object. This is the fundamental cause of the diffraction limit.

If the high spatial frequency information of a sample is to preserved, therefore, a mechanism is required to somehow reduce the values of k_x and k_y of the evanescent waves so that the condition $k_x^2 + k_y^2 < k^2$ is met and they can propagate away from the sample surface. This is the role of the s-SNOM probe.

3.2.3 Detection of Evanescent Waves

An s-SNOM probe can be considered a shorter wavelength analogue of a radio antenna; its job is to channel electromagnetic energy from the far-field to the near-field as evanescent waves (a receiver) and then to re-emit these evanescent waves into the far-field (a transmitter). It is intuitive to consider the following sequence of events when describing the probe-sample interaction (figure 3.3):

1. Propagating light hits both the probe and the sample forming evanescent waves with high spatial frequencies, k_A.

2. Those evanescent waves that exist between the probe and the sample mediate an interaction between the two, k_B.

3. This interaction effects the way in which light is scattered back from the probe into the far-field, k_C.

4. The scattered light is detected by the pseudo-heterodyne setup.

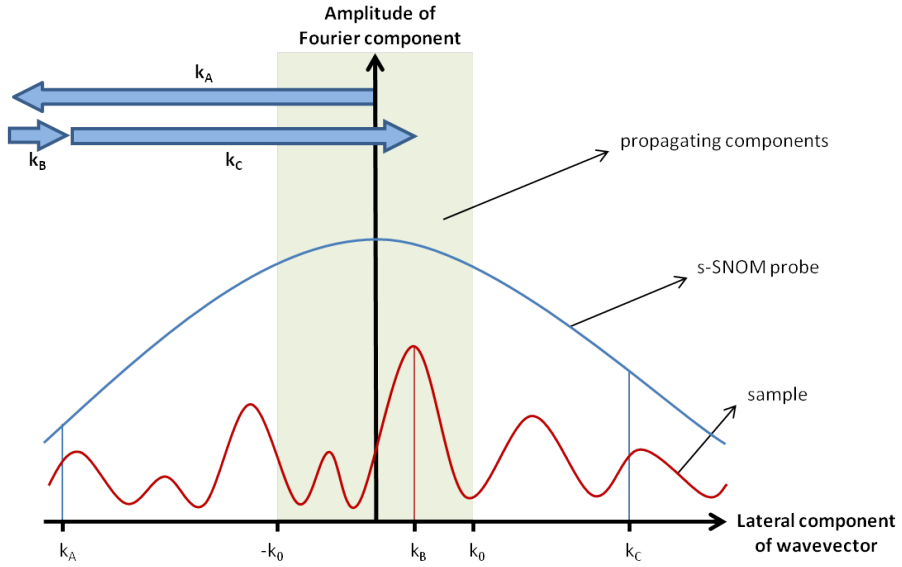

Figure 3.3: The spread of lateral wavevectors created by an s-SNOM probe (blue) and an example sample surface (red). Incoming light is converted into evanescent waves with lateral wavevector k_A. This then interacts with the surface where it receives a "kick" of k_B before being converted back into a propagating wave upon scattering from the probe again, k_C, after which it can be picked up in the far-field by a detector.

The following section looks more closely at the interaction between the probe and the sample, and in particular what causes the magnitude and phase of the scattered light to vary from pixel to pixel.

3.3 Signal Contrast

There are two distinct mechanisms by which s-SNOM image contrast is formed, and these are dependent on the type of sample being look at. In an "active" sample - one which exhibits a plasmonic resonance, for example - the incoming light is somehow manipulated to form a new electric field distribution on the surface. In this case, the s-SNOM probe is sensitive to electric field, and image contrast is formed by differences in field strength across the sample.

In "passive" samples, however, where the surface plays no part in manipulating the field distribution, the scattered light from the probe is instead sensitive to changes in the sample's chemical constituents. The underlying causes of dark and bright regions are now rather more complicated, but are fundamentally based on differences in the sample's dielectric constant. This will be looked at in more detail in section 3.3.2.

Although the signal mechanisms are distinct for the two types of sample, care must be taken not to record both types of signal - i.e. field strength and chemical differences - from the same sample as this leads to difficulties in image interpretation. This is particularly problematic for plasmonic samples, where special experimental configurations must be implemented to ensure that only the electric field caused by the plasmon resonance is detected and that no material contrast is picked up. This setup will be fully explained in chapter 4.

3.3.1 Plasmonic Samples

Plasmonics

Strictly speaking, the term *plasmon* refers to the quanta of collective density oscillations of electrons within a free electron gas (e.g. within a metal). These oscillations occur within the bulk of the material and are often known as volume or bulk plasmons. In common parlance, however, the word "plasmon" has come to mean a *surface plasmon-polariton* (SPP) which is formed when the electron density oscillations couple to an electromagnetic wave at the interface between a dielectric and a metal (see figure 3.4). The word *polariton* refers to any quasi-particle formed between light and one of the solid state quanta; phonons, plasmons and excitons, for example.

SPPs have been the subject of much research in recent years due to their promise in a variety of fields such as metamaterials, biosensing and optical computing. Despite the quantum mechanically based nomenclature, however, SPPs can largely be described by classical electromagnetism [102]. A more detailed presentation of their theory will be given in chapter 4.

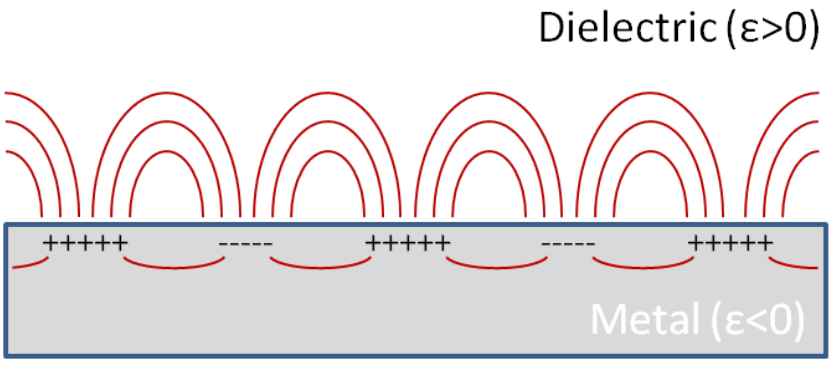

Figure 3.4: Schematic representation of a surface plasmon-polariton. Incoming light couples to electron density oscillations in the metal to form a field which is strongly confined in the directions normal to the interface but free to propagate along the boundary. The wavelength of the SPP can be much smaller than the free space wavelength of the light hence the interest for applications which require length scales below the diffraction limit.

The fields of an SPP are confined to the interface between a metal and a dielectric, and this confinement can be described by examining the dependence of the SPP's wavelength on the dielectric constants of the materials that support it, ϵ_m and ϵ_d:

$$\lambda_{\mathrm{SPP}} = \frac{\lambda_0}{\sqrt{\frac{\epsilon_m \epsilon_d}{\epsilon_m + \epsilon_d}}} \tag{3.5}$$

The square root term can be considerably larger than one, and thus light can be "squeezed" into volumes much smaller than the free-space wavelength λ_0. The ability of plasmonics to manipulate light at the nanoscale has opened up a large range of potential applications including enhanced photodetectors, new opto-electronic interchanges and cancer treatments, amongst others [103].

From an imaging point of view, it is important to make a distinction between two types of surface plasmons; those which propagate and those which are confined by the structures that support them. The former, once launched, ripple along the metal-dielectric interface indefinitely (or until they are attenuated by losses). The latter, however, exist only within the lateral confines of a supporting structure and are therefore known as localised surface plasmon resonances (LSPRs) or, more colloquially, just as plasmon resonances. It is these latter types of surface plasmon which are primarily studied in this thesis.

Contrast from Plasmonic Samples

Recalling equation 2.1, we can split the scattering coefficient σ into two constituent parts; one that is sensitive to the field on a surface, and one that is sensitive to the chemical differences:

$$\sigma = \sigma_{\mathrm{plasmon}} + \sigma_{\epsilon} \tag{3.6}$$

When imaging plasmonic samples, the chemical contribution σ_ϵ can vary significantly across the sample as the nanostructures which support LSPRs are typically made of gold on a glass or silicon substrate (see section 3.3.2). σ_ϵ must therefore be suppressed if the image is going to be an accurate representation of just the plasmonic resonance and not a combination of the plasmon and chemical contrasts. This suppression can be engineered into the experimental setup by choosing the polarisation of incident light such that it does not have a component along the probe's main axis (i.e. s-polarisation rather than p-polarisation). This limits the field enhancement beneath the probe and hence the scattering signal related to the chemical contrast. This experimental setup is fully explained in section 4.2. If the suppression of σ_ϵ can be achieved, the scattering from the probe is dominated by the local strength of the plasmon resonance:

$$E_{out} = \sigma_{\mathrm{plasmon}} E_{in} \tag{3.7}$$

In images from plasmonic samples therefore, the measured values of s_n and ϕ_n (see equations 2.5 and 2.6) correspond to the LSPRs magnitude and phase at each particular pixel.

3.3.2 Chemical Mapping

When s-SNOM was first developed, it was not with plasmonic samples in mind; the primary goal was to use the technique's high resolution to optically image different materials [20, 21, 104]. This "chemical mapping" is unique among optical techniques because the achievable resolution is not at all dependent on the wavelength; it depends on the size of the probe. As a result, samples exhibiting a contrast in the mid-infrared or even terahertz regimes (where the resolution approaches $\lambda/10,000$) can be studied with nanometric precision.

In experiments set up for chemical mapping, the contrast arises from the the dependence of the scattering coefficient σ on the local value of the sample's dielectric constant. The dielectric constant itself is, of course, dependent on the frequency of the light. This means that s-SNOM is also capable of spectroscopically identifying materials provided that the wavelength is tuneable over a suitably broad range (or a broadband source is used - see section 2.7) [31].

3.4 Modelling the Probe-Sample Interaction

To understand what causes changes in s and ϕ - the magnitude and phase of the scattering coefficient σ - the probe-sample interaction must be modelled. Unfortunately it is impossible to do so with complete accuracy using analytical methods. There are simply too many variables on which the interaction depends; the precise tip geometry, the topography of the surface, the beam profile of the incoming light, the probe and sample materials, the wavelength of the light and the probe height to name a few. Several attempts have been made to use numerical methods [105–107], but this approach gives us no intuitive understanding as to why the scattering coefficient changes from one material to another, nor how the relationship between the near-field and background scattering, σ_N and σ_B, varies at different harmonics.

There are two widely used analytical approaches for modelling the interaction between light, an s-SNOM probe and a sample surface; the point-dipole model [28, 84, 108, 109] and the more complex monopole model [110]. The point-dipole model describes the s-SNOM probe as a sphere with a point-dipole located at its centre. The monopole model, on the other hand, models the charge distribution inside the probe as a single charge near the apex with the rest of the (opposite) charge distributed amongst the body of the probe. For the purposes of this thesis, the point-dipole model is better for gaining an intuitive understanding of the source of material contrast, whereas the monopole model is better for quantitatively predicting experimental results. The differences between the two largely stem from the accuracy of the modelled electric field around the apex of the probe; the point-dipole model assumes the field to be somewhat weaker than it actually is [107, 110].

3.4.1 The Point-Dipole Model

In the point-dipole model, two major simplifications are made. Firstly, it is assumed that only the very tip of the s-SNOM probe is responsible for the near-field interaction with the surface. The probe is therefore approximated as a small sphere. Secondly, the tip dipole caused by the

illuminating light is assumed to be infinitely small (hence the name of the model) and located at the centre of the sphere as shown in figure 3.5.

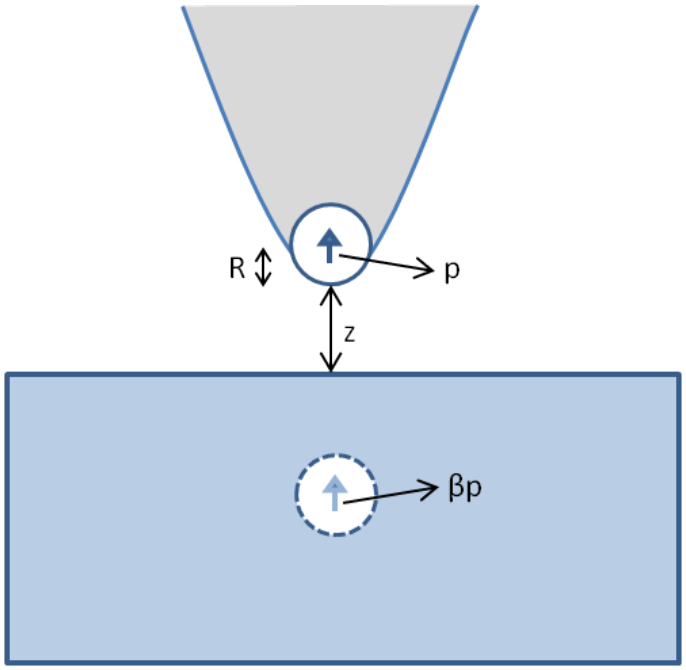

Figure 3.5: In the point-dipole model the s-SNOM probe is described as a small sphere of radius R at height z above the surface with an infinitely small dipole of dipole moment p at its centre. This induces a "mirror dipole" of dipole moment βp beneath the surface which in turn acts upon the tip dipole to change the way in which light is scattered by the probe.

The tip dipole acts upon the sample and induces a "mirror dipole" an equal distance beneath the surface as the tip is above it. The strength of this dipole is given by βp where p is the original tip dipole moment and β is given by:

$$\beta = \frac{\epsilon_{\text{surface}} - 1}{\epsilon_{\text{surface}} + 1}. \tag{3.8}$$

This is simply the Fresnel reflection coefficient for TM or p-polarised radiation in the electrostatic limit, or in other words the limit of the transverse wavevector components becoming much larger than the overall wavevector k, as is the case for evanescent waves (see section 3.2.1). The form that β takes has some important consequences; when $|\epsilon_{\text{surface}}|$ is less than zero, β is greater than one. Such a result may initially seem to contradict energy conservation as it appears that more energy is reflected from the sample than is incident upon it. This, however, is not the case as β only applies to evanescent waves which do not transport energy. For propagating waves, equation 3.8 is replaced by the full Fresnel reflection coefficient expression, which can itself never be greater than unity.

A convergent geometric series can be used to mathematically describe how the two dipoles (the mirror dipole and the initial tip dipole) interact; the newly created mirror dipole acts back upon the initial tip dipole which, in turn, affects the mirror dipole again. This back and forth

continues until an equilibrium is reached (note that this represents a mathematical description of the interaction rather than a real, physical time series). To understand what effect the mirror dipole has on the scattered light, we first need to describe the tip's dipole moment in terms of the polarisability, α, of the sphere at its apex without the presence of a sample, given by:

$$p = \alpha E_0 \tag{3.9}$$

where E_0 is the incident field and the polarisability is given by [111]:

$$\alpha = 4\pi R^3 \frac{\epsilon_{\text{tip}} - 1}{\epsilon_{\text{tip}} + 2}. \tag{3.10}$$

After the interaction with the mirror dipole in the sample has been taken into account, however, the polarisability in equation 3.9 can be replaced with an *effective polarisability*:

$$\alpha_{\text{eff}} = \frac{\alpha}{1 - \frac{\alpha\beta}{16\pi(R+z)^3}}. \tag{3.11}$$

At large distances from a dipole, the scattered field is directly proportional to its polarisability, $E_s \propto \alpha_{\text{eff}} E_0$. When this is compared to the expression for the scattering coefficient σ (equation 2.1), it is clear that:

$$\sigma \propto \alpha_{\text{eff}}. \tag{3.12}$$

Note that both of these parameters are complex and so have both a magnitude and phase. As equation 3.12 is not an equality, one can only use the effective polarisability to calculate the *relative* s-SNOM scattering signal strength from different materials. However, as long as relative contrasts are always used, calculating α_{eff} gives us some understanding of what values σ takes for different materials.

3.4.2 The Monopole Model

In the monopole model, the probe is modelled as an extended spheroid. The field around the spheroid (excluding, for now, any nearby surface) is described by placing two charges, Q_0 and $-Q_0$ - thereby creating an extended dipole - at a distance W_0 from the apexes. When the probe is placed close to a surface, the interaction between the two can be described by inserting another extended dipole into the probe - Q_i at a distance W_i from the probe apex while $-Q_i$ is distributed among the rest of its length. Figure 3.6 shows the parameters of the model.

The monopole model's equivalent of the effective polarisability α_{eff} is the near-field contrast factor η (i.e. the expression that is proportional to the probe's scattering σ). This is formed by dividing the dipole moment related to the near-field interaction, p_i, by the dipole moment that would exist without the presence of a sample, p_0. The near-field contrast factor, $\eta = p_i/p_0$, can therefore be regarded as a measure of the s-SNOM signal that does not depend on the intensity of the incident light field. The rather complicated expression for η is given by:

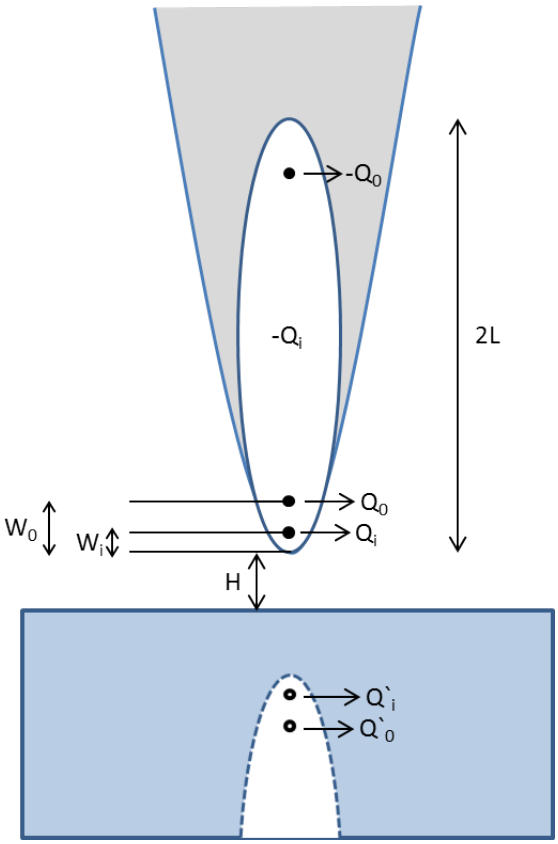

Figure 3.6: Diagram of an s-SNOM probe defining the various parameters of the monopole model. In contrast to the point-dipole model, the probe is modelled as an elongated spheroid.

$$\eta(t) = \frac{\beta(2Lg - 2H(t) - W_0 - R)\ln(\frac{4\mathrm{L}}{4\mathrm{H(t)}+2\mathrm{W_0}+\mathrm{R}})}{4L\ln(\frac{4\mathrm{L}}{\mathrm{R}}) - \beta(4\mathrm{Lg} - 4\mathrm{H(t)} - 3\mathrm{R})\ln(\frac{4\mathrm{L}}{4\mathrm{H(t)}+2\mathrm{R}})} \tag{3.13}$$

where R is the radius of curvature of the probe apex and g is an experimentally deduced value that can be used to correct for phenomena that are not explicitly included in the model - finite conductivity and radiation resistance, for example. For the modelling in this thesis, the value of g is set at $0.7e^{0.06i}$ [112]. $H(t)$ describes the height of the probe above the sample as it oscillates, and is given by:

$$H(t) = A(1 + \cos(\Omega \mathrm{t})) \tag{3.14}$$

where A is the amplitude of the oscillation.

Although the monopole model is capable of quantitative scattering predictions, the near-field contrast factor can - like the effective polarisability in the point-dipole model - be used to predict the *relative* contrast differences between different materials in a sample. This is because it depends on β which, in turn, depends on $\epsilon_{\mathrm{surface}}$. For two materials A and B, the contrast is given by the ratio of the scattering coefficients at a particular harmonic:

$$\frac{\sigma_{A,n}}{\sigma_{B,n}} = \frac{\eta_{A,n}}{\eta_{B,n}} \tag{3.15}$$

where the Fourier component of the near-field contrast at a given harmonic η_n is given - similarly to equation 2.4 - by:

$$\eta_n = \frac{1}{T}\int_0^T \eta(t)e^{-in\Omega t} \tag{3.16}$$

A similar expression can be used to find $\alpha_{\mathrm{eff,n}}$ with the point-dipole model, the effective polarisability at a given harmonic n.

As an example of the predictive power of the near-field contrast factor, a comparison between an experimental image of a nanotransistor (taken from [78]) and the contrast ratios calculated by the monopole model is shown in figure 3.7. It is clear that the experimental and predicted contrast show good quantitative agreement. Values for the dielectric constants used in the calculation of η_n (where the second harmonic, n=2, was used to match the experiment) were as follows:

- Si: ϵ = 11.63 + 0.01i [113]
- Amorphous SiO_2: ϵ = -4.27 + 8.29i [114]
- Amorphous Si_3N_4: ϵ = 3.24 + 0.01i [115]
- NiSi: ϵ = -809.67 + 236.33i [115]

3.4.3 Including the Far-Field Contribution

The effective polarisability, α_{eff}, and the near-field contrast factor, η, both describe the near-field interaction between the probe and the sample, but to improve on either model some far-field aspects

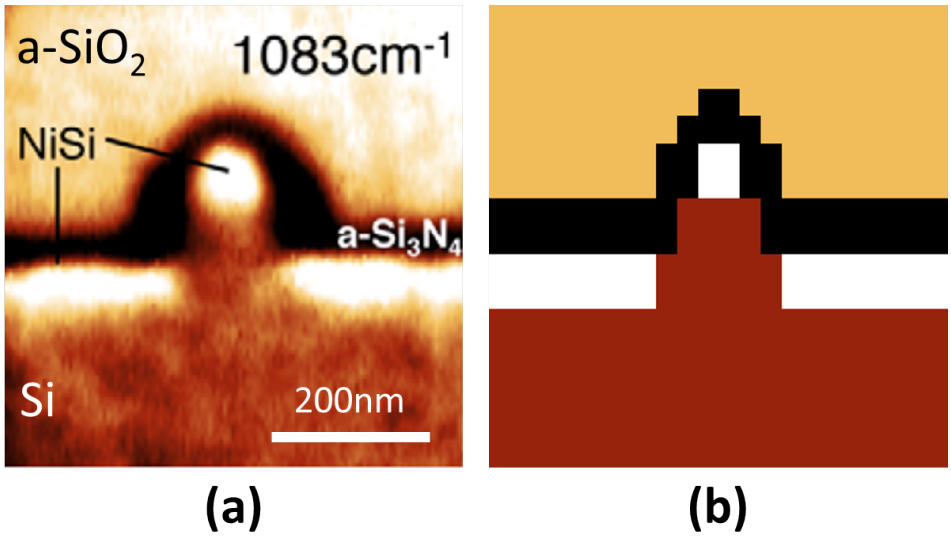

Figure 3.7: (a) An experimental image of a nanotransistor taken at 1083cm^{-1} and (b) the predicted image contrast using $\eta_{n=2}$ from the monopole model. Direct comparison between the two images is made possible by use of the same pixel value colour gradient (*Gwyddion* default). Image (a) taken from [78].

of the probe's illumination and scattering must also be included. Not all of the light incident upon the probe comes directly from the light source, and likewise not all of the scattered light reaching the detector comes straight from the probe; some of the light is reflected from the surface as shown in figure 3.8.

The scattering coefficient, therefore, must be multiplied by a factor of $1+r_p$ for both the incident and scattered pathways (where r_p is the Fresnel reflection coefficient for TM or p-polarised light). The quantities relating to the scattering coefficient - α_{eff} and η in the point-dipole and monopole models respectively - can therefore be rewritten as:

$$\alpha'_{\text{eff}} = (1 + r_p)^2 \alpha_{\text{eff}} \text{ and } \eta' = (1 + \text{r}_\text{p})^2 \eta \tag{3.17}$$

Care must be taken, however, when applying the factor of $(1 + r_p)^2$ to account for material differences across the surface. When the probe is at the boundary between two materials, the dielectric constant used to calculate r_p may not be the same as that used to calculate α_{eff} or η.

3.4.4 Different Materials, Different Signals

Using the models presented in sections 3.4.1 and 3.4.2, we can use the dielectric constant of the sample to predict its relative near-field scattering strength. The dielectric constant has both a real and imaginary part, so we can calculate a "signal surface" on which the near-field scattering signals for all materials lie. Focusing on the point-dipole model, the value of the signal surface is found by calculating the Fourier component at a given harmonic of α_{eff} - which outputs a complex number - and then finding the magnitude or the phase, the result of which is then multiplied by the far-field factor $(1+r_p)^2$ to give α'_{eff}. This output should then be proportional to the experimentally

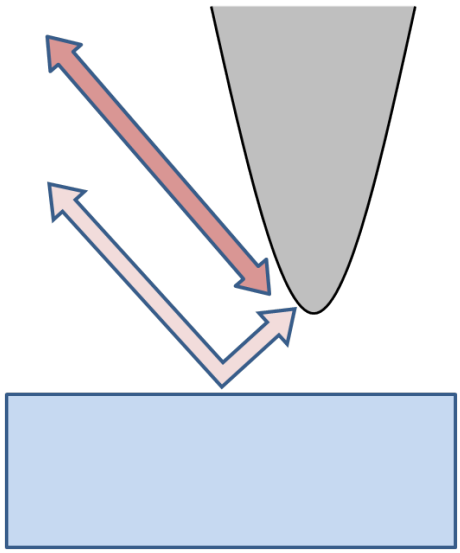

Figure 3.8: An s-SNOM probe's illuminating and scattered light doesn't only take the direct path; there is also an element reflected from the sample surface which needs to be taken into account by any signal strength models.

measured values of s_n and ϕ_n. Figure 3.9 shows just the far-field contribution to the measured signal, while figure 3.10 shows the magnitude and phase of α'_{eff} at the second harmonic of the probe oscillation frequency:

The signal surface is useful for determining why different chemicals produce different signal values. It is easiest to explore this idea by looking at the signal value in various limiting cases of the dielectric constant of the sample surface:

- $\epsilon_s \to \pm\infty$: In this case, β tends to 1, and $|\alpha_{\text{eff}}|$ tends to a constant value. The difference between these extremes in figure 3.10 lies in the far-field element $(1 + r_p)^2$. For metals ($\epsilon_s \to -\infty$) the reflection coefficient is larger and so the overall signal is greater.

- $\epsilon_s \to -1$: Now β tends to infinity and for lossless materials (i.e. $\text{Im}(\epsilon) = 0$), α_{eff} tends to zero. When losses are included, $\epsilon_s = -1$ corresponds to a minimum in scattered intensity.

- $\epsilon_s \to -5$: For lossy materials this corresponds to a minimum in $(1 + r_p)^2$ and hence a minimum in the signal strength too.

- $\epsilon_s \approx -2$: The most prominent feature on the surface is a large peak occurring when $\frac{\alpha\beta}{16\pi(R+z)^3} \approx 1$. In this case α_{eff} becomes extremely large. The effect of losses is to dampen this resonance.

The peak seen around $\epsilon = -2$ is known as a near-field resonance - it is not just a mathematical oddity. Its presence in silicon carbide (SiC) substrates has been used to enhance the near-field scattering signal of whatever is placed on its surface, allowing for more sensitive measurements [116–118]. The effect, in this instance, is caused by exciting optical phonons in the crystal lattice.

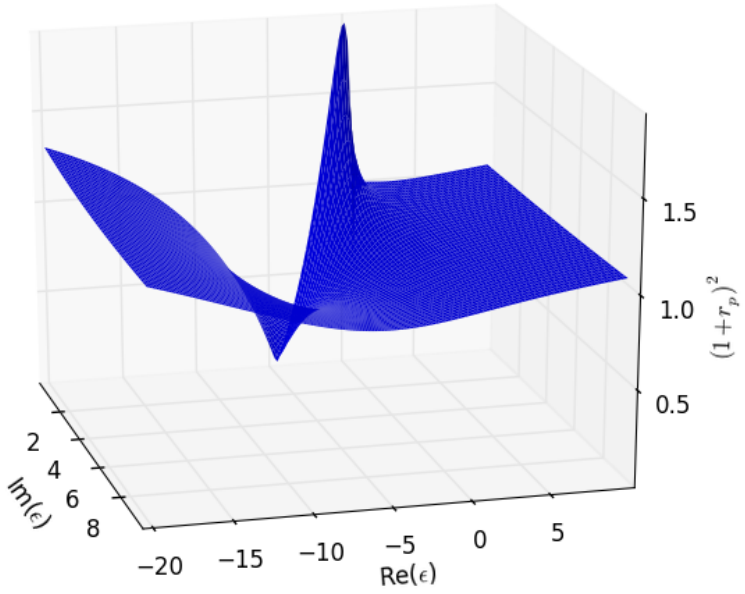

Figure 3.9: The far-field contribution to the signal "surface" $(1 + r_p)^2$ for different values of the real and imaginary parts of ϵ. The angle of incidence is set to 60° from the surface normal to match the experimental geometry used in this thesis.

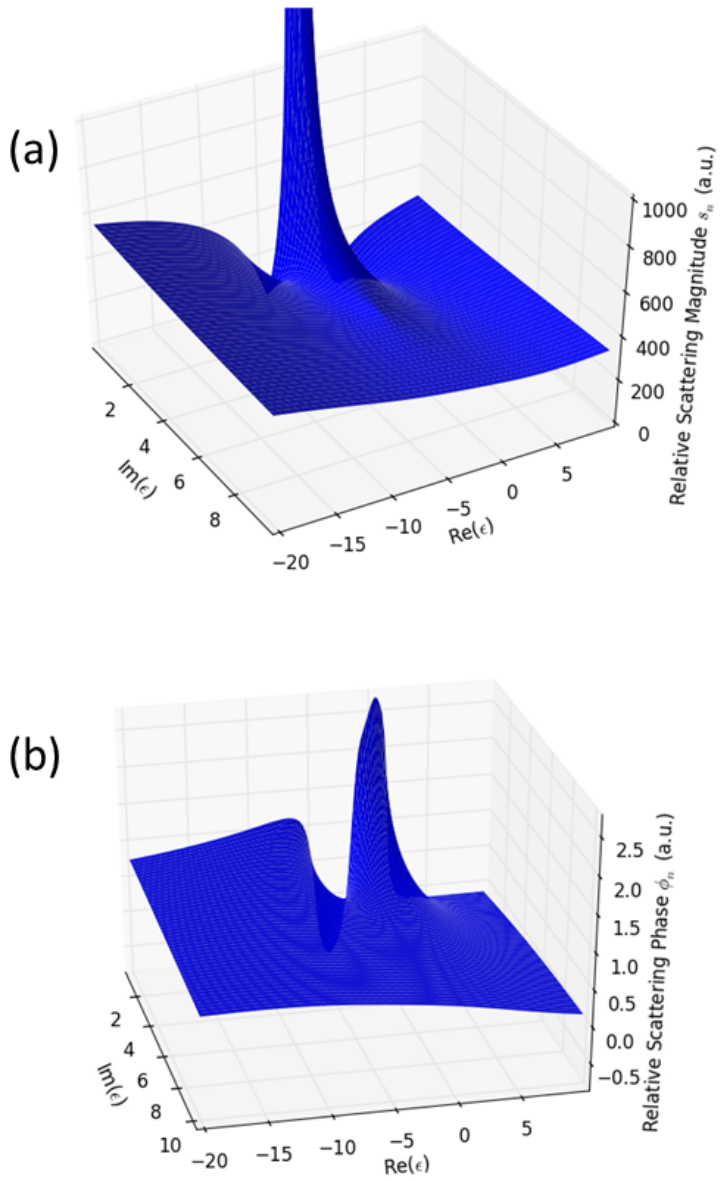

Figure 3.10: The signal "surface" representing the relative s-SNOM signal strengths for different values of the real and imaginary parts of a sample surface. It is found by demodulating $(1 + r_p)^2 \alpha_{eff}$ at the second harmonic of the probe frequency.

For comparison with the point-dipole model, the signal surfaces as calculated by the monopole model are shown in figure 3.11. Although there is a good qualitative match between the two models, there are some significant quantitative differences. The peak positions are slightly shifted to more negative values of $\mathrm{Re}(\epsilon)$ in the monopole model, for example, and the peak heights and shapes - in the phase diagram in particular - show differences between the two models.

3.4.5 Near-Field Spectroscopy

The dependence of the near-field scattering signal on the complex dielectric constant makes s-SNOM a powerful tool for spectroscopic measurements; as the dielectric constant is wavelength dependent, so is the near-field scattering signal. Each chemical, therefore, will have it's own unique spectral "fingerprint". If the appropriate range of wavelengths can be accessed (by use of a tuneable laser, for example), this allows for unambiguous material identification with nanometric spatial resolution.

It has also been shown that the far-field absorption coefficient of a material is proportional to the imaginary part of the scattering coefficient, $\alpha \propto \mathrm{Im}\{\sigma\}$, for non-polar materials (i.e. materials which do not exhibit a near-field resonance the peaks in figures 3.10 and 3.11). This means that existing absorption spectra databases can be used for material identification at the nanoscale without the need to build up new databases containing specifically near-field scattering spectra.

3.5 Noise in Measurements

Noise is an issue that affects all experimental measurements, and s-SNOM is no exception. In practical terms, the noise level sets a lower limit to the detectable signal power; if the noise is larger than the near-field scattering signal at a given harmonic, no measurements can be taken.

Noise comes from a variety of sources; mechanical, optical and electrical. These include detector noise, laser noise, thermal changes in the optical alignment and the effects of mechanical vibrations. A fuller analysis of the impact of different noise sources is provided in chapter 7, but for the purposes of this section they will be considered together as a "disturbance" term D.

The effects of D can be mitigated through good experimental design (mechanically isolating the optical bench, for example) and using low-noise equipment (the laser and the detector). Fundamentally, however, the measurement will always be constrained by the shot-noise limit (see section 7.2.3); however perfectly the experiment is set up, the disturbances will always have to be considered. Their biggest impact - aside from defining whether the measurement is even possible - is on the optimum choice of demodulation harmonic, n.

