Arca, Ahmet (2010) The design and optimisation of nanophotonic devices using the Finite Element Method. PhD thesis, University of Nottingham.

The design and optimisation of nanophotonic devices using the Finite Element Method

by Ahmet Arca

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy, June 2010

The University of
Nottingham

Abstract

The aim of this thesis is to develop a technique which can be used in the reliable modelling, design and optimisation of practical suboptical wavelength sized photonic/plasmonic devices, which may involve arbitrary geometries on various scales. The technique involves the application of numerical electromagnetic simulation led by theoretical knowledge and physical insight to determine, design and optimise the operating mechanism of such devices.

The work in this thesis contains a variety of problems/devices which involve arbitrary structures of different scales. This poses difficulties in both the fabrication and the modelling aspects of the design. The problems range in difficulty from those which can be simply and perfectly described via an analytical solution, to those which would be impractical to design using any other technique. The nature of the problems considered, i.e the complicated geometry and the range of scales, necessitates the use of a flexible modelling technique. Finite Element Method (FEM) was found to be a valuable tool in the design and optimisation of the devices throughout this thesis, owing its success mainly to its versatility and flexible meshing abilities which allowed its operation in different length scales in an efficient manner.

Three nanophotonic/plasmonic devices are considered in an effort to demonstrate the implementation and the application of the developed technique. The devices considered in this thesis demonstrate different challenges in the modelling and design while being of considerable interest in their own right as nanostructures for sensing and measurement. These devices are: A self-calibrated plasmon sensor, a plasmon resonator and an ultrahigh frequency optical acoustic surface wave detector. Whilst the first two devices are important as an application of plasmonics, the third device links the mechanical and optical processes together.

Acknowledgements

I would like to thank my supervisors, Dr. M Clark for his constant guidance and support and Prof. M. Somekh for his guidance and suggestions throughout this research. My appreciation goes to everyone in Applied Optics group especially Dr. T. Stratoudaki and Dr. S. Sharples for their willingness to be there with their advice. I would like to thank Jasbinder Chauhan for giving me the necessary training to use clean room facilities and his expertise for my sample preparation.

I couldn't thank enough to my parents and my sister who supported me through thick and thin in every way possible. Above all, I would like to thank Soraya Bent for tolerating me through the stress, hard times and anticipation and for going through them all with me.

Contents

5

List of Figures

8

Chapter 1

Introduction

1.1 Objective of thesis

This thesis aims to describe the development and application of a technique which can be used to model, design and optimise subwavelength sized photonic/plasmonic devices in the optical frequencies. The work mainly focuses on the numerical modelling aspect of the design and optimisation process, yet fabrication and experimental aspects are also covered.

In the developed technique, numerical electromagnetic modelling is led by theoretical knowledge and inductive reasoning to design, analyse and optimise the operating mechanism of nanophotonic devices in light of practical limitations such as fabrication and experimental methods. Numerical modelling is applied to analyse the effects of various design parameters by varying and resolving the problem repeatedly via a script, while an appropriate output metric, such as reflected or transmitted light is recorded. Tables are built relating the parameters to the results, from which optimum parameters are picked according to fabrication limitations.

In order to meet the needs of the industry to build smaller devices, it is necessary to be able to deal with the understanding, design and optimisation of suboptical wavelength sized devices efficiently. The design and optimisation of such devices are difficult since there is usually no analytical solution to many problems in those dimensions and geometries, which means that theoretical analysis is out of reach. These devices are difficult

to fabricate and visualise with high levels of certainty due to their small size, arbitrary geometry and different scales. Therefore, the problem is to efficiently design and optimise devices that cannot be built with certainty, viewed efficiently with appropriate resolution, or be analysed theoretically. The way to design and optimise these devices is to therefore carry out numerical analysis, as an approximation to the exact solutions.

Modelling aspect of the design and optimisation process requires the use of a flexible tool, which can represent the complicated structures efficiently, within practical computational and time limits. Finite Element Method (FEM) is a versatile modelling tool that is found to be advantageous in these applications. A commercially available software, Comsol Multiphysics ®is used to carry out the low-level FEM implementation.

The application of the developed technique is demonstrated with the successful design and optimisation of three photonic/plasmonic devices/problems of increasing difficulty. These devices are (listed from simple to difficult): a self-calibrated surface plasmon sensor, a surface plasmon resonator, and an optical transducer that can generate/detect ultrahigh frequency surface elastic waves. The confidence on the developed technique and the choice of modelling tools is established through verification with experimental results and other software.

The devices range from a relatively simple micrometer sized structure which has analytical solutions and can be easily designed/optimised using many modelling tools, to a novel device with suboptical wavelength dimensions and complicated geometry/mechanism that would be impractical to design or optimise without the application of the developed technique with FEM.

1.2 Thesis outline

This chapter will give an introduction to the main concepts discussed throughout this thesis. It starts with the introduction to main concepts, such as photonics, plasmonics, Nondestructive Testing and Evaluation (NDT&E) and Computational ElectroMagnetics(CEM). Different methods in computational electromagnetics are introduced and some qualitative descriptions are given regarding their mechanisms and performance.

This chapter is followed by the description of methods and theoretical background,

where the description of the developed optimisation technique is followed by a theoretical review in electromagnetics, surface plasmons, surface acoustic waves. The workings of the modelling methods used in this thesis, namely the analytical solutions, Finite-Difference Time-Domain (FDTD) method and Finite Element Method are described. The theoretical background is followed by the features of the FEM models.

In chapter 3, the experimental techniques and the applications of the devices considered in this thesis are described. The different methods of exciting surface plasmons, are followed by the applications of surface plasmons in the current photonics and sensor industry, where the first two devices can be applied. Afterwards, ultrasonics in nondestructive testing is considered, and different methods to optically generate/detect surface acoustic waves (SAWs) are reviewed. Then the issues with the ultrahigh frequency SAWs are discussed as the application of the third device considered.

The fourth chapter will demonstrate the use of FEM in conjunction with the developed technique in the design and optimisation of three interesting nanophotonic devices. In the first part, three increasingly difficult test problems are presented where the FEM results are compared with the analytical solutions and a commercially available software, employing FDTD, in terms of speed and accuracy. In the second section, a novel self-calibrated plasmonic refractive index sensor is analysed. Afterwards, another plasmonic device, namely surface plasmon resonator is considered. The operation of this device is analysed in detail and verified. This device has not been fabricated and experimentally tested before to the author's knowledge. Several design parameters are analysed considering the fabrication limitations and optimised using FEM. The resonators are then fabricated using photolithography, measured using AFM and experimentally tested using an objective-based SP coupling optical system. The experimental results are shown to agree with the model predictions and the discrepancies between the predictions of the model and the practical results are related to the uncertainties induced by the fabrication method and measurement techniques. This proves that it is possible to predict practical device behaviour and the effects of fabrication using the developed technique. In the final section, a novel optical ultrasonic transducer is designed and analysed. This device involves a complicated sinusoidal geometry, which has not been considered in the literature to author's knowledge. Additionally, an approximate rectangular model is built, represent-

ing a simplified version of this situation, which could also be modelled using conventional regular rectangular grids, such as the ones employed in FDTD (see section 2.6.1). Then, it is shown that such simplified models fail when very small dimensions are considered, and it is necessary to model such situations involving the complicated geometry using nonorthogonal grids, such as the ones employed in FEM. This demonstates that complicated geometries/devices involved in applications such as optical ultrasound detection, are impractical to model, design and optimise using other numerical methods such as FDTD. It is shown that ultrahigh frequency surface acoustic waves with wavelengths shorter than the optical wavelengths can be detected with enhanced signal levels and for a wider parameter range at dimensions where the traditional diffraction theory would predict low signals. Besides, this device represents a small perturbation to the acoustic field which makes the device more applicable experimentally. The design and optimisation of this device is done by carefully adjusting the transducer parameters and properties to exploit the subwavelength effects, simultaneously keeping the device practical to fabtricate and operate.

Chapter 5 is the conclusion and summary of the thesis, where the obtained results are reviewed briefly. In this chapter, the whole thesis is summarised and achieved objectives are reviewed. The importance of the topics covered and the application of this work in science and engineering is briefly discussed. This is followed by a description of the work that will follow this thesis, and the outlook of this research.

1.3 Introduction to photonics and plasmonics

Optics is a branch of science that was developed before the knowledge of light could be quantised into photons. Therefore it deals with light and its manipulation at a very wide range of wavelengths. Classical optics theory does not necessarily depend on the fact that light is quantised into photons. Photonics, as the name suggests also relate to photons. It deals with aspects of light such as the generation, manipulation, and detection of light for different purposes. It involves the wave and the particle nature of light, and is usually referred to as the technical, application based aspect of light in the areas such as telecommunications, spectroscopy and nanocircuits.

In optics and photonics, the size of the devices is generally associated with the "diffraction limit". This limit states that light cannot be squeezed into a space with dimensions below its wavelength, and it determines the lower limit of the possible device size and the resolution of the viewed objects.

1.3.1 Surface plasmons and plasmonics

Surface plasmons (theoretical background given in section 2.3), also known as surface plasmon polaritons, are surface electromagnetic waves that propagate in a direction parallel to a metal/dielectric (or metal/vacuum) interface. Surface plasmons were first reported in 1957 by R.H. Ritchie, after which they immediately received a great interest from the scientific community due to their interesting characteristics and possible applications.

Surface plasmons are candidates to decrease optics into two dimensions, below the diffraction limit [1, 2, 3](see section 3.2.5 for more details). The advances in micro-fabrication and nano-fabrication techniques, together with the advances in visualisation technology, such as SNOM (Scanning Nearfield Optical Microscope), led to further developments and research areas in photonics, namely nano-photonics and plasmonics, which involve devices on the orders of nanometers.

Plasmonics is a thought by some as the future of electronics [4]. Transfer of light can work at very high frequencies compared to electronics, therefore very large bandwidth is obtainable. However, optics require bulky "wires", due to diffraction limit, which limits their use in small circuitry. On the other hand, electronics can be squeezed in very small wire dimensions but the wires get very lossy at very high frequencies. Therefore, electronics is limited in terms of bandwidth. Electronic circuits suffer from extended delays in the interconnects, when transferring data from one place to another in a microprocessor, which put a limitation on the speed of the circuits. Optical interconnects can carry digital data with much higher capacity but they are bulky, hence cannot be miniaturised to nanometre dimensions [4].