3.5.1 Signal, Background and Noise

It is useful to define the *Near-Field to Background and Disturbances Ratio* (NFBDR) [112]. This is simply a ratio of the useful near-field scattering signal ($\sigma_{N,n}$) to the unwanted background signal and the disturbances ($\sigma_{B,n}$ and D_n):

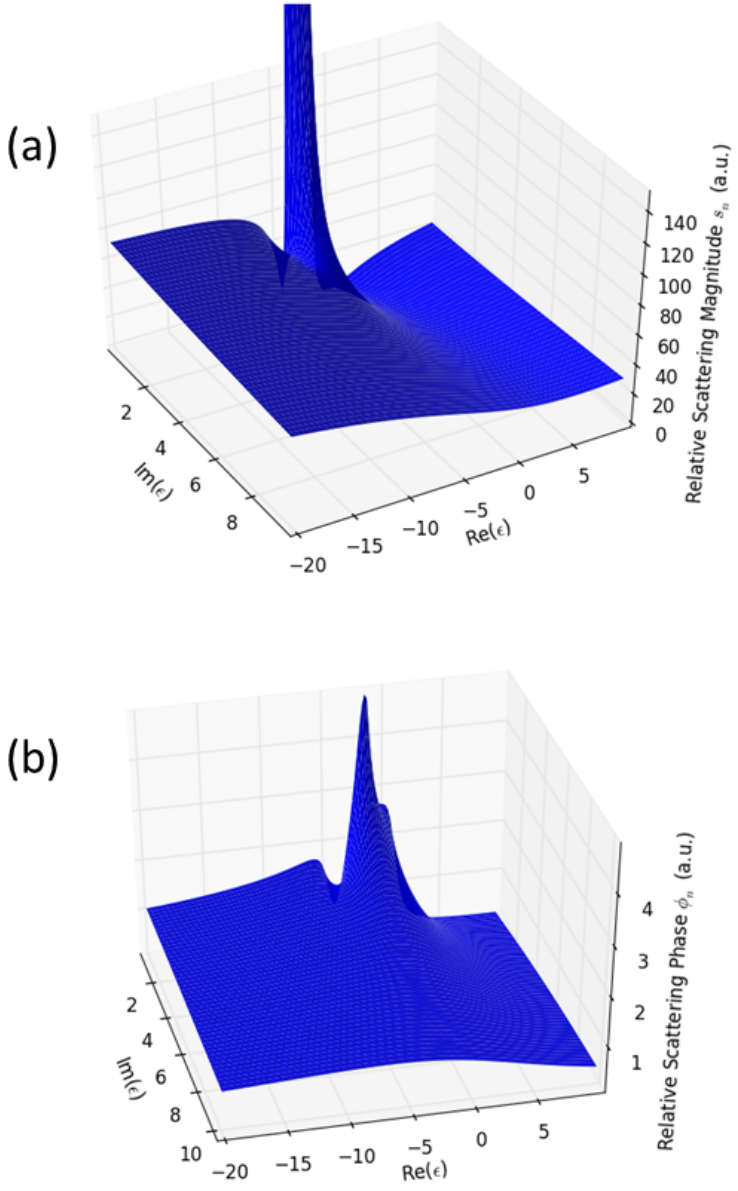

Figure 3.11: The signal surface for the monopole model, η', at the second harmonic of the probe oscillation frequency.

$$\mathrm{NFBDR} = \frac{\sigma_{\mathrm{N,n}}}{\sigma_{\mathrm{B,n}} + \mathrm{D_n}}. \tag{3.18}$$

To make use of this equation, the monopole model can be used to find $\sigma_{N,n}$ and $\sigma_{B,n}$. The value of D_n - the Fourier component of D at the frequency of the n$^{\mathrm{th}}$ harmonic - can then be varied to observe its effects on the measurement. It is instructive, however, to first set D_n to zero to compare the effects of higher harmonic demodulation on the near-field and background scattering coefficients (see figure 3.12).

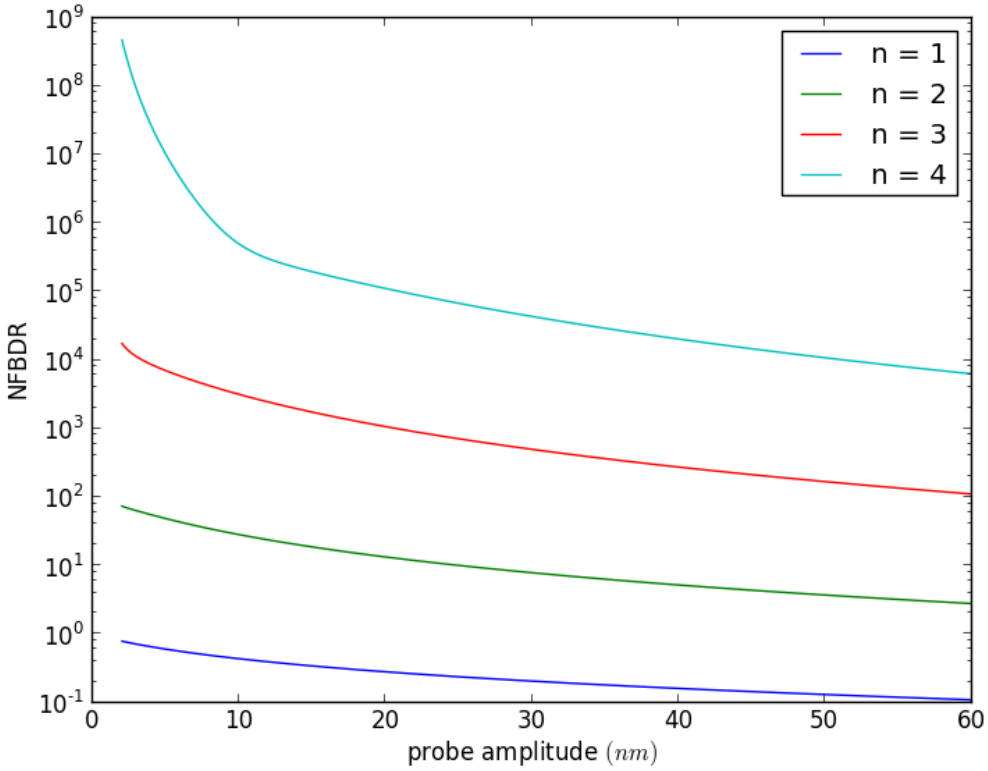

Figure 3.12: The NFBDR at different harmonics for light at 10μm on a gold surface with D_n set to zero. As one moves up the harmonics, the useful near-field signal provides a bigger fractional contribution to the overall scattering than does background signal. This is the premise on which the pseudo-heterodyne detection technique works (section 2.5).

Without noise, therefore, it is clear that the higher the harmonic, the better the NFBDR of the measurement. However, it is important to note that the NFBDR is a *ratio* and so hides the fact that the *absolute* signal value drops as n increases. When the noise contribution D_n is considered, this becomes apparent as a minimum required near-field scattering signal $\sigma_{N,n}$ is set by the noise floor.

For the purposes of modelling, the disturbance contribution D_n is set as a fraction of the largest contribution to the overall scattering coefficient σ. This is the DC component of the background

scattering, $\sigma_{B,0}$. If the system is assumed to have a dynamic range of 100dB, where the dynamic range is defined by the detected photovoltage as:

$$dB = 20 \log_{10} \left(\frac{V_{max}}{V_{noise}} \right), \tag{3.19}$$

this corresponds to a $\sigma_{B,0} : D_n$ ratio of 10^5. The disturbance level D_n is therefore set at $10^{-5}\sigma_{B,0}$ for figure 3.13.

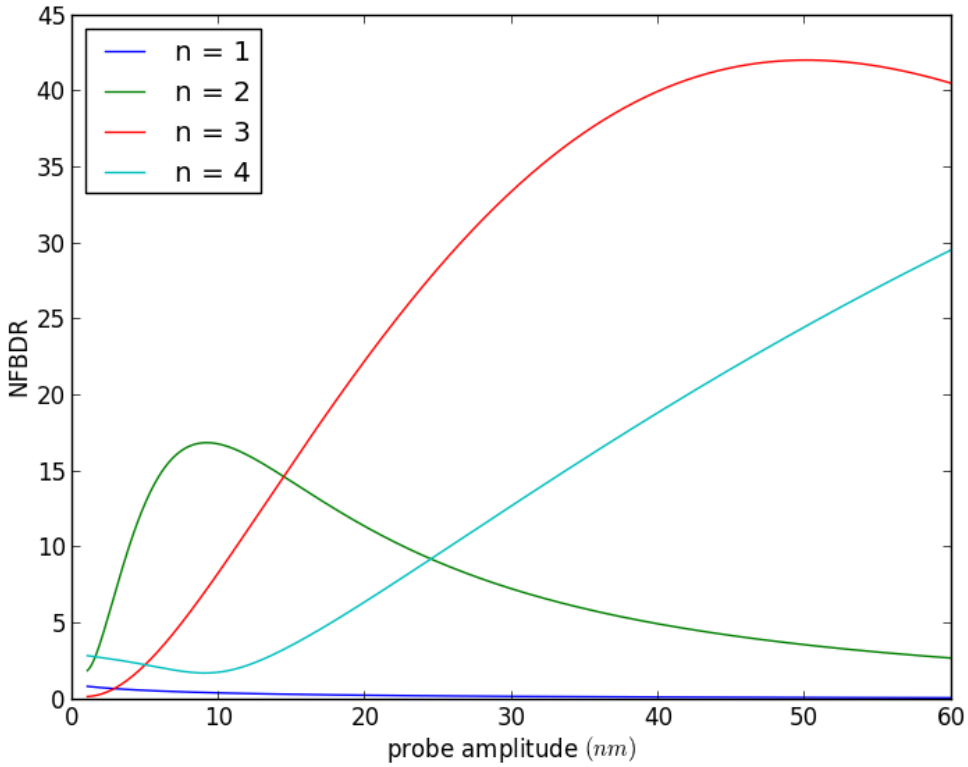

Figure 3.13: The NFBDR at different harmonics for light at 10μm on a gold surface where D_n has been set to 10^{-5} the value of the largest contribution to the overall scattering coefficient, $\sigma_{B,0}$. In contrast to figure 3.12, the highest harmonic no longer provides the largest NFBDR (also note the change from a logarithmic to a linear scale). Instead, the interplay between the necessity to be above an absolute signal level (as defined by the noise floor) and the requirement to suppress the background scattering in favour of the near-field signal means that the second or third harmonics now provide the best options. Which of the two depends on the size of the probe oscillation amplitude.

Depending on the probe oscillation amplitude, it is clear that the second or third harmonics are the best choices for maximising the NFBDR. In reality, the achievable dynamic range is often somewhat less than 100dB and so figure 3.13 is an ideal case. The principle still stands, however, that for every type of sample surface, illumination wavelength and probe amplitude configuration there is a harmonic that will provide the best near-field signal. As a rule of thumb, the third harmonic is best for visible and near-infrared applications while the second harmonic is sufficient

in the mid-infrared. This wavelength dependence of the NFBDR is largely due to the wavelength dependence of the background scattering $\sigma_{B,n}$:

$$\sigma_{B,n} \propto \left(\frac{A}{\lambda}\right)^n \tag{3.20}$$

where A represents the probe vibration amplitude. As the wavelength increases, therefore, the background scattering strength rapidly decreases.

3.5.2 Defining the SNR for s-SNOM Measurements

The NFBDR, although related to the measurement signal-to-noise ratio (SNR), differs in one important way - the measured signal in an s-SNOM measurement is not the near-field scattering $\sigma_{N,n}$ directly, but rather one of the sidebands created by the pseudo-heterodyne detection system $C_{n,m}$. Assuming that the harmonic n is large enough that the background scattering is negligibly small, the SNR can therefore be defined as:

$$\mathrm{SNR} = \frac{\mathrm{C_{n,m}}}{\mathrm{D_n}} \tag{3.21}$$

where $C_{n,m}$ represents the sideband value at the frequency of the $\mathrm{n^{th}}$ harmonic plus m times the vibration mirror frequency, and $D_{n,m}$ represents the disturbances at the same point in frequency space.

Alternatively, the SNR can be defined using the pixel values - s_n and ϕ_n - which are produced from the values of two sidebands using equations 2.5 and 2.6. Since the value of $D_{n,m}$ does not vary over frequency separations of a few hundred Hertz, the noise at any given harmonic n is the same at all the sidebands m (i.e. $D_{n,1} = D_{n,2} ... = D_n$). Given that the noise, in effect, represents an uncertainty in the measurement of the sidebands, the propagation of errors formula can be used to find how the uncertainty in a measurement of a single sideband affects the uncertainty in s_n or ϕ_n. For a function f of two variables, $f(x, y)$, the error in f, written e_f can be found by:

$$e_f^2 = \left(\frac{\partial f}{\partial x}\right)^2 e_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 e_y^2 \tag{3.22}$$

Setting x and y to $C_{n,1}$ and $C_{n,2}$ respectively, and e_x and e_y both to $D_{n,m}$, the resulting error in the near-field scattering's amplitude and phase is also given by D_n. The SNR of the pixel values - denoted $\mathrm{SNR_s}$ and SNR_ϕ - are therefore given by:

$$\mathrm{SNR_s} = \frac{\mathrm{s_n}}{\mathrm{D_n}} \text{ and } \mathrm{SNR}_\phi = \frac{\phi_\mathrm{n}}{\mathrm{D_n}} \tag{3.23}$$

3.5.3 Increasing the SNR

In essence, the pseudo-heterodyne detection technique is a method for suppressing the unwanted components in a measurement; by only measuring a very small frequency range around a sideband, a huge amount of extraneous noise is excluded (although at the expense of increasing the

measurement time). In general, the voltage trace in time from the detector $u(t)$ can be described with two parts; the useful signal $x(t)$ and a noise (or disturbance) contribution $d(t)$:

$$u(t) = x(t) + d(t) \tag{3.24}$$

In figure 3.14, the typical Fourier transform of the noise contribution to a measurement has been overlaid onto the spectrum from a pseudo-heterodyne setup (see figure 2.5). When the Fourier component of the noise at a given frequency, $D[f]$, becomes larger than that from the detection system, measurements are no longer possible. Further details on the causes of this noise contribution are given in chapter 7.

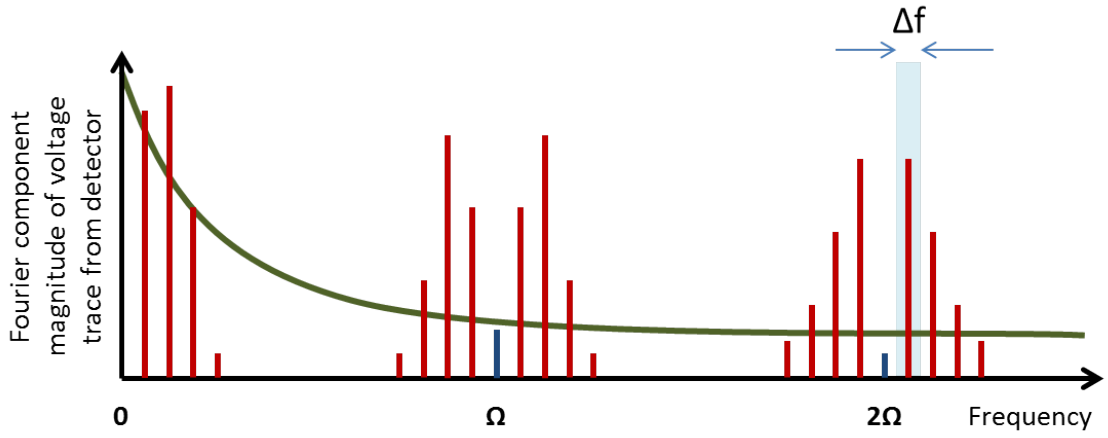

Figure 3.14: Representation of the frequency components from different sources: **blue** lines showing the signal at the tip vibration frequencies; **red** lines showing tip frequencies splitting into sidebands upon mixing with the phase modulated reference beam; **green** line representing the typical Fourier transform of a measurement's noise contribution.

In real measurements - due to a non-infinite integration window - each point on the spectrum in fact represents a small spread of frequencies Δf, the size of which is defined by the integration period at each pixel as $\Delta f = 1/T$. The larger this integration time is made, the smaller the frequency spread and therefore the lower the noise contribution $D[f]$ as the measurement bandwidth has been reduced. The larger integration time effectively splits the spectrum up into more frequency "buckets" and, given that the amount of energy in the noise spectrum does not change, the noise contribution in each is reduced. The simplest way to increase the SNR, therefore, is simply to increase the integration period at each pixel.

Chapter 4

Plasmonics in the Near-Infrared

"Look the beast in the eyes..." **Prof. Chris Phillips**

4.1 Introduction

The field of plasmonics is a vibrant area of research with an enormous range of potential applications; to date it has found roles in solar cells, chemical catalysis and cancer treatments to name a few [119–121]. This chapter will focus on the imaging of localised surface plasmon resonances (LSPRs - see section 3.3.1), which, due to their ability to "squeeze" light into substantially sub-wavelength volumes, are of particular interest in fluorescent imaging, emission enhancement and optical sensing applications (see section 3.3) [103, 122–124]. As s-SNOM is capable of imaging electric fields on a surface with 10nm resolution, it is one of the few techniques capable of directly detecting LSPR field distributions [45].

The chapter starts with a brief overview of the experimental setup required for plasmonic imaging (so that contrast caused by chemical differences across the sample is suppressed - see section 3.3) before looking at three different types of structures: the first is a simple gold colloid, used as a proof-of-principle measurement; the second is a dolmen structure, for which the difficulties in using s-SNOM in reflection (rather than transmission) mode become apparent; and the final structure is a cruciform pentamer oligomer, which exhibits polarisation invariance in its far-field response but not in its near-field distribution.

4.2 Experimental Setup for Plasmonic Samples

As mentioned in section 3.3, "active" samples, which manipulate the incoming light, require a particular experimental setup to prevent an s-SNOM imaging the chemical contrast as well as the field caused by the plasmon resonance. This is usually achieved by using the *cross-polarisation* technique [77, 125–131], which exploits the fact that an s-SNOM probe - due to its antenna-like properties - is a better scatterer for electric fields in the out-of-plane (z) direction [32]. The scheme itself is shown in figure 4.1. The principle of the technique is to illuminate the sample with s (or

TE) polarised light. The probe's antenna function is very weak in this direction, so the probe itself does not scatter the incident light directly (which would lead to chemical contrast in the image). Upon interaction with the sample's structure, however, the plasmon resonance forms some field component in the z (vertical) direction, which corresponds to p (or TM) polarised light. The probe is much more sensitive to this newly formed field, and so it is scattered into the far-field and onto the detector. Hence it is the z-component of the plasmon's electric field distribution that is mapped.

The cross-polarisation scheme (s in, p out) can be modified slightly if the in-plane plasmonic field enhancement is very large (such as for gap modes between antennae [132]). In this case - despite the weak antenna properties of the probe for this field direction - the in-plane field components are scattered effectively enough to be detectable in the far field. For this setup (s in, s out), therefore, it is the in-plane components of the plasmon resonance that are imaged. This technique is used in section 4.5 later in this chapter.

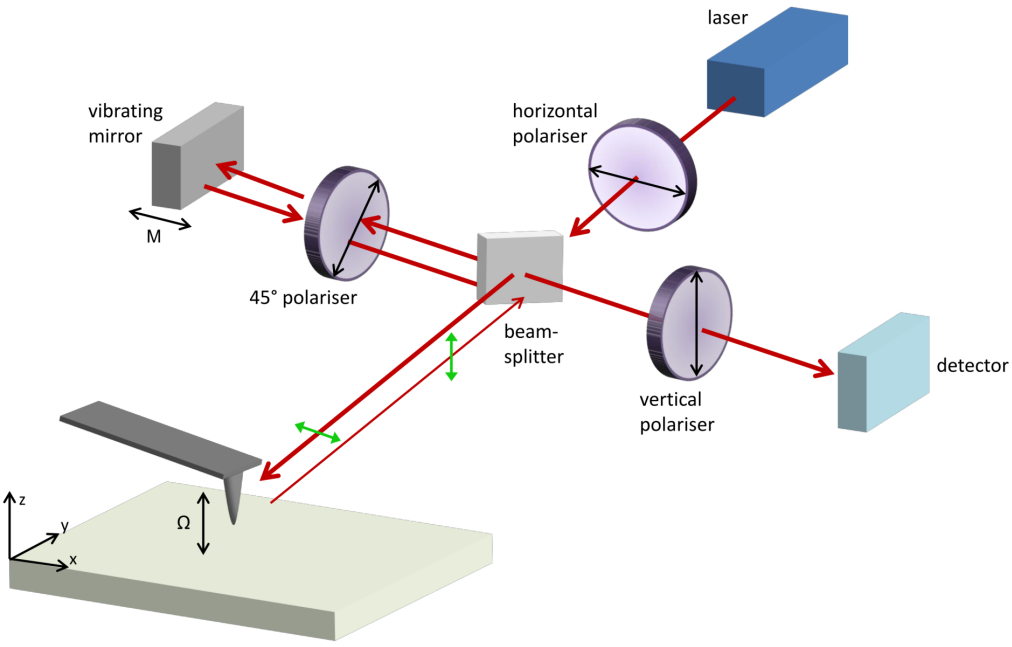

Figure 4.1: Optical setup for imaging plasmonic samples using the cross-polarisation scheme. This illuminates the sample with s-polarised light and detects only p-polarised light. In this way, any light scattered by the probe has necessarily been manipulated by the structure on the sample surface, and hence chemical contrast is suppressed. The incident and backscattered polarisation states are represented by the green arrows.

For the experiments which follow, the light source was a titanium-sapphire laser (*Spectra-Physics, 3900S*) [133]. This was optically pumped with a frequency-doubled solid state laser (itself

diode pumped) at 532nm (*Coherent, Verdi V5*). The titanium-sapphire laser provides continuous tuning between 700 and 1000nm. Its light was delivered to the s-SNOM setup via a single mode optical fibre. This process rotates the polarisation of the light by some arbitrary angle, so a $\lambda/2$ waveplate was used to bring it to the horizontal. The light then passes through a linear polariser to remove all but the s-polarised light. Once it reaches the beam splitter, it follows one of two paths; one goes through to the probe. Anything scattered back from the probe will have a component in the vertical direction for the reasons outlined at the beginning of this section. The light which goes into the reference arm, however, passes through a polariser set at 45° between horizontal and vertical. This introduces a component of the reference beam in the vertical direction which can then interfere with the p-polarised signal scattered from the probe, allowing the pseudo-heterodyne detection technique to work. Just before the detector there is a third polariser which is set in the vertical direction, making sure that only the p-polarised light - the source of which can only be the plasmon - is detected. The s-SNOM probe itself was platinum coated with a mechanical resonance frequency of approximately 250kHz (*Arrow NCPt, NanoWorld*).

To demonstrate the necessity of the cross-polarisation setup (or at least the need for s-polarised illuminating light), figure 4.2 shows an image of a cruciform pentamer structure (see section 4.5) that is illuminated and detected using p-polarised light. As this sample supports a plasmon resonance, there are two sources of contrast; the enhanced electric field of the plasmon (seen as "halos" around the discs) and the material differences of the gold structures on a glass substrate (seen as the bright areas within the halos). This can lead to difficulties in unambiguous image interpretation (although not in this particular case).

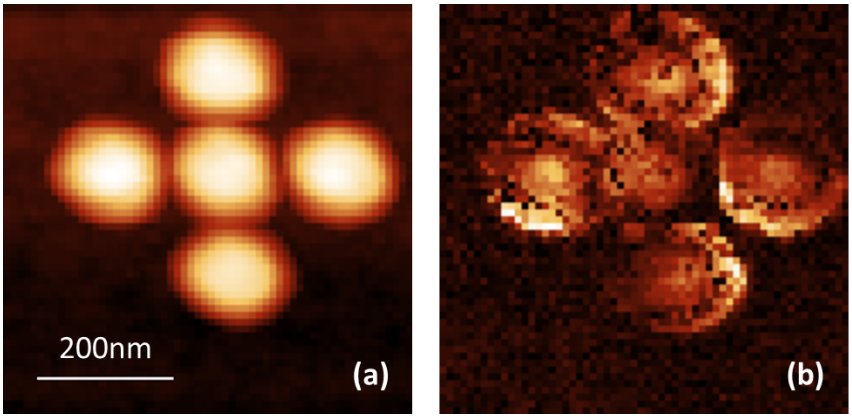

Figure 4.2: (a) AFM and (b) third harmonic of the magnitude of the near-field scattering s_3 images of a cruciform pentamer (see section 4.5) illuminated and detected with p-polarised light. As this sample exhibits a plasmonic resonance there is some difficulty in interpreting the image; some of the contrast comes from the material differences of the gold nanostructures on a glass substrate, whereas the second source of contrast is the enhanced electric fields of the plasmon resonance.

The key to the cross-polarisation scheme's success is the s-polarisation of the incident light. The mapping of chemical differences relies on a strong field enhancement in the small volume between the probe and the sample. When the incident field is polarised parallel to the sample, the "lightning rod" function of the tip is lost (section 2.4) and so the scattered light is very weakly

affected by any probe-sample interaction which would lead to material contrast.

4.3 Gold Colloids

Gold colloid samples were used as an initial test of the s-SNOM in conjunction with the titanium-sapphire laser (set up in the cross-polarisation scheme) [134]. These are approximately spherical particles of gold on a silicon substrate that fulfilled the requirements of such a proof-of-principle experiment - they are easily prepared, and they support the simplest of LSPRs, namely a dipole resonance.

4.3.1 Modelling and Far-Field Spectra

Using a commercially available finite-difference time-domain (FDTD) method software package (*Lumerical*), the field around a 200nm diameter colloid was simulated. The output of an xy-field intensity monitor placed in the centre of the sphere is shown in figure 4.3. A simple dipole mode - where charges oscillate from one side of the sphere to the other at the same frequency as that of the incident light - is visible as two bright lobes along the polarisation axis at both the simulated wavelengths of 633 and 735nm.

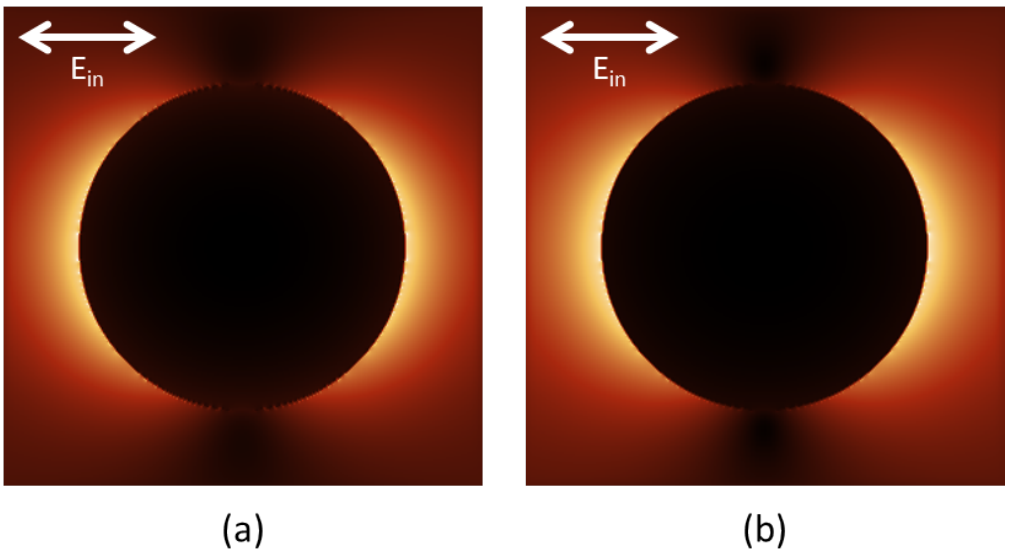

Figure 4.3: Simulated electric field magnitude around a colloidal sphere of 200nm diameter on a silicon substrate at (a) 633nm and (b) 735nm. The dipole mode supported by the structure is clearly visible at both wavelengths. The incident light's polarisation is represented by a white arrow in both parts.

The simulations are supported by dark field spectra [135] of different colloidal particles shown in figure 4.4. These spectra give a measure of the scattering strength of the particles at different frequencies and show that approximately 680nm is the central wavelength of the dipole mode. Each particle, however, is a slightly different shape and size, so no two spectra are exactly alike. The

simulated wavelengths of figure 4.3 (633 and 735nm), therefore, fall on either side of the central 680nm peak. Despite this, their field profile is very similar.

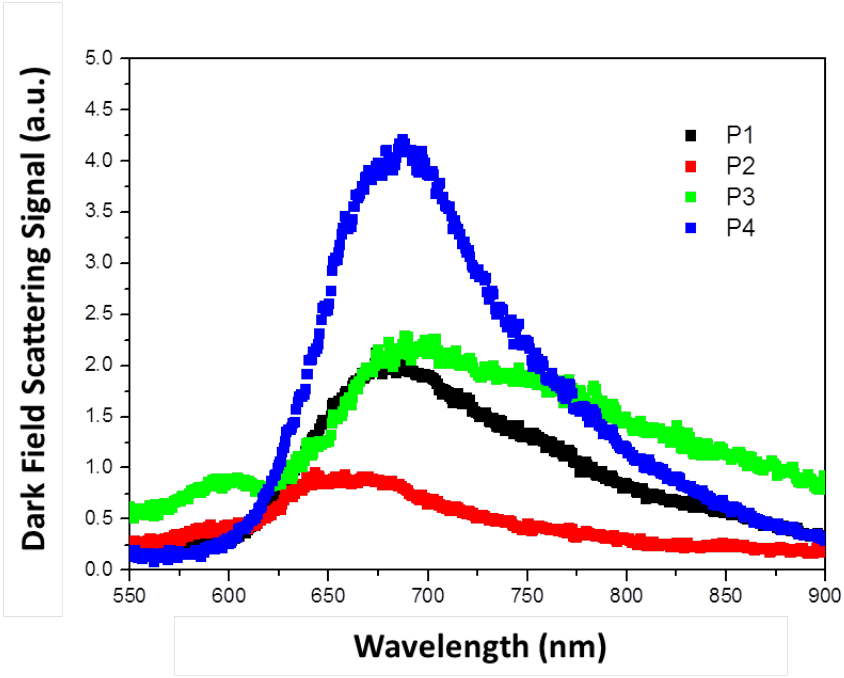

Figure 4.4: Dark field spectra (which give a measure of the scattering strength at different wavelengths) of four different gold colloidal particles (labelled P1-4) on a silicon substrate. These show that the dipole mode supported by the structure is centred around approximately 680nm, with variations in scattering strength and spectral position caused by slight discrepancies in the size and shape of the particles.

4.3.2 Near-Field Images of Colloidal Spheres

Figure 4.5 shows the shows the topography (a,d) as well as the near-field scattering magnitude s_3 (b,e) and phase ϕ_3 (c,f) recorded from two colloidal particles (which are approximately equal in size) at the two wavelengths presented previously; 633nm with a helium-neon laser and 735nm with a titanium-sapphire laser. In both cases, the experimental setup was as shown in figure 4.1.

At both wavelengths, the dipole mode is clearly present, with two areas of stronger near-field scattering signal s_3 at each edge of the particle along the axis of the incident radiation's polarisation. The optical phase, ϕ_3, shows a 180° shift from one side of the particle to the other. This can be interpreted by looking at the charge distribution throughout the particle, as will be explained in section 4.3.3.

Experimental deviations from a perfect dipole mode - where each lobe is equally strong - are likely caused by non-spherical particles. In image (b) of figure 4.5 where this imperfection is most pronounced, the AFM image (a) clearly shows a particle that is not perfectly round. Note that for the phase images, the scale loops - a pixel value of 0 has the same colouring as 2π to represent the cyclical nature of the phase information.

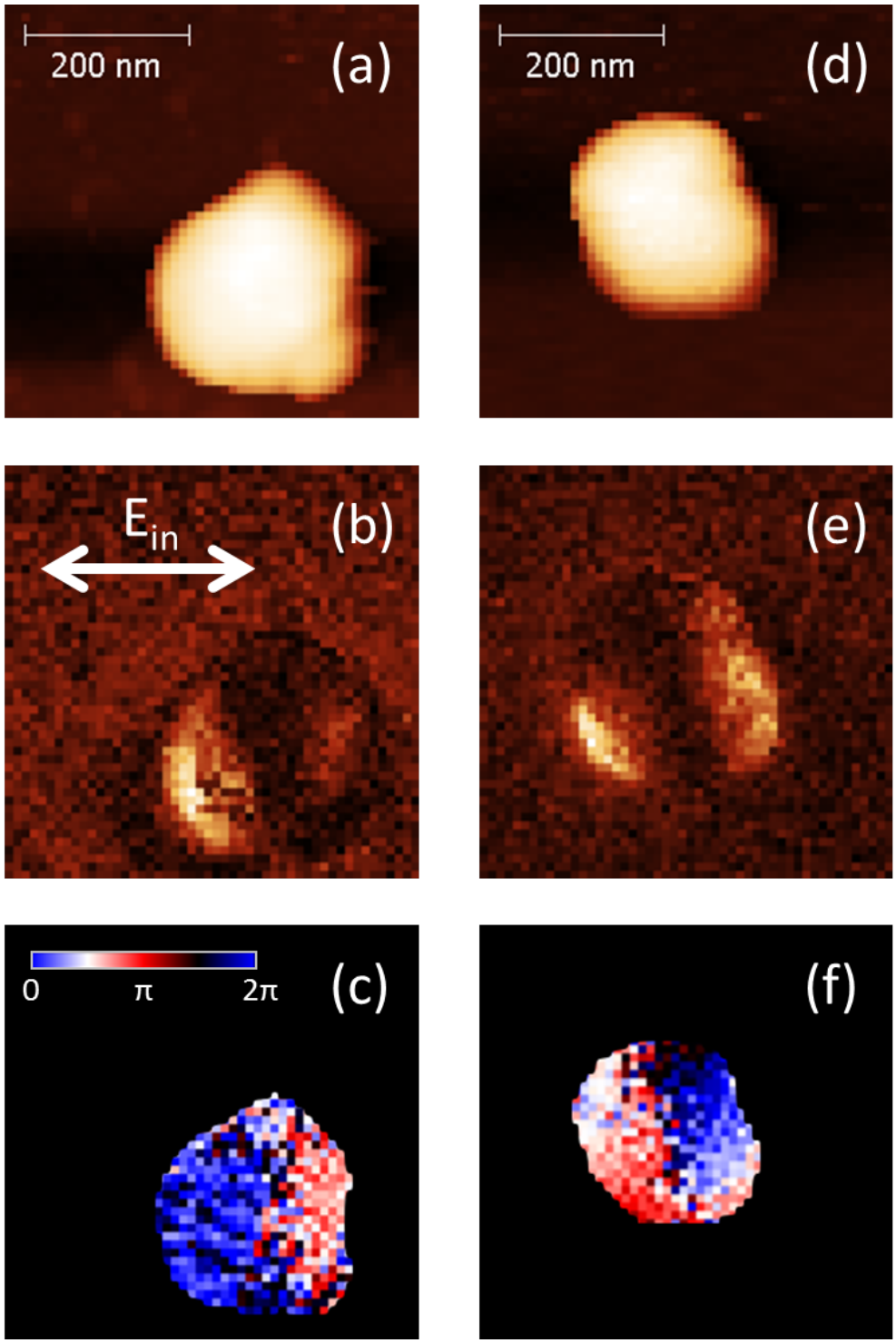

Figure 4.5: (a,b,c) AFM, optical magnitude and optical phase (n=3) for incident light at $\lambda = 633$nm. (d,e,f) Similar images taken at $\lambda = 735$nm. The incident radiation is horizontally polarised relative to the particles as indicated by the white arrow in figure (b).