Plasmonics offer the bandwidth of optics at dimensions smaller than the diffraction limit [5, 6, 7]. Therefore, it can shrink optics, a very large bandwidth, onto very small wires, thus combining and utilising the strongest points of optical and electronic data transfer.

This technology needs further improvement, since surface plasmons have short propagation lengths, due to their dissipative nature. There is research toward negative refractive index materials and gain materials [8, 9], which are thought to potentially have the ability to support plasmon propagation over longer distances. There is also research to find a plasmonic device analogous to transistors [10].

1.4 Introduction to laser ultrasonics and NDT&E

Ultrasonics has been a branch of study since the times of Lord Rayleigh [11]. However it has only been since World War I that ultrasonics have started to make significant impact in the practical world. The beginning point of substantial research in ultrasound dates back to World War I, when underwater transducers were developed [12]. The pioneering work in ultrasonics belongs to Langevin, who is considered as the father of ultrasonics [13]. The immediate application was in submarine technology of the 1910s [12]. Sokolov [12] was the first to suggest the use of ultrasound to detect discontinuities in metals. In 1930s, he developed the transducers to generate and detect ultrasound.

1.4.1 Ultrasound

Ultrasound is the part of the audio spectrum which corresponds to frequencies above the human hearing range, generally taken at higher frequencies than 20kHz. Ultrasonics, have similar properties to electromagnetics, and they have followed similar paths in their development [12]. Most phenomena in acoustics have an equivalent in optics, such as reflection, dispersion etc. Acoustics have additional wavemodes, for instance, a longitudinal part in bulk media, which leads to density changes during propagation. All of the acoustic phenomena occur the same way throughout the acoustic spectrum, apart from the extreme limits when abnormal behaviour is observed such as nonlinear effects [12]. Acoustic Spectrum can be demonstrated as in Figure 1.1. The related work in this thesis is based around frequencies such as 10GHz.

The power of ultrasound comes from its ability to visualise the interior of optically opaque objects. This property makes ultrasound find its place in medicine, and medical analysis [14]. Also, high power ultrasound has been used in microbiology as a sterilisation
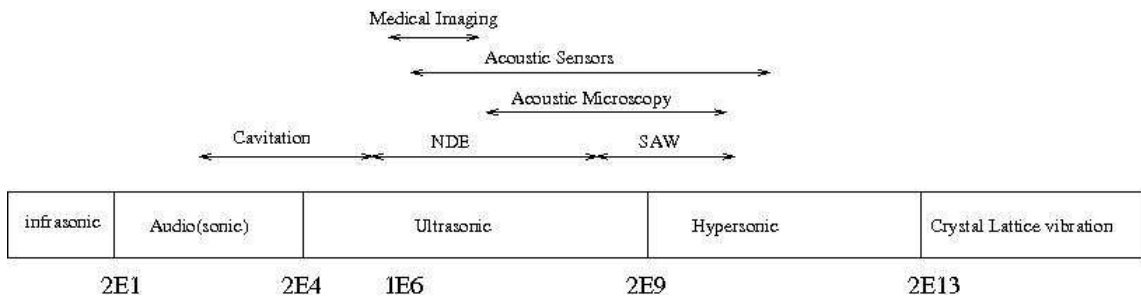
Figure 1.1: The acoustic spectrum with common applications of sound

technique [15, 16]. High power ultrasonic cleaner is used in laboratories to shake contamination such as debris and small particles off the samples. Sonochemistry is another branch where high power ultrasound is used to facilitate chemical reactions [16, 14]. Ultrasound has been used in the manufacture of microwave electronic devices, non-invasive plant measurements. The application of ultrasound focused on in this thesis is non-destructive evaluation (NDE).

1.4.2 NDT&E

NDT&E (also referred to as NDT or NDE) is the examination, test or evaluation of any type of object without changing it. NDE is performed in order to assess the presence of certain features or faults on the object of interest, which may be of crucial importance to the usefulness or quality of that object. Reliable NDE techniques that have been developed include thermography, radiography, visual inspection, magnetic techniques, in applications such as aircrafts, bridges, oil rigs and nuclear reactors [17].

The power of ultrasonics is revealed in NDE, since interior defects can be imaged in optically opaque structures. It is of interest to take measurements, such as detection and characterisation of defects and materials in the manufacturing line. This is the key application where the shortcomings of many present devices are realised, due to the adverse conditions involved [18].

Interior defects can easily be detected by bulk ultrasonic waves, however such waves cannot be utilised if the defect is small and close to or at the surface, since this would require very high frequencies. In order to detect such defects, a mode that travels on or close to the surface is more useful. Surface acoustic waves (SAW), or Rayleigh waves

(details given in section 2.4), can also be used to detect surface defects in NDE applications.

1.4.2.1   Contact vs noncontact methods

The systems which generate and detect ultrasound can be classified under two groups depending on whether or not they are in contact with the sample under investigation. Importance of such a classification comes into play in industrial applications. When NDE is to be carried out in industrial applications, there may be situations where contact techniques may not be practical [19]. Contact transducers, are usually bigger, need wires etc, and thus could be awkward to use in hard to reach places in industrial applications. The main drawback of the contact transducers, is the requirement of a couplant medium, which couples the device onto the sample [20]. Such couplant often cannot be used for a very long time, or at elevated temperatures, and may induce sensitivity variations due to uneven thickness. Besides, there may be inaccessible, hard to reach areas, in an industrial application, which could make the use of contact transducers very awkward.

Ultrasonic surface waves can be commonly generated or detected using three general methods. These methods include electromagnetic, piezoelectric and optical methods. When generated electromagnetically, via eddy currents and Lorentz forces, the transducers are called Electromagnetic Acoustic Transducers (EMATs)[21, 22]. When voltage is applied to a "piezoelectric material" to cause mechanical vibrations, the transducer is called piezoelectric [23]. These include the SAW devices, such as Interdigital transducers (IDT) [23]. The main idea in IDTs is to lithographically pattern electrodes on to piezoelectric crystals [24], to create periodic disturbances in the piezoelectric material, when alternating voltage is applied to the electrodes.

1.4.2.2   Introduction to laser ultrasonics

In adverse conditions, such as elevated temperatures, hazardous environments, it becomes important to be able to generate and detect surface waves remotely. Laser ultrasonics is a branch where the ultrasound is generated and/or detected remotely via optical means. Laser ultrasonics is considered to be a potential solution to taking measurements remotely on the manufacturing line [14], when possibly hazardous environments are considered. (More information is given in section 3.3.1).

Optical transducers generally suffer from reduced sensitivity [18, 14] and generally are very complex and expensive to build and use (Optical ultrasound generation techniques are further reviewed in section 3.3).

1.5 Computational electromagnetics and modelling

It is necessary to be able to predict the device behaviour at very small dimensions, especially at extreme fabrication tolerance limits. However, Maxwell's equations (introduced later in section 2.2) do not have closed form solutions to many real world problems. The only way that can get close to a solution for complex situations is numerical analysis, which approximate the exact solutions to the governing equations.

This issue has only been partially solved until now, in oversimplified cases using computational electromagnetics (CEM). CEM is an important and effective way to understand, design and optimise complicated electromagnetic phenomena underlying certain devices. This makes numerical modelling, hence CEM, an important field in the design of nanophotonic devices.

1.5.1 The power of CEM

Computational Electromagnetics is a branch of theoretical analysis which has become much more popular, due to its versatility, ease of use, and availability and the increase in computing power. There are a number of major methods used in CEM, however, there is no one method that will solve all kinds of problems efficiently. Engineers usually look for a method according to a well specified problem.

The power of CEM, i.e the situations that can be accurately and efficiently modelled, depend not only on the situation but also ultimately on the computing resources available. A lot of these methods are known to be time consuming, and/or heavy on computing resources. With the advances in computing technology, the range of problems that can be addressed is growing rapidly and this is becoming a smaller concern [25], hence CEM is becoming more available and widespread. Still, care must be taken to reduce the simulation times and increase the accuracy.

1.5.2 Commercial software in CEM

Some CEM methods can be more cumbersome to implement than others. Besides, building a general environment for a generic problem to be solved is a fairly large project. Time that needs to be invested in the understanding and implementation of some CEM methods is usually unattractive to the research community [26]. Easy to implement methods are usually more available in the literature. Even though, some numerical methods are cumbersome to understand, implement and optimise for speed and accuracy, today commercially available software are present which make these methods more accessible than before.

Commercial software reduce the need for low-level programming of the numerical method implementation and optimisation. This makes numerical modelling and computational electromagnetics more readily available and accessible. The ease of use, in addition to the increased computational power nowadays available, makes computational electromagnetics become more powerful and popular.

Commercial packages include, CST Microwave Studio ©, REMCON's XFDTD, Ansoft's HFSS, Comsol Multiphysics, Ansys, EM-explorer. Although there are some problems associated with commercial packages, many researchers are choosing them. This is due to the time and effort saved by not dealing with coding, debugging and testing every aspect of the self-designed software, which falls under other branches of sciences rather than applied optics or electromagnetics in general. In this thesis, the low-level programming and the implementation related topics in the application of FEM is not mentioned, since a commercial software is employed for such issues.

1.6 Common techniques in CEM and theoretical analysis

There are two conceptual means to analyse electromagnetic problems, experimental and theoretical. In this thesis devices are designed using the latter, and the former remains as a tool of verification for certain cases. Numerical analysis is referred to as modelling throughout this thesis. In this section, some common modelling methods are introduced.

CEM problems typically involve the computation of the E (Electric), and H (Magnetic) fields across the domain of the problem. Power flow direction (Poynting vector), normal

modes of a waveguide, dispersion of wave due to media, and scattering are quantities of interest, that can be computed from the knowledge of the E and H fields. CEM methods conventionally make assumptions to simplify real world structures, and extensively make use of symmetry, so as to solve for reduced dimensions from 3D to 2D and even 1D.