4.3.3 Aside: Interpretation of Plasmonic s-SNOM Images

In the literature, many papers compare experimental s-SNOM images to that of the simulated charge distribution in a plasmonic structure [125, 127, 136]. The aim of this section is to provide a justification for doing this. It has been shown that, for gold colloidal spheres, there is variation in the particles' near-field extinction cross-section around the maximum of the dipole resonance's spectral position - it is smaller at lower energies (longer wavelengths) and greater at higher energies (shorter wavelengths) [137, 138]. This is not detectable in our dark-field scattering measurement (figure 4.4) because the incident and scattered radiation is not coherent - interference between the excitation beam and light subsequently scattered from the particles does not occur meaning that the variation in the extinction is not picked up.

The explanation for this variation in the extinction lies with the charge distribution - on low energy side of the resonance the charges align themselves along the direction of the incident field and on the high energy side they align themselves opposite the field (see figure 4.6). For wavelengths longer than the central peak of the dipole resonance, therefore, the scattered field constructively interferes with the incident field causing a decrease in the extinction. For shorter wavelengths, the scattered field destructively interferes with the incoming radiation and the extinction is increased.

The optical contrast of an s-SNOM image of a plasmonic sample can be related to this charge distribution. A plasmon resonance is not stationary in time; the charges oscillate backwards and forwards throughout the structure as the electric field of the incident light changes direction (around 10^{15}Hz). In the simple case of a gold colloidal sphere, therefore, each position on its surface has a time varying charge distribution, as shown in figure 4.7.

The variation of the charge distribution with time at two points, $P1$ and $P2$, is shown in the plot. The s-SNOM optical signals depend on this in two ways:

- The magnitude of the near-field scattering signal σ_n depends on the amplitude of the charge density oscillations. From Gauss' Law, the greater the charge density, the greater the field around the structure, and this leads to a larger scattering signal s_n. In the colloidal sphere example, the charge density is greatest at the edges of the spheroid, and so the scattering magnitude is greatest there i.e. s_n at $P1$ is greater than s_n at $P2$.

- The phase of the scattering signal ϕ_n depends on the relative phase difference between charge oscillations at different points within the sphere. In figure 4.7 for example, $P1$ and $P2$ do not experience their maximum positive charge density at the same time - there is a phase delay between them (180 degrees in this case). Thanks to the interferometric nature of an s-SNOM measurement, this delay is detectable in a way that wouldn't be possible for incoherent light [139].

Note that, due to the cyclical nature of the phase measurement (i.e. a phase difference of 2π amounts to the same thing as a phase difference of 0), the s-SNOM phase contrast can only be interpreted within a single image. The differing appearances of the phase measurements ϕ_n in the $\lambda = 633$nm and $\lambda = 735$nm (images (c) and (f) of figure 4.5) have been presented as such to illustrate the changing charge distribution relative to the incident field at frequencies either side of

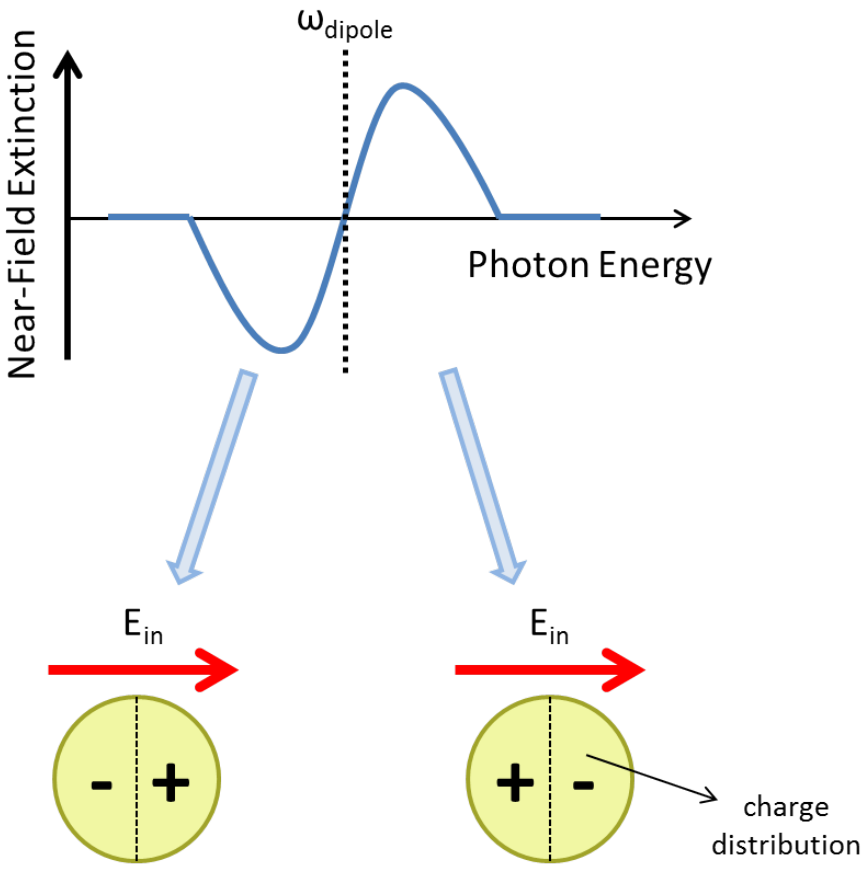

Figure 4.6: The near-field extinction for colloidal particles, adapted from [138]. On the low energy side of the dipole resonance, the charges within the structure align themselves with the incident field causing a decrease in the extinction, because the scattered radiation constructively interferes with the incident light. On the high energy side, however, the charges are aligned opposite to the incident field and the extinction is increased due to destructive interference.

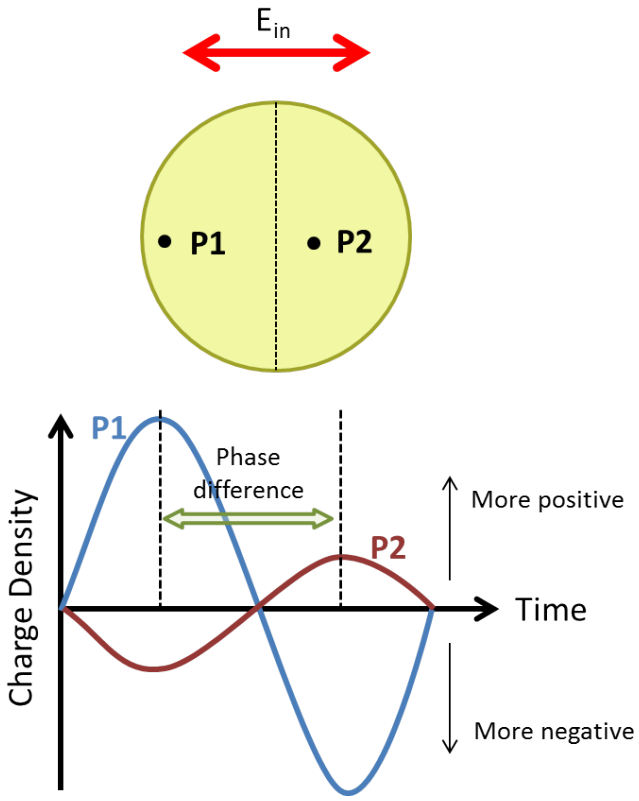

Figure 4.7: Comparison of the charge density with time at two points on a colloid's surface. The phase difference between two points on opposite sides of the colloid is measurable due to the interferometric nature of the pseudo-heterodyne detection setup. Also, the closer the point to the edge of the sphere, the greater its maximum charge density.

the dipole resonance. It is not possible, however, to draw any conclusions about the actual charge distribution (i.e. which areas of the sphere are positively charged and which areas are negatively charged) from these images alone - all that can be taken from them is that the phase difference between the two sides of the sphere is 180 degrees at both wavelengths. It is only via the extinction measurement that the alignment of the charges relative to the incident field can be deduced.

4.4 Retardation Effects

One of the limitations of our experimental setup (figure 4.1) is the angle at which the incoming beam illuminates the surface. It is not at normal incidence; instead it arrives at an oblique angle - 60 degrees from the surface normal and the light which is backscattered along the same path is collected (reflection mode). As a result, the electric field experienced by the structure is not uniform in the direction of the light's propagation because of retardation, as shown in figure 4.8. This is due to a path difference between the rays striking the near and far sides of the sample.

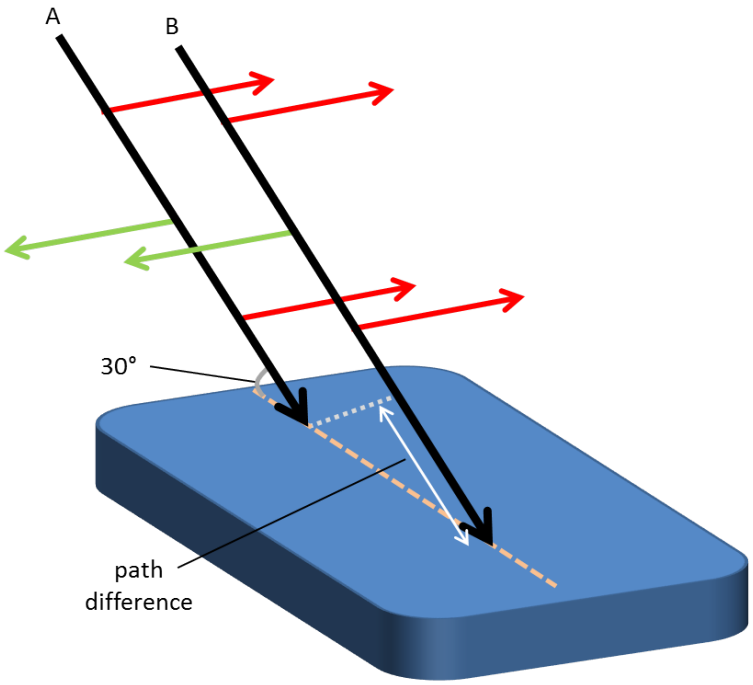

Figure 4.8: Diagram showing the retardation effect of two rays, labelled A and B, due to the path difference between then. This is caused by the non-normal angle of incidence (30° from the substrate's plane in our setup), where the red and green arrows represent the electric field directions of the rays. Plasmonic structures, therefore, may not experience a uniform field when imaged in reflection mode. The uniformity of the field illuminating a given structure depends on its size and can be calculated by simple geometry.

If we decide, for example, that the maximum allowed path difference for the rays hitting the outermost parts of the structure is $\lambda/10$, we can use simple trigonometry to calculate the maximum structure size x for which retardation effects are not problematic. With the angle of incidence (from

the surface normal) written θ, this is given by:

$$x = \frac{\lambda}{10\sin(\theta)} \tag{4.1}$$

For a typical set of parameters; θ set at 60 degrees and λ at 750nm, this yields a maximum structure size of 85nm. For all but the simplest of plasmonic structures, therefore, retardation effects are likely to enter into reflection mode imaging. The problem can be solved, in principle, by imaging in transmission mode, where the sample is illuminated at normal incidence from below. The s-SNOM probe would then scatter the fields created by the plasmon resonance into the far-field as it does in reflection mode.

4.4.1 Dolmen Structures: A Cautionary Tale

As an example of the influence of retardation effects, simulations and experimental images of the field around dolmen structures are presented. Dolmens are composite nanostructures made up of two parts; bar dimers and a monomer. They are of particular interest for the plasmonic Fano resonance they support. A spectrally narrow quadrupole mode (on the dimer) interferes with a much broader dipole mode on the dimer [136, 140–145]. A scanning electron microscope (SEM) image of the structure is displayed in figure 4.9.

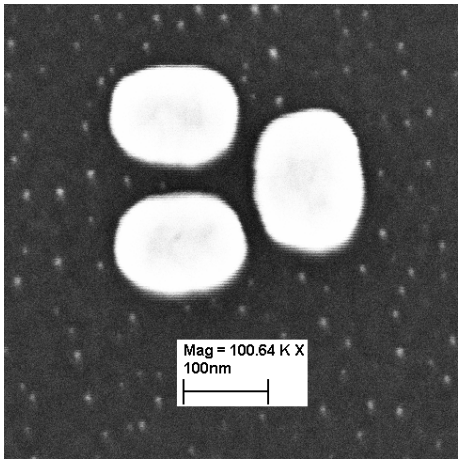

Figure 4.9: Scanning Electron microscope (SEM) image of a dolmen structure (made of gold) on a glass substrate.

The far-field scattering response of a series of dolmen structures with increasing dimer size is shown in figure 4.10. As the dimer becomes larger, the spectral position of the sharp quadrupole shifts to longer wavelengths, and this is seen in the dark field scattering response by a shift in the position of the characteristic Fano dip. The particular structure that is presented in this section is the uppermost (green) trace in figure 4.10, with its Fano dip minimum at $\lambda = 760$nm.

Dolmen structures are particularly interesting to image with an s-SNOM because the near-field distribution varies considerably as the excitation wavelength changes and the structure exhibits its Fano resonance [146]. This can clearly be seen in parts (a) and (b) of figure 4.11, which show

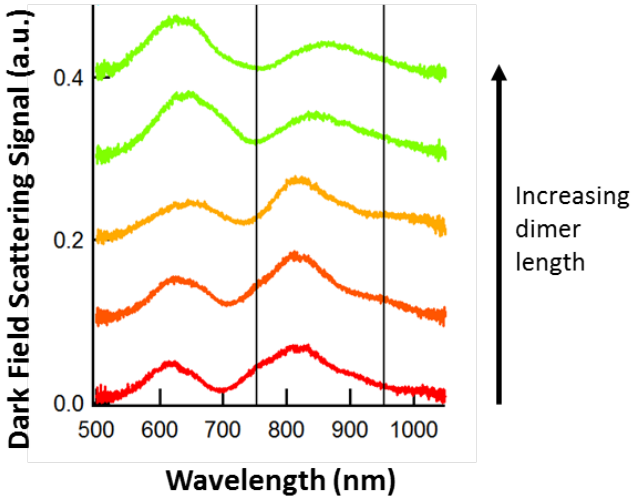

Figure 4.10: Dark field scattering signals from dolmen structures with increasing dimer length. The larger the dimer, the longer the wavelength of the quadrupole resonance, and this is seen as a shift in the position of the characteristic Fano dip. The curves from each dolmen have been separated along the y-axis for clarity. The dolmen is made up of a two bar dimer (left) and a single bar monomer perpendicular to the dimer (right).

the z-component of the electric field, when the incoming light is normally incident, at 10nm above the structures at $\lambda = 750$ and $\lambda = 925$nm respectively. In an s-SNOM experiment in reflection mode, however, the light is not normally incident. For broadband simulations, the FDTD method cannot be used to replicate angled beams, so instead parts (c) and (d) of figure 4.11 show the z-component of the electric field, at 10nm above the structure, for light at grazing incidence (i.e. parallel to the substrate). In this case, it is clear that there is not a significant redistribution of the field between the different wavelengths. Physically, this is due to the fact that the two modes - the quadrupole on the dimer and the dipole on the monomer - do not interfere with each other for light incident at this angle, and so no Fano resonance is created.

The corresponding experimental images are shown in figure 4.12. As is expected due to the oblique angle of incidence, the experiment shows more similarity with the simulations at grazing incidence; the field distribution is not significantly changed between the two excitation wavelengths as the Fano interference effect does not take place.

These dolmen structures, therefore, highlight the problem of retardation in the reflection mode of s-SNOM imaging; it is clear that larger and more complex plasmonic structures cannot be examined without changing to a transmission mode imaging setup.

4.5 Pentamer Structures

Please note, this section borrows heavily from a paper titled "Plasmonic Cavities with Rotational Symmetry: Polarization Invariant Far-Field Properties vs. Changing Near Field" submitted to ACS Nano in September 2013 with Mohsen Rahmani, Ben Hopkins, Chris Phillips, Stefan Maier

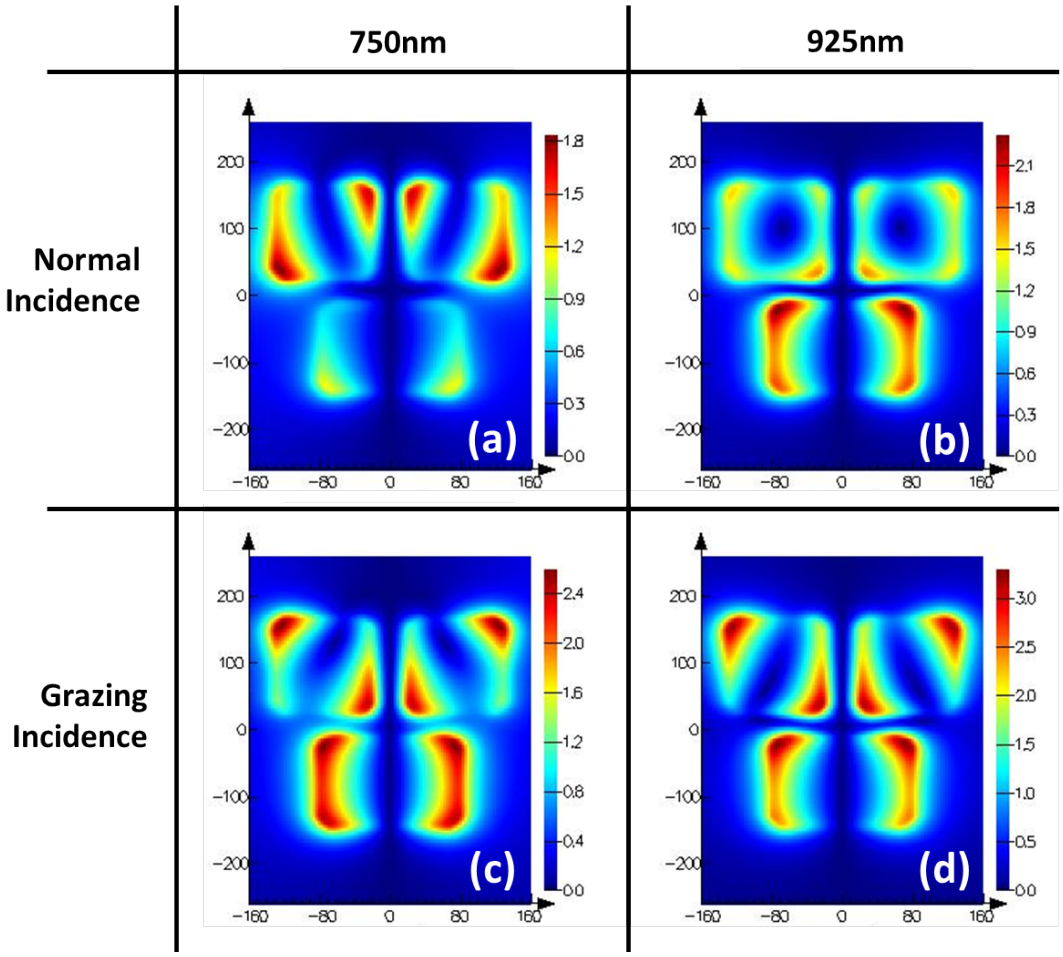

Figure 4.11: Simulated profiles of the z-component of the electric field 10nm above the sample surface. The scale of the x and y axes is nanometres. For light at grazing incidence, the Fano resonance is not excited and so a significant redistribution of the field profile on the structure is not seen.

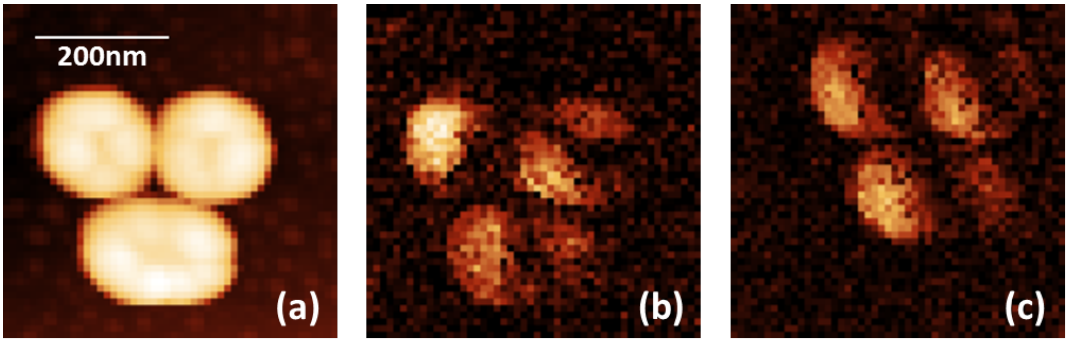

Figure 4.12: (a) AFM, (b) measured optical magnitude (n=3) at $\lambda = 750$nm and (c) optical magnitude (n=3) at $\lambda = 925$nm of a gold dolmen sample on a glass substrate. The image in (c) is displaced relative to (a) and (b) due to drift - it is, however, a scan of the same structure.

and Andrey Miroschnichenko.

The final structure to be presented in this chapter is the cruciform pentamer oligomer. An oligomer is the general term for a small cluster of individual nanostructures that have been symmetrically grouped together [147–149]. Complex plasmonic structures generally owe their far-field and near-field properties to the coupling of the resonances of their constituent particles [150] - this can lead to some extraordinary behaviour such as subradiance, superradiance and asymmetric lineshapes [146, 151, 152]. These properties also render them highly susceptible to their environment - and hence well suited to sensing applications - but the coupling of different resonances also means that complex plasmonic structures are highly dependent on the polarisation of the incident light.

Oligomers, among other designs [153, 154], are promising candidates for bucking this trend of polarisation dependence. Unlike other complex structures, their far-field response is not necessarily dictated by their near-field distribution. Instead, this section looks at how their far-field response can be designed to remain polarisation-independent, while their near-field distribution varies significantly with the field direction of the incident light [155]. The key to this behaviour is in the rotational symmetry of the structure - indeed it has been theoretically predicted that there is a direct link between the polarisation independence of the far-field response of a structure (i.e. its extinction, absorption and scattering cross-sections) and the rotational symmetry of the structure [156]. Cruciform pentamers are one example of a much larger family of structures that fulfil this requirement.

It is worth noting that although previous papers have looked at the near-field distribution of oligomers in the mid-infrared [136], this is the first report of direct experimental measurements in the near-infrared.

4.5.1 Theory: Oligomer Responses

The optical response of the cruciform pentamer structures can be studied using the coupled dipole approximation (CDA). This model - which describes a complex structure as a sum of polarisable points - can be used to explain why the far-field response is polarisation-invariant and why the near-field distribution can vary while still eliciting the unchanging far-field behaviour. In fact, two approaches can be taken to do so; a top-down method, which starts with the relationship between the incident field and the far-field cross-sections, and a bottom-up method, which breaks down the incident field as a set of eigenmodes of the system.

Top-Down Approach

For any scattering system described with the CDA, the far-field cross sections can be described as the expectation value of a corresponding far-field matrix $\widehat{M}$ with the incident field $\mathbf{E}_0$:

$$\sigma_f = \mathbf{E}_0^t \widehat{M_f} \mathbf{E}_0, \tag{4.2}$$

where σ_f stands for any of the three far-field cross-sections; extinction, scattering and absorption. It has been shown that, in this form, $\widehat{M_f}$ must commute with the group theory symmetry operators

that correspond to the structure's symmetry [156]. $\widehat{M_f}$, therefore, provides a link between the far-field properties of a structure and its symmetry. For the particular case of the cruciform pentamer, the symmetry group is the cyclic tetragonal group $C_{4\nu}$ where the symmetry operations are based around the normal axis of the structure. Applying Schur's Lemma then says that the only way $\widehat{M_f}$ can commute with all the elements of the $C_{4\nu}$ is if $\widehat{M_f}$ is a multiple of the identity matrix:

$$\widehat{M_f} = \begin{pmatrix} k_f & 0 \\ 0 & k_f \end{pmatrix} \tag{4.3}$$

As such, the cross-sections of the pentamer must be proportional to the magnitude of the incident field only, and are subsequently completely independent of the polarisation. Note that this result doesn't place any requirements on the structure's near-field distribution.

Bottom-Up Approach

To consider the near-field origins of the far-field polarisation invariance, the incident field must be analysed in terms of the eigenmodes of the structure's optical response [157]. These eigenmodes can be investigated analytically without having to explicitly solve for them. If there exists some basis of linearly independent eigenmodes of the cruciform pentamer structure $\{|\nu_i\rangle\}$, the incident field polarisation can be described as a unique combination of these eigenmodes. An x polarised field, for example, can be described as:

$$|\mathbf{E}_x\rangle = \sum_i a_i \, |\nu_i\rangle \tag{4.4}$$

Any response of the system $|F(\mathbf{E}_x)\rangle$ (where $F(\mathbf{E}_i)$ represents the response F to an incident field $\mathbf{E}_i$), therefore, is given by:

$$|F(\mathbf{E}_x)\rangle = \sum_i a_i \lambda_i \, |\nu_i\rangle \tag{4.5}$$

where a_i are the coefficients of the combination of eigenmodes that describe the x-polarised incident field and λ_i are the eigenvalues of $|\nu_i\rangle$. The x and y polarisations of the incident field can be linked by a 90° rotation of the whole system. This rotation can be described by an operator $\widehat{R}$ which means that a y polarised incident field $|\mathbf{E}_y\rangle$ can be written in terms of the eigenmode decomposition of the x-polarised field:

$$|\mathbf{E}_y\rangle = \widehat{R} \, |\mathbf{E}_x\rangle = \sum_i a_i \widehat{R} \, |\nu_i\rangle \tag{4.6}$$

As a result of the degeneracy of the eigenmodes under symmetry operations (i.e. $|\mathbf{E}_y\rangle$ has the same eigenvalues as $|\mathbf{E}_x\rangle$), any eigenmode that is excited by an x-polarised field must have a degenerate partner that is excited by a y-polarised field. It then follows that any weighted sum of the x and y eigenmodes must itself be an eigenmode which shares the same eigenvalues as the original eigenmodes:

$$|\nu_i'\rangle = x\,|\nu_i\rangle + y\widehat{R}\,|\nu_i\rangle \tag{4.7}$$

where x and y are the weighting coefficients. This new eigenmode, therefore, allows for an arbitrarily polarised incident field to be described:

$$|\mathbf{E}_0\rangle = x\,|\mathbf{E}_x\rangle + y\,|\mathbf{E}_y\rangle = \sum_i a_i\,|\nu_i'\rangle \tag{4.8}$$

This is, therefore, the origin of the polarisation invariance of the pentamer structure; eigenmodes excited by any arbitrary polarisation $|\mathbf{E}_{0,a}\rangle$ are degenerate with those excited by any other polarisation $|\mathbf{E}_{0,b}\rangle$ as both can be described by combinations of $|\mathbf{E}_x\rangle$ and $|\mathbf{E}_y\rangle$ which share the same eigenvalues λ_i, and thus the same far-field response.

4.5.2 Experimental Images and Discussion

Figure 4.13 shows the simulated and experimental extinction cross-sections of the cruciform pentamer structures. The structures each consist of a central, circular disk surrounded by triangular-shaped components (both made of gold) on a quartz substrate, and were fabricated by electron beam lithography. The individual elements of the pentamer are 145nm in width and are separated by sub-20nm gaps (so the entire pentamer structure has a width of approximately 470nm). An SEM image of a pentamer array is shown in image (a). The match between simulation and experiment in parts (b) and (c) is good, with the far-field polarisation invariance clearly visible in the extinction spectra.

The experimental setup for recording the s-SNOM images of the cruciform pentamer structures differs slightly from that of the cross-polarisation scheme shown in figure 4.1. The cross-polarisation scheme uses s-polarisation in excitation and p-polarisation in detection to measure the z component of the plasmon's field distribution. In the case of the pentamer structures, however, s-polarised light was used in both excitation and detection [158–160], as shown in figure 4.14. As mentioned in section 4.2, the key to suppressing chemical contrast is the use of s-polarised light to excite the plasmon resonance. This polarisation is only weakly concentrated by the s-SNOM probe and so the dominant scattering signal is due to the plasmon's field profile rather than the dielectric differences across the sample. Detection with the p-polarisation is usually required because the probe is a stronger scatterer for these field directions.

With the cruciform pentamer structures, however, the use of p-polarised light in detection would mask the "hot-spots" - areas of high field enhancement in the small gaps between the elements of the oligomer - because here the fields are largely in-plane. In fact, for the particular case of the cruciform pentamers, there are two reasons that the near-field scattering signal strength is no weaker than if the z components were measured by using the cross-polarisation setup. Firstly, the in-plane field enhancement is approximately 20 times that of the z component enhancement and secondly, the heterodyning reference beam does not need to be attenuated by a 45° polariser as it would be in the cross-polarization scheme, so the signal amplification is larger. The experimental images of the near-field scattering magnitude taken in this s-in, s-out manner with a wavelength

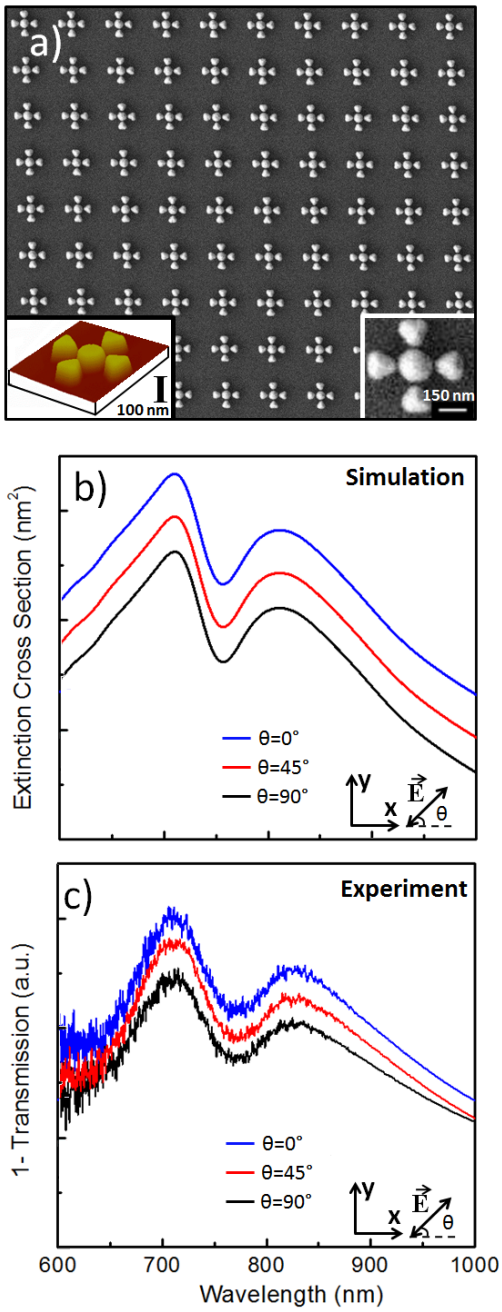

Figure 4.13: (a) SEM image of periodic array of pentamers with (left inset) AFM image and (right inset) close-up of single structure. The pitch is 1μm. (b) Simulated extinction cross-sections and (c) experimental extinction cross-sections (found by 1 - Transmission) spectra at 0°, 45° and 90°. The spectra have been separated vertically to avoid overlap.

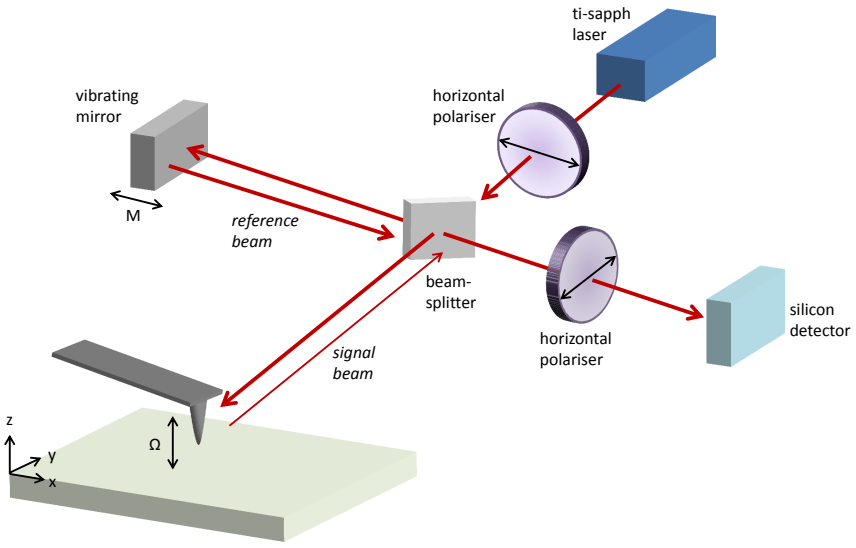

Figure 4.14: In contrast to the cross-polarisation scheme (see figure 4.1), the imaging setup for the cruciform pentamer structures uses s-polarised light for excitation *and* detection, meaning that the in-plane components of the plasmon's field are measured, rather than the out-of-plane component.

of $\lambda = 850$nm are shown in figure 4.15.

As the polarisation is rotated from 45° to 0° with respect to one of the structure's axes, the positions of the hot spots change. This is in contrast to the far-field extinction cross-sections (figure 4.13) which remain constant. Again, the simulations and the experimental images agree well with only slight discrepancies. These arise from slight imperfections in the fabrication process (leading to variations in shape or refractive index across the structure) or from retardation effects (see section 4.4).

Further evidence of the changing near-field distribution is shown by the phase images of the near-field scattering, ϕ_3, as shown in figure 4.16. Again, simulation and experiment show good agreement; the near-field properties of the cruciform pentamer structure - its magnitude and phase - are highly dependent on the direction of the illuminating beam's polarisation.

In conclusion, the optical properties of cruciform pentamers (namely, far-field polarisation independence and near-field dependence) offer the ability to move hot spots into different sub-20nm gaps by rotating the polarisation of the incident light without the loss of excitation efficiency for any configuration. This should be a general property of all plasmonic oligomers exhibiting rotational symmetry and could lead to applications requiring nanoscale control of hot spot positioning (sensing, for example).

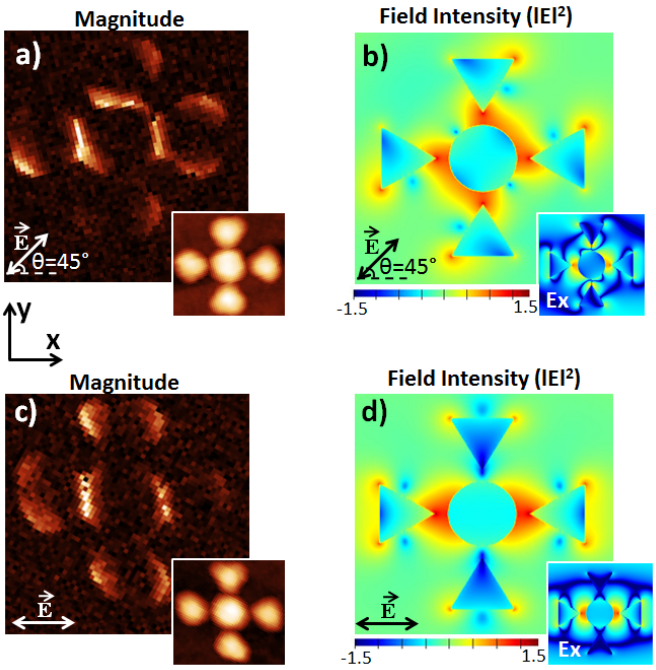

Figure 4.15: (a,c) Measured near-field scattering magnitude at the third harmonic of the probe oscillation s_3. Insets show the AFM image. (b,d) Simulated field intensity $|E|^2$ for incident fields of 45° and 0° respectively. The inset shows the in-plane E_x component. The excitation wavelength is $\lambda =$ 850nm.