Different methods in CEM may be favourable for different problems and applications. Transient response of a structure may be modelled by simulating electromagnetic pulses in time domain, by methods such as FDTD. Treating curved geometrical objects is done more accurately by using finite elements FEM, or non-orthogonal grids [27].

In the analysis of nanophotonic devices, numerical models are used for complicated structures and analytical solutions are used for simpler problems or for the verification of the numerical methods in simpler situations. Throughout this thesis a commercial software is mainly used, namely Comsol Multiphysics, for the numerical models. Also, a second commercial software called EM Explorer, is utilised to verify certain models/results produced using Comsol.

The main idea behind the numerical methods is to discretise an unknown quantity [25]. The process of discretisation is called meshing or gridding. Meshing is done by dividing the entire model describing the problem into smaller areas. The shape, distribution and the use of these sub-areas are dependent on the technique under consideration. Within each element, a functional dependence is assumed on the spatial variation of the unknown [25], which is known as basis functions.

Generally, the accuracy of the problem is related to the amount of discretisation. Therefore, the smaller the grid, or mesh elements, the more accurate the solution, which tends to the exact solution for infinitely small mesh elements. However, the computational resources necessary to solve the problem will increase with the mesh size. Therefore, a trade off is sometimes to be made between accuracy and simulation time/memory requirements.

1.6.1 Analytical models

Analytical methods provide very accurate and useful design information, but they are very limited in availability since they depend on the existence of Green's functions and closed-form solutions, which are very rare in practical electromagnetics problems. Also,

Description of Analytical Model Geometry
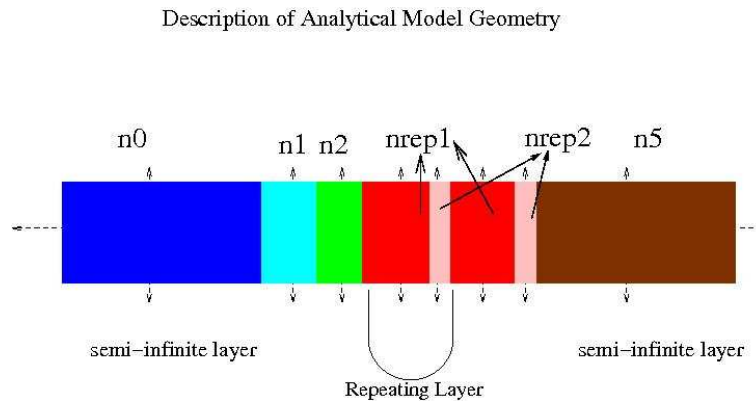
Figure 1.2: Demonstration of the analytical model geometry. The excitation is assumed to come from the left through n0 and travel through till n5. The reflection coefficient in n0 and the transmission coefficient in n5 are obtained through Fresnel equations.

they could be derived only for regular, simpler geometries. Variational methods and perturbation methods can be used to increase the applicability of some analytical methods, at the expense of accuracy [28]. Additionally, model based approaches can be made by simplifying more complicated problems with certain assumptions, however these assumptions make the solution very problem-specific, and reduce versatility [28], which is very important in designing nanophotonic devices.

Common analytical methods include method of separation of variables [29, 30, 31], point matching method [32], perturbation method [33], orthogonal function expansion method [34], method of images [35], Green's function [36, 37], conformal mapping [38] and variational methods [28]. Detailed information about these methods is outside the scope of this report, however such details can be found in [28, 39, 40].

The analytical model/solution used in this thesis [1] is not as complicated as the methods described above. It is a one dimensional model, which uses the Fresnel equations to calculate the total reflected and transmitted fields of a system of planar stratified media, sandwiched between two semi-infinite layers. It calculates the reflection into the first semi-infinite layer and the transmission into the final semi-infinite layer given the wavelength of the incoming plane wave, refractive indices and the thicknesses of different layers, and the angle of incidence. This models the geometry shown in Figure 1.2.

---

[1]is written by Mike Somekh (see appendix A.1)

The problem with analytical models is that a very small change in the problem, can make a significant change in the way every aspect is programmed and invalidate the whole model and solution, making it necessary to build another one from scratch. The code for this analytical program can be found in Appendix A.1, the necessary equations are given in section 2.5. Since this specific analytic code is built on simple Fresnel coefficients, the model could be easily invalidated if it was necessary to look at a particle embedded in one of the layers. Another case specific analytic code would have to be written from scratch to model the new situation, using one of the techniques mentioned above.

1.6.1.1 Rigorous coupled-wave analysis(RCWA)

RCWA was popularised by Moharam and Gaylord in 1980s [41]. This method was developed to overcome the shortcomings of methods such as "coupled-wave approach" and "modal approach", in analysing diffraction of electromagnetic waves by planar gratings [42]. It has been used in the literature to analyse different types of gratings for different applications [43, 44].

When using RCWA the Maxwell's equations in the diffraction grating region are solved by expanding the periodic dielectric in the diffraction grating region in a Fourier series, expanding the EM fields in the diffraction grating region in a set of Floquet harmonics whose amplitudes are functions of the longitudinal coordinate(normal to the grating surface), substituting these expansions in Maxwell's equations and organising the resulting equations into state variable form where eigen solutions to the state variable system can be found [45]. In RCWA the solution accuracy is obtained by including however many Floquet harmonics is necessary for convergence of solution.

In this method, generally the grating structure is divided into many thin gratings, and the profile of the grating is approximated by a rectangle. Therefore, the actual grating profile is approximated by staircase approximations [46]. Due to the slicing, it becomes difficult to represent shallow grating profile angles. The RCWA is known to converge slowly for deep gratings with TM polarisation [45, 41].

1.6.2 Method of Moments(MoM)

Method of moments, often used as a synonym to the boundary element method (BEM) in electromagnetics, is a numerical technique that can solve a system of linear partial differential equations, using integral form of Maxwell's equations [47]. This method is commonly used to solve scattering problems, where the scattering/radiating body is conventionally represented as equivalent surface currents [25]. Equations are set as boundary integrals, which are discretised into a linear system of equations by projecting the whole system onto a set of basis and testing functions [48, 49], relating the elements. The technique is also referred to as the *Weighted-residual method.* The origins of this method extends to the early works of Rayleigh, Ritz and Galerkin. It is considered as a general prescription for discretisation of the integral form of Maxwell's equations. [49]

The solution to the integral equation is obtained by discretising the integral surface and representing the continuous integral around the whole surface as a system of linear weighted sums of integrals over small discretised surfaces. The whole system is then represented as a matrix equation and the integrals over the small elements are approximated in some cases without the need for actual integration. The resulting system of linear equations are dealt with using matrix equations. All discretisation processes can be thought of approximately similar to this method. The method of moments is frequently used especially in electromagnetics scattering problems. It has also been used in a hybrid Finite Element-Boundary Integral method [50, 51].

Method of moments is traditionally applied in the frequency domain and yield matrices with complex elements. Time domain integral equations have been used on occasions, but proved difficult so far [25]. The method may be more efficient than other methods, in terms of computational resources, for problems where there is a small surface/volume ratio. It is ideal for radiation type of problems, since it automatically incorporates the far field. The method treats high/perfectly conducting surfaces efficiently, since no "air" region has to be meshed in these situations [25].

Radiation and scattering by wires, rods, 3D scattering from arbitrary shaped metal or dielectric scatterers [2] [52] have been successfully modelled. However, for surfaces with

---

[2]less efficiently than other methods

cavity like behaviour, where multiple ray behaviour is to be modelled, this method became inappropriate. Besides, MoM is applicable to problems for which Green's functions can be calculated [48]. This places considerable restrictions on the range and generality of problems to which boundary elements can usefully be applied. Nonlinearities can be included in the formulation, although they will generally introduce volume integrals hence volume discretisation. This removes one of the most often cited advantages of BEM [48].

For many applications, including resonant features, the method of moments, however applicable, is significantly less efficient than volume-discretisation methods (Finite element method, Finite difference method) [25]. This is due to the fact that, however elegant and straight-forward, the formulation of the method of moments, yields system of linear equations whose matrices tend to be fully populated with complex elements [53, 52]. Therefore, the storage and computations requirements can be significant for some problems. Compression techniques such as multipole expansions can be used to get around these problems, though at the cost of added complexity [52].

1.6.3   Finite-Difference Time-Domain (FDTD) method

Finite Difference Time Domain is a common technique in modelling electromagnetics problems. It is considered to be easy to understand and implement [25]. It is a time-domain method, so depending on the excitation type used, it could cover a wide range of frequencies in a single run, hence it is usually the method of choice for wideband systems [54, 25] . The technique was introduced first by Kane Yee [55], and the acronym was given to the technique by Allen Taflove in 1980 [56]. FDTD does not need the existence of Green's function like Method of Moments. It directly approximates the differential operators in Maxwell's equations on a grid staggered in time and space. FDTD is an explicit finite difference approach, i.e. no matrix equation is setup, stored and solved. The field values at the next time step are given entirely in terms of the field at the current and the previous time steps [25].

The details of FDTD method and the basic formulation in 2D are shown in section 2.6.1. FDTD discretises the domain in which the solution is sought into unit cells called "Yee cells" named after its inventor [55, 54]. In such Yee cells, Electric Fields are conventionally represented by the edges of a cube, where the magnetic field is represented by the

faces of the cubic cell. Given the offset (in space) of the magnetic fields from the electric fields, the values of the field with respect to time are also offset. Time is divided into small steps, which are equivalent to the amount of time necessary for the fields to travel from one Yee cell to the next or less. FDTD method solves Maxwell's equations using the relationship between the partial time and space derivatives. Yee's algorithm solves both **E** and **H** in time and space using Maxwell's curl equations [54].

The biggest advantage of FDTD is its ease of use and quick application . There are a lot of open source scripts that can be used as building blocks towards a big and complicated model [57, 58]. Many modelling papers have been published in design and optimisation of photonic devices using FDTD [59, 60, 61].

Disadvantages of the FDTD method which apply to this thesis are mostly related to meshing. Rectangular grids conventionally employed in FDTD algorithms are inefficient in resolving curved surfaces [62, 63], or small geometries in the basic formulation. Several multi level adaptive meshing algorithms and multi resolution algorithms have been considered in the literature [64], but these often imply added complexity to the simple hence attractive nature of FDTD. It is worth mentioning that there has been attempts to use FDTD in the frequency domain [65], but these methods were found to be problematic due to ill-conditioning of matrices and slow convergence [25, 66].

1.6.4 Transmission Line Modelling (TLM) method

Transmission Line modelling is a time domain technique that can be used to solve electromagnetic problems. It is based on an equivalent transmission line representation of electromagnetic waves. A discrete model is formed by conceptually filling space with a network of transmission-lines in such a way that the voltage and current give information on the electric and magnetic fields. The point at which the transmission-lines intersect is referred to as a node. At each timestep, voltage pulses are incident upon the node from each of the transmission-lines. These pulses are then scattered to produce a new set of pulses which become incident on adjacent nodes at the next time step. The relationship between the incident pulses and the scattered pulses is determined by the scattering matrix, which is set to be consistent with Maxwell's equations. Additional elements, such as transmission-line stubs, can be added to the node so that different material properties can

be represented. TLM can be used to analyse a wide range of electromagnetic problems [67].

An initial excitation is defined and the impulses are propagated throughout the mesh using the scattering theory of simple TEM transmission lines. The method makes all the spatial field components at the centre of each node in the mesh. The TLM analysis is typically carried out on a cartesian grid or mesh, but unstructured tetrahedral mesh has been considered [68]. Each TLM node contains two series or shunt connected ideal transmission lines. One of the advantages of TLM is that the core algorithm is very straightforward. Each timestep can be divided into two processes, scatter and connect.

During the "scatter" stage, voltage pulses incident on a node are scattered to produce a new set of outgoing voltage pulses. During the "connect" stage, voltage pulses are transferred to neighbouring nodes. Therefore, a pulse is input in the beginning and all parts of the model is radiated with the repeated application of the scattering matrix forming a solution [69]. From the theoretical point of view, TLM formulation of the symmetric condensed node model is exactly equivalent to FDTD formulation [70, 71].

Some TLM applications include Electromagnetic Compatibility (EMC)[72], interaction of lightning and EMP with structures, electrical and thermal modelling of lasers and semiconductor devices [73].

1.6.5 Multi-Resolution Time-Domain (MRTD) method

The multiresolution time-domain (MRTD) method is an efficient tool for the analysis of time-domain electromagnetic fields, which in some cases can save computational resources while maintaining accuracy, when compared to traditional FDTD method [74]. This technique has been introduced as an answer to the inability of FDTD to model very small but important geometries with the available computing resources [75, 76].

The MRTD update equations are typically derived via application of the Method of Moments, which utilises field expansions as summations of wavelets and scaling functions [77]. MRTD possesses higher order accuracy and multi-resolution capability [78] when compared with FDTD. The use of scaling and wavelet functions and the application of the multiresolution analysis, in conjunction with the Method of Moments based discretisation of Maxwell's equations, forms the cornerstone of the MRTD technique [77]. Multiresolution

is achieved by using denser resolution at the field locations where the field varies faster [74]. MRTD technique has higher order field expansion structure than both the FDTD and MoM techniques. If pulse functions are used as the basis functions in the MRTD method, the use of MoM and MRTD yields the FDTD update equations. FDTD method can be seen as a basic subgroup of the MRTD technique [77].

MRTD has been applied to a big variety of microwave and antenna problems, photonic devices and nonlinear optics [79, 80, 81]. Also, it has been used in microwave cavities and circuits [79], scattering by general targets [78]. The multiresolution time domain (MRTD) method shows an excellent capability to approximate the exact solution, even for the near Nyquist sampling limit [82].

One of the drawbacks of this method has to do with the nonlocalised nature of the basis functions. Such basis functions cannot resolve localised boundary conditions, hence it is not possible to have localised boundaries and descriptions of material properties [79]. As a workaround, Perfect Electric Conductor Boundary Condition and Perfect Magnetic Conductor Boundary condition can be modelled using the image method. Perfectly Matched Layers have been proposed to deal with the truncation of the simulation domain and they may be applied with the price of added complexity [77].

1.6.6 Finite Element Method(FEM)

The finite element method (FEM) is a numerical technique which approximates the solutions of differential equations [83]. In electromagnetics, FEM generally approximates the solution to the Maxwell's equations in the frequency domain, hence is usually used with time-harmonic situations. It is also capable of time domain simulations.

FEM can be derived using two methods. First, is the variational method, which finds a variational functional whose minimum/maximum/stationary point corresponds to the solution of the PDE subject to certain boundary conditions (a brief formulation is given in section 2.7.1). Second method is sometimes called the "weak formulation" in the literature. It works by introducing a weighted residual error to one of the differentials in the PDE form of Maxwell's equations [25] and equating the sum of the error to zero. If the weighting functions are Dirac delta functions, the resulting procedure is similar to finite difference method . If the weighting function is the basis functions, then the method is called the

Galerkin's method. [50]

The physical interpretation of FEM is closely related to the discovery and extensive use of the method in the field of structural mechanics in early 1940s [84, 85]. The basic concept in the physical interpretation is the breakdown of a complex mechanical system into simpler, disjoint components called finite elements [86]. The physical response of an element is characterised in terms of a finite number of "degrees of freedom". These degrees of freedom are represented as the values of the unknown functions as a set of node points. The element response is defined by algebraic equations constructed from physical or energy specifications. The response of the original system is approximated by that of the discrete model constructed by connecting the elements. Therefore, a physical look would indicate that, electromagnetic fields would be solved for each element, and connected to form a whole physical model, by applying the correct boundary conditions and physical approximations/assumptions.

The main advantage of FEM, especially over grid based methods such as FDTD, is in discretisation/meshing. Mesh elements in FEM can be of triangular shape and irregular in arrangement[50]. Mesh will give details about where every element is, to which material in the simulation domain it associates with, and approximates the geometrical shape of the domain. How well the geometrical shapes are resolved in the domain will contribute to the stability and accuracy of the solution, in addition to the resources used. Mesh characteristics of FEM modelling is known to be more efficient than conventional FDTD mesh [28]. This is because, the non-regular triangular FEM mesh can resolve different, arbitrary, small geometries better than the regular rectangular mesh conventionally employed by FDTD.

1.7 Summary

This chapter serves as a brief introduction to the motivation and objective of the thesis as well as the main concepts. A statement of the motivation and objective of this work is followed by some background in optics photonics, plasmonics, ultrasonics and computational electromagnetics.

The motivation of this work is to develop a technique to understand, efficiently de-

sign and optimise the operating mechanism of subwavelength sized photonic/plasmonic devices which may be difficult to model, fabricate, and visualise, due to their small size, arbitrary shapes and the range of scales. The work in this thesis is novel in the devices that the developed technique is applied to. Inductive reasoning combined with theoretical knowledge and experience in fabrication and experimental methods was used to lead numerical simulations in order to efficiently understand, design and optimise three devices, each with unique significance in their own application area. From the modelling, design, and fabrication point of view, these can be viewed as problems of increasing difficulty, ranging from a problem that can be exactly and quickly designed and optimised using analytical solutions, to a problem that would be impractical to design and optimise using other techniques and CEM method.

Chapter 2

Theoretical Background and Methods

2.1 Introduction

The main objective of this thesis is to develop a technique to efficiently design and optimise nanophotonic/plasmonic devices of arbitrary geometries and scales. This chapter introduces the theoretical background for the important subjects that are required in the design process.

The conceptual explanation of the technique used to design the devices is briefly introduced before the necessary theoretical background. Steps of this technique are demonstrated in figure 2.1. The first few steps are towards defining the problem to be solved and picking the appropriate candidates for modelling. This chapter aims to give sufficient background in the concepts involved in these steps.

The first step is to determine the expected operating principle of the device under consideration, i.e the underlying mechanism and the basic equations. Therefore, this chapter starts with sections giving a brief background on electromagnetic theory, surface plasmons, and Rayleigh waves, which are essential to the operation of the devices considered in this thesis.

This is followed by a brief overview of the underlying mechanism of the modelling methods and the model features. Theoretical background on the mechanism of the an-

alytical solutions, FDTD and FEM are given. Model features include concepts such as truncation issues (open air or periodic problems), source types (how to describe plane waves in the software), data extraction etc.

2.1.1 Overview of the technique developed

Suboptical wavelength sized nanophotonic/plasmonic devices are known to be difficult to model, due to the evanescent nature of the fields and the arbitrary nature of the geometries involved in realistic situations.

In order to efficiently understand in detail, design and optimise nanophotonic/plasmonic devices of arbitrary geometries, a technique is developed. A guideline of the technique that can be applied in the design process is demonstrated in figure 2.1.

In the design process the first step is to determine the operating principle of the device under consideration. Although this may not be known exactly, the inputs and the outputs and the general underlying mechanism of the device under consideration can be determined. This step is followed by the determination of the model approach and the problem definition in terms of modelling. This relates to predicting initial modelling issues that may be important and the basic features of the model to be built. Model features include, truncation issues(open air or periodic problems), linearity of problem (linear, nonlinear, coupled), the type of analysis (transient analysis, time harmonic analysis), the type of mode to be considered (Transverse Electric , Transverse Magnetic, etc).

These features aid in the determination of and appropriate modelling method. Some common modelling methods were discussed in section 1.6. Some mathematical theory behind the operation of the common modelling tools considered in this thesis (FEM and FDTD) are given in sections 2.6.1 and 2.7.1. The modelling tool selection for the devices considered in this thesis is described in section 4.1.1 and is mostly based on speed and versatility. After the modelling tool selection, a base model is built to test the basic operation principle of the device. The operating mechanism of the device is then analysed in detail to gain insight into the key features of the device, and how they can be improved.