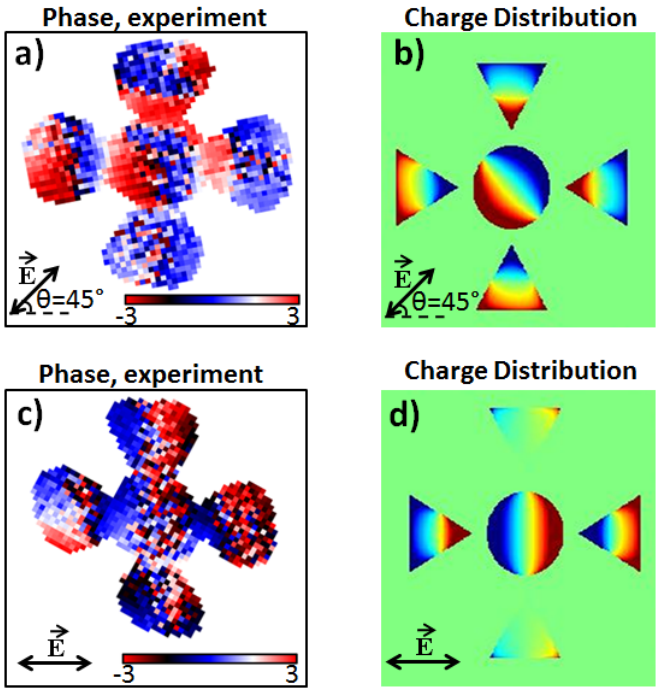

Figure 4.16: (a,c) Measured near-field scattering phase at the third harmonic of the probe oscillation ϕ_3. (b,d) Simulated charge distributions for incident fields of 45° and 0° respectively.

4.6 Conclusion

Although s-SNOM measurements have clearly been demonstrated to be successful in imaging plasmon resonances in this chapter, it should be noted that it is not yet fully understood how - or even if - the cross-polarisation scheme (or at least s-polarised incident light) prevents the probe from affecting the plasmon resonance in all possible cases. Although the probe itself is very sharp, its body is often orders of magnitudes larger than the structures it is imaging - particularly in the visible and near-infrared. It is also not clear how much the optical scattering of the structures themselves affect the detected signal [161], or what role the in-plane components of the plasmon resonance have on the scattering of the out-of-plane components (and vice versa) [86, 162].

There are also some more practical difficulties in the use of s-SNOM. The fabrication methods for many plasmonic samples leave a resist on the surface of the metallic structures. Even though these are typically less than 10nm in depth, they can have a significant damping effect on the measurement, rendering many of the near-field properties undetectable. As such, samples often have to be specially prepared for near-field study. The s-SNOM probe itself can also be too large to physically fit into the smallest gaps (less than 10nm) between structures which exhibit the largest field enhancements.

Our particular setup, using reflection mode, also has significant drawbacks for imaging plasmonic samples (see section 4.4). An implementation of transmission mode is of critical importance for the study of more complicated, coupled structures. For both modes, however, the nature of

the s-SNOM probe as a scatterer means that the measurement will always be more sensitive to the out-of-plane components than the in-plane components of the plasmons's field. Measurements of these different components, therefore, are not directly comparable.

Finally, the implementation of a widely tuneable titanium-sapphire laser as the light source for measurement holds much promise for near-field spectral measurements. Many plasmonic phenomena exhibit different responses at different wavelengths, and if these properties are to be studied by an s-SNOM, a suitable laser is required to do so. The rather stringent requirements on lasers (see section 7.3) means that most s-SNOM studies of plasmonic phenomena have been carried out with a CO_2 laser in the mid-infrared or a helium-neon laser in the visible. Neither of these offer the relative tuneability afforded by the titanium-sapphire laser, and neither cover the spectral range for which many plasmonic structures are designed - the near-infrared.

Chapter 5

The Use of Pulsed QCL Sources

"It's like having sex in a hammock standing up." **Prof. Chris Phillips**

This chapter borrows from a paper entitled "Widely Tuneable s-SNOM Using Pulsed Quantum Cascade Lasers" submitted to Applied Physics Letters in September 2013 with Miguel Navarro-Cia, Mohsen Rahmani, Stefan Maier and Chris Phillips.

5.1 Introduction

At present, one of the major limiting factors in the development of s-SNOM as a tool for single wavelength nanoscale imaging is the lack of suitable sources in the mid-infrared. As mentioned in section 3.5, there are stringent requirements on a suitable light source, namely a minimum power of around 1mW, a coherence length of greater than 1cm and a signal-to-noise ratio (SNR) of around 10^4 (see section 7.3). If other, more practical, considerations are included such as cost, portability and ease-of-use, one is left with a narrow selection of gas lasers, and the one that has been overwhelmingly used in the literature is the CO_2 laser.

Recent developments in the field of quantum cascade lasers (QCLs), however, seem set to revolutionise the choice of light sources in the mid-IR [163–165]. These are narrow linewidth, solid state lasers which can be tuned through a hundreds of wavenumbers when arranged in an external cavity configuration - considerably more than is possible with a gas laser [166]. Unfortunately, with the current generation of commercially available devices, this tuneability often comes at the cost of the laser only working in a pulsed mode rather than continuous wave (CW) mode. This can present a problem if QCLs are to be used in conjunction with an s-SNOM - its pulsed nature must be accounted for. If it is not, a substantial amount of digitisation noise can be added to the measurement.

5.2 Quantum Cascade Lasers

QCLs offer narrow linewidths, broad tuneability, compact size and high power outputs with no requirement for extensive cooling. In operation, they do not function like standard diode lasers in which electrons transition from the conduction band to the valence band and in doing so emit a photon. Instead, they operate on transitions purely within the conduction band of the device which has been engineered to contain a series of "mini bands" (see figure 5.1).

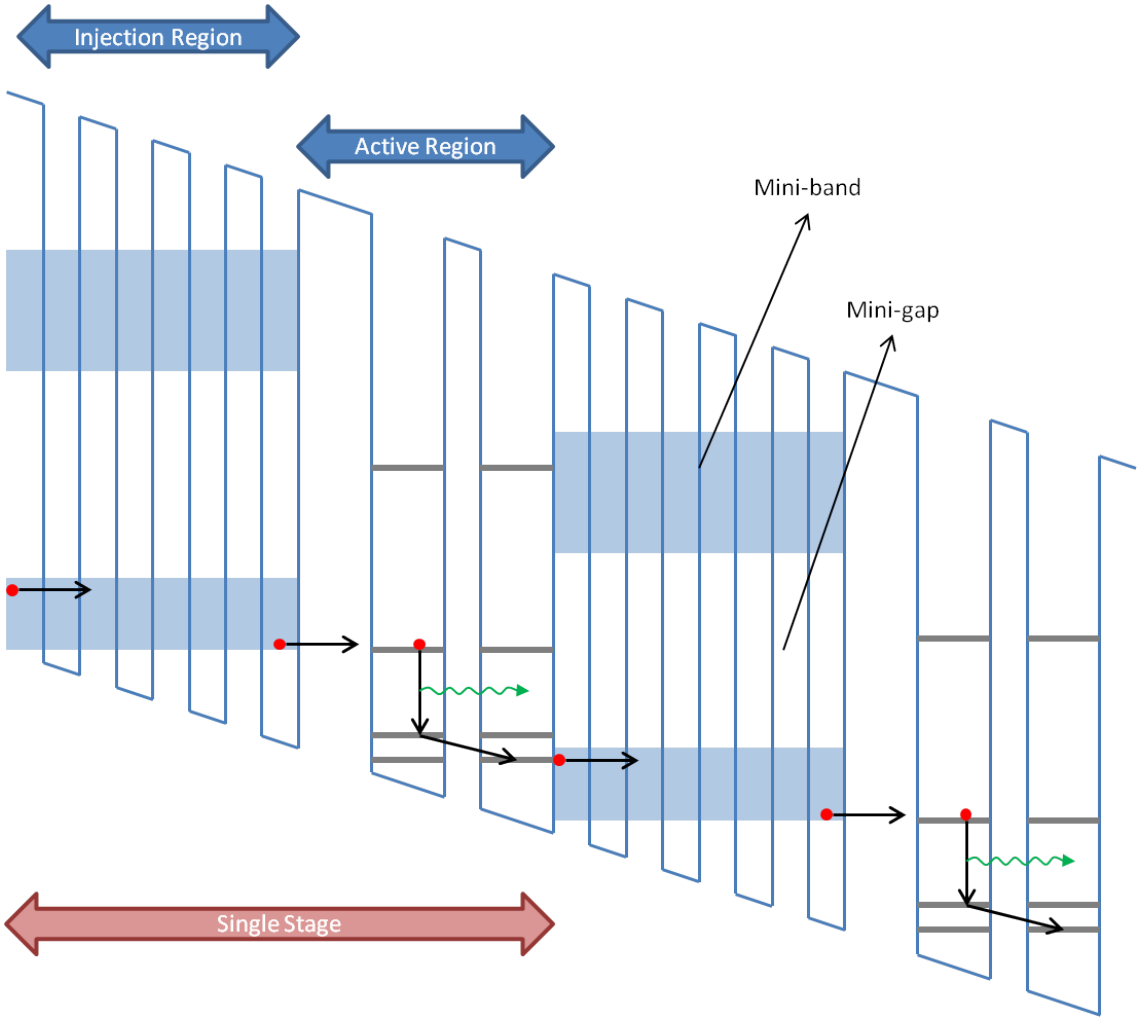

Figure 5.1: Diagram of energy versus distance in the conduction band of a QCL along the beam axis. The red dot indicates an electron as it passes through the band structure emitting a photon at each stage and the green line represents the emitted photon at each stage.

A QCL is made up of many "stages" stacked on top of each other. Each of these stages consists of an injector region and an active region. The former has a series of closely spaced quantum wells, and, just as in a standard crystal lattice, these quantum wells have allowed and forbidden energy levels known as "mini-bands" and "mini-gaps". The lowest energy of the mini-band is designed to equal the third energy level of the active region. The latter consists of much broader quantum

wells allowing only discrete states. After electrons tunnel from the injector region into the active region, they fall to the second energy level and in doing so emit a photon.

For the device to lase, an electric field is applied across the structure, meaning that the electrons "cascade" through the staircase structure. Each electron can, therefore, create as many photons as there are stages - typically around fifty. As a result, the quantum efficiency of QCLs can be greater than unity.

Tuneability is introduced by using the device in an external cavity configuration. This uses a diffraction grating to select the wavelength that is propagated within the optical cavity; the diffraction grating is angled such that the first order reflection is passed back into the cavity while the zeroth order is reflected as the outgoing beam as shown in figure 5.2. This is also known as the Littrow configuration.

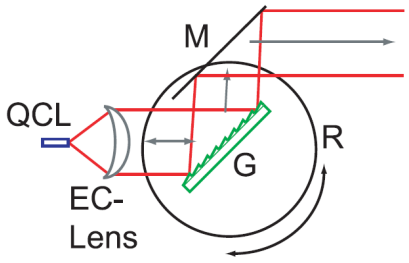

Figure 5.2: Diagram of the external cavity (EC) in a Littrow configuration that provides the QCL's tuneability. The first order reflection from a grating G is passed back into the cavity, whereas the zeroth order continues via a mirror M to the laser output. Tuning is managed by rotation R of the grating. Taken from [166].

5.2.1 Lasertune

For the experiments reported in this thesis in the mid-infrared, a commercially available QCL was used (*Lasertune, Block Engineering*). It consists of three separate QCL chips and provides continuous tuning between $\lambda \sim 6 - 10\mu$m with an average power output of around 5mW (which varies with wavelength). It has a maximum duty cycle of 5% and can be triggered either externally, at repetition rates up to 1.25MHz, or by an internal signal at roughly 200kHz. As will be shown later in this chapter, the ability to externally trigger the device is critical for success in combining it with s-SNOM.

5.3 Pulsed Sources and s-SNOM

There is nothing fundamentally problematic about using a pulsed laser source for s-SNOM measurements. Steps must be taken, however, to ensure that the pulsed nature of the laser doesn't create a large amount of noise when the voltage waveform from the detector is digitised. To demonstrate this point, figure 5.3 shows the waveforms of two perfect (i.e. fluctuation free) laser sources - one CW and one pulsed at 185kHz - and figure 5.4 shows their corresponding Fourier transforms.

Given that s-SNOM measurements take place in frequency space at the harmonics of the probe oscillation frequency, the only stipulation for using a pulsed laser is that the measurements do not take place at one of the harmonics of the laser repetition rate (i.e. 185kHz, 370kHz, 555kHz etc.)

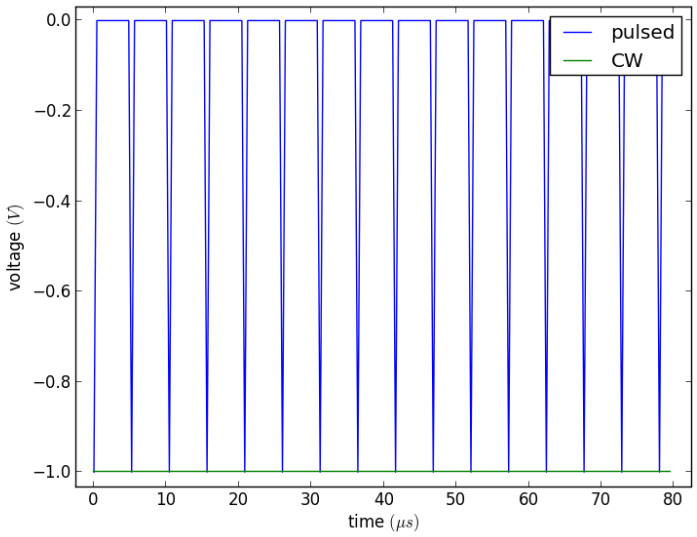

Figure 5.3: Modelled detector response to fluctuation-free CW and pulsed laser beams.

There is no physical reason, therefore, that s-SNOM images taken with a pulsed laser should be inferior to those taken with a CW light source. Achieving this parity in practise, however, involves careful instrumentation to make sure that the sampling circuitry records a value from the detector when the laser is actually producing light.

5.3.1 Synchronisation of Sampling and Laser Pulse

When working with pulsed lasers, it is important to consider the Nyquist limit - and in the case of s-SNOM it must be considered at two levels. The Nyquist limit states that the maximum bandwidth of the system is half that of the sampling frequency. In other words, if a waveform is to be accurately replicated digitally, the analogue-to-digital conversion (ADC) rate must be at least two times faster than the waveform's maximum frequency. For the s-SNOM and pulsed QCL combination, this must be applied twice:

1. The sampling rate of the s-SNOM must be at least twice that of the repetition rate of the laser to accurately replicate the pulses. As the sampling rate is detector-bandwidth limited at 2.5MHz, it follows that the maximum allowed repetition rate of the laser is 1.25MHz. The experiments which follow used a repetition rate of 625kHz.

2. The repetition rate of the laser must in turn be fast enough to sample the signal at the second or third harmonic of the probe oscillation frequency. There is considerable choice in the frequency of the latter, with a wide range of commercially available probes which vibrate

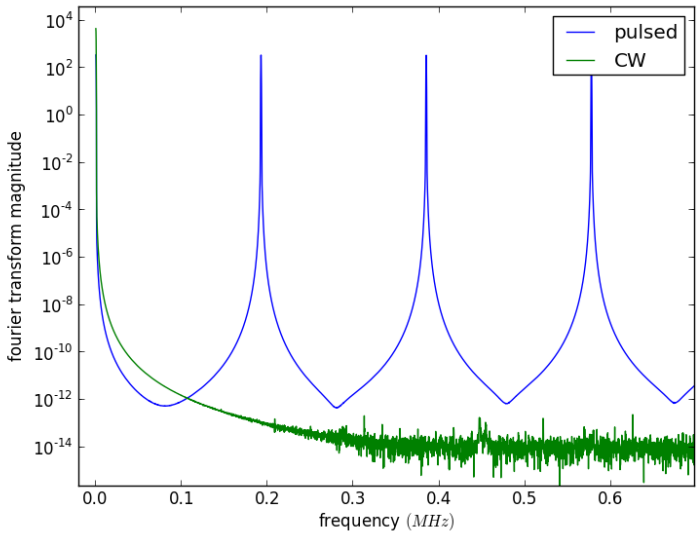

Figure 5.4: Fourier transforms of the waveforms shown in figure 5.3 in a 6.5ms integration period. The traces do not drop to zero off the peaks because of windowing; although the signals were Blackman windowed to prevent leakage, the fact that the signal is not infinitely long in time means that there are non-zero values between the peaks in frequency space.

at frequencies from 10 to 500kHz. We chose a probe with a resonance close to 75kHz (*Budget Sensors, Multi75GB-G*). Given that, in the mid-infrared, using the second harmonic of the probe frequency is sufficient to suppress the background scattering signal σ_B (see section 2.5), our measurement frequency was approximately 150kHz.

With a laser repetition rate of 625kHz and an ADC rate of 2.5MHz, many samples are taken when the laser is off. If the sampling and the laser pulses are not well synchronised, therefore, it is possible that the sampling will record a rising or falling edge (instead of the maximum of the QCL output peak) resulting in a randomly measured pulse height (figure 5.5). The digitisation of the detector voltage, therefore, is potentially a major source of "artificial" noise - these laser fluctuations do not exist physically but are an artefact of the electronic waveform sampling process.

If the laser pulses and the digitisation circuitry are properly synchronised, one can be sure that the sampling will fall on a QCL peak maximum. In practise we achieve this by triggering the laser with an integer divisor of the sampling frequency - in our case 4. This corresponds to a laser repetition rate of 625kHz with a cycle length of $\sim$ 1600ns and a lasing period of $\sim$ 80ns.

As shown in section 3.5, the SNR of an s-SNOM measurement is defined in Fourier space as:

$$\mathrm{SNR_s} = \frac{\mathrm{s_n}}{\mathrm{D_n}} \tag{5.1}$$

where s_n - the near-field scattering magnitude - is calculated from the heights of the sidebands $C_{n,m}$ at a harmonic n according to equation 2.5, and D_n is the Fourier component caused by the disturbances (laser fluctuations, detector noise etc.) at that same frequency. When an unsynchro-

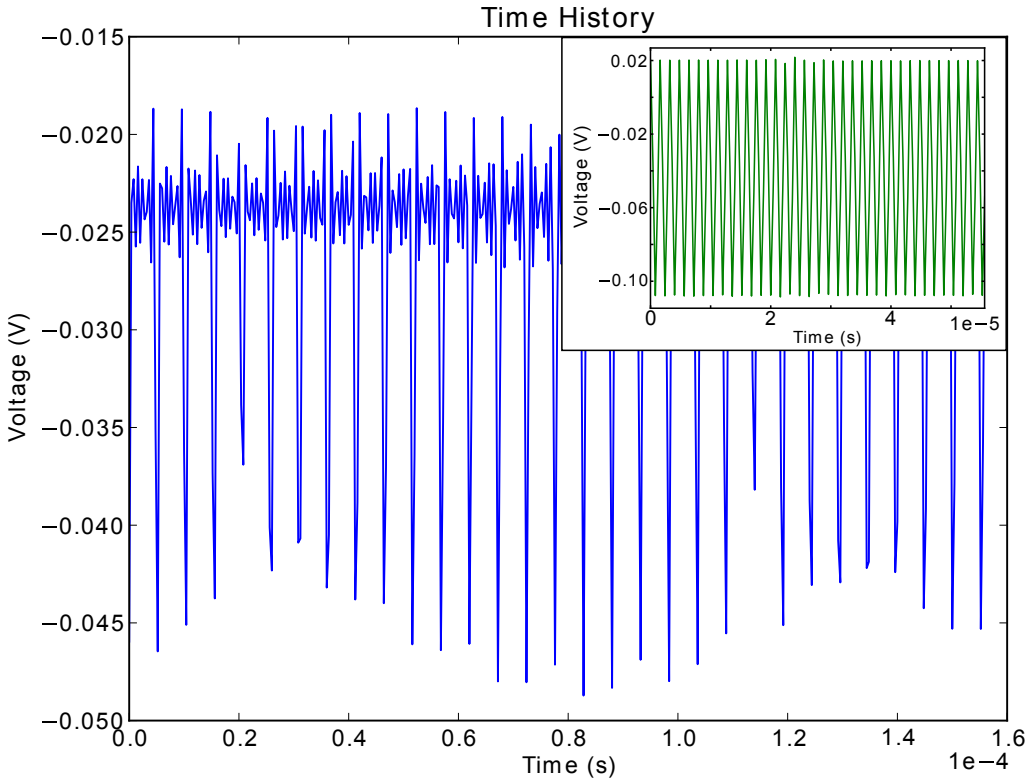

Figure 5.5: An example of a pulse train measured at the detector when the laser is internally triggered (at 185kHz) and therefore not synchronised with the ADC sampling circuit. The lack of synchronisation between the laser pulses and the sampling leads to a recorded measurement with a very large apparent pulse-to-pulse variation. The inset box shows the waveform when the sampling and pulses are well synchronised (at 625kHz).

nised waveform (such as that in the main part of figure 5.5) from the pseudo-heterodyne detection scheme is Fourier transformed, the dominant contribution to D_n comes from the artificial laser fluctuations caused by the digitisation. This is clearly visible in figure 5.6 where the Fourier transforms of the detector waveform (showing sidebands at first and second harmonics of the probe oscillation frequency) are displayed for both the synchronised and unsynchronised cases.

Given that the Fourier components of the laser fluctuations which dominate D_n are, at these frequencies, white (i.e. equal across the range of measured frequencies - see section 7.3), D_n can easily be calculated by finding the average Fourier component for a range of frequencies immediately adjacent to the sideband cluster. In the unsynchronised case (b), D_n is $\sim 1 \times 10^{-5}$V. The sideband heights at the second harmonic $C_{2,m}$ are clearly considerably smaller than this, and the SNR is therefore less than 1, making imaging impossible. In the synchronised case (a), however, D_n is much smaller, at $\sim 3 \times 10^{-7}$V, and by combining the sideband heights according to equation 2.5 gives $s_2 \sim 8 \times 10^{-6}$V, corresponding to an SNR of 26. In both cases the measurement bandwidth is 153Hz. D_n can be further reduced by decreasing the measurement bandwidth or, equivalently, by increasing the integration time at each pixel. As always, the trade-off for lower noise is a longer image acquisition time.

5.3.2 Inherent Pulse-To-Pulse Variation

As well as the pulse-to-pulse variation introduced by unsynchronised digitisation of the detector waveform, the Lasertune device exhibits a significant spread of inherent pulse-to-pulse variations as the wavelength is scanned. The variation is extremely sensitive to the wavelength, and can change from 1% to 50% pulse-to-pulse differences within a single wavenumber. Figure 5.7 shows the Fourier transforms of two synchronised pulse trains in a measurement bandwidth of 153Hz; one at $\lambda = 1008.1\text{cm}^{-1}$ with a pulse-to-pulse variation of less than 1% and the other at $\lambda = 1565.5\text{cm}^{-1}$ with a variation of around 5%.

The resulting Fourier components at the typical measurement frequency in the mid-infrared (around 150kHz) differ by an order of magnitude (see figure 5.7). This significantly affects the SNR, and when the pulse-to-pulse variation becomes too large, the Fourier components of the laser fluctuations drown out the sidebands from the pseudo-heterodyne detection scheme. This link between the fluctuations in the time and frequency domains can be explicitly expressed by Parseval's theorem which states that the amount of energy contained within the two domains must be equal. This relationship is discussed in further detail in section 7.3.1. As a result of the way that the fluctuations change with wavelength, the Lasertune device is only suitable for use in near-field measurements at a select number of wavelengths, and this limits its ability to continuously tune through a wide spectral range. Explicitly, our particular QCL lases stably enough for making s-SNOM measurements at a few select positions within $8 \pm 0.2\mu$m and $9.6 \pm 0.3\mu$m.

There are a number of possible causes for this inherent variation [167, 168]. The first is mode-hopping; a QCL arranged in an external cavity can support a large number of modes, and as the device is lasing, the optical energy can be redistributed among these modes leading to differing power outputs. The second - and related - cause is thermal effects. As current passes through the chip its temperature rapidly increases. This changes the optical properties of the cavity which

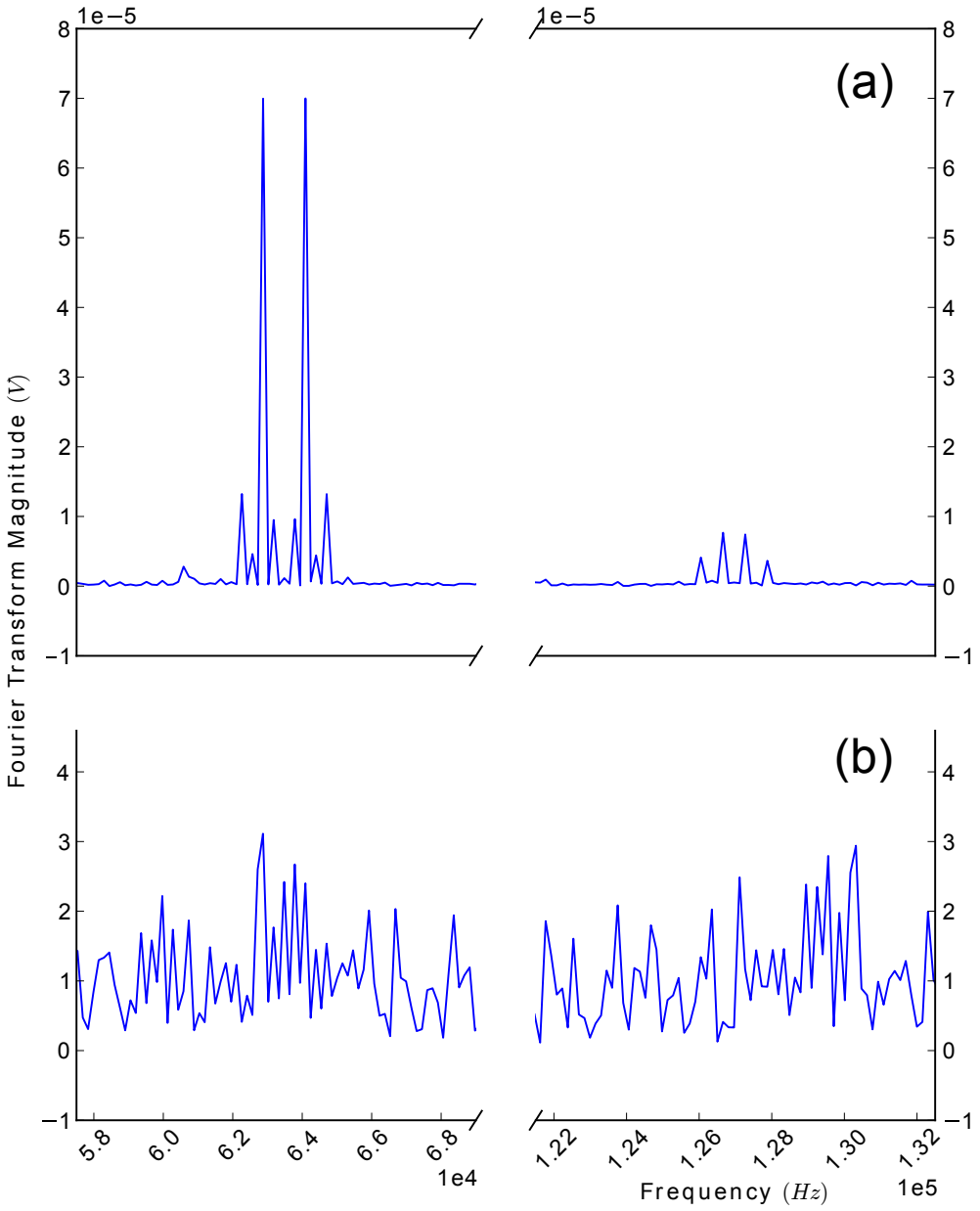

Figure 5.6: Fourier transform of (a) synchronised and (b) unsynchronised (b) digitised pulse trains for an s-SNOM measurement on a gold substrate using a 6.5ms signal integration period. The probe vibration frequency is $\Omega = 63.5$kHz and the reference mirror oscillation frequency is $M = 300$Hz. The first and second harmonic sidebands are clearly visible in the (synchronised) low-noise case, whereas with the laser unsynchronised, they are buried in the extra noise artefacts generated in the digitisation process.

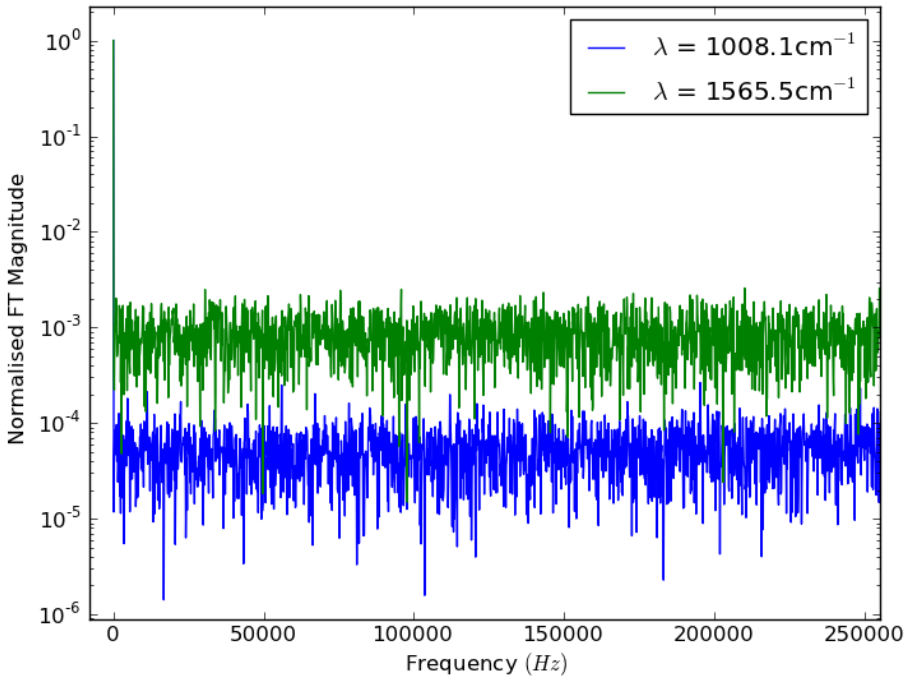

Figure 5.7: The Fourier transforms of the waveform at the detector for two QCL wavelengths; $\lambda = 1008.1\text{cm}^{-1}$, with a pulse-to-pulse variation of less than 1% and $\lambda = 1565.5\text{cm}^{-1}$, with a variation of 5%. The measurement bandwidth is 153Hz. The traces have been normalised to their DC level to allow for direct comparison.

can lead to instabilities in the output. Finally, the quantum noise of the laser - fluctuations due to the discrete nature of the QCL's allowed energy states - depends on the gain and the cavity losses [169]. These in turn depend on the wavelength, and so there is likely to be some wavelength dependence in the fundamental lower limit of the laser's noise spectrum.

5.4 Broadband Antennas

To demonstrate the ability of the pulsed QCL and s-SNOM to take near-field images over a range of wavelengths, this section presents images of a broadband mid-infrared log-periodic nanoantenna [170,171]. This sample exhibits plasmonic resonances over a broad range of wavelengths and, as the wavelength is scanned, the electric field distribution across the nanoantenna changes significantly, thus providing a simple proof-of-principle experiment.

As the images of a plasmonic sample need to measure the electric field distribution (rather than chemical contrast), the cross-polarisation scheme was used. The setup was identical to that shown in figure 4.1, with the only adjustments being to replace the beamsplitter and polarisers with their counterpart components that function in the mid-infrared (a zinc selenide window and holographic wire grid polarisers respectively).

5.4.1 Results: Simulation and Experiment

The results comparing the numerical simulations and the experimental images are shown in figure 5.8. For the FDTD simulations $\big($images (d) and (e)$\big)$ - again performed with *Lumerical* - a 40nm thick gold nanoantenna, lying over a 2nm chromium adhesion layer and a semi-infinite BaF2 substrate, was discretised with a non-uniform mesh. As a standard throughout, the mesh had a volume grid of 7nm $\times$ 7nm $\times$ 5nm within the total-field scattered field (TFSF) source volume. However, at the nanoantenna gap a finer mesh, of 4nm $\times$ 4nm $\times$ 5nm was used, and in the chromium, the length of the discretisation in the z-direction was 0.5 nm.

The incident illumination, generated by the TFSF, was polarized parallel to the teeth of the nanoantenna, and was incident, as in the experiment, at an angle of 60° with respect to the sample normal. The optical properties of gold and chromium were modelled with multi-coefficient functions to fit experimental tabulated data [113], whereas the BaF_2 was assumed to be lossless with a constant refractive index of n = 1.465 [172]. Note that in the simulation, the presence of the s-SNOM probe is not considered. To record the field distribution in the xy-cross-sectional plane 20 nm above the gold-air interface, a 2D field profile monitor was used.

The pixel resolution for each experimental image $\big($parts (a),(b) and (c)$\big)$ was set at 100nm, corresponding to a value of approximately $\lambda/80$. The acquisition time for each image was less than 90 seconds. Despite the 180° rotational symmetry of the structures, the field profile itself is not rotationally symmetrical, again due to retardation effects. As the size of the nanoantenna is comparable to the wavelength, there is a significant difference in the phase of the exciting field across it, so the field strength and direction across the sample is not uniform.

In terms of the image quality, the $\lambda = 9.5\mu$m image is superior due to the lower intrinsic pulse-to-pulse intensity variation of the laser at this wavelength. At the position of greatest field

enhancement the SNR is 17 at $\lambda = 9.5\mu$m and 9 at $\lambda = 7.9\mu$m.

There is a good qualitative match between the simulation and experiment on the lower half of the antenna; the field profile is similar between the two images at $\lambda = 7.9\mu$m and $\lambda = 9.5\mu$m in the experiment, and this similarity between the two wavelengths is mirrored in the simulations. On the upper half of the antenna at both wavelengths there is a better quantitative match. At $\lambda = 9.5\mu$m a hot spot, P1, is observed on the outermost left tooth. In the image at 7.9μm, this area of high intensity has clearly shifted to the right onto the second tooth from the top, P2, in agreement with the simulations.

5.5 Conclusion

For the reasons outlined in the introduction, QCL's show enormous potential for use with s-SNOM; together they can provide a versatile imaging system, offering high spatial and spectral resolution coupled with a wide wavelength tuning range, fast image acquisition times and simple turn-key operation. Their tuneability is the key to spectroscopic imaging (and therefore chemical identification) at the nanoscale.

Unfortunately, the inherent pulse-to-pulse variation of our particular commercially available device (*Lasertune, Block Engineering*) limits its usefulness in gathering full spectroscopic information about a sample; its fluctuations are too large (which introduces noise that drowns out the signal at the measurement frequency of approximately 150kHz) for all but a select few wavelengths within $8\pm0.2\mu$m and $9.6\pm0.3\mu$m. It is here that the laser is at its most stable, and so it is only within these narrow spectral ranges that images can be recorded.

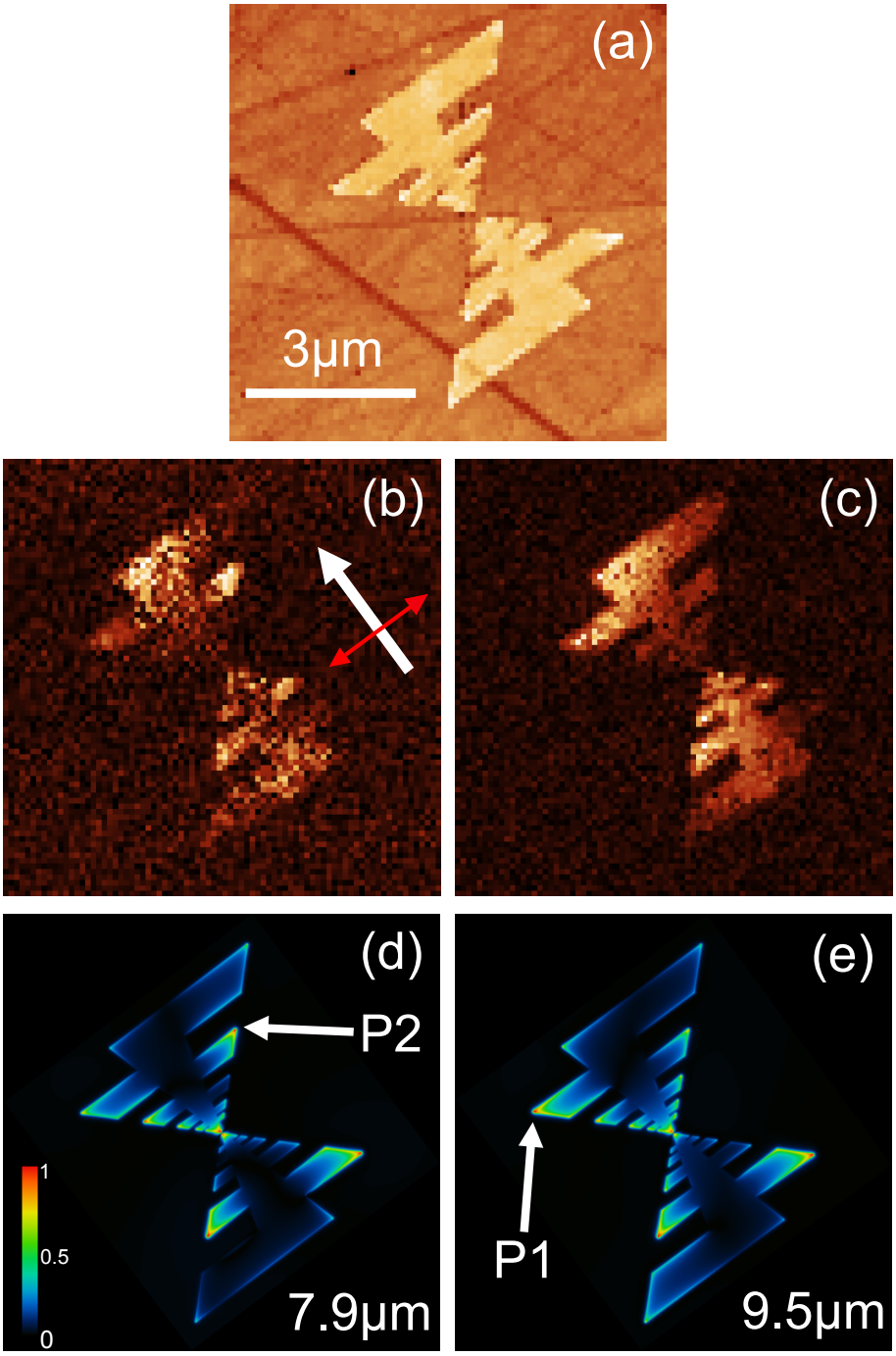

Figure 5.8: (a) AFM, (b,c) s-SNOM amplitude images and (d,e) simulation of E_{rmz} field at 20nm above the structure at $\lambda = 7.9\mu$m (left) and 9.5μm (right) for broadband antenna samples. The in-plane direction of the incident light (at 60° to the sample normal) is indicated by the white arrow and its polarisation by the red arrow. The pixel size is 100 x 100nm (i.e. $\sim \lambda/80$) and the integration time at each pixel is 13ms, corresponding to an image acquisition time of less than 90 seconds. A scale bar showing the relative magnitude of the E_{rmz} field is shown in image (d).

Chapter 6

Cellular Imaging with s-SNOM

"The proof is in the pudding." **Prof. Chris Phillips**

6.1 Introduction

The greatest benefit that comes with coupling a tuneable mid-infrared laser with an s-SNOM is the ability to perform chemical spectroscopy at nanometric resolutions. Mapping material composition at such high levels of detail has a huge range of potential applications, perhaps none greater than in the biological sciences. Infrared absorption spectroscopy has - until now - always been limited by diffraction, and given that the wavelength of light used in these experiments is comparable to the size of a typical cell (around 10μm), the technique has been limited to bulk tissue measurement. However, with the resolution improvement of three orders of magnitude offered by s-SNOM, deeply intra-cellular features would be observable, and further, chemically identifiable [80, 173, 174].

6.2 Biological Imaging

Although s-SNOM promises a big impact in terms of spectroscopic biological imaging, there are several technical challenges which must first be met before the technique is to become more mainstream. The biggest of these is one that faces all types of scanning probe microscopies (SPMs) - using probes with biological material *in vivo* [175–178]. Live samples throw up many difficulties: they tend to be wet and viscous, introducing hysteresis into the motion of the probe which makes s-SNOM measurements difficult given the requirement for sinusoidal tip motion; they are often very rough, requiring the travel of the probe's position to be large; finally, live samples move, requiring fast scan times.

Many of these problems have been solved, and indeed atomic force microscopy (AFM) on which s-SNOM is based is often employed in the biological sciences - particularly "high-speed" AFM which is capable of observing dynamic processes in real time (typically producing images at tens of frames per second with a 100 $\times$ 100 pixel density) [179, 180]. The solution to the problems facing *in vivo* measurement, however, is often to perform the measurements with the probe and

sample completely submerged. This is not an option for s-SNOM measurements where a beam must be focused at the probe's apex - mid-infrared light is strongly absorbed by water.

The simplest path, therefore, at this stage in the development of s-SNOM as a tool for the biological sciences, is to image samples *in vitro*. In this way, tissue "slices" can be specifically prepared by a variety of techniques to be inert and flat, without changing their physical chemistry. While dynamic processes are clearly no longer visible, material composition can readily be studied. The first, proof-of-principle step for such an approach is to try to match what is already known about cellular composition - that lipids are found in the cell walls, for example - with the well-known biological spectral absorption peaks. A typical biological spectrum is shown in figure 6.1 with the most prominent peaks labelled.

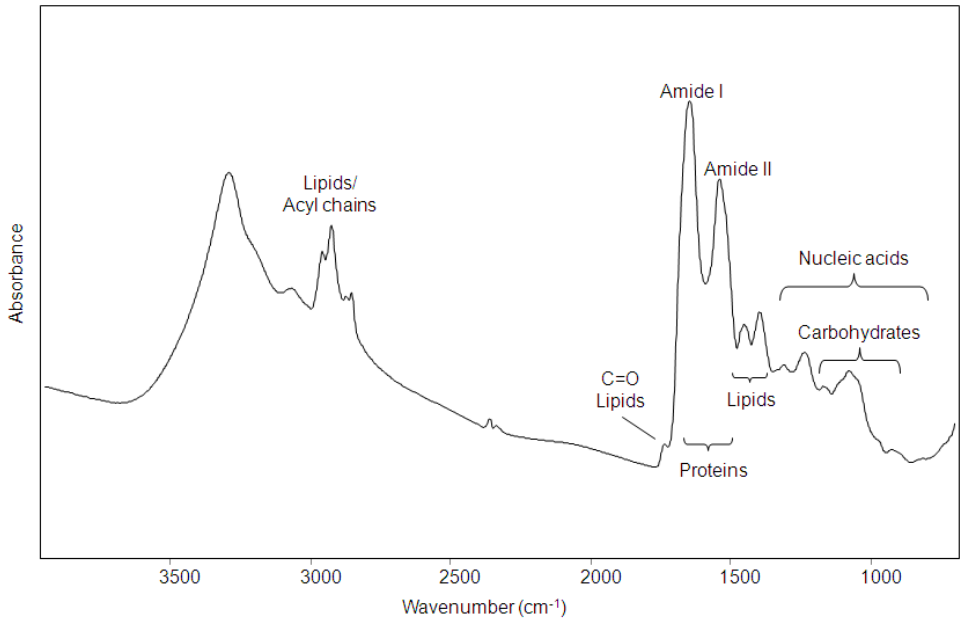

Figure 6.1: A typical mid-infrared spectrum of biological tissue, where the major absorption peaks caused by the vibrations of molecular bonds have been labelled . Taken from [181].

The peaks in figure 6.1 correspond to specific vibration or stretching frequencies in organic bonds. As these vibration modes are quantized, they can only absorb specific photon energies, and at these particular wavelengths there is a maximum in the absorption spectrum. The amide I peak, for example, centred at 1650cm^{-1}, is governed by stretching vibrations of C=O and C-N bonds in covalent molecules. There are many of these types of bonds in proteins, so a sample which exhibits high absorption in the amide I band can be deduced as being protein rich. The resolution of traditional IR spectroscopy is diffraction limited to approximately 10μm (i.e. the approximate size of a typical cell), however, meaning that it is impossible to say (using an IR absorption spectrum alone) whether the amide I absorption peak comes from a particularly strong area of absorption within a cell - such as the nucleus or cell wall - or whether it is spread equally throughout the 10μm pixel.

The ability of s-SNOM to gather spectra at much higher resolutions - down to groups of just a

few hundred molecules [182] - means that the precise location of the absorption peak's source can be located within the cell. As such, s-SNOM could, in time, prove to be an important technique in the study of intracellular chemistry.

6.2.1 Oesophageal Cancer Sample

The tissue used in this chapter to demonstrate the ability of s-SNOM to image biological material was collected in the course of a medical oesophageal biopsy. It was made available for imaging research under a Local Research Ethics Committee consent (07-Q1604-17). The biopsy sample was processed conventionally by fixing in formalin solution for 24 hours, followed by 24 hours in 70% ethanol. The material is then embedded in paraffin wax. This process stabilises and preserves the specimen without unduly distorting its microanatomy. Shaved slices are then taken from the paraffin block with a microtome block at a thickness of 5μm, and were then placed onto barium fluoride (BaF_2) discs. These were then immersed in xylene to remove the paraffin.

These samples were originally prepared for far-field imaging purposes [183], and as such the physical thickness of the samples (5μm) was considerably larger than would be ideal for s-SNOM measurements. Cellular samples are not flat; they do not cover the entire substrate. The 5μm thickness, therefore, is the distance through which the s-SNOM probe must be able to scan vertically. Unfortunately, our particular probe has a travel of less than 3μm, meaning that some areas of the sample are out of range. Microtomes are capable of much thinner slices ($\sim$ 50nm [184]), which would be more suitable when preparing samples specifically for near-field imaging.

Microscope images at two magnifications are shown in figure 6.2. Image (a) shows a large scale view of the sample, where the different types of tissue - the epithelium, the muscle and the stroma - have been labelled. Image (b) shows a higher magnification picture of the epithelium, where the approximate area scanned by the s-SNOM in figure 6.5 has been highlighted. The large, dark areas are likely to be mucous glands or lymph tissue.

6.2.2 Far-Field Spectroscopy

Figure 6.3 shows the absorbance spectrum of a 5mm^2 area of the biopsy, which was taken using a commercial Fourier transform infrared (FTIR) spectrometer (*FT/IR 460 Plus, Jasco*). Several peaks are clearly visible; the amide I peak at 6μm, the amide II peak at 6.5μm and a lipid peak at around 3μm are the most prominent.

Bulk far-field spectra of tissue biopsies are useful in making comparisons between samples - they may even prove to be useful in cancer diagnosis [185] - as peak positions and heights can be analysed. Given that the peaks contain information about the quantity and types of, for example, proteins and lipids in the sample, they provide a means by which different types of tissue can be distinguished.

6.2.3 Calculating the Dielectric Constant

If the monopole model (see section 3.4) is to be used to predict the relative near-field scattering signals across a range of wavelengths, the far-field absorbance spectrum shown in figure 6.3 must

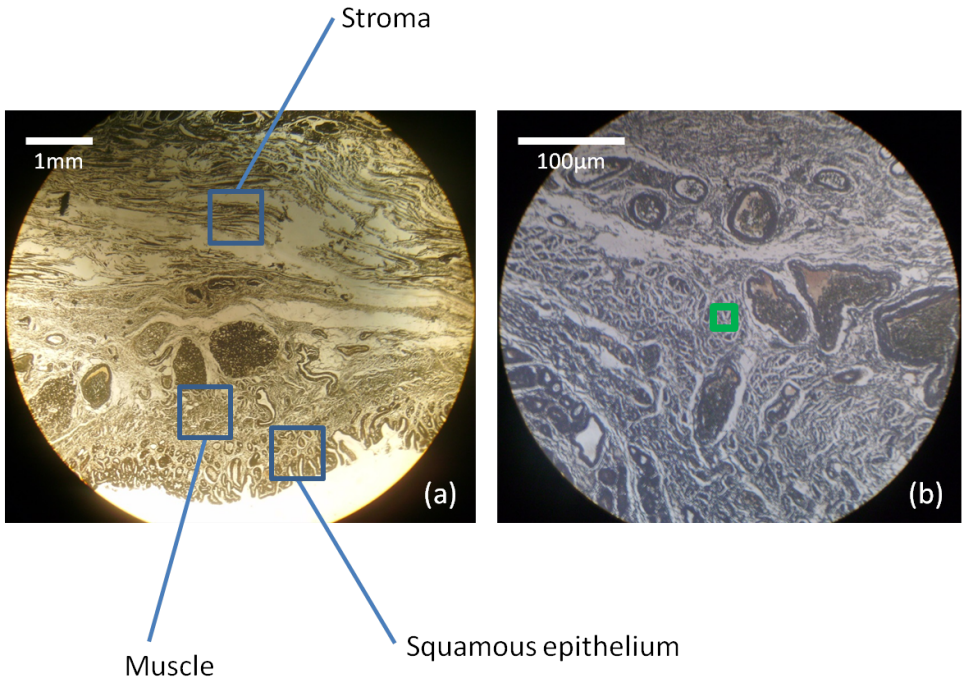

Figure 6.2: Conventional light microscope images of the oesophagial cancer sample, taken at (a) 2.5x magnification and (b) 40x magnification. Image (a) shows the structure of the entire sample, while image (b) shows a close-up of the epithelial cells. The green box shows the approximate area which was scanned during the s-SNOM imaging (see figure 6.5).

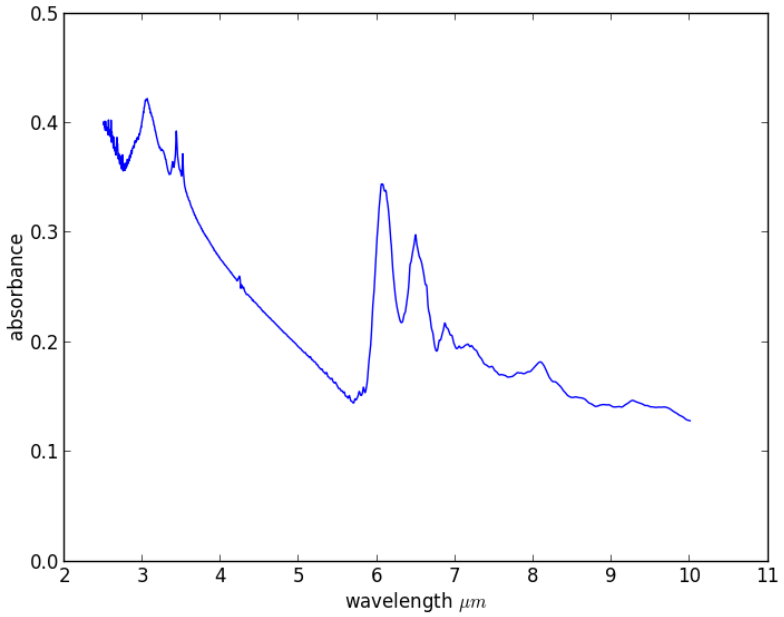

Figure 6.3: The absorbance spectra of a bulk 5mm^2 area of the tissue biopsy, capturing portions of the epithelial, muscle and stroma cells visible in figure 6.2. Typical biological absorption peaks, as labelled in figure 6.1, are clearly visible.

be used to calculate the real and imaginary parts of the sample's dielectric constant, $\epsilon_s(\lambda)$. The first step is to calculate the absorption coefficient, α, from the dimensionless absorbance A. The absorbance A is expressed as a base-10 logarithmic ratio of the incident light intensity I_0 and the transmitted light intensity I:

$$\frac{I}{I_0} = 10^{-A} \tag{6.1}$$

The absorption coefficient α, meanwhile, is found from the Beer-Lambert law, usually expressed as:

$$\frac{I}{I_0} = \mathrm{e}^{-\alpha z} \tag{6.2}$$

where z is the thickness of the sample, taken to be 5μm in our case. The absorption coefficient α can then simply be found by:

$$\alpha = \frac{A \ln(10)}{z} \tag{6.3}$$

The next step in finding $\epsilon_s(\lambda)$ is to convert the absorption coefficient to the imaginary part of the sample's refractive index, κ. This is done by:

$$\kappa = \frac{\alpha\lambda}{4\pi} \tag{6.4}$$

Finding the real part of the sample's refractive index, n is more complicated. It is achieved by using the Kramers-Kronig relationship, which links the real and imaginary parts of any function that is analytic for positive imaginary numbers (or in other words, that the physical system being modelled is causal) - a condition which the refractive index meets. The relationship between α and n, therefore, is as follows:

$$n(\lambda) = 1 + \frac{1}{2\pi^2}\mathcal{P}\int_0^\infty \frac{\alpha(\lambda')}{1 - \frac{\lambda'^2}{\lambda^2}}\,\mathrm{d}\lambda' \tag{6.5}$$

where $\mathcal{P}$ represents the Cauchy principle part of the integral. The real and imaginary parts of the dielectric constant, ϵ_s' and ϵ_s'', are then found by combining the real and imaginary parts of the refractive index, n and κ, using:

$$\epsilon_s' = n^2 - \kappa^2 \tag{6.6}$$

$$\epsilon_s'' = 2n\kappa \tag{6.7}$$

Note that in equation 6.5, the integral is taken over all wavelengths. The far-field absorbance data in figure 6.3 covering 2 to 10μm are, therefore, incomplete. To compensate for this, a constant value, n_∞, of 0.5 is added to the values of n to take into account the contribution of the higher energy oscillators to the refractive index [186] so that the average value of n is approximately 1.5 for the tissue sample in accordance with reference [187]. The resulting values of ϵ_s' and ϵ_s'' against the wavelength are shown in figure 6.4.

6.2.4 Imaging the Sample

Figure 6.5 shows a scan of a group of squamous epithelial cells taken at $\lambda = 9.3\mu$m. The material visible between the cells is likely to be extraneous protein, released during the fixing process, or paraffin wax that was not completely removed during the xylene wash. In the near-field scattering magnitude image, s_2, the cells themselves clearly show an optical signal, whereas the extracellular material does not.

Figure 6.6 shows the near-field scattering magnitude s_2 and phase ϕ_2 of a smaller area of the epithelial cells (the blue box in figure 6.5) at a range of wavelengths. These images have a spatial resolution of 50nm - a level of detail that is both deeply sub-wavelength and sub-cellular. The image wavelengths were dictated by the frequencies at which our QCL lases stably enough for near-field measurements (see section 5.3.2). The measurements show that there is a near-field scattering signal peak close to 9.3μm, and that the signal strength decays away from this wavelength.

Note that the images at different wavelengths in figure 6.6 are not directly comparable - factors such as varying laser power and changing system alignment mean that pixel values are only comparable within a single image. The SNRs of each image, however, do provide a method to scrutinise the relative signal strengths as systematic differences from recording to recording cancel in the ratio. This takes advantage of the fact that the ratio between the laser's contribution to the disturbances, D_2, and the laser power is constant between the wavelengths - in other words, the laser fluctuations relative to its DC level are the same for each image. Table 6.1 summarises the

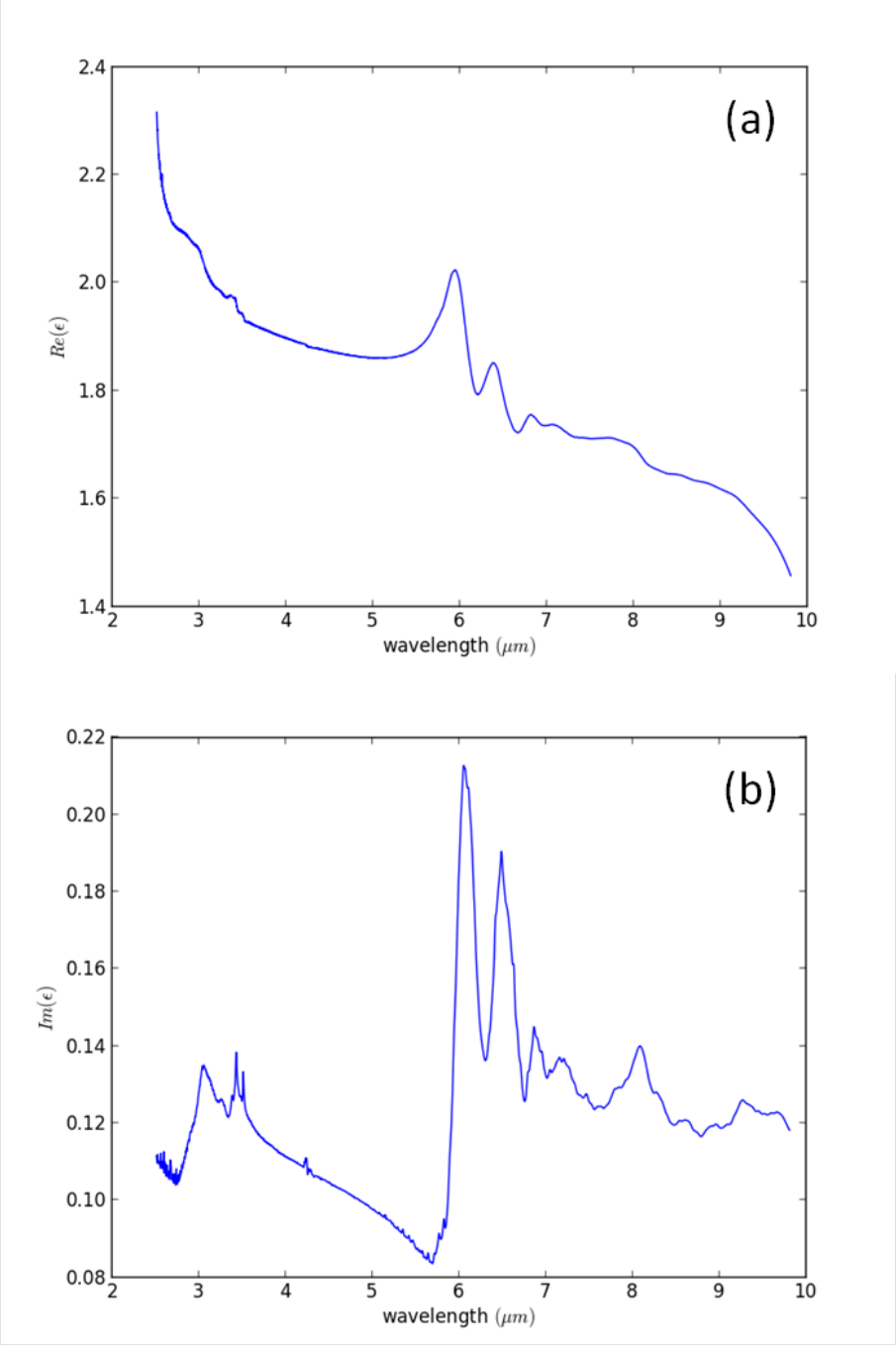

Figure 6.4: (a) The real part of the bulk tissue sample's dielectric constant ϵ_s' and (b) the imaginary part ϵ_s'' found from the far-field absorbance spectrum in figure 6.3 using the series of steps outlined in section 6.2.3. The values of ϵ_s' and ϵ_s'' can be used in the monopole model (see section 3.4) to calculate the near-field scattering signal's magnitude s_n and phase ϕ_n.

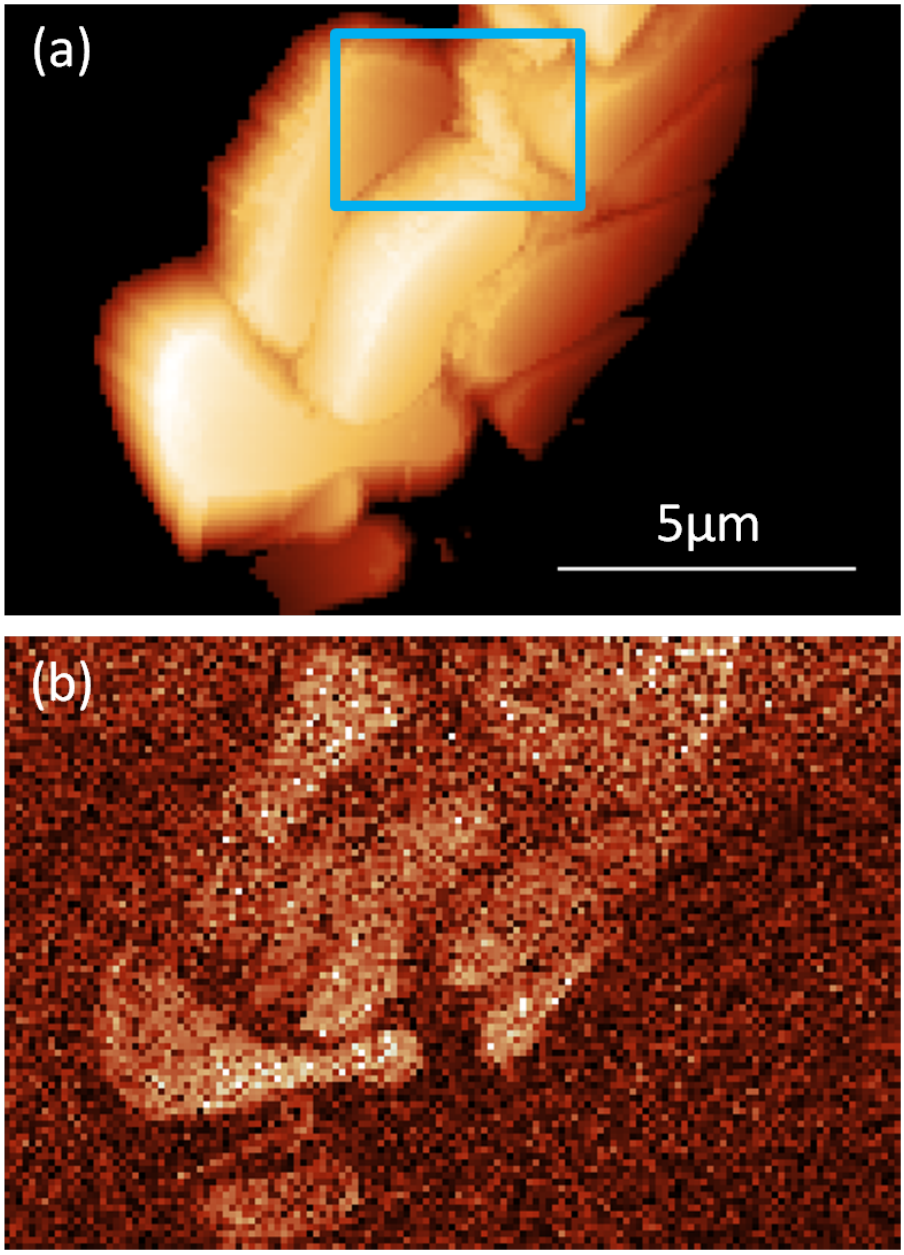

Figure 6.5: Large scale (a) AFM and (b) near-field scattering magnitude s_2 of squamous epithelial cells taken at $\lambda = 9.3\mu$m. The pixel resolution is 100nm. The black areas are where the sample is lower than the travel of the s-SNOM probe, and the blue rectangle highlights the area scanned at a variety of wavelengths in figure 6.6.

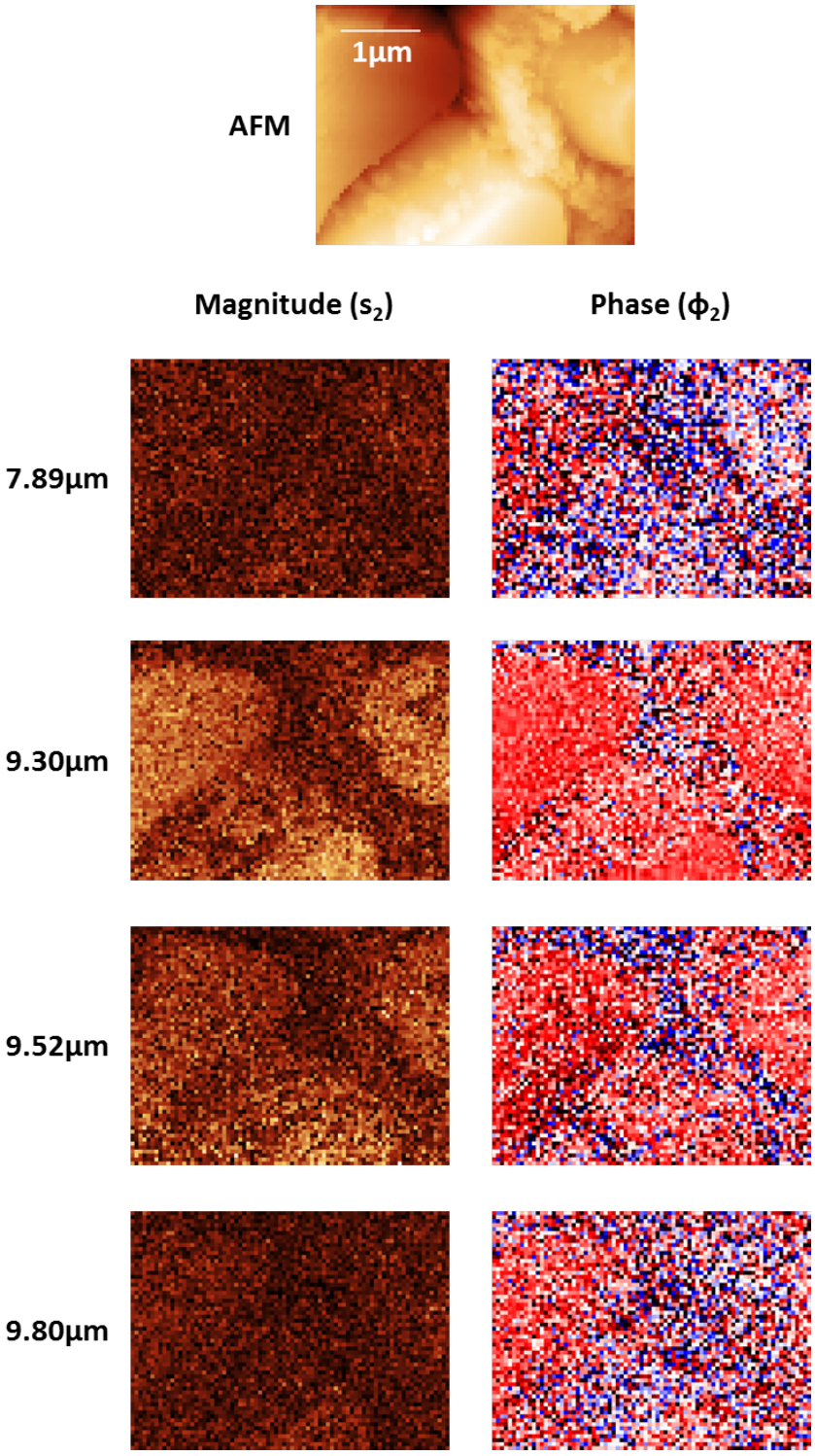

Figure 6.6: 50nm resolution images of the near-field scattering magnitude s_2 and phase ϕ_2 of epithelial cells at a range of wavelengths.The image acquisition time was approximately 2 minutes. The contrast comes from differing signal strengths between the cells themselves (bright areas) and the extracellular material (dark areas), which is either protein or paraffin wax.

SNRs of the near-field scattering magnitude (s_2/D_2) at each wavelength calculated in the same way as equation 5.1. The absolute signal values cannot be directly compared without a concurrent normalisation measurement - a silicon surface is often the most convenient [98, 99].

Wavelength	**SNR**
7.89μm	1.5
9.30μm	6
9.52μm	4
9.80μm	2

Table 6.1: Table showing the SNRs of the individual near-field scattering magnitude s_2 images of figure 6.6. These provide a measure of the relative contrast strengths which is free from systematic differences between the different imaging wavelengths.

6.2.5 Discussion of Results

The experimental results gathered in figure 6.6 can be compared against the predicted values by using the monopole model (section 3.4) and the dielectric constant values deduced from the far-field absorbance spectra. Table 6.2 summarises the relative contrast predictions of s_2 and ϕ_2.

Wavelength	**Predicted s_2**	**Predicted ϕ_2**
7.89μm	1	0 rads
9.30μm	0.84	0.026 rads
9.52μm	0.78	0.035 rads
9.80μm	0.64	0.061 rads

Table 6.2: Predicted values of the near-field scattering magnitude and phase using the monopole model and the values of the dielectric constant calculated from the bulk tissue absorbance spectrum in figure 6.3. Note that the values of s_2 and ϕ_2 have been normalised by the predicted values at 7.89μm.

There is clearly very little correlation between the predicted near-field scattering strengths (shown in table 6.2) and the experimental images (whose relative signal magnitudes are summarised in table 6.1). This is not particularly surprising - the predicted results are based on the average dielectric constant of a bulk tissue sample ϵ_s, which may have a very different value to that of the locally measured epithelial cells, ϵ_l.