After verifying the operation principle of the device, it is necessary to determine which parameters may play an important role in the operation of the device and which of those parameters can be optimised in light of the limitations of fabrication and experimental
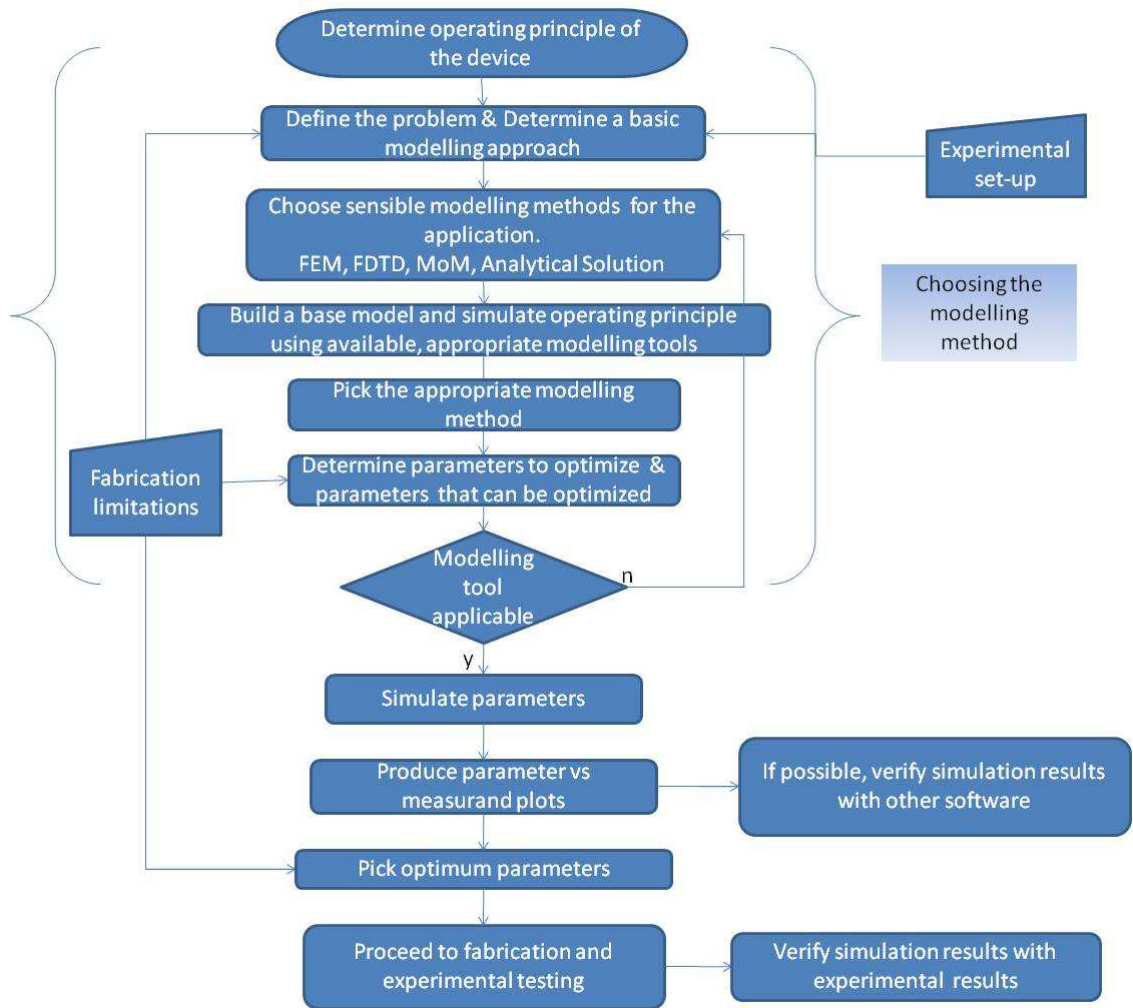
Figure 2.1: The flowchart describing the guideline of the actions taken in the modelling, design and optimisation of suboptical wavelength sized devices

methods available. Fabrication limitations dictate the practical aspect of the process, whether or not a device can be manufactured in practice. For example a 50nm wide and $1\mu m$ thick golden rod may not be possible to fabricate with the current fabrication techniques by thermal evaporation. After the determination of the parameters that can be optimised, the applicability of the modelling technique is reconsidered. The parameters to optimise may exceed the limits of certain modelling techniques. For example very thin metal films, highly slanted grating profiles, or complicated geometries involved in a relatively big simulation domain, may rule out the modelling techniques which use rectangular grids.

If the modelling tool of choice is appropriate, then the effect of these parameters on the device performance is investigated. This is done by resolving the model at a range of parameter values and recording an output metric. This output metric could be the reflected or transmitted light from the structure, as the case in this thesis. Associated tables are built for parameters, or combination of parameters to see which combination could yield better performance. At this stage it is a good idea to verify the modelling results with that of another software if possible, at least for a few parameter values.

After the tables are built and appropriate plots are produced, the optimum parameters are picked with fabrication limitations in mind. Generally, when pushing the limits of a fabrication method, some tolerance must be allowed. Therefore, the optimum parameters are picked to yield an operating region wide enough to accommodate these tolerances. After sensible design parameters are picked, the device is fabricated and its operation is experimentally verified.

2.2 Theoretical background in electromagnetics

2.2.1 Maxwell's equations in the integral form

Light can be viewed from two aspects: it is an Electromagnetic Wave and it has particle-like properties. Electromagnetic Wave expressions can be expressed as a solution to a set of equations called Maxwell's Equations [87].

Maxwell's Equations can be written as a set of integral equations [1]. Maxwell's Equa-

---

[1]These equations can also be written as differential form

tions can be written in integral form as follows.

$$\oiint_{\partial V} \mathbf{E}.d\mathbf{A} \quad = \quad \frac{Q_{enc}}{\epsilon} \tag{2.1}$$

$$\oiint_{\partial V} \mathbf{B}.d\mathbf{A} \quad = \quad 0 \tag{2.2}$$

$$\oint_{\partial S} \mathbf{E}.d\mathbf{l} \quad = \quad -\frac{\partial \Phi_{B,S}}{\partial t} \tag{2.3}$$

$$\oint_{\partial S} \mathbf{B}.d\mathbf{l} \quad = \quad \mu I_s + \mu \epsilon \frac{\partial \Phi_{E,S}}{\partial t} \tag{2.4}$$

Where, $\mathbf{E}$: Electric Field $[V/m]$

ϵ : Electric Permittivity$[F/m]$

μ : Magnetic Permeability $[N/A^2]$

ρ : Volume Charge Density $[C/m^3]$

$\mathbf{J}$: Electric Current Density $[A/m^2]$

Q_{enc} : Net charge $[C]$

$\mathbf{D}$: Electric Flux Density $[C/m^2]$

$\mathbf{B}$: Magnetic field (magnetic flux Density) $[T]$

$\mathbf{H}$: Magnetic Field Intensity $[A/m]$

∂V : through any closed surface S

∂S : along the boundary of a closed surface S

$d\mathbf{A}$: differential vector element of surface area A

with infinitesimally small magnitude

and direction normal to surface S

$d\mathbf{l}$: differential vector element of path length

tangential to the path/curve

The magnetic and the electric flux through a surface S can be expressed as follows.

$$\Phi_{B,S} \quad = \quad \iint_s \mathbf{B}.d\mathbf{A} \tag{2.5}$$

$$\Phi_{E,S} \quad = \quad \iint_s \mathbf{E}.d\mathbf{A} \tag{2.6}$$

2.2.1.1 Maxwell's equations in the time-domain(differential form)

Equations 2.3, 2.1, 2.2 ,2.4 are restatements of Faraday's Induction Law, Gauss's Law-Electric, Gauss's Law-Magnetic and Ampere's Circuital Law [87]. These equations can be expressed as vector expressions [87, 88]. For sourceless media, J and ρ are equal to 0. These equations describe all classical electromagnetic phenomena. Maxwell's equations in the time domain are given as:

$$
\begin{aligned}
\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon} \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \times \mathbf{B} &= \mu(\mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t})
\end{aligned}
\tag{2.7}
$$

Together with the assumptions of linear isotropic, time-invariant, nondispersive media.

$$
\begin{aligned}
\mathbf{D} &= \epsilon \mathbf{E} \\
\mathbf{B} &= \mu \mathbf{H}
\end{aligned}
\tag{2.8}
$$

Equations 2.7 can be written as in equation 2.9, as a hyperbolic system of coupled PDEs that are first order in space and time. They can also be combined in a single PDE to form a second order in both space and time (equation 2.10).

$$
\begin{aligned}
\epsilon \frac{\partial \mathbf{E}}{\partial t} &= \nabla \times \mathbf{H} - \mathbf{J} \\
\mu \frac{\partial \mathbf{H}}{\partial t} &= -\nabla \times \mathbf{E}
\end{aligned}
\tag{2.9}
$$

$$
\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \nabla^2 \mathbf{E}
\tag{2.10}
$$

2.2.1.2 Maxwell's equations in the frequency domain

Maxwell's equations are written in time-harmonic form (in the frequency domain) by replacing the time derivative with $j\omega$ [87].

$$\nabla \times \mathbf{E} \;=\; -j\omega\mathbf{B} \tag{2.11}$$

$$\nabla \times \mathbf{H} \;=\; \mu\mathbf{J} + j\omega\mu\epsilon\mathbf{E} \tag{2.12}$$

$$\nabla \cdot \mathbf{E} \;=\; \frac{\rho}{\epsilon} \tag{2.13}$$

$$\nabla \cdot \mathbf{B} \;=\; 0 \tag{2.14}$$

Equation 2.10 can be written in time-harmonic form (in the frequency domain) by replacing the time derivative with $-j\omega$. This is shown in equation 2.15

$$\nabla^2\mathbf{E} - \frac{1}{c^2}(-jw)(-jw)\mathbf{E} \;=\; 0$$
$$k \;=\; \frac{\omega}{c} \tag{2.15}$$
$$\nabla^2\mathbf{E} + k^2\mathbf{E} \;=\; 0$$

This equation has three vector field components that can satisfy Helmholtz scalar wave equation $\nabla^2\psi + k^2\psi = 0$. ψ to be replaced with E_x, E_y or E_z.