Looking back at the predicted signal surfaces (section 3.4.4), it is clear that the largest signals are found for small, negative values of $\mathrm{Re}(\epsilon_\mathrm{s})$ with concurrently small values of $\mathrm{Im}(\epsilon_\mathrm{s})$. In naturally occurring materials, these conditions only exist close to absorption peaks [186], so we can deduce that the wavelengths at which the strongest signals are measured in figure 6.6 are approximately the same as those which are most strongly absorbed by the sample. Indeed, it has been shown that the imaginary part of the near-field scattering coefficient $\mathrm{Im}(\sigma)$ is proportional to the far-field absorption α [99]. Using the dielectric constant values of figure 6.4 and the monopole model, this relationship is demonstrated by figure 6.7. Note its similarity to the far-field absorbance spectra as shown in figure 6.3.

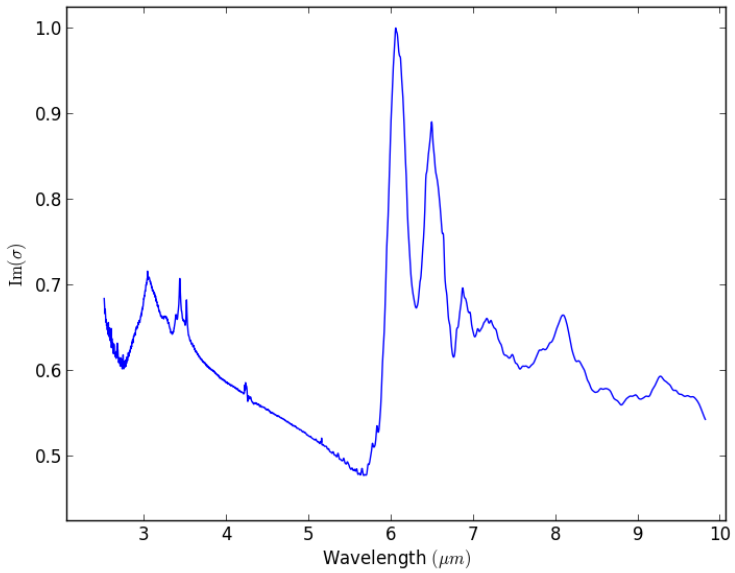

Figure 6.7: Plot of the predicted imaginary part of the near-field scattering coefficient $\mathrm{Im}(\sigma)$ using the monopole model (normalised to its maximum value). As expected [99], this looks very similar to the far-field absorbance spectrum of figure 6.3.

If stronger s-SNOM scattering signals are related to absorption, therefore, it is easy to find what might be causing the stronger scattering at 9.3μm. According to [188], the C-O bond of glycogen has a resonance at 9.3μm. While glycogen is usually found in liver rather than oesophagial cells, the epithelial cells may contain another compound rich in C-O bonds, or alternatively the stronger signal may be an artefact of the fixing process. Either way, it is clear that the extra-cellular material does not exhibit the same resonance, hence the contrast within the image.

Without access to a wider range of wavelengths, it is difficult to say with any certainty what the source of the stronger scattering signal is; there is another C-O resonance at 8.65μm for example. The lack of signal at $\lambda = 7.89\mu$m can only tell us that there are few phosphate ions (PO_3^-) on the sample. These are found in high numbers within cell nuclei, so either a nucleus has not been scanned in figure 6.6 or the scattering signal caused by the absorption of the phosphate is not sufficiently strong to overcome the noise.

6.3 Conclusion

The results presented in this chapter provide a glimpse of what s-SNOM can offer to the biological sciences. Given suitable sample preparation, we have shown that near-field measurements can produce meaningful contrast on tissue samples. We are, however, rather restricted by our available light source; the pulsed QCL is simply not stable enough at enough wavelengths to provide images that are suitable for truly spectroscopic interpretation.

Beyond this, we have also shown that bulk tissue spectra are not suitable for predicting near-

field scattering signals; the average dielectric constant of a sample ϵ_s can be very different to that of the local dielectric constant ϵ_l which is measured by s-SNOM. As such, the near-field scattering is highly dependent on the different absorption peaks of different kinds of tissues and cells.

In future work, much thinner tissue slices ($\sim$ 50nm if possible) should be imaged at a wider range of wavelengths. This will be a feasible option as QCL technology matures and more stable lasers become commercially available. Further, different types of tissue should be imaged to demonstrate the change in local absorption peaks - potentially at highly intracellular resolutions. This would introduce infrared chemical spectroscopy to the biosciences at a level of detail three orders of magnitude finer than is available with traditional far-field measurements.

Chapter 7

Towards Low Noise s-SNOM

"Suck it and see." **Prof. Chris Phillips**

7.1 Introduction

S-SNOM is an exquisitely sensitive technique capable of identifying chemical differences at resolutions approaching a few hundreds of molecules [182]. As it tries to pick out such small signals, however, it is also highly susceptible to the influence of noise, which can easily drown out any useful photovoltages in the measurement. To take the highest quality images, therefore, noise in s-SNOM measurements must be well understood and characterised so that its effects can be limited; doing so is critical for the development of s-SNOM as a useful spectroscopic tool.

The pseudo-heterodyne detection scheme (see section 2.6) is capable of achieving the gold standard in optical measurements - the shot noise limit. This is a fundamental quantum effect related to the fact that photons (and electrons) arrive at detectors in discrete lumps. In most experimental situations (including our setup), however, other sources of noise also play a role [189, 190]. The biggest contributor is usually laser intensity fluctuations, but Johnson (thermal) noise and amplifier noise can also play a significant part. The challenge, therefore, is to tease apart the noise contributions from different sources so that the effects of the most significant can be isolated and mitigated.

7.2 Electronic Sources of Noise

In a sense, all sources of noise are electronic; light is converted into electricity at the detector and the raw measurements are simply a series of voltages in time that are then Fourier transformed to find the particular frequency components of interest. Although the source of laser intensity fluctuations is not electronics, they ultimately enter the measurement as a voltage. It is useful, therefore, to characterise all sources of noise - laser and electronic - in terms of voltages. In this way they can be directly related to the voltages of the measured signals, which is useful for assessing the impact of any particular noise source.

7.2.1 Noise Power Spectral Density

The impact of different noise sources can vary significantly between different frequency ranges. This is, perhaps, unsurprising - silicon detectors used in the visible and near-infrared have very different properties to the mercury-cadmium-telluride (MCT) detectors that are used in the mid-infrared. The characteristics of different lasers are, of course, also different. What is constant between the different measurement wavelengths, however, is the spectral nature of the noise. Noise is often analysed in terms of its power spectral density - in other words, how much noise power is contained within a given frequency interval. A typical noise power spectral density graph from a detector looks like figure 7.1. At lower frequencies, the power spectral density drops as $1/f$. This is usually known as "flicker" noise. Its sources are not well defined - overhead lighting, building vibration and thermal fluctuations are all likely to play a role. Usually at around several kilohertz, the curve flattens out. In this region, the noise is "white" - or spectrally flat - and this defines the noise floor (its lower limit). For s-SNOM measurements, the contributions to the white region usually come from shot noise, Johnson noise and laser intensity fluctuations. Furthermore, the white noise region is the most important to understand as it is here that s-SNOM measurements take place (between 100 and 1000kHz). Finally, at some upper limit dependent on the bandwidth of the measurement electronics, there is a cutoff frequency. For our particular instrument (*NeaSNOM, Neaspec, Germany*), this is designed to occur at 1.25MHz.

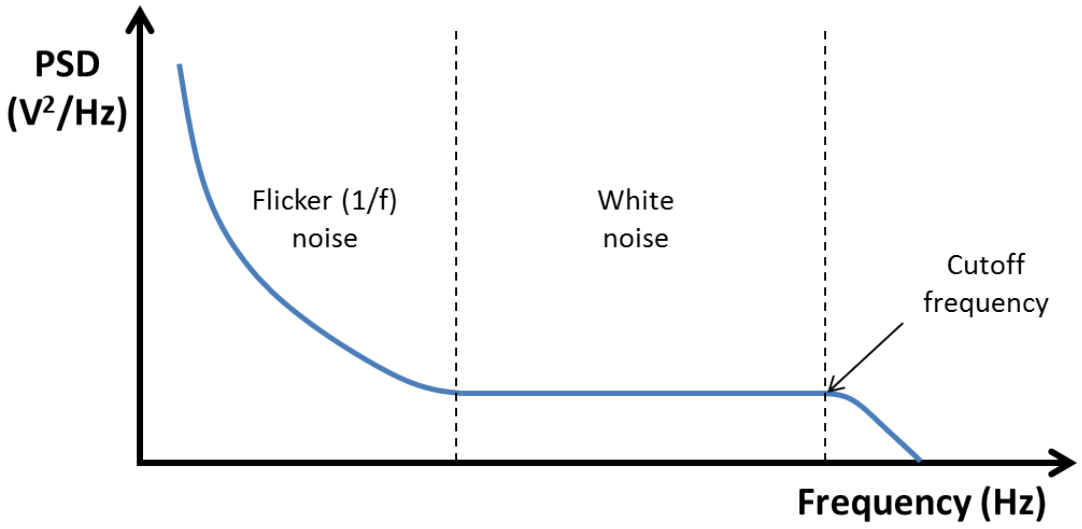

Figure 7.1: A typical noise power density spectrum (not to scale). At low frequencies (typically less than a few kHz) the curve has a $1/f$ characteristic. At middling frequencies, the power is spread equally resulting in a flat line. This "white" noise defines the noise floor - the lower limit. Finally, at some upper cutoff frequency dependent on the system bandwidth, the noise will drop to zero. For our s-SNOM, this occurs at 1.25MHz.

7.2.2 Equivalent Noise Circuit for Photodetection

A typical photodetection circuit will couple a reverse biased photodiode to a current-to-voltage amplifier (known as a transimpedance amplifier [191]). This particular setup strikes a good balance between sensitivity, linearity and speed [192]. The actual experimental circuit for a transimpedance amplifier is shown in figure 7.2. As the op-amp draws almost no current through its inputs, any photocurrent created at the photodiode I_{PC} must flow through the feedback resistor, labelled R_f, to the op-amp's output. Given that the both inputs must also be at zero volts, the voltage out V_{out} must be given by $-I_{PC}R_f$. If the current-to-voltage gain was to be 10,000V/A, one would choose R_f as a 10kΩ resistor.

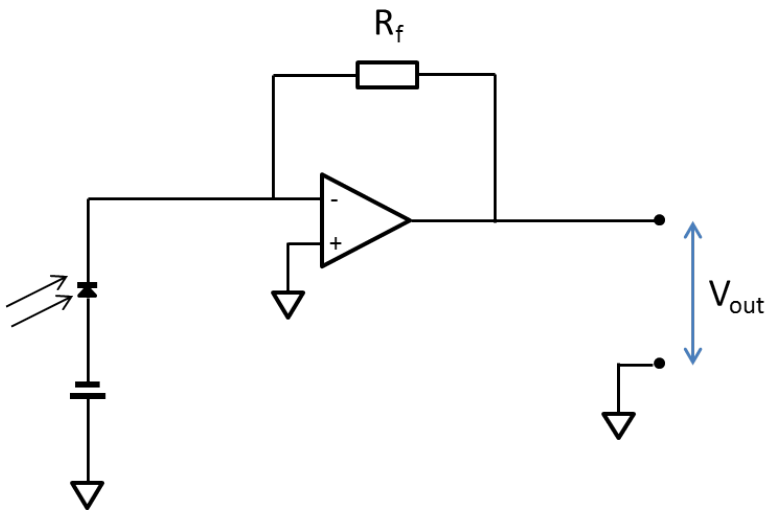

Figure 7.2: A simple photodetection circuit using a transimpedance amplifier. Current from the photodiode flows through the feedback resistor R_f resulting in an output voltage $V_{out} = -I_{PC}R_f$. The feedback resistor therefore controls the current-to-voltage gain.

Although an actual circuit diagram is useful for design purposes, it is often more convenient to study electronic noise by using an equivalent noise circuit. This simplifies analysis by breaking down complicated components into simpler ones combined with a number of current (or voltage) sources. The equivalent noise circuit for a photodiode and transimpedance amplifier combination is shown in figure 7.3.

The photodiode is broken down into its representative components: three distinct current sources for the photocurrent, the dark current and the shot noise; a capacitor to represent the parasitic junction capacitance of the diode; and a resistor to represent the shunt resistance. The dark current is the relatively small current which exists even when the photodiode is not illuminated. It is caused by the fact that the diode is reverse biased, and for this reason it can also be called the "reverse bias leakage current". The parasitic capacitance exists because a photodiode acts like a parallel plate capacitor; positive and negative charges are separated by the depletion region. If this capacitance is large it can significantly affect the operation of the detector at high frequencies. Finally, all photodiodes behave as if they have a shunt resistor across them. A low shunt resistance

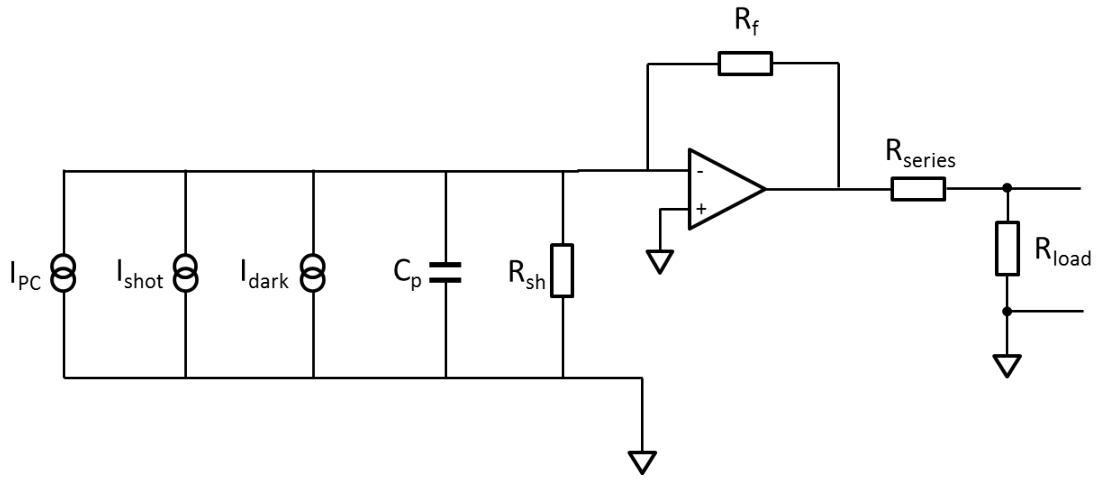

Figure 7.3: The equivalent noise circuit for the simple photodetection setup in figure 7.2. The photodiode has been replaced with three current sources representing the photocurrent, the shot noise and the dark current, plus a capacitor representing the parasitic capacitance and a resistor representing the shunt resistance. The series resistance (output impedance seen by the load) and the load resistance (input impedance seen by the photodetector) are also shown. The load resistance of our detection system is 2kΩ.

is caused by defects in the manufacturing process; values also vary enormously between different detector materials. For good silicon photodiodes it is typically in the gigaohm region whereas for MCT detectors it is often just a few hundred ohms. Low values can contribute significantly to the thermal noise.

The following sections look at the different sources of electronic noise individually, isolating their causes and giving typical values for both the near-infrared and mid-infrared setups. Most noise sources are dependent on the measurement bandwidth (see figure 3.14), so the following values are given for a detection bandwidth of 153Hz. This corresponds to 16384 (2^{14}) samples taken over a 6.5ms integration period - a good compromise between noise and measurement time for our particular system. Finally, as measurements of noise are inherently random, the values given for the different noise sources are actually the root mean square (rms) voltage fluctuations. As such they are displayed in lower case, v.

7.2.3 Shot Noise

Shot noise is present in both optical beams and electrical currents, and represents the uncertainty in a measurement of power and current respectively. It arises from the discrete nature of matter; photons and electrons arrive at their destination as individual particles meaning that Poisson statistics must be applied. It also represents the absolute noise floor; the shot noise limit cannot be overcome without exploiting quantum entanglement effects of multiple beams [193]. The formula for calculating the shot noise is given by:

$$v_{\text{shot}} = G_{\text{TI}}\sqrt{2qI_{\text{PC}}B} \tag{7.1}$$

where q is the electronic charge, I_{PC} is the photocurrent, B is the measurement bandwidth (153Hz in this case) and G_{TI} is the transimpedance gain of the detector. Shot noise is spectrally flat, so it does not matter at what frequency one is measuring - it is the same everywhere. For both the near and mid-IR setups, a typical DC value of the detected photocurrent is 0.25mA. Our MCT detector for the mid-IR has transimpedance gain of 10,000V/A. Our silicon detector for the near-IR has a variable gain, but is also usually run at a value of 10,000V/A. These values correspond to a shot noise voltage of 1.1μV for both frequency ranges. Although these values may look small compared to the DC photovoltages of around 3V, they can become significant during s-SNOM measurements, where the useful near-field scattering signal can also be of the order of microvolts.

7.2.4 Thermal Noise

Thermal noise (otherwise known as Johnson noise) is the second fundamental source of noise. It arises in all resistors at temperatures above absolute zero as random thermal movement of electrons cause voltage fluctuations. Like shot noise it is spectrally flat, and together with shot noise it makes up the white noise area of the typical noise power spectral density curve (see figure 7.1). The formula for thermal noise depends on whether it is coming from a resistance that is before or after the transimpedance amplifier. For a resistor before the amplifier it is given by:

$$v_{\mathrm{thermal}} = G_{\mathrm{TI}}\sqrt{\frac{4k_BTB}{R}} \tag{7.2}$$

and in the case of a resistor after the amplifier:

$$v_{\mathrm{thermal}} = \sqrt{4k_BTBR} \tag{7.3}$$

where T represents the temperature, k_B is Boltzmann's constant and R is the resistance in question. The load resistor, across which our s-SNOM system samples the voltage from the photodiode and amplifier combination, has a value of 2kΩ. Using equation 7.3, this corresponds to an rms voltage fluctuation of just 0.07μV. The shunt resistance of an MCT detector, however, can be rather low, resulting in a large thermal noise - particularly after amplification. For our particular detector the shunt resistance is 2800kΩ. The detector is cooled with liquid nitrogen to reduce T from room temperature at 293K to 77K. Applying equation 7.2 to these values yields a thermal noise of 0.15μV. Note that this only applies to measurements in the mid-IR - the shunt resistance of silicon detectors is typically gigaohms.

7.2.5 Other Noise Sources

There are many sources of electrical noise, but shot noise and thermal noise usually dominate. Other notable examples are the shot noise of the dark current, lattice generation-recombination noise and amplifier noise. The shot noise of the dark current can become significant when very small photocurrents are registered; if the dark current is larger, then its shot noise will be the dominant noise source. It is readily calculated using equation 7.1 with I_{PC} replaced with I_{dark}. The lattice generation-recombination noise is due to random fluctuations in the number of carriers

in a photodiode. Unfortunately its calculation requires knowledge of the majority carrier lifetime and mobility, neither of which is available in the data sheets of either our MCT detector or our silicon detector. The amplifier noise, however, is included in the information provided with the silicon detector and amounts to 0.04μV in our measurement bandwidth of 153Hz. We can assume that our MCT detector gives a similar value given that it also uses a transimpedance amplifier. Regardless, it is negligibly small compared to the shot noise of the photocurrent.

7.2.6 Summary of Electronic Noise Sources

In many senses, a discussion about noise without mentioning the signal levels is meaningless - what matters is the signal-to-noise ratio. We have mentioned that the DC photovoltage for both near-IR and mid-IR measurements is typically around 3V. This however, is not the *signal* level. The signal in this case refers to the size of the sidebands created by the pseudo-heterodyne detection system (see figure 2.5) as these are the fundamentally measured quantities.

The relationship between the DC level and the height of these sidebands is not trivial to calculate analytically; there a large number of parameters which are difficult to determine. The relationship also depends on the strength of the near-field scattering, and this in turn depends on the surface that is being scanned. For simple comparative purposes, therefore, the following values have been empirically determined from a typical measurement on a gold sample. For the mid-infrared (where near-field measurements are possible at the second harmonic of the probe frequency), the maximum sideband height is typically 85μV and in the near-infrared (where the third harmonic must be used), the maximum height is around 25μV.

	Mid-IR (MCT detector)	**Near-IR** (Si detector)
DC voltage	3V	3V
Sideband peak	85μV (n=2)	25μV (n=3)
Shot noise	1.1μV	1.1μV
Thermal noise	0.15μV (shunt) 0.07μV (load)	0.07μV (load)
Combined noise	1.1μV	1.1μV
SNR on gold	75	23

Table 7.1: Typical electrical noise values for s-SNOM measurements for a DC voltage of 3V and a 153Hz measurement bandwidth in the mid and near-infrared. Note that the different noise sources (corresponding to rms voltage fluctuations) are combined in quadrature.

The results from this section are summarised in table 7.1. It is clear that - electronically speaking - the limiting factor on the SNR is the shot noise. Unfortunately, there is little that can be done to improve upon the shot noise limit aside from increasing the integration period (thus reducing the measurement bandwidth but increasing image acquisition times). In practice, however, it is rare that these signal-to-shot noise ratios are actually reached (75 and 23 for a measurement on gold in the mid and near-infrared respectively). The real limiting factor for s-SNOM measurements lies in laser intensity fluctuations, as will be seen in the following section.

7.3 Laser Noise

Laser noise - or relative intensity noise (RIN) - is caused by fluctuations in the power output of a laser in time [194]. If a perfect laser measured by a perfect detector has a completely steady output, a noisy laser has a time-varying component:

$$V(t) = V_0 + x(t) \tag{7.4}$$

Where $V(t)$ represents the overall voltage measured by the detector and V_0 is the steady (or average) value. The laser fluctuations $x(t)$ become particularly important when measurements are carried out in frequency space. If they have Fourier components $X(f)$ at the same frequency as the signal of interest - the sidebands in an s-SNOM measurement - they will manifest themselves as an uncertainty in the height of the peaks. If the values of $X(f)$ are larger than the sidebands, then measurement is impossible (see figure 7.4).

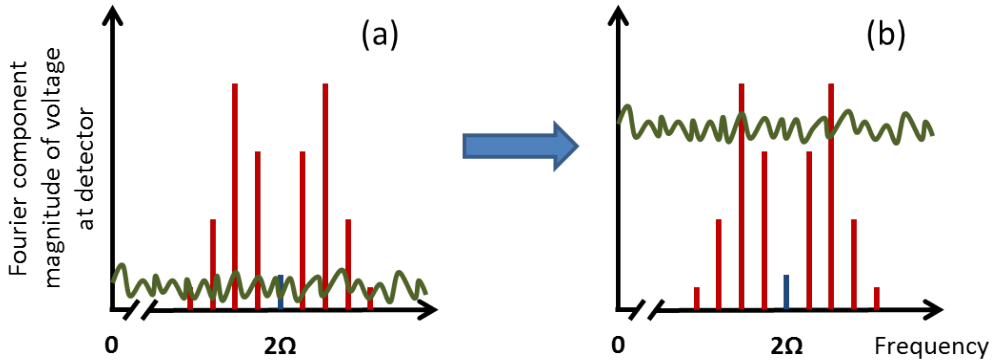

Figure 7.4: Representation of the frequency components at the probe vibration frequency (**blue**) and the sidebands created by the pseudo-heterodyne detection scheme (**red**). The Fourier components of the laser fluctuations $X(f)$ are represented by the **green** line. When the laser fluctuations are small - case (a) - their frequency components are also small and measurement of the sidebands is possible. When they are large however - case (b) - the sidebands are drowned out and measurement is impossible.

Laser noise manifests itself differently in time and frequency domains. In time it is simply a non-stationary component to the measured voltage whereas in the frequency domain it shows up as non-zero Fourier components over the measurement bandwidth of interest. There are many possible physical causes of laser noise; some of the most common include cavity vibration, thermal fluctuations in the gain medium and noise in the pump beam. These physical effects can cause mode hopping or small changes in alignment, either of which affects the power output.

7.3.1 Relating the Time and Frequency Domains

If we want to quickly establish which lasers are suitable for use with s-SNOM, we need to relate laser fluctuations in the time domain (which are readily measured) to their Fourier components in the frequency domain, where they will ultimately affect the measurements [195]. This is done using Parseval's theorem, which simply states that the total amount of power in the time domain must

be equal to the total amount of power in the frequency domain. This is, perhaps, unsurprising as a Fourier transform doesn't add any new information - it simply lets us think about the time domain signal in a different way. The power is found by summing the squares of the measurements in time or frequency, so Parseval's theorem can be expressed (in discrete form) by the following:

$$\sum_{n=0}^{N} |x[n]|^2 = \frac{N}{2} \sum_{k=1}^{N/2} |X[k]|^2 + NX[0]^2 \tag{7.5}$$

where N is the number of samples taken, $x[n]$ are the voltages at the detector in time and $X[k]$ represents the corresponding Fourier component magnitudes. Note that $x[n]$ and $X[k]$ correspond to $x(t)$ and $X(f)$ respectively after discretisation by the ADC sampling circuitry. The factors of N and 2 on the right hand side of the equation account for the normalisation of the Fourier transform. The variance of the time domain values, $x[n]$, can be written as follows:

$$\sigma^2 = \frac{1}{N} \sum_{n=0}^{N} (x[n] - V_0)^2 \tag{7.6}$$

which looks very similar to the left hand side of equation 7.5. In fact, if we exclude the DC value in the frequency domain (which is a valid thing to do as s-SNOM measurements never take place there), we can also remove the average voltage V_0 from the time domain. This means that we can directly substitute equation 7.6 into equation 7.5 after setting V_0 equal to zero:

$$N\sigma^2 = \frac{N}{2} \sum_{k=1}^{N/2} |X[k]|^2 \tag{7.7}$$

We are now left with a relationship between the variance of the voltage samples taken in time - a measure of the laser's noise - with the values of Fourier components $X[k]$. To simplify matters further, we can make the assumption that the distribution of the Fourier components is white throughout the spectrum (i.e. $\langle X[1] \rangle = \langle X[2] \rangle ... = \langle X[N/2] \rangle$). In this case:

$$\sum_{k=1}^{N/2} |X[k]|^2 = \frac{N}{2} \left\langle X[k]^2 \right\rangle \tag{7.8}$$

and we can rewrite equation 7.7 as:

$$\langle X[k] \rangle = \frac{2\sigma}{\sqrt{N}} \tag{7.9}$$

Equation 7.9 now gives us a simple way of relating the standard error $\sigma/\sqrt{N}$ of the laser fluctuations to the average value of their Fourier components $\langle X[k] \rangle$. If, for example, we require the Fourier components of the laser fluctuations to be less than 10μV (which would correspond to an SNR of around 2:1 for a near-infrared measurement on gold - see table 7.1), the standard deviation of the laser fluctuations in a 6.5ms window - 16384 samples at a sampling frequency of 2.5MHz - must be less than 0.64mV. If the impact of the laser fluctuations is to be less than that of the shot noise limit, the standard deviation must be less than 70μV in the same integration period.

Note, however, that these standard deviation limits on the laser fluctuations σ assume that the Fourier components $X[k]$ are indeed spectrally white (which is not necessarily the case). This issue is addressed in the following section.

7.3.2 Comparison of Laser Sources

Two of the most commonly used laser sources in the literature are helium-neon (HeNe) lasers in the visible (633nm) and carbon dioxide lasers in the mid-infrared (around 10μm). This is due to their stable lasing characteristics - they exhibit very small fluctuations - and as such they can be considered something of a "gold standard" in s-SNOM measurements. The experiments in this thesis have, for the most part, used either a titanium-sapphire laser in the near-infrared or a pulsed quantum cascade laser (QCL) in the mid-infrared. To directly compare these sources, figure 7.5 shows the full spectrum of their Fourier components during a near-field measurement on a gold sample. Each trace has been normalised to its DC level.

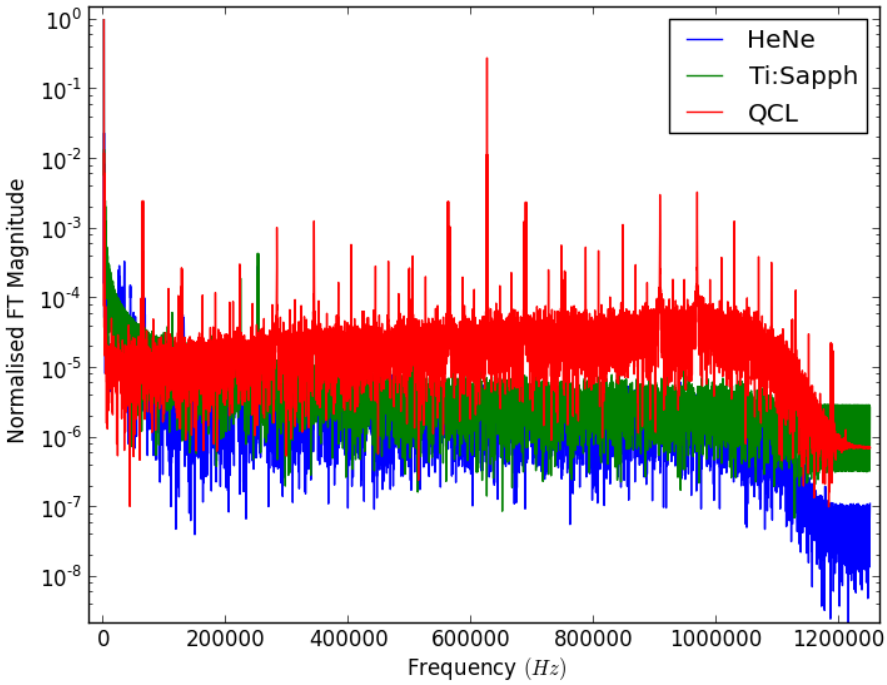

Figure 7.5: The full spectra of the Fourier transforms made during an s-SNOM measurement for three different lasers. The spectra have each been normalised to their DC level to allow for direct comparison. The probe was scanning a gold sample and the integration time was 6.5ms (corresponding to 16384 samples at a 2.5MHz sampling rate).

It is clear from these traces that the laser exhibiting the largest relative noise is the QCL, followed by the titanium-sapphire, followed by the HeNe. The un-normalised traces can also be used to test the assumption that the frequency spectrum of the laser fluctuations is white, or in other words, the accuracy of equation 7.9. The results are shown in table 7.2.

Laser Type	**σ of Laser Fluctuations**	**Predicted $\langle X[k]\rangle$**	**Actual $\langle X[k]\rangle$**
HeNe	0.87mV	13.6μV	2.5μV
Ti:sapph	1.30mV	20.3μV	3.8μV
QCL	3.47mV	108.4μV	22.1μV

Table 7.2: Typical electrical noise values for s-SNOM measurements in the mid and near-infrared. The number of samples used to calculate $X[k]$ was 16384 in an integration time of 6.5ms, corresponding to a detection bandwidth of 153Hz. The actual values of $\langle X[k]\rangle$ were found by averaging the values of $X[k]$ for each laser between 150 and 800kHz (excluding the sideband peaks). Note that, for the QCL, only the data points where the laser was "on" were included (1 in 4), meaning that the number of samples used to calculate $\langle X[k]\rangle$ in its case was just 4096.

It is clear from the approximately factor 5 difference between the predicted values of $\langle X[k]\rangle$ and the measured values of $\langle X[k]\rangle$ that equation 7.9 does not hold particularly accurately - it is useful for finding the acceptable order of magnitude of allowable laser fluctuations σ, but not for anything further. The assumption that the entire $X[k]$ spectrum is white, therefore, does not hold (which is, in fact, clear from figure 7.5), and this is the source of the factor of 5; considerably more power is concentrated within the lower frequencies (approximately $\leq$ 50kHz) than it would be for a truly white spectrum. The fact that the three lasers all have a similar factor between the predicted and measured values of $\langle X[k]\rangle$ is due to their similar power spectral densities.

The values from table 7.2 can be used to rewrite an empirical version of equation 7.9 by taking into account the factor of 5 between the predicted and actual values of $\langle X[k]\rangle$. If we wish, therefore, that the contribution from the laser at the measurement frequency $\langle X[k]\rangle$ is less than that of the shot noise (1.1μV for a typical DC photovoltage of 3V - see table 7.1), the following empirical condition must be satisfied:

$$\sigma < 2.75\sqrt{N}\ \mu\text{V} \tag{7.10}$$

In a 6.5ms integration window (16384 samples), this requires that σ is less than 0.35mV - roughly 0.01% of the DC value. None of the lasers presented in this thesis have such a low standard deviation, but it provides an ideal level by which to judge future light sources.

7.3.3 Improving the SNR: Attenuating the Reference Beam

The pseudo-heterodyne detection technique (see section 2.6) relies on the very weak scattered light from the probe being amplified by a much stronger reference beam. The majority of the detected light, therefore, is from the reference beam, not the probe. This means that the majority of the spectral "tail" that is seen in figure 7.5 for each laser is caused by light from the reference beam, not light scattered by the probe [196, 197]. If we want to reduce the impact of this tail, therefore, we need to reduce the amount of power in the reference beam, and we can do this by attenuation.

Remembering that the size of the sidebands created by the pseudo-heterodyne detection scheme are dictated by $\sqrt{V_{\text{REF}}V_{\text{SIG}}}$ (see equation 2.11), it is clear that the measured quantity scales with $\sqrt{V_{\text{REF}}}$. The Fourier components $X[k]$ caused by the laser fluctuations, on the other hand, scale

linearly (i.e. with V_{REF}). Attenuating the reference beam, therefore, affects the spectral tail more than it does the sidebands, and so can be used to improve the SNR. The limitation to this simple technique lies in the electronic noise - the sidebands *must* be larger than the shot noise, setting a limit to the amount the beam can be attenuated. In practise there is an optimum attenuation level which can be experimentally determined by using a variable attenuator in the reference beam path.

7.3.4 Additive vs Multiplicative Noise

Until now, the discussion of laser noise has centred on *additive noise*, or in other words the unwanted Fourier components caused by laser fluctuations at the same frequency as the measured sidebands. These cause an uncertainty in the value of the sidebands resulting in a noisy measurement. There is, however, a second way in which the laser fluctuations can affect the measurement. The values of the sidebands themselves also vary with changes in the DC level of the laser; the more powerful the beam, the stronger the near-field scattering signal. This is *multiplicative noise*. It tends to have considerably less impact than the additive noise - the integration windows are relatively long (at least 6.5ms) and the average DC level changes relatively little from window to window.

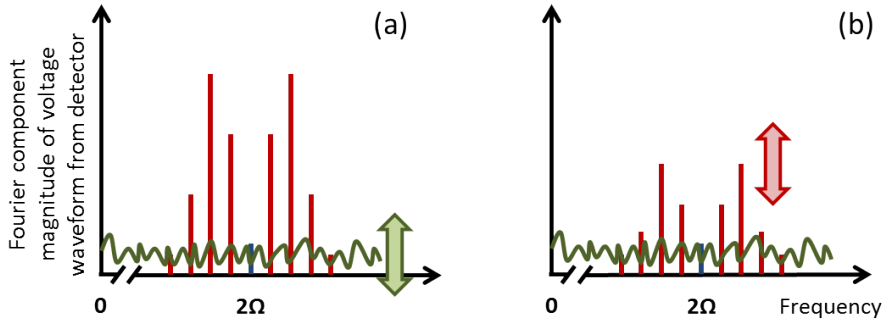

Figure 7.6: Diagram showing the difference between *additive* and *multiplicative* noise. The former, caused by laser fluctuations within an integration window, presents itself as Fourier components at all frequencies (green trace), some of which will be at the same frequencies as the sidebands to be measured. Changes in the standard deviation of the laser fluctuations show up as larger (or smaller) Fourier components across all frequencies as represented by the green arrow in figure (a). With multiplicative noise, however, changes in the average laser intensity between different integration windows show up as variations in the sideband heights, represented by the red arrow in figure (b).

7.4 Adding a Reference Channel

A more complex method for increasing the SNR (than simple attenuation of the reference beam) involves using a second detector as a reference channel. The benefit to this approach is that the signal amplification need not be affected by the noise reduction: as well as decreasing the noise, the signal can also be increased. The purpose of a reference channel is to collect information purely about the intensity of the laser beam. This can be used in a number of schemes to stop the laser fluctuations from affecting the measurement. Note that the term *reference channel* refers to an

additional beam path ending in a detector. This is a separate entity to the *reference beam* which is one of the interferometric arms of the pseudo-heterodyne detection scheme.

The setup for adding a reference channel to an s-SNOM measurement is shown in figure 7.7. A portion of the incident laser beam is "picked off" by a second beamsplitter (placed before the original beamsplitter used in the pseudo-heterodyne setup) and sent to a second detector. This new beam path is the reference channel. The two detectors register separate pieces of information; the original detector (labelled detector 2 in figure 7.7) contains the sidebands from the pseudo-heterodyne setup while the new detector contains information solely regarding the intensity of the laser. The voltages from the detectors are then combined - by subtraction or division - and the result then fed into the standard s-SNOM signal processing sequence (i.e. Fourier transform and measurement of the sideband peaks).

For our particular setup, the combination of the two detector signals must be done with analogue electronics. A simpler solution would digitise both signals individually and then combine them digitally prior to the Fourier transform. Our data acquisition (DAQ) circuitry, however, is only capable of digitising and transforming one channel.

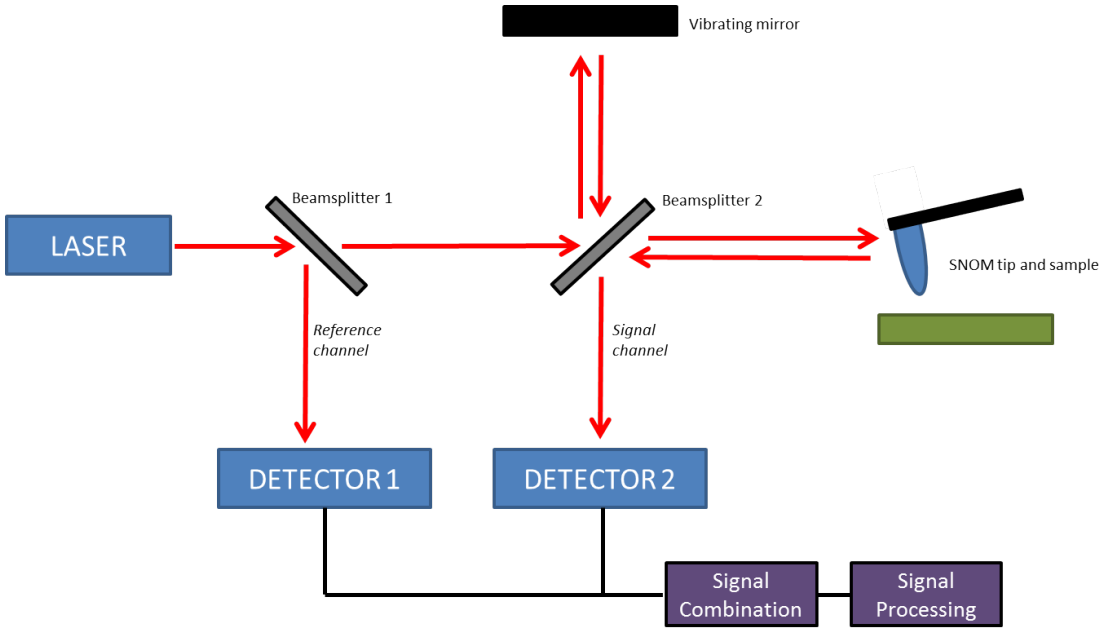

Figure 7.7: Setup for s-SNOM measurements with a reference channel. Another beamsplitter (labelled beamsplitter 1) is added to the original setup to pick off a portion of the incoming laser beam. This new beam path is the reference channel itself. The voltage signals from the two detectors are combined in some manner - usually by subtraction or division - and then passed on for s-SNOM signal processing.

7.4.1 Subtraction vs Division

The reference channel is used to remove the effects of laser fluctuations from the measurement. Two common ways of doing this electronically are through subtraction and division. In a subtractive scheme, the *reference photocurrent* - the photocurrent from Detector 1 in 7.7 - is subtracted from

the *signal photocurrent* (the photocurrent from Detector 2). In this setup, the beam intensities of each channel must be precisely adjusted such that the laser fluctuations (and the DC level) are exactly equal. If this is the case, the only remaining photocurrent is due to the pseudo-heterodyne scheme; the spectral tail caused by the fluctuations is eliminated.

In a division scheme, on the other hand, the signal photocurrent is divided by the reference photocurrent. Instead of removing the absolute laser fluctuations, this ratios them out of the measurement. Dividers have several benefits over subtractors; they do not require precise matching of the beam intensities in the experimental setup and they can also account for laser drift over long periods of time (they remove multiplicative as well as additive noise). For these reasons, the rest of this section will be focused on implementing a reference channel based on voltage division.

The major problem with electronic noise suppression is the relatively narrow bandwidth within which divider components work [198]. The performance of a feedback loop (on which dividers are based) is dictated by the gain error - the accuracy with which the output matches the mathematical operation performed on the input. The gain error increases with the measurement frequency, and so this sets a limit on the frequency at which the divider is effective. As a general rule of thumb, this occurs at around 1% of the device's cutoff frequency. For a divider based on components with a cutoff at 1MHz, therefore, the maximum frequency at which the gain error does not affect the noise suppression is about 10kHz.

7.4.2 Circuit Design for Voltage Division

Most voltage division circuits are based on voltage multipliers (figure 7.8) which are readily available as pre-packaged chips. Two inputs - the numerator and the denominator - are taken from the detectors. The output V_{out} is then given by $-V_{\text{num}}/V_{\text{den}}$. For our particular setup, an AD539 linear multiplier (60MHz bandwidth) was used in conjunction with an AD811 high performance op-amp (120MHz bandwidth). Both are manufactured by *Analog Devices*.

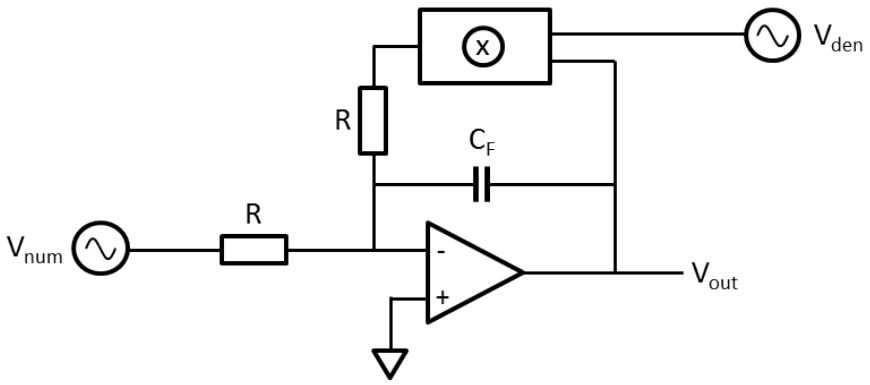

Figure 7.8: Circuit diagram for a voltage divider. As the op-amps input draws a negligibly small current, the current flowing through the two resistors must be equal. Furthermore, because the non-inverting input is pinned at ground, the feedback loop tries to keep the inverting input at 0V as well. As a result, $V_{\text{num}}/R = -V_{\text{out}}V_{\text{den}}/R$, or rearranging, $V_{\text{out}} = -V_{\text{num}}/V_{\text{den}}$. The op-amp input and outputs are decoupled by a feedback capacitor C_F. C_F is adjustable to account for variations in the high speed responses of different load conditions.

The chips were mounted in a "dead bug" configuration on a circuit board. General recommendations regarding the use of a good ground plane and power supply decoupling (with 0.47μF capacitors) were followed to try and prevent a deterioration in the divider's performance at high frequencies. Practically, there were limitations on the input voltages; the numerator had to be kept within ±6.3V and the denominator between 0.01 and 3.1V to prevent damage. Further, the divider's bandwidth (rather than the bandwidths of the individual components) was strongly affected by the value of the denominator as shown in figure 7.9. This further limited the value of the denominator to between 1 and 3.1V (to allow for a bandwidth greater than 1MHz).

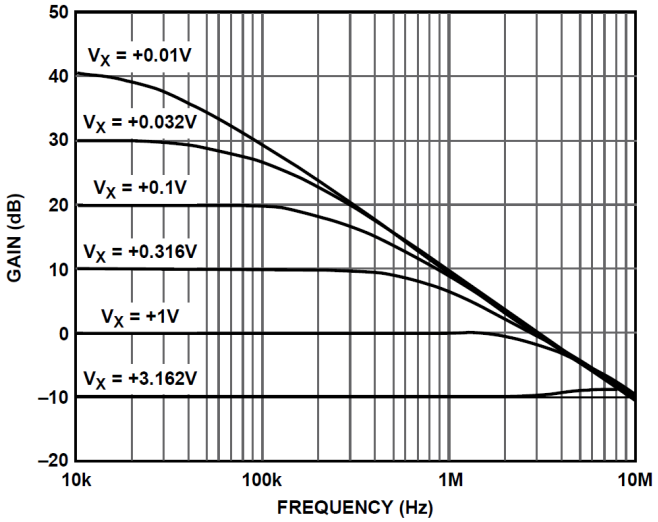

Figure 7.9: Dependence of the divider's bandwidth on the value of the denominator (labelled V_x) with C_F = 7pF. For the bandwidth to be larger than 1MHz, the denominator must be larger than 1V. Figure taken from the *AD539* datasheet.

7.4.3 Testing the Divider

To examine how well our voltage divider worked, it was placed in our s-SNOM setup as shown in figure 7.7. The titanium-sapphire laser - which we know to be reasonably close to shot-noise limited (see section 7.3.2) - was used in conjunction with two identical silicon detectors (*PDA36A, Thorlabs*). To maximise the similarity between the two channels, both beam paths were designed to have the same path length and identical lengths of cable were used to connect the detectors to the divider. Variable attenuators were used to ensure that the signal channel recorded approximately 1V and the reference channel around 2V. The divider output would therefore be close to 0.5V. As before, the integration period was 6.5ms, or 16384 samples.

Figure 7.10 compares the time traces and their corresponding Fourier transforms for both a single channel and the divider outputs. It is clear that instead of reducing the laser fluctuations, the voltage divider adds a substantial amount to them. In keeping with this, across much of the spectrum (greater than approximately 200kHz) the voltage divider's Fourier components are an order of magnitude larger than the single channel's. Although this may not seem particularly

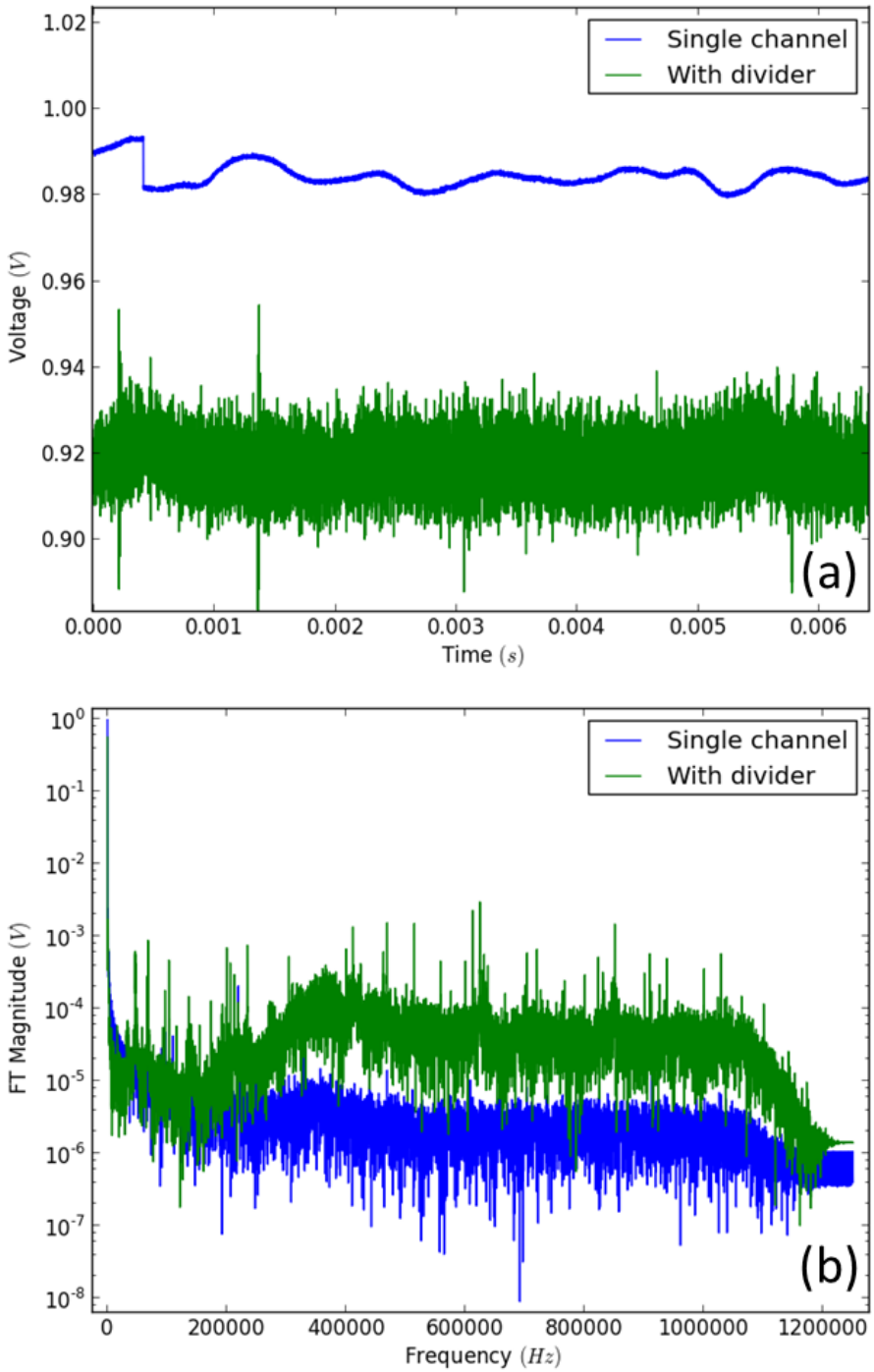

Figure 7.10: (a) Time traces from a single channel (**blue**) and from the voltage divider (**green**) and (b) their Fourier transforms. The integration period was 6.5ms. The green time trace from the voltage divider in (a) has been offset by 0.4V for easier comparison with the single channel trace.

promising as a noise reduction technique, there are some situations where this would it would indeed offer an SNR improvement.

Average DC value of single channel	1V
Shot noise limit of single channel	0.6μV
Electronic noise of single detector	1.8μV
Electronic noise of two detectors	2.5μV
Electronic noise of divider chip	2.5μV
Typical sideband value (n=3)	25μV
Typical sideband value (n=2)	85μV
Single channel average value at 750kHz	2.3μV
Voltage divider average value at 750kHz	59μV
Single channel average value at 130kHz	3.2μV
Voltage divider average value at 130kHz	5.4μV

Table 7.3: Table summarising the different contributions to the recorded voltage at a given frequency based on figure 7.10. The measurement bandwidth is 153Hz. The average Fourier component values are found by averaging within a ±10kHz window. The values of 750 and 130kHz are based on typical measurement frequencies in the near-infrared (n=3 using 250kHz resonant frequency probes) and the mid-infrared respectively (n=2 using 65kHz resonant frequency probes).

Table 7.3 summarises the important values for assessing the performance of the voltage divider. The first thing to note is that the titanium-sapphire laser is quiet enough that the laser noise from a single channel at 750kHz (2.3μV) is already less than the electronic noise of two detectors combined (2.5μV). For this particular laser, therefore, using two channels will always result in a reduced SNR.

Looking at (b) in figure 7.10 it is clear that the low frequency response of the divider is not dissimilar to that of the laser alone. The main difference is the presence of sharp peaks in the divider's spectrum which are likely caused by electrical interference or radio frequency pickups. It is only after 200kHz that the Fourier components increase dramatically in size in comparison with the single channel. This rise is caused by the gain-error problem outlined in section 7.4.1. Clearly the divider is unusable at frequencies above 200kHz - the average Fourier component at 750kHz (59μV) is more than double that of a typical sideband height that would be measured here (25μV).

7.4.4 Using the Voltage Divider

It is in the lower frequency region (less than 200kHz), therefore, that the voltage divider has potential applications. The average Fourier component of the voltage divider at 130kHz is 5.4μV - considerably lower than a typical sideband value of a measurement in the mid-infrared (i.e. n=2) at 85μV. This lower frequency is easily reached by using probes resonant at 65kHz - in fact most of the experiments in chapters 5 and 6 were carried out using just such probes. The voltage divider would improve the SNR of the measurement if the average Fourier component of a single channel was larger than 5.4μV. Using equation 7.10 as an order of magnitude guide, this corresponds to any laser producing time domain fluctuations with a standard deviation larger than 1.75mV (in a 6.5ms integration window).

Despite the SNR problems due to pulse-to-pulse variation of our own QCL (*Lasertune*, Block

Engineering), we were unable to test it with our voltage divider. The pulsed nature of the laser presents some challenging problems for voltage division using analogue electronics. When the laser is off - and this is around 95% of the time - the voltage divider has to deal with ratioing 0/0. This, unfortunately, did not give a consistent result. Efforts were made to construct a "sample and hold" circuit for the denominator which would sample a pulse and then hold its value until the next pulse arrived (thus avoiding the dividing by zero problem), but there were further complications in making this switch fast enough to capture each pulse reliably. This is an area for future work, and one with the potential to greatly increase the usefulness of our pulsed QCL.

7.5 Conclusion

Noise affects all measurements, and particularly those where the signal is very weak as in s-SNOM. This chapter has shown, however, that the pseudo-heterodyne detection technique is capable of producing close to electronic-noise limited results *when the laser noise is low*, as is the case for our helium-neon and titanium-sapphire lasers. When the laser fluctuations are large, however, the additive laser noise can dominate and render near-field measurement impossible. The empirically determined equation 7.10 shows us that for the spectral tail to contribute less than the shot noise, the standard deviation of the laser fluctuations must be less than approximately 0.01% of the DC value for an integration windows of 6.5ms.

Little can be done to improve upon the electronic noise. For both the near and mid-infrared the biggest contributor is the shot noise, and this is a fundamental statistical limit. With our nitrogen cooled MCT detector, the thermal noise is already an order of magnitude smaller than the shot noise - equipment or experimental design changes will not make drastic improvements here.

By far the biggest impact on the SNR, therefore, is the choice of laser. A steady output with low intensity fluctuations is critical. If a laser which exhibits significant fluctuation must be used, there are steps that can be taken to mitigate its effect on the measurement. The simplest is to attenuate the reference beam. The attenuation affects the noise more than it does the signal, resulting in a larger SNR. A second, more involved method requires the use of a voltage divider. This uses a second beam path and detector to ratio out laser fluctuations. There are, however, several caveats to using this approach: firstly, it will only be effective within a narrow bandwidth (less than 200kHz); secondly, for all but relatively noisy lasers the voltage divider will actually add noise due to the additive effects of using two separate channels; and finally, it will only function with CW beams. Future work in this area might take advantage of more complex - but also larger bandwidth - laser noise cancellation schemes [199, 200].

Chapter 8

Conclusion and Further Work

"Back to the coal face." **Prof. Chris Phillips**

8.1 Conclusion

The purpose of the work presented in this thesis was to explore the applications of s-SNOM in the infrared part of the spectrum. These applications fall into two main categories; measurements of plasmonic resonances (largely in the near-infrared) and the early implementation of single wavelength nanoscale spectroscopy by use of a tuneable laser (in the mid-infrared). For the former, a number of novel plasmonic structures have been examined, including broadband log-periodic nanoantennas, dolmens and cruciform pentamers. For all but the broadband antennas, this was made possible by coupling a titanium-sapphire laser (tuneable between 700 and 1000nm) to our s-SNOM system.

In terms of the development of nanoscale spectroscopy, a significant instrumentation barrier was overcome by making use of a pulsed quantum cascade laser (tuneable between 6 and 10μm). Although this proved to be experimentally viable, our particular device (*Lasertune, Block Engineering*) was not an ideal light source due to the large pulse-to-pulse variations it produced at all but a select number of wavelengths, thus hampering any attempts at truly spectroscopic measurements. Despite this, near-field images of epithelial tissue samples were taken for the first time at a number of frequencies allowing for at least a rudimentary interpretation of their chemical composition.

One constant issue in s-SNOM measurements - at any wavelength range - is that of noise. As seen in chapter 7, noise enters the system in a variety of ways, but the dominant source here tends to be laser fluctuations. Choosing a suitable laser, therefore, is on the critical path for the development of the system as it has significant knock-on effects; not only does it define whether measurements are even possible, but it also dictates their sensitivity, or conversely, the speed at which they can be carried out. Although the technique for implementing a pulsed laser presented in chapter 5 opens the door for a wider choice of lasers, it is absolutely critical that the fluctuations of any source be as low as possible; a standard deviation that is less than 0.01% of the DC value

is ideal for measurements in a detection bandwidth of around 150Hz.

8.2 Further Work

If our s-SNOM is to be used to image complex plasmonic samples - in either the near or mid-infrared - it is critical that a transmission mode setup is implemented. This will prevent retardation effects due to non-normally incident beam angles from entering the measurement, which render many interesting phenomena unobservable (as seen in sections 4.4 and 5.4). Once this barrier has been overcome, the tuneability of both the titanium-sapphire laser and the pulsed QCL can be used to study the wavelength dependence of different structures; this is particularly applicable to linewidth narrowing effects such as Fano resonances.

Theoretically, work needs to be done to clarify the precise relationship between the incident field's polarisation, the presence of the tip and the sample structures themselves on the light scattered by the probe. It still remains to be established that the cross-polarisation technique (see section 4.2), which is used to suppress chemical contrast, is viable under all experimental conditions - what role, for example, do the structures play in scattering light? And what role does the tip play in enhancing the scattering of certain polarisations above others? These are all questions that must be answered before s-SNOM can be used with absolute confidence in the direct measurement of a plasmonic resonance supported by any arbitrarily shaped structure.

Much of this thesis has been based on developing the prerequisite tools for nanoscale infrared spectroscopy. Chapter 6 showed the first glimpse of these tools in action, and there is huge scope for further applications and instrumentation enhancements. The first major improvement would be to replace our current QCL source with a more stable laser. If this is not possible, a digital - rather than analogue - voltage division scheme could be enacted (see section 7.4), which would prevent the divider from becoming a noise source (instead of a noise solution) as it is for the majority of cases at the moment.

With these instrumentation problems solved, the images of biological material (section 6.6) could be expanded upon - a wider range of wavelengths and different types of tissues could be explored. Gathering chemical information at such an intracellular level could lead to the ability to distinguish healthy and unhealthy cells, or examine how drugs are dispersed within the body. This thesis has established that near-field imaging is possible on tissue samples; future work will need to find the limitations and sensitivities of the technique, and find whether or not it is applicable to real-world problems.

Beyond the biosciences, infrared spectroscopy may initially find roles in the semiconductor industry for its ability to map dopant concentrations, or as part of a superlens imaging system which would completely remove the need for a probe to come into contact with a sample. In this regard, s-SNOM is as useful a tool as the applications to which it is applied. Many of these applications are, as yet, unforeseen, so to some extent s-SNOM remains a "solution waiting for a problem". What is certain, however, is that the future of s-SNOM is tightly bound to that of laser development; as lasers improve in both stability and tuneability, single wavelength s-SNOM will become an increasingly powerful means of imaging and identifying the nanoscale, and that will

only serve to broaden its appeal.

Bibliography

[1] L. Novotny and B. Hecht, *Principles of nano-optics*. Cambridge: Cambridge University Press, 2012.

[2] R. Erni, M. D. Rossell, C. Kisielowski, and U. Dahmen, "Atomic-resolution imaging with a sub-50-pm electron probe," *Physical Review Letters*, vol. 102, p. 096101, Mar. 2009.

[3] L. Gross, F. Mohn, N. Moll, P. Liljeroth, and G. Meyer, "The chemical structure of a molecule resolved by atomic force microscopy," *Science*, vol. 325, pp. 1110–1114, Aug. 2009.

[4] S. V. Kalinin and A. Gruverman, *Scanning Probe Microscopy: Electrical and Electromechanical Phenomena at the Nanoscale*. Springer, 2007 edition ed., Dec. 2006.

[5] B. Bhushan, *Scanning Probe Microscopy in Nanoscience and Nanotechnology*. Springer, 2012 edition ed., Oct. 2012.

[6] S. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated-emission - stimulated-emission-depletion fluorescence microscopy," *Optics Letters*, vol. 19, pp. 780–782, June 1994.

[7] T. A. Klar and S. W. Hell, "Subdiffraction resolution in far-field fluorescence microscopy," *Optics Letters*, vol. 24, pp. 954–956, July 1999.

[8] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, pp. 1642–1645, Sept. 2006.

[9] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature Methods*, vol. 3, pp. 793–795, Oct. 2006.

[10] E. H. Synge, "A suggested method for extending microscopic resolution into the ultra-microscopic region," *Philosophical Magazine Series 7*, vol. 6, no. 35, p. 6, 1928.

[11] E. Ash and G. Nicholls, "Super-resolution aperture scanning microscope," *Nature*, vol. 237, no. 5357, pp. 510–&, 1972.

[12] D. W. Pohl, W. Denk, and M. Lanz, "OPTICAL STETHOSCOPY - IMAGE RECORDING WITH RESOLUTION LAMBDA/20," *Applied Physics Letters*, vol. 44, no. 7, pp. 651–653, 1984.

[13] "http://www.lovalite.com."

[14] J. Gersten and A. Nitzan, "ELECTROMAGNETIC THEORY OF ENHANCED RAMAN-SCATTERING BY MOLECULES ADSORBED ON ROUGH SURFACES," *Journal of Chemical Physics*, vol. 73, no. 7, pp. 3023–3037, 1980.

[15] P. F. Liao and A. Wokaun, "LIGHTNING ROD EFFECT IN SURFACE ENHANCED RAMAN-SCATTERING," *Journal of Chemical Physics*, vol. 76, no. 1, pp. 751–752, 1982.

[16] L. Novotny, A. Bouhelier, M. R. Beversluis, and A. Hartschuh, "Near-field optical spectroscopy with 20nm spatial resolution," *Optical Society of America*, 2003.

[17] F. Festy, A. Demming, and D. Richards, "Resonant excitation of tip plasmons for tip-enhanced raman SNOM," *Ultramicroscopy*, vol. 100, no. 3-4, pp. 437–441, 2004.

[18] J. Wessel, "SURFACE-ENHANCED OPTICAL MICROSCOPY," *Journal of the Optical Society of America B-Optical Physics*, vol. 2, no. 9, pp. 1538–1541, 1985.

[19] W. Denk and D. W. Pohl, "NEAR-FIELD OPTICS - MICROSCOPY WITH NANOMETER-SIZE FIELDS," *Journal of Vacuum Science & Technology B*, vol. 9, no. 2, pp. 510–513, 1991.

[20] F. Zenhausern, M. P. Oboyle, and H. K. Wickramasinghe, "APERTURELESS NEAR-FIELD OPTICAL MICROSCOPE," *Applied Physics Letters*, vol. 65, no. 13, pp. 1623–1625, 1994.

[21] F. Zenhausern, Y. Martin, and H. K. Wickramasinghe, "SCANNING INTERFEROMETRIC APERTURELESS MICROSCOPY - OPTICAL IMAGING AT 10 ANGSTROM RESOLUTION," *Science*, vol. 269, no. 5227, pp. 1083–1085, 1995.

[22] G. Binnig, C. F. Quate, and C. Gerber, "Atomic force microscope," *Physical Review Letters*, vol. 56, pp. 930–933, Mar. 1986.

[23] G. Binnig and H. Rohrer, "PATENT: tunnel-effect scanning electron microscope - has needle electrode close to surface of sample in high vacuum and carrying measurable tunnel current," 1981.

[24] G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, "Surface studies by scanning tunneling microscopy," *Physical Review Letters*, vol. 49, pp. 57–61, July 1982.

[25] B. Knoll, F. Keilmann, A. Kramer, and R. Guckenberger, "Contrast of microwave near-field microscopy," *Applied Physics Letters*, vol. 70, no. 20, pp. 2667–2669, 1997.

[26] H. T. Chen, R. Kersting, and G. C. Cho, "Terahertz imaging with nanometer resolution," *Applied Physics Letters*, vol. 83, no. 15, pp. 3009–3011, 2003.

[27] H. T. Chen, S. Kraatz, G. C. Cho, and R. Kersting, "Identification of a resonant imaging process in apertureless near-field microscopy," *Physical Review Letters*, vol. 93, no. 26, 2004.

[28] R. Hillenbrand and F. Keilmann, "Complex optical constants on a subwavelength scale," *Physical Review Letters*, vol. 85, no. 14, pp. 3029–3032, 2000.

[29] T. Taubner, R. Hillenbrand, and F. Keilmann, "Nanoscale polymer recognition by spectral signature in scattering infrared near-field microscopy," *Applied Physics Letters*, vol. 85, no. 21, pp. 5064–5066, 2004.

[30] M. B. Raschke, L. Molina, T. Elsaesser, D. H. Kim, W. Knoll, and K. Hinrichs, "Apertureless near-field vibrational imaging of block-copolymer nanostructures with ultrahigh spatial resolution," *Chemphyschem*, vol. 6, no. 10, pp. 2197–2203, 2005.

[31] J. M. Stiegler, Y. Abate, A. Cvitkovic, Y. E. Romanyuk, A. J. Huber, S. R. Leone, and R. Hillenbrand, "Nanoscale infrared absorption spectroscopy of individual nanoparticles enabled by scattering-type near-field microscopy," *Acs Nano*, vol. 5, no. 8, 2011.

[32] F. Keilmann and R. Hillenbrand, "Near-field microscopy by elastic light scattering from a tip," *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences*, vol. 362, no. 1817, pp. 787–805, 2004.

[33] M. Brehm, H. G. Frey, R. Guckenberger, R. Hillenbrand, D. Kazantsev, F. Keilmann, N. Ocelic, and T. Taubner, "Consolidating apertureless SNOM," *Journal of the Korean Physical Society*, vol. 47, pp. S80–S85, 2005.

[34] R. Hillenbrand, "Scattering-type near-field microscopy: From nanoscale infrared material recognition to superlens studies," *2007 Pacific Rim Conference on Lasers and Electro-Optics, Vols 1-4*, pp. 39–40, 2007.

[35] R. Hillenbrand and F. Keilmann, "Material-specific mapping of metal/semiconductor/dielectric nanosystems at 10 nm resolution by backscattering near-field optical microscopy," *Applied Physics Letters*, vol. 80, no. 1, pp. 25–27, 2002.

[36] A. J. Huber, A. Ziegler, T. Kock, and R. Hillenbrand, "Infrared nanoscopy of strained semiconductors," *Nature Nanotechnology*, vol. 4, no. 3, pp. 153–157, 2009.

[37] F. Keilmann, A. J. Huber, and R. Hillenbrand, "Nanoscale conductivity contrast by scattering-type near-field optical microscopy in the visible, infrared and THz domains," *Journal of Infrared Millimeter and Terahertz Waves*, vol. 30, no. 12, pp. 1255–1268, 2009.

[38] J. M. Stiegler, A. J. Huber, S. L. Diedenhofen, J. G. Rivas, R. E. Algra, E. P. A. M. Bakkers, and R. Hillenbrand, "Nanoscale free-carrier profiling of individual semiconductor nanowires by infrared near-field nanoscopy," *Nano Letters*, vol. 10, no. 4, 2010.

[39] T. Taubner, F. Keilmann, and R. Hillenbrand, "Nanoscale-resolved subsurface imaging by scattering-type near-field optical microscopy," *Optics Express*, vol. 13, pp. 8893–8899, 2005.

[40] R. Bachelot, G. Lerondel, S. Blaize, S. Aubert, A. Bruyant, and P. Royer, "Probing photonic and optoelectronic structures by apertureless scanning near-field optical microscopy," *Microscopy Research and Technique*, vol. 64, no. 5-6, pp. 441–452, 2004.

[41] S. C. Schneider, S. Grafstrom, and L. M. Eng, "Scattering near-field optical microscopy of optically anisotropic systems," *Physical Review B*, vol. 71, Mar. 2005.

[42] A. Charnukha, A. Cvitkovic, T. Prokscha, D. Prpper, N. Ocelic, A. Suter, Z. Salman, E. Morenzoni, J. Deisenhofer, V. Tsurkan, A. Loidl, B. Keimer, and A. V. Boris, "Nanoscale layering of antiferromagnetic and superconducting phases in rb_{2}fe_{4}se_{5} single crystals," *Physical Review Letters*, vol. 109, p. 017003, July 2012.

[43] N. Ocelic and R. Hillenbrand, "Subwavelength-scale tailoring of surface phonon polaritons by focused ion-beam implantation," *Nature Materials*, vol. 3, no. 9, pp. 606–609, 2004.

[44] A. Huber, N. Ocelic, D. Kazantsev, and R. Hillenbrand, "Near-field imaging of mid-infrared surface phonon polariton propagation," *Applied Physics Letters*, vol. 87, no. 8, p. 3, 2005.

[45] R. Vogelgesang and A. Dmitriev, "Real-space imaging of nanoplasmonic resonances," *Analyst*, vol. 135, no. 6, pp. 1175–81, 2010.

[46] J. Chen, M. Badioli, P. Alonso-Gonzlez, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenovi, A. Centeno, A. Pesquera, P. Godignon, A. Zurutuza Elorza, N. Camara, F. J. G. de Abajo, R. Hillenbrand, and F. H. L. Koppens, "Optical nano-imaging of gate-tunable graphene plasmons," *Nature*, vol. 487, pp. 77–81, July 2012.

[47] Z. Fei, A. S. Rodin, G. O. Andreev, W. Bao, A. S. McLeod, M. Wagner, L. M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, M. M. Fogler, A. H. C. Neto, C. N. Lau, F. Keilmann, and D. N. Basov, "Gate-tuning of graphene plasmons revealed by infrared nano-imaging," *Nature*, vol. 487, pp. 82–85, July 2012.