2.2.2 Plane waves

Plane waves are the simplest example of three dimensional waves. Plane waves satisfy the 1D wave equation.

$$\frac{\partial^2\psi}{\partial t^2} = v^2\frac{\partial^2\psi}{\partial z^2} \tag{2.16}$$

To obtain plane waves, the position vector must remain perpendicular to a given plane. Pick a point on the plane $\mathbf{r}_O$. In order for the wave to be planar $(\mathbf{r} - \mathbf{r_0}).\mathbf{k} = 0$ and $\mathbf{r}.\mathbf{k} = \mathbf{r_0}.\mathbf{k} = a$ where a is a constant. The wave then satisfies the generalisation of the one-dimensional equation.

$$\psi = \psi_0\cos(\omega t - \mathbf{k}.\mathbf{r} + \phi) \tag{2.17}$$

Where, ω : Angular Frequency

ϕ : Phase

$\mathbf{k}$: Wave vector

$$\mathbf{k} = k_x \mathbf{i} + k_y \mathbf{j} + k_z \mathbf{l} \tag{2.18}$$

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2} = \frac{2\pi}{\lambda_m} \tag{2.19}$$

Where, $\quad$ k: Wave number

λ_m: Wavelength of light in the medium

All points on a given plane are of constant phase. They are called wave fronts. Simply put, a plane wave in two or three dimensions is like a sine wave in one dimension except that constant phase loci are not points, but form lines (2-D) or planes (3-D) perpendicular to the direction of wave propagation [89]. Plane waves are generic sources for illumination in the monochromatic simulations in this thesis.

2.2.2.1 Poynting vector

The Poynting vector can be thought of as representing the energy flux (in W/m^2) of an electromagnetic field. It is named after its discoverer John Henry Poynting. It is defined as [89, 87]

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \tag{2.20}$$

where $\mathbf{E}$ is the electric field and $\mathbf{H}$ the magnetic field.

In a propagating sinusoidal electromagnetic plane wave of a fixed frequency, the Poynting vector oscillates, always pointing in the direction of propagation. The time-averaged magnitude of the Poynting vector is

$$\langle S \rangle = \frac{1}{2\mu_0 c} E_0^2 = \frac{\epsilon_0 c}{2} E_0^2, \tag{2.21}$$

where E_0 is the maximum amplitude of the electric field and c is the speed of light in free space. This time-averaged value is also called the irradiance or intensity I.

2.3 Background in Surface Plasmon Resonance(SPR)

Surface plasmons (SPs) are coupled Transverse-Magnetic electromagnetic waves [2] that propagate at an interface between two media [90]. Surface plasmons can be excited by light or by electrons [90, 91]. Surface plasmons are surface waves, therefore are bound to the surface, i.e. their amplitude decays exponentially moving away from the surface.

2.3.1 Dispersion of SPs

The EM fields associated with surface plasmons can be found by solving Maxwell's equations on a planar surface between a metal and dielectric medium. The incoming light enters this system through the dielectric medium and when certain conditions are met, such as appropriate boundary conditions and material properties, the SPs are generated at the interface between the two media and propagate at this interface exponentially decaying as they penetrate into the media . Some metals can be used for SP excitation, such as Aluminium, Silver or most commonly Gold due to its resistance to oxidation. "Bound to surface" specifically means that these waves have their intensity maximum in the surface and their fields perpendicular to the surface decays exponentially away from the surface [90]. The details of the derivation of the surface plasmon dispersion is beyond the scope of this thesis. The dispersion of surface plasmons is given by [90]:

$$k_x = k'_x + ik''_x = \left[\frac{\omega}{c} \left(\frac{\varepsilon'_1 \varepsilon_2}{\varepsilon'_1 + \varepsilon_2} \right)^{1/2} \right] + i \left[\frac{\omega}{c} \left(\frac{\varepsilon'_1 \varepsilon_2}{\varepsilon'_1 + \varepsilon_2} \right)^{3/2} \frac{\varepsilon''_1}{2(\varepsilon'_1)^2} \right]. \qquad (2.22)$$

Where, ω : angular frequency (rad)

c : speed of light

ε_1 : dielectric constant of metal

ε'_1 : real part of ε_1

ε''_1 : imaginary part of ε_1

ε_2 : dielectric constant of ambient medium

Observing figure 2.2, it can be seen that the surface plasmon dispersion curve starts out like the photon curve but becomes asymptotic as the k-increases. It must be noted

---

[2] they can also be considered as charge density oscillations
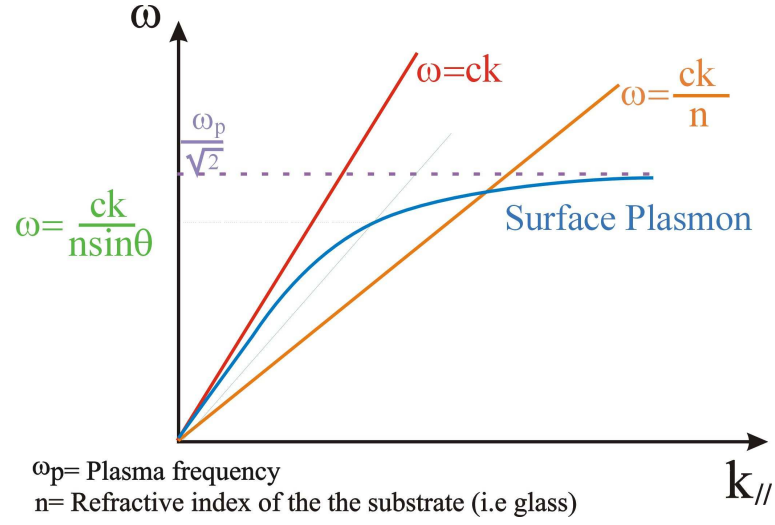
Figure 2.2: Dispersion relationship of surface plasmons. The horizontal axis represents the parallel component of the k-vector. The vertical axis represents the angular frequency. Plasmon curve lies to the right of the light in air curve. Hence they cannot be excited by light in air. A second curve of dispersion relation in glass(refractive index n) crosses the plasmon curve(labelled as $\omega = \frac{ck}{n}$), enabling excitation from glass.(ω_p=plasma frequency of metal, n=refractive index of the substrate, i.e. glass)

that the k-vector of the plasmon is always greater than the k-vector of photons in free-space. Therefore, surface plasmons cannot be excited in air on a plane surface due to the fact that the k-vector of the plasmon is greater at all frequencies than that of light in air [90].(refer to figure 2.2). In other words, on a plane surface they are non-radiative modes, which means that they cannot be generated directly by or decay spontaneously into photons [90, 92].

2.3.1.1 Propagation length of SPs

As a surface plasmon propagates along the surface, it quickly loses its energy to the metal due to absorption. The intensity of the surface plasmon decays with the square of the electric field, so at a distance x, the intensity has decreased by a factor of $\exp(2k_x'' x)$ [90]. The propagation length is defined as the distance for the surface plasmon to decay by a factor of $1/e$. This condition is satisfied at a length

$$L = \frac{1}{2k_x''}. \tag{2.23}$$

The typical propagation length for plasmons is around $22\mu m$ and $10\mu m$ for solver and gold respectively at 639nm incident wavelength.

2.3.1.2 Penetration of SPs

The penetration of SPs depends on the surrounding material and is useful in sensing applications. It is given by [90]:

$$z_i = \frac{\lambda}{2\pi} \left(\frac{|\varepsilon_1'| + \varepsilon_2}{\varepsilon_i^2} \right)^{1/2} \tag{2.24}$$

where z_i is the skin depth into the ith material (i=2 for air , i=1 for metal)).

2.4 Theoretical background in Rayleigh waves

An elastic medium behaves as a disturbed mass-spring system in which displacement of a single element results in the propagation of a disturbance throughout the medium. A particle at a surface will have different behaviour from a particle located at the interior of a substance, since the former is free from one side and the latter is surrounded by other particles in all directions. Former in this case will support surface waves whereas the latter will support bulk waves.

There are a few different modes of ultrasonic waves. These include longitudinal or compressional waves, transverse or shear waves, surface waves such as Rayleigh waves and lamb waves [93]. In this thesis, surface waves more specifically Rayleigh waves are of interest. Penetration depth of the Rayleigh waves is approximately equal to their wavelength, which makes them appealing in the NDE of surface defects. Rayleigh waves were used in the 1950s for material investigations [94, 95]. There were studies to detect Rayleigh waves in 1970s using scanning detection schemes [96]. Rayleigh waves have remained as an interest of the NDE community [11], and has been used for detecting surface-breaking fatigue cracks [97, 98], measuring stress [20], assessing surface roughness [99],surface nonlinearity [100] and imaging surface features [19, 101].

In an isotropic medium, A Rayleigh wave has both longitudinal and shear contributions [12] which are 90° out of phase with each other, thus producing elliptical motion. At a depth of 0.2λ the longitudinal part changes sign. The motion changes direction as we go deeper, to practically dissappear at 2λ. The derivation of the particle displacements are given in [102].

$$U_R = Ak \left(e^{-qz} - \frac{2qs}{k^2 + s^2} e^{-sz} \right) sin(kx - \omega t) \tag{2.25}$$

$$W_R = Aq \left(e^{-qz} - \frac{2k^2}{k^2 + s^2} e^{-sz} \right) cos(kx - \omega t) \tag{2.26}$$

In these equations the unknowns q and s can be expressed as follows,

$$q^2 = k^2 - k_l^2 \tag{2.27}$$

$$s^2 = k^2 - k_t^2 \tag{2.28}$$

$$k_l = \omega \sqrt{\rho/(\lambda + 2\mu)} \tag{2.29}$$

$$k_t = \omega \sqrt{\rho/\mu} \tag{2.30}$$

Where, U_R: Perpendicular component of displacement

W_R: Parallel component of displacement

k: Wave number

λ: bulk modulus

μ: shear modulus

ρ: modulus density

k_l: longitudinal k-vector

k_t: transverse k-vector

ω: angular frequency

A: amplitude

As can be seen from equation 2.25 and 2.26, The amplitude of the particle displacement is modified with an exponential, indicating the surface-bound nature of the wave.

2.5 Theoretical background in the reflection and transmission of polarised light by stratified planar structures (the analytical solution)

The analytical solution considers a stratified structure that consists of a stack of m parallel layers, sandwiched between two semi-infinite layers. The first layer is layer 0 and the last layer is layer m+1. The analysis assumes that the layers are all linear, homogeneous and isotropic. For a layer with index j, the thickness is noted as d_j, and the refractive index is noted as N_j. A monochromatic plane wave incident on the layer 0, yields a resultant reflected field in layer 0 and a resultant transmitted field in layer m+1. In each layer, there is a resultant field travelling forward(denoted as positive) and backwards(denoted as negative) directions. The geometry and the direction conventions are shown in figure 2.3. In this section the main focus is on the p-polarised waves since p-polarised waves can excite plasmons. (For details and the derivation of these equations and for expressions in s-polarisations, refer to [103]).

The total field in layer j can be expressed as a column vector as in equation 2.31.

$$\mathbf{E}(z) = \begin{bmatrix} E^+(z) \\ E^-(z) \end{bmatrix} \tag{2.31}$$

The total field at z' is related to the total field at z" by equation 2.32, or more concisely as 2.33.

$$\begin{bmatrix} E^+(z') \\ E^-(z') \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} E^+(z'') \\ E^-(z'') \end{bmatrix} \tag{2.32}$$

$$\mathbf{E}(z') = \mathbf{S}\mathbf{E}(z'') \tag{2.33}$$

When z' and z'' are chosen to be on opposite sides of a boundary between two adjacent media, then $\mathbf{S}$ relates the two boundaries together. A general equation representing this
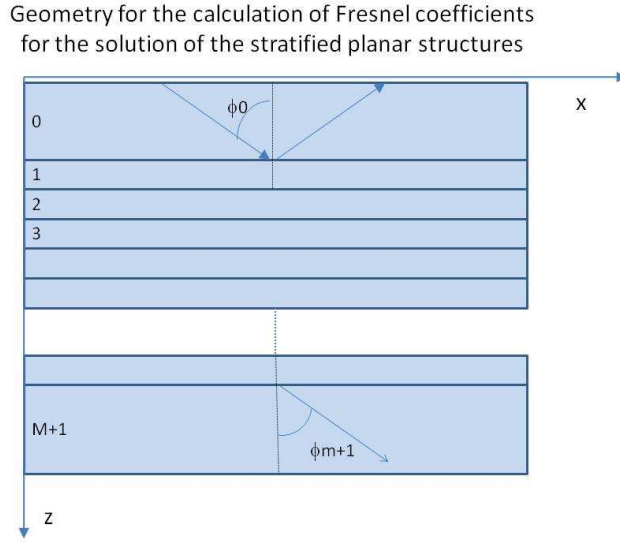
Figure 2.3: The geometry for the calculation of the Fresnel coefficients of planar stratified media. The field comes in from layer 0 and exits from layer m+1

relationship, the jth boundary between layer j and $j+1$ is

$$\mathbf{E}(z_j - 0) = \mathbf{I}_{(j-1)j}\mathbf{E}(z_j + 0) \tag{2.34}$$

In this notation, $\mathbf{I}_{(j-1)j}$ is a 2X2 matrix representing the properties of the boundary j. The propagation of the plane waves in a layer can be expressed by taking the parallel lines at the two layer boundaries as,

$$\mathbf{E}(z_j + 0) = \mathbf{L}_j\mathbf{E}(z_j + d_j - 0) \tag{2.35}$$

Therefore, the relationship between the fields at planes $z_1 - 0$ and $z_{m+1} + 0$ can be expressed as

$$\mathbf{E}(z_1 - 0) = \mathbf{S}\mathbf{E}(z_{m+1} + 0) \tag{2.36}$$

Plugging equation 2.35 into equation 2.34, a definition for $\mathbf{S}$ can be expressed as

$$\mathbf{S} = \mathbf{I}_{01}\mathbf{L}_1\mathbf{I}_{12}\mathbf{L}_2 \cdots \mathbf{L}_m\mathbf{I}_{m(m+1)} \tag{2.37}$$

I and **L** are given by equations 2.39 and 2.38.

$$\mathbf{I}_{ij} = \begin{bmatrix} 1/t_{ij} & r_{ij}/tij \\ r_{ij}/t_{ij} & 1/t_{ij} \end{bmatrix} \qquad (2.38)$$

$$\mathbf{L}_j = \begin{bmatrix} e^{j\beta_j} & 0 \\ 0 & e^{-j\beta_j} \end{bmatrix} \qquad (2.39)$$

In equation 2.39,

$$\beta_j = \frac{2\pi d_j n_j cos(\theta_{in,j})}{\lambda} \qquad (2.40)$$

θ_{in_j} is the angle the forward travelling total field makes with the normal to the boundaries in layer j. n_j and d_j are refractive index and the thickness of layer j, respectively. θ_{in_j} or shortly θ_j, can be expressed in terms of the incidence angle θ_0 from Snell's law.

$$n_0 sin(\theta_0) = n_j sin(\theta_j) \qquad (2.41)$$

The elements of $\mathbf{I}_{ij}$, namely r_{ij} and t_{ij} can be expressed as

$$r_{ij} = \frac{n_j cos(\theta_i) - n_i cos(\theta_j)}{n_j cos(\theta_i) + n_i cos(\theta_j)} \qquad (2.42)$$

$$t_{ij} = \frac{2n_i cos(\theta_i)}{n_j cos(\theta_i) + n_i cos(\theta_j)} \qquad (2.43)$$

r_{ij} and t_{ij} are the reflection and transmission coefficients relating the layers i and j together.

2.6 EM–Explorer and theoretical background in FDTD

EM-Explorer is a commercial software which employs Finite Difference Time Domain (FDTD) method. FDTD was described in section 1.6.3. Even though the coding of FDTD is fairly simple and relatively easy to develop, building a complete FDTD model from scratch is a fairly large project, hence it was decided to search for a readily available commercial code. Commercial FDTD packages are available such as EM-Explorer, which

collect some boundary conditions etc. into their libraries and make them available for electromagnetics modelling, without the need for low level programming. FDTD scripts can be programmed for short duration pulses to yield wideband results. Monochromatic simulations are also possible if plane wave sources are used, as it is in EM-Explorer. This thesis will only involve monochromatic problems.

EM Explorer sets up the source of incident waves automatically, with source type, polarisation and incidence angle as user inputs. In the simulations considered, plane waves were swept through the incidence angles via a script. The cubic mesh in EM Explorer was set up to have one mesh element in transverse directions, i.e the directions that are out of plane, and many in the normal direction, within the plane, so as to be able to resolve the interfaces of different materials. The periodic boundaries were set up according to the angle of incidence. A gauge was set up to measure the reflected amplitudes. The location of the gauge was selected behind the source boundary so as to exclude the amplitude of the incoming wave. The source was set as a sinusoid with known amplitude and its direction was chosen to be left to right. Knowledge of the incoming wave amplitude, and the reflected amplitude yielded the reflectivity results. Next, some theoretical background to FDTD method is given to demonstrate its most attractive feature, simplicity.

2.6.1 Theoretical background in FDTD

The starting point of the Finite-Difference Time-Domain Mechanism is the Maxwell's time domain equations in the differential form (see equations 2.7). Equations 2.7 can be written as in equation 2.44, as a hyperbolic system of coupled PDEs that are first order in space and time. They can also be combined in a single PDE to form a second order in both space and time (equation 2.45). They are repeated here for demonstration.

$$
\begin{aligned}
\epsilon \frac{\partial \mathbf{E}}{\partial t} &= \nabla \times \mathbf{H} - \mathbf{J} \\
\mu \frac{\partial \mathbf{H}}{\partial t} &= -\nabla \times \mathbf{E}
\end{aligned}
\tag{2.44}
$$

$$
\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \nabla^2 \mathbf{E}
\tag{2.45}
$$

2.6.1.1 Finite difference approximation

In the finite difference approximation, the derivative of a smooth function with respect to a variable is approximated by a two-point finite difference over the sufficiently discretised variable. Increased accuracy is achieved when the finite difference is shifted to be centred. In the core of FDTD, both space and time derivatives are approximated by centred two-point differences [48]. Equation 2.46 shows the central finite difference equation. The term $O(\Delta z^2)$ indicates that the remaining terms are in the order of Δz^2, which will vanish more quickly as Δz goes to zero.

$$\frac{df}{dz}\mid_{z=z_0} = f' = \frac{f(z_0 + \frac{\Delta z}{2}) - f(z_0 - \frac{\Delta z}{2})}{\Delta z} + O(\Delta z^2) \tag{2.46}$$

2.6.1.2 Finite-Difference Time-Domain (FDTD) mechanism

Here, the wave propagation in two dimensions is briefly described. It will be assumed that the only components present in a 2D model are H_z, E_x and E_y for an isotropic medium in the xy-plane. FDTD analysis starts with discretising geometry into a number of cells each of size $\Delta x \times \Delta y$. Each of these unit cells are characterised by the medium properties, i.e ϵ, μ, σ. etc.

Maxwell's curl equations simplify to equation 2.47 for TE modes since H_x, H_y and E_z are made zero.

$$\begin{aligned}
\epsilon_0 \frac{\partial E_x}{\partial t} &= \frac{\partial H_z}{\partial y} \\
\epsilon_0 \frac{\partial E_y}{\partial t} &= -\frac{\partial H_z}{\partial x} \\
\mu_0 \frac{\partial H_z}{\partial t} &= \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}
\end{aligned} \tag{2.47}$$

The FDTD spatial discretisation is performed by positioning the field variables such that all spatial derivatives can be approximated by two point, centred differences as described above (see figure 2.4).