[48] K. Taniguchi and Y. Kanemitsu, "Development of an apertureless near-field optical microscope for nanoscale optical imaging at low temperatures," *Japanese Journal of Applied Physics Part 1-Regular Papers Brief Communications & Review Papers*, vol. 44, no. 1B, pp. 575–577, 2005.

[49] H. U. Yang, E. Hebestreit, E. E. Josberger, and M. B. Raschke, "A cryogenic scattering-type scanning near-field optical microscope," *Review of Scientific Instruments*, vol. 84, Feb. 2013.

[50] J. Sun, J. C. Schotland, R. Hillenbrand, and P. S. Carney, "Nanoscale optical tomography using volume-scanning near-field microscopy," *Applied Physics Letters*, vol. 95, no. 12, 2009.

[51] N. I. Zheludev, "What diffraction limit?," *Nature Materials*, vol. 7, no. 6, pp. 420–422, 2008.

[52] S. Jaeger, A. M. Kern, M. Hentschel, R. Jaeger, K. Braun, D. Zhang, H. Giessen, and A. J. Meixner, "Au nanotip as luminescent near-field probe," *Nano Letters*, vol. 13, pp. 3566–3570, Aug. 2013.

[53] Z. Wang, W. Guo, L. Li, B. Luk'yanchuk, A. Khan, Z. Liu, Z. Chen, and M. Hong, "Optical virtual imaging at 50 nm lateral resolution with a white-light nanoscope," *Nature Communications*, vol. 2, Mar. 2011.

[54] T. Kalkbrenner, U. Hakanson, A. Schadle, S. Burger, C. Henkel, and V. Sandoghdar, "Optical microscopy via spectral modifications of a nanoantenna," *Physical Review Letters*, vol. 95, no. 20, 2005.

[55] D. R. Honigstein, J. Weinroth, M. Werman, and A. Lewis, "Noniterative exact solution to the phase problem in optical imaging implemented with scanning probe microscopy," *Acs Nano*, vol. 6, pp. 220–226, Jan. 2012.

[56] T. Schmid, L. Opilik, C. Blum, and R. Zenobi, "Nanoscale chemical imaging using tip-enhanced raman spectroscopy: a critical review," *Angewandte Chemie (International ed. in English)*, vol. 52, pp. 5940–5954, June 2013.

[57] P. L. Stiles, J. A. Dieringer, N. C. Shah, and R. R. Van Duyne, "Surface-enhanced raman spectroscopy," in *Annual Review of Analytical Chemistry*, vol. 1, pp. 601–626, Palo Alto: Annual Reviews, 2008.

[58] J. B. Pendry, "Negative refraction makes a perfect lens," *Physical Review Letters*, vol. 85, no. 18, pp. 3966–3969, 2000.

[59] D. O. S. Melville and R. J. Blaikie, "Super-resolution imaging through a planar silver layer," *Optics Express*, vol. 13, no. 6, pp. 2127–2134, 2005.

[60] N. Fang, H. Lee, C. Sun, and X. Zhang, "Sub-diffraction-limited optical imaging with a silver superlens," *Science*, vol. 308, no. 5721, pp. 534–537, 2005.

[61] J. B. Pendry, D. Schurig, and D. R. Smith, "Controlling electromagnetic fields," *Science*, vol. 312, pp. 1780–1782, June 2006.

[62] V. M. Shalaev, "Optical negative-index metamaterials," *Nature Photonics*, vol. 1, no. 1, pp. 41–48, 2007.

[63] X. Zhang and Z. W. Liu, "Superlenses to overcome the diffraction limit," *Nature Materials*, vol. 7, no. 6, pp. 435–441, 2008.

[64] P. Chaturvedi and N. X. Fang, "Sub-diffraction-limited far-field imaging in infrared," *Frontiers of Physics in China*, vol. 5, pp. 324–329, Sept. 2010.

[65] V. A. Podolskiy and E. E. Narimanov, "Near-sighted superlens," *Optics Letters*, vol. 30, pp. 75–77, Jan. 2005.

[66] T. Taubner, D. Korobkin, Y. Urzhumov, G. Shvets, and R. Hillenbrand, "Near-field microscopy through a SiC superlens," *Science*, vol. 313, no. 5793, pp. 1595–1595, 2006.

[67] S. C. Kehr, Y. M. Liu, L. W. Martin, P. Yu, M. Gajek, S.-Y. Yang, C.-H. Yang, M. T. Wenzel, R. Jacob, H.-G. von Ribbeck, M. Helm, X. Zhang, L. M. Eng, and R. Ramesh, "Near-field examination of perovskite-based superlenses and superlens-enhanced probe-object coupling," *Nature Communications*, vol. 2, Mar. 2011.

[68] W. Cai, U. K. Chettiar, A. V. Kildishev, and V. M. Shalaev, "Optical cloaking with metamaterials," *Nature Photonics*, vol. 1, pp. 224–227, Apr. 2007.

[69] J. Valentine, J. Li, T. Zentgraf, G. Bartal, and X. Zhang, "An optical cloak made of dielectrics," *Nature Materials*, vol. 8, pp. 568–571, July 2009.

[70] J. Zhang, L. Liu, Y. Luo, S. Zhang, and N. A. Mortensen, "Homogeneous optical cloak constructed with uniform layered structures," *Optics Express*, vol. 19, pp. 8625–8631, Apr. 2011.

[71] R. A. Shelby, D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, vol. 292, no. 5514, pp. 77–79, 2001.

[72] H. J. Lezec, J. A. Dionne, and H. A. Atwater, "Negative refraction at visible frequencies," *Science*, vol. 316, no. 5823, pp. 430–432, 2007.

[73] A. J. Hoffman, L. Alekseyev, S. S. Howard, K. J. Franz, D. Wasserman, V. A. Podolskiy, E. E. Narimanov, D. L. Sivco, and C. Gmachl, "Negative refraction in semiconductor metamaterials," *Nature Materials*, vol. 6, no. 12, pp. 946–950, 2007.

[74] J. R. Plumridge, R. J. Steed, and C. C. Phillips, "Negative refraction in anisotropic waveguides made from quantum metamaterials," *Physical Review B*, vol. 77, no. 20, 2008.

[75] C. M. Soukoulis, J. F. Zhou, T. Koschny, M. Kafesaki, and E. N. Economou, "The science of negative index materials," *Journal of Physics-Condensed Matter*, vol. 20, no. 30, 2008.

[76] M. A. Noginov, G. Dewar, M. W. McCall, and N. I. Zheludev, eds., *Tutorials in Complex Photonic Media*. 1000 20th Street, Bellingham, WA 98227-0010 USA: SPIE, Dec. 2009.

[77] R. Esteban, R. Vogelgesang, J. Dorfmuller, A. Dmitriev, C. Rockstuhl, C. Etrich, and K. Kern, "Direct near-field optical imaging of higher order plasmonic resonances," *Nano Letters*, vol. 8, pp. 3155–3159, Oct. 2008.

[78] A. J. Huber, J. Wittborn, and R. Hillenbrand, "Infrared spectroscopic near-field mapping of single nanotransistors," *Nanotechnology*, vol. 21, no. 23, 2010.

[79] R. Hillenbrand and F. Keilmann, "Optical oscillation modes of plasmon particles observed in direct space by phase-contrast near-field microscopy," *Applied Physics B-Lasers and Optics*, vol. 73, no. 3, pp. 239–243, 2001.

[80] F. Ballout, H. Krassen, I. Kopf, K. Ataka, E. Bruendermann, J. Heberle, and M. Havenith, "Scanning near-field IR microscopy of proteins in lipid bilayers," *Physical Chemistry Chemical Physics*, vol. 13, no. 48, 2011.

[81] "Still feeling the force," *Nature Nanotechnology*, vol. 3, pp. 241–241, May 2008.

[82] A. Bek, R. Vogelgesang, and K. Kern, "Optical nonlinearity versus mechanical anharmonicity contrast in dynamic mode apertureless scanning near-field optical microscopy," *Applied Physics Letters*, vol. 87, no. 16, p. 3, 2005.

[83] R. Esteban, R. Vogelgesang, and K. Kern, "Tip-substrate interaction in optical near-field microscopy," *Physical Review B*, vol. 75, May 2007.

[84] B. Knoll and F. Keilmann, "Near-field probing of vibrational absorption for chemical microscopy," *Nature*, vol. 399, no. 6732, pp. 134–137, 1999.

[85] D. Haefliger, J. M. Plitzko, and R. Hillenbrand, "Contrast and scattering efficiency of scattering-type near-field optical probes," *Applied Physics Letters*, vol. 85, pp. 4466–4468, Nov. 2004.

[86] N. Behr and M. B. Raschke, "Optical antenna properties of scanning probe tips: Plasmonic light scattering, TipSample coupling, and near-field enhancement," *The Journal of Physical Chemistry C*, vol. 112, pp. 3766–3773, Mar. 2008.

[87] B.-S. Yeo, J. Stadler, T. Schmid, R. Zenobi, and W. Zhang, "Tip-enhanced raman spectroscopy - its status, challenges and future directions," *Chemical Physics Letters*, vol. 472, no. 1-3, 2009.

[88] S. Berweger, J. M. Atkin, R. L. Olmon, and M. B. Raschke, "Light on the tip of a needle: Plasmonic nanofocusing for spectroscopy on the nanoscale," *The Journal of Physical Chemistry Letters*, vol. 3, pp. 945–952, Apr. 2012.

[89] F. Huth, A. Chuvilin, M. Schnell, I. Amenabar, R. Krutokhvostov, S. Lopatin, and R. Hillenbrand, "Resonant antenna probes for tip-enhanced infrared near-field microscopy," *Nano Letters*, vol. 13, pp. 1065–1072, Mar. 2013.

[90] B. Knoll and F. Keilmann, "Enhanced dielectric contrast in scattering-type scanning near-field optical microscopy," *Optics Communications*, vol. 182, no. 4-6, pp. 321–328, 2000.

[91] P. G. Gucciardi and G. Bachelier, "Far-field background suppression in tip-modulated apertureless near-field optical microscopy," *Journal of Applied Physics*, vol. 99, no. 12, 2006.

[92] R. Hillenbrand, B. Knoll, and F. Keilmann, "Pure optical contrast in scattering-type scanning near-field microscopy," *Journal of Microscopy-Oxford*, vol. 202, pp. 77–83, 2001.

[93] I. Stefanon, S. Blaize, A. Bruyant, S. Aubert, G. Lerondel, R. Bachelot, and P. Royer, "Heterodyne detection of guided waves using a scattering-type scanning near-field optical microscope," *Optics Express*, vol. 13, pp. 5553–5564, July 2005.

[94] A. Zayats and D. Richards, *Nano-Optics and Near-Field Optical Microscopy*. Artech House, 2009.

[95] N. Ocelic, A. Huber, and R. Hillenbrand, "Pseudoheterodyne detection for background-free near-field spectroscopy," *Applied Physics Letters*, vol. 89, no. 10, 2006.

[96] S. Amarie, T. Ganz, and F. Keilmann, "Mid-infrared near-field spectroscopy," *Optics Express*, vol. 17, no. 24, pp. 21794–21801, 2009.

[97] S. Amarie, P. Zaslansky, Y. Kajihara, E. Griesshaber, W. W. Schmahl, and F. Keilmann, "Nano-FTIR chemical mapping of minerals in biological materials," *Beilstein Journal of Nanotechnology*, vol. 3, 2012.

[98] F. Huth, M. Schnell, J. Wittborn, N. Ocelic, and R. Hillenbrand, "Infrared-spectroscopic nanoimaging with a thermal source," *Nature Materials*, vol. 10, no. 5, 2011.

[99] F. Huth, A. Govyadinov, S. Amarie, W. Nuansing, F. Keilmann, and R. Hillenbrand, "Nano-FTIR absorption spectroscopy of molecular fingerprints at 20nm spatial resolution," *Nano letters*, vol. 12, no. 8, 2012.

[100] N. Hayazawa, A. Tarun, A. Taguchi, and S. Kawata, "Development of tip-enhanced near-field optical spectroscopy and microscopy," *Japanese Journal of Applied Physics*, vol. 48, no. 8, 2009.

[101] D. Courjon, *Near-Field Microscopy and Near-Field Optics*. Imperial College Press, 2003.

[102] S. A. Maier, *Plasmonics: Fundamentals and Applications*. Springer, 2007.

[103] J. A. Schuller, E. S. Barnard, W. Cai, Y. C. Jun, J. S. White, and M. L. Brongersma, "Plasmonics for extreme light concentration and manipulation," *Nature Materials*, vol. 9, pp. 193–204, Mar. 2010.

[104] L. Novotny, "The history of near-field optics," in *Progress in Optics, Vol 50*, vol. 50, pp. 137–184, Amsterdam: Elsevier Science Bv, 2007.

[105] R. Fikri, D. Barchiesi, F. H'Dhili, R. Bachelot, A. Vial, and P. Royer, "Modeling recent experiments of apertureless near-field optical microscopy using 2D finite element method," *Optics Communications*, vol. 221, pp. 13–22, June 2003.

[106] S. V. Sukhov, "Role of multipole moment of the probe in apertureless near-field optical microscopy," *Ultramicroscopy*, vol. 101, pp. 111–122, Nov. 2004.

[107] J. Renger, S. Grafstrom, L. M. Eng, and R. Hillenbrand, "Resonant light scattering by near-field-induced phonon polaritons," *Physical Review B*, vol. 71, Feb. 2005.

[108] A. Dereux, C. Girard, and J. C. Weeber, "Theoretical principles of near-field optical microscopies and spectroscopies," *Journal of Chemical Physics*, vol. 112, no. 18, pp. 7775–7789, 2000.

[109] M. B. Raschke and C. Lienau, "Apertureless near-field optical microscopy: Tip-sample coupling in elastic light scattering," *Applied Physics Letters*, vol. 83, pp. 5089–5091, Dec. 2003.

[110] A. Cvitkovic, N. Ocelic, and R. Hillenbrand, "Analytical model for quantitative prediction of material contrasts in scattering-type near-field optical microscopy," *Optics Express*, vol. 15, no. 14, pp. 8550–8565, 2007.

[111] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*. Wiley-Interscience, 1998.

[112] N. Ocelic, *Quantitative Near-Field Phonon-Polariton Spectroscopy*. PhD thesis, 2007.

[113] E. D. Palik, *Handbook of Optical Constants of Solids: Index*. Academic Press, 1991.

[114] M. K. Gunde, "Vibrational modes in amorphous silicon dioxide," *Physica B*, vol. 292, pp. 286–295, Nov. 2000.

[115] "RefractiveIndex.INFO - refractive index database."

[116] R. Hillenbrand, T. Taubner, and F. Keilmann, "Phonon-enhanced light-matter interaction at the nanometre scale," *Nature*, vol. 418, no. 6894, pp. 159–162, 2002.

[117] A. Cvitkovic, N. Ocelic, J. Aizpurua, R. Guckenberger, and R. Hillenbrand, "Infrared imaging of single nanoparticles via strong field enhancement in a scanning nanogap," *Physical Review Letters*, vol. 97, no. 6, p. 4, 2006.

[118] A. Cvitkovic, N. Ocelic, and R. Hillenbrand, "Material-specific infrared recognition of single sub-10 nm particles by substrate-enhanced scattering-type near-field microscopy," *Nano letters*, vol. 7, pp. 3177–3181, Oct. 2007.

[119] H. A. Atwater, "The promise of PLASMONICS," *Scientific American*, vol. 17, pp. 56–63, Sept. 2007.

[120] M. L. Brongersma and V. M. Shalaev, "The case for plasmonics," *Science*, vol. 328, pp. 440–441, Apr. 2010.

[121] M. I. Stockman, "Nanoplasmonics: The physics behind the applications," *Physics Today*, vol. 64, pp. 39–44, Feb. 2011.

[122] S. Zhang, K. Bao, N. J. Halas, H. Xu, and P. Nordlander, "Substrate-induced fano resonances of a plasmonic: Nanocube: A route to increased-sensitivity localized surface plasmon resonance sensors revealed," *Nano Letters*, vol. 11, pp. 1657–1663, Apr. 2011.

[123] J. Zhang and L. Zhang, "Nanostructures for surface plasmons," *Advances in Optics and Photonics*, vol. 4, pp. 157–321, June 2012.

[124] J. M. Hoffmann, B. Hauer, and T. Taubner, "Antenna-enhanced infrared near-field nanospectroscopy of a polymer," *Applied Physics Letters*, vol. 101, Nov. 2012.

[125] R. L. Olmon, P. M. Krenz, A. C. Jones, G. D. Boreman, and M. B. Raschke, "Near-field imaging of optical antenna modes in the mid-infrared," *Optics Express*, vol. 16, pp. 20295–20305, Dec. 2008.

[126] A. C. Jones, R. L. Olmon, S. E. Skrabalak, B. J. Wiley, Y. N. Xia, and M. B. Raschke, "Mid-IR plasmonics: near-field imaging of coherent plasmon modes of silver nanowires," *Nano letters*, vol. 9, pp. 2553–2558, July 2009.

[127] M. Schnell, A. Garca-Etxarri, A. J. Huber, K. Crozier, J. Aizpurua, and R. Hillenbrand, "Controlling the near-field oscillations of loaded plasmonic nanoantennas," *Nature Photonics*, vol. 3, pp. 287–291, May 2009.

[128] M. Schnell, A. Garcia-Etxarri, A. J. Huber, K. B. Crozier, A. Borisov, J. Aizpurua, and R. Hillenbrand, "Amplitude- and phase-resolved near-field mapping of infrared antenna modes by transmission-mode scattering-type near-field microscopy," *Journal of Physical Chemistry C*, vol. 114, no. 16, pp. 7341–7345, 2010.

[129] J. Chen, P. Albella, Z. Pirzadeh, P. Alonso-Gonzlez, F. Huth, S. Bonetti, V. Bonanni, J. kerman, J. Nogus, P. Vavassori, A. Dmitriev, J. Aizpurua, and R. Hillenbrand, "Plasmonic nickel nanoantennas," *Small*, vol. 7, no. 16, p. 23412347, 2011.

[130] M. Esslinger, J. Dorfmuller, W. Khunsin, R. Vogelgesang, and K. Kern, "Background-free imaging of plasmonic structures with cross-polarized apertureless scanning near-field optical microscopy," *The Review of scientific instruments*, vol. 83, no. 3, p. 033704, 2012.

[131] M. Schnell, P. Alonso-Gonzalez, L. Arzubiaga, F. Casanova, L. E. Hueso, A. Chuvilin, and R. Hillenbrand, "Nanofocusing of mid-infrared energy with tapered transmission lines," *Nature Photonics*, vol. 5, pp. 283–287, May 2011.

[132] R. Hillenbrand, F. Keilmann, P. Hanarp, D. S. Sutherland, and J. Aizpurua, "Coherent imaging of nanoscale plasmon patterns with a carbon nanotube optical probe," *Applied Physics Letters*, vol. 83, no. 2, pp. 368–370, 2003.

[133] J. Dorfmller, R. Vogelgesang, R. T. Weitz, C. Rockstuhl, C. Etrich, T. Pertsch, F. Lederer, and K. Kern, "Fabry-prot resonances in one-dimensional plasmonic nanostructures," *Nano letters*, vol. 9, pp. 2372–2377, June 2009.

[134] S. Zeng, K.-T. Yong, I. Roy, X.-Q. Dinh, X. Yu, and F. Luan, "A review on functionalized gold nanoparticles for biosensing applications," *Plasmonics*, vol. 6, pp. 491–506, Sept. 2011.

[135] D. Y. Lei, A. I. Fernndez-Domnguez, Y. Sonnefraud, K. Appavoo, R. F. Haglund, J. B. Pendry, and S. A. Maier, "Revealing plasmonic gap modes in particle-on-film systems using dark-field spectroscopy," *ACS Nano*, vol. 6, pp. 1380–1386, Feb. 2012.

[136] P. Alonso-Gonzalez, M. Schnell, P. Sarriugarte, H. Sobhani, C. Wu, N. Arju, A. Khanikaev, F. Golmar, P. Albella, L. Arzubiaga, F. Casanova, L. E. Hueso, P. Nordlander, G. Shvets, and R. Hillenbrand, "Real-space mapping of fano interference in plasmonic metamolecules," *Nano Letters*, vol. 11, pp. 3922–3926, Sept. 2011.

[137] A. A. Mikhailovsky, M. A. Petruska, M. I. Stockman, and V. I. Klimov, "Broadband near-field interference spectroscopy of metal nanoparticles using a femtosecond white-light continuum," *Optics Letters*, vol. 28, pp. 1686–1688, Sept. 2003.

[138] A. A. Mikhailovsky, M. A. Petruska, K. Li, M. I. Stockman, and V. I. Klimov, "Phase-sensitive spectroscopy of surface plasmons in individual metal nanostructures," *Physical Review B*, vol. 69, p. 085401, Feb. 2004.

[139] P. S. Carney, B. Deutsch, A. A. Govyadinov, and R. Hillenbrand, "Phase in nanooptics," *Acs Nano*, vol. 6, pp. 8–12, Jan. 2012.

[140] B. Luk'yanchuk, N. I. Zheludev, S. A. Maier, N. J. Halas, P. Nordlander, H. Giessen, and C. T. Chong, "The fano resonance in plasmonic nanostructures and metamaterials," *Nature Materials*, vol. 9, no. 9, pp. 707–715, 2010.

[141] Z.-J. Yang, Z.-S. Zhang, L.-H. Zhang, Q.-Q. Li, Z.-H. Hao, and Q.-Q. Wang, "Fano resonances in dipole-quadrupole plasmon coupling nanorod dimers," *Optics Letters*, vol. 36, pp. 1542–1544, May 2011.

[142] B. Gallinet and O. J. F. Martin, "Relation between near-field and far-field properties of plasmonic fano resonances," *Optics Express*, vol. 19, pp. 22167–22175, Oct. 2011.

[143] V. Giannini, Y. Francescato, H. Amrania, C. C. Phillips, and S. A. Maier, "Fano resonances in nanoscale plasmonic systems: A parameter-free modeling approach," *Nano Letters*, vol. 11, pp. 2835–2840, July 2011.

[144] Z. Ye, S. Zhang, Y. Wang, Y.-S. Park, T. Zentgraf, G. Bartal, X. Yin, and X. Zhang, "Mapping the near-field dynamics in plasmon-induced transparency," *Physical Review B*, vol. 86, Oct. 2012.

[145] Y. Francescato, V. Giannini, and S. A. Maier, "Plasmonic systems unveiled by fano resonances," *Acs Nano*, vol. 6, pp. 1830–1838, Feb. 2012.

[146] N. Verellen, Y. Sonnefraud, H. Sobhani, F. Hao, V. V. Moshchalkov, P. Van Dorpe, P. Nordlander, and S. A. Maier, "Fano resonances in individual coherent plasmonic nanocavities," *Nano Letters*, vol. 9, pp. 1663–1667, Apr. 2009.

[147] J. A. Fan, C. Wu, K. Bao, J. Bao, R. Bardhan, N. J. Halas, V. N. Manoharan, P. Nordlander, G. Shvets, and F. Capasso, "Self-assembled plasmonic nanoparticle clusters," *Science*, vol. 328, pp. 1135–1138, May 2010.

[148] M. Hentschel, D. Dregely, R. Vogelgesang, H. Giessen, and N. Liu, "Plasmonic oligomers: The role of individual particles in collective behavior," *Acs Nano*, vol. 5, pp. 2042–2050, Mar. 2011.

[149] M. Rahmani, D. Y. Lei, V. Giannini, B. Lukiyanchuk, M. Ranjbar, T. Y. F. Liew, M. Hong, and S. A. Maier, "Subgroup decomposition of plasmonic resonances in hybrid oligomers: Modeling the resonance lineshape," *Nano Letters*, vol. 12, pp. 2101–2106, Apr. 2012.

[150] E. Prodan, C. Radloff, N. J. Halas, and P. Nordlander, "A hybridization model for the plasmon response of complex nanostructures," *Science*, vol. 302, pp. 419–422, Oct. 2003.

[151] Y. Sonnefraud, N. Verellen, H. Sobhani, G. A. E. Vandenbosch, V. V. Moshchalkov, P. Van Dorpe, P. Nordlander, and S. A. Maier, "Experimental realization of subradiant, superradiant, and fano resonances in Ring/Disk plasmonic nanocavities," *Acs Nano*, vol. 4, pp. 1664–1670, Mar. 2010.

[152] A. E. Miroshnichenko, S. Flach, and Y. S. Kivshar, "Fano resonances in nanoscale structures," *Reviews of Modern Physics*, vol. 82, pp. 2257–2298, Aug. 2010.

[153] W. Liu, A. A. Sukhorukov, A. E. Miroshnichenko, C. G. Poulton, Z. Xu, D. N. Neshev, and Y. S. Kivshar, "Complete spectral gap in coupled dielectric waveguides embedded into metal," *Applied Physics Letters*, vol. 97, July 2010.

[154] K. Aydin, V. E. Ferry, R. M. Briggs, and H. A. Atwater, "Broadband polarization-independent resonant light absorption using ultrathin plasmonic super absorbers," *Nature Communications*, vol. 2, Nov. 2011.

[155] V. Giannini, R. Rodriguez-Oliveros, and J. A. Sanchez-Gil, "Surface plasmon resonances of metallic Nanostars/Nanoflowers for surface-enhanced raman scattering," *Plasmonics*, vol. 5, pp. 99–104, Mar. 2010.

[156] B. Hopkins, W. Liu, A. E. Miroshnichenko, and Y. S. Kivshar, "Optically isotropic responses induced by discrete rotational symmetry of nanoparticle clusters," *Nanoscale*, vol. 5, no. 14, pp. 6395–6403, 2013.

[157] S. N. Sheikholeslami, A. Garcia-Etxarri, and J. A. Dionne, "Controlling the interplay of electric and magnetic modes via fano-like plasmon resonances," *Nano Letters*, vol. 11, pp. 3927–3934, Sept. 2011.

[158] Z. H. Kim and S. R. Leone, "Polarization-selective mapping of near-field intensity and phase around gold nanoparticles using apertureless near-field microscopy," *Optics Express*, vol. 16, pp. 1733–1741, Feb. 2008.

[159] M. Schnell, A. Garcia-Etxarri, J. Alkorta, J. Aizpurua, and R. Hillenbrand, "Phase-resolved mapping of the near-field vector and polarization state in nanoscale antenna gaps," *Nano Letters*, vol. 10, pp. 3524–3528, Sept. 2010.

[160] R. L. Olmon, M. Rang, P. M. Krenz, B. A. Lail, L. V. Saraf, G. D. Boreman, and M. B. Raschke, "Determination of electric-field, magnetic-field, and electric-current distributions of infrared optical antennas: A near-field optical vector network analyzer," *Physical Review Letters*, vol. 105, Oct. 2010.

[161] M. Brehm, A. Schliesser, F. Cajko, I. Tsukerman, and F. Keilmann, "Antenna-mediated back-scattering efficiency in infrared near-field microscopy," *Optics Express*, vol. 16, no. 15, pp. 11203–11215, 2008.

[162] D.-S. Kim and Z. H. Kim, "Role of in-plane polarizability of the tip in scattering near-field microscopy of a plasmonic nanoparticle," *Opt. Express*, vol. 20, p. 86898699, Apr. 2012.

[163] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser," *Science*, vol. 22, pp. 553–556, Apr. 1994.

[164] N. Anscombe, "Quantum leap," *Electro Optics*, 2011.

[165] I. M. Craig, M. C. Phillips, M. S. Taubman, E. E. Josberger, and M. B. Raschke, "Infrared scattering scanning near-field optical microscopy using an external cavity quantum cascade

laser for nanoscale chemical imaging and spectroscopy of explosive residues," in *Quantum Sensing and Nanophotonic Devices X* (M. Razeghi, E. Tournie, and G. J. Brown, eds.), vol. 8631, Bellingham: Spie-Int Soc Optical Engineering, 2013.

[166] A. Hugi, R. Maulini, and J. Faist, "External cavity quantum cascade laser," *Semiconductor Science and Technology*, vol. 25, no. 8, 2010.

[167] F. Rana and R. J. Ram, "Current noise and photon noise in quantum cascade lasers," *Physical Review B*, vol. 65, Mar. 2002.

[168] K. Knabe, P. A. Williams, F. R. Giorgetta, C. M. Armacost, S. Crivello, M. B. Radunsky, and N. R. Newbury, "Frequency characterization of a swept- and fixed-wavelength external-cavity quantum cascade laser by use of a frequency comb," *Optics Express*, vol. 20, pp. 12432–12442, May 2012.

[169] C. H. Henry and R. F. Kazarinov, "Quantum noise in photonics," *Reviews of Modern Physics*, vol. 68, pp. 801–853, July 1996.

[170] M. Navarro-Cia and S. A. Maier, "Broad-band near-infrared plasmonic nanoantennas for higher harmonic generation," *ACS nano*, vol. 6, pp. 3537–3544, Apr. 2012.

[171] H. Aouani, H. Sipova, M. Rahmani, M. Navarro-Cia, K. Hegnerova, J. Homola, M. Hong, and S. A. Maier, "Ultrasensitive broadband probing of molecular vibrational modes with multifrequency optical antennas," *Acs Nano*, vol. 7, pp. 669–675, Jan. 2013.

[172] "Crystran website."

[173] M. Brehm, T. Taubner, R. Hillenbrand, and F. Keilmann, "Infrared spectroscopic mapping of single nanoparticles and viruses at nanoscale resolution," *Nano Letters*, vol. 6, pp. 1307–1310, 2006.

[174] M. Paulite, Z. Fakhraai, I. T. S. Li, N. Gunari, A. E. Tanur, and G. C. Walker, "Imaging secondary structure of individual amyloid fibrils of a beta(2)-microglobulin fragment using near-field infrared spectroscopy," *Journal of the American Chemical Society*, vol. 133, pp. 7376–7383, May 2011.

[175] A. Simon and M. C. Durrieu, "Strategies and results of atomic force microscopy in the study of cellular adhesion," *Micron*, vol. 37, no. 1, pp. 1–13, 2006.

[176] O. L. Mooren, E. S. Erickson, N. E. Dickenson, and R. C. Dunn, "Extending near-field scanning optical microscopy for biological studies," *JALA*, 2006.

[177] P. Parot, Y. F. Dufrene, P. Hinterdorfer, C. Le Grimellee, D. Navajas, J. L. Pellequer, and S. Scheuring, "Past, present and future of atomic force microscopy in life sciences and medicine," *Journal of Molecular Recognition*, vol. 20, no. 6, pp. 418–431, 2007.

[178] J. Y. Wang, Z. F. Wan, W. M. Liu, L. Li, L. Ren, X. Q. Wang, P. Sun, L. L. Ren, H. Y. Zhao, Q. Tu, Z. Y. Zhang, N. Song, and L. Zhang, "Atomic force microscope study of tumor cell membranes following treatment with anti-cancer drugs," *Biosensors & Bioelectronics*, vol. 25, no. 4, pp. 721–727, 2009.

[179] T. Ando, T. Uchihashi, and T. Fukuma, "High-speed atomic force microscopy for nano-visualization of dynamic biomolecular processes," *Progress in Surface Science*, vol. 83, pp. 337–437, Nov. 2008.

[180] T. Ando, "High-speed atomic force microscopy coming of age," *Nanotechnology*, vol. 23, Feb. 2012.

[181] D. Ami, P. Mereghetti, and S. Maria, "Multivariate analysis for fourier transform infrared spectra of complex biological systems and processes," in *Multivariate Analysis in Management, Engineering and the Sciences* (L. Freitas, ed.), InTech, Jan. 2013.

[182] X. G. Xu, M. Rang, I. M. Craig, and M. B. Raschke, "Pushing the sample-size limit of infrared vibrational nanospectroscopy: From monolayer toward single molecule sensitivity," *The Journal of Physical Chemistry Letters*, vol. 3, pp. 1836–1841, July 2012.

[183] H. Amrania, G. Antonacci, C.-H. Chan, L. Drummond, W. R. Otto, N. A. Wright, and C. Phillips, "Digistain: a digital staining instrument for histopathology," *Optics Express*, vol. 20, pp. 7290–7299, Mar. 2012.

[184] "Microscopes and microscopy imaging solutions: Leica microsystems."

[185] G. Bellisola and C. Sorio, "Infrared spectroscopy and microscopy in cancer research and diagnosis," *American Journal of Cancer Research*, vol. 2, pp. 1–21, Nov. 2011.

[186] M. Fox, *Optical properties of solids*. Oxford; New York: Oxford University Press, 2010.

[187] S. L. Jacques, "Optical properties of biological tissues: a review (vol 58, pg r37, 2013)," *Physics in Medicine and Biology*, vol. 58, pp. 5007–5008, July 2013.

[188] D. E. Maziak, M. T. Do, F. M. Shamji, S. R. Sundaresan, D. G. Perkins, and P. T. Wong, "Fourier-transform infrared spectroscopic study of characteristic molecular structure in cancer cells of esophagus: An exploratory study," *Cancer Detection and Prevention*, vol. 31, no. 3, pp. 244–253, 2007.

[189] A. Van der Ziel, *Noise in measurements*. New York: Wiley, 1976.

[190] R. H. Kingston, *Detection of Optical and Infrared Radiation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1978.

[191] P. Horowitz and W. Hill, *The Art of Electronics*. Cambridge University Press, 2 edition ed., July 1989.

[192] M. Johnson, *Photodetection and Measurement : Making Effective Optical Measurements for an Acceptable Cost (Professional Engineering)*. McGraw-Hill Professional, 1 edition ed., July 2003.

[193] G. Brida, M. Genovese, and I. R. Berchera, "Experimental realization of sub-shot-noise quantum imaging," *Nature Photonics*, vol. 4, pp. 227–230, Apr. 2010.

[194] R. Paschotta, "Noise in laser technology," *Optik & Photonik*, vol. 4, no. 2, p. 4850, 2009.

[195] J. G. Proakis and M. Salehi, *Fundamentals of Communication Systems*. Prentice Hall, 2 edition ed., June 2013.

[196] H. Yuen and V. Chan, "Noise in homodyne and heterodyne-detection," *Optics Letters*, vol. 8, no. 3, pp. 177–179, 1983.

[197] J. Holmes and B. Rask, "Optimum optical local-oscillator power levels for coherent detection with photodiodes," *Applied Optics*, vol. 34, pp. 927–933, Feb. 1995.

[198] P. C. D. Hobbs, *Building Electro-optical Systems: Making it All Work*. Wiley-Blackwell, 2nd edition edition ed., Sept. 2009.

[199] P. C. D. Hobbs, "REACHING THE SHOT NOISE LIMIT FOR $10," *Optics and Photonics News*, vol. 2, pp. 17–23, Apr. 1991.

[200] P. C. D. Hobbs, "Ultrasensitive laser measurements without tears," *Applied Optics*, vol. 36, pp. 903–920, Feb. 1997.