Therefore a unit Yee cell and the field variables around it can be expressed as equation 2.48.
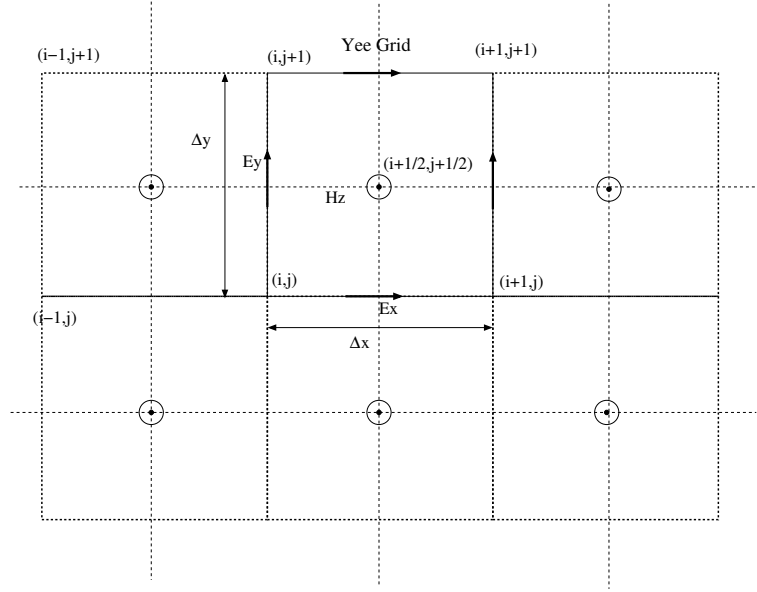
Figure 2.4: The demonstration of the Yee Cell, the basis of the FDTD algorithm. The cell edges are set to be the Electric field components whereas the centre of the cells are set to represent the Magnetic field. This rule of thumb is due to the general approach of implementing the boundary conditions based on the Electric field, rahter than the magnetic field.

$$
\begin{aligned}
\epsilon_0 \, \dot{E}_x(i + \frac{1}{2}, j) &= \frac{H_z(i + \frac{1}{2}, j + \frac{1}{2}) - H_z(i + \frac{1}{2}, j + \frac{1}{2})}{\Delta y} \\
\epsilon_0 \, \dot{E}_y(i, j + \frac{1}{2}) &= -\frac{H_z(i + \frac{1}{2}, j + \frac{1}{2}) - H_z(i - \frac{1}{2}, j + \frac{1}{2})}{\Delta x} \\
\mu_0 \, \dot{H}_y(i + \frac{1}{2}, j + \frac{1}{2}) &= \frac{E_x(i + \frac{1}{2}, j + 1) - E_x(i + \frac{1}{2}, j)}{\Delta y} \\
&\quad - \frac{E_y(i + 1, j + \frac{1}{2}) - E_y(i, j + \frac{1}{2})}{\Delta x}
\end{aligned}
\tag{2.48}
$$

The dots in equation 2.48 indicate the time derivatives. The time derivatives are achieved via the time-discretisation scheme of FDTD. The field variables at a given time step are represented as equation 2.49.

$$
E_x^n \equiv E_x(t = t_0 + n\Delta t)
\tag{2.49}
$$

This scheme requires that the **E** and **H** are staggered in time by one half of a time step. If **E** occupies the integer time steps, then **H** must be allocated the half-integer time

points. Equation 2.50 represents the time derivative expressions as to be plugged into equation 2.48. Equation 2.51 represents the field expressions at time steps as required by equation 2.50.

$$\begin{aligned}
\epsilon_0 \dot{\mathbf{E}}^{n-\frac{1}{2}} &\sim& \epsilon_0 \frac{\mathbf{E}^n - E^{n-1}}{\Delta t} = \nabla \times \mathbf{H}^{n-\frac{1}{2}} \\
\mu_0 \dot{\mathbf{H}}^n &\sim& \mu_0 \frac{\mathbf{H}^{n+\frac{1}{2}} - \mathbf{H}^{n-\frac{1}{2}}}{\Delta t} = -\nabla \times \mathbf{E}^n
\end{aligned} \tag{2.50}$$

where,

$$\begin{aligned}
\mathbf{E}^n &=& \mathbf{E}^{n-1} + \frac{\Delta t}{\epsilon_0} \nabla \times \mathbf{H}^{n-\frac{1}{2}} \\
\mathbf{H}^{n+\frac{1}{2}} &=& \mathbf{H}^{n-\frac{1}{2}} - \frac{\Delta t}{\mu_0} \nabla \times \mathbf{E}^n
\end{aligned} \tag{2.51}$$

It is seen that FDTD is an explicit finite difference approach, i.e. no matrix equation is setup, stored and solved. The field values at the next time step are given entirely in terms of the field at the current and the previous time steps [25].The electric and magnetic fields are updated using a leapfrog scheme where first the electric, then the magnetic field is computed at each step in time.

2.7 Comsol and theoretical background in FEM

In this section, Finite Element Method and Comsol are viewed. A theoretical background is followed by the explanation of various features of the models presented in this thesis. This section aims to provide a description of FEM and Comsol as a method and tool in the design and optimisation of the devices mentioned in this thesis.

2.7.1 Theoretical background in finite element method

FEM is a general variational method, which finds a variational functional whose minimum/maximum/stationary point corresponds to the solution of the PDE subject to certain boundary conditions. Its application to EM problems can be done through the Maxwell's equations in the frequency domain (see equation 2.15). Finite Element Method analysis of a problem consists of few general steps

1. meshing the simulation domain so that the simulated material is homogeneous on each element

2. deriving the governing equation for an element

3. assembly of all elements into a system of equations

4. the solution of the system of equations

2.7.1.1 Segmentation of geometry: meshing

A given problem is divided into many subproblems, that are easier to solve, via a process called meshing. The problem is said to be divided into a finite number of elements [28]. Different shapes may be used for elements depending on the geometry to be resolved and the computational resources available. Throughout this thesis, triangular elements will be considered. A simple conceptual example of the use of triangular mesh elements are shown in figure 2.5. General tendency in 2D modelling is towards variable sized triangular elements [104].

Consider that the circle in figure 2.5 represents an isotropic source free area. Assuming that the field distribution of a general element e is φ^e, the distribution throughout the domain is a linear combination of the distribution in each element. N_e in equation 2.52 represents the number of elements, which is directly related to the accuracy of the solution. A rule of thumb is to choose N_e so that the maximum element edge length is less then $\lambda_o/10$, where λ_o is the optical wavelength. This is to ensure proper discretisation of the fields. In the models considered in this report, the minimum element size will be dependent on the geometry specifications. For some of the very small structures described in this thesis, the mesh elements in the order of $\lambda/600$ were used.

$$\varphi(x,y) = \sum_{e=1}^{N_e} \varphi^e(x,y) \tag{2.52}$$

2.7.1.2 Derivation of element matrix

There are a few methods to derive the element matrix. First, is the variational method, which finds a variational functional whose minimum/maximum/stationary point corre-

Discretization of a circular geometry with simple triangular FEM elements
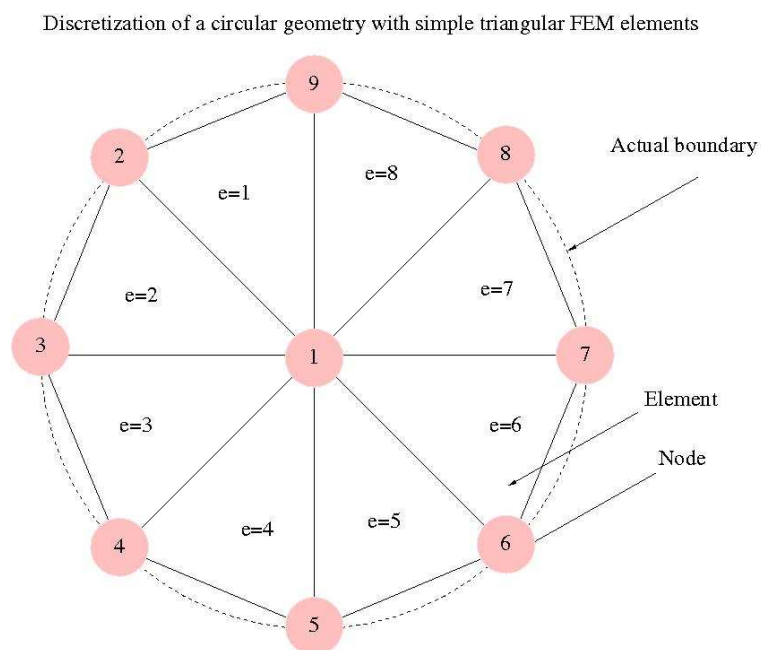
Figure 2.5: The demonstration of the triangular mesh elements used in FEM. In this example the 8 elements are used to approximate a circle as an octagon. smaller mesh elements could be used for better approximation

sponds to the solution of the PDE subject to certain boundary conditions. Second method is sometimes called the "weak formulation" in the literature. It works by introducing a weighted residual error to one of the differentials in the PDE form of Maxwell's equations [25]. In this section, the former will be discussed due to its simplicity.

The field distribution variable φ for each element is expanded in suitable basis functions ψ_i^e. ψ_i^e can be node based or edge based functions. In this section, ψ_i^e are considered to be node-based for simplicity.

Basis functions only have a value over the element. They vanish outside of the element e. They influence the field value in the neighbouring elements through boundary conditions. The basis functions should form a complete set in order to describe φ^e efficiently. Due to the limitation of the computing resources, the number of basis functions are limited to a few, determined conventionally by the order of the governing differential equation, via a continuity requirement.

Assuming an element e with p nodes, this expansion can be achieved as in equation 2.53.

$$\varphi^e(x,y) = \sum_{i=1}^{p} c_i^e \psi_i^e(x,y) \tag{2.53}$$

In equation 2.53, c_i are unknown complex coefficients, and ψ_i^e are the basis functions. By enforcing the governing differential equation 2.15, the coefficients c_i are determined. The differential equation cannot be satisfied fully, due to the computational requirements. However, an average satisfaction of the differential equation may be obtained by using variational functionals as in equation 2.54.

$$F^e[\varphi^e] = \frac{1}{2} \iint_S [|\nabla \varphi^e|^2 - k^2 (\varphi^e)^2] dx dy \tag{2.54}$$

Equation 2.53 and 2.54 can be joined in matrix form as in equation 2.55. This form represents a set of equations to be solved for each element.

$$[F^e] = [A^e][\varphi^e] \tag{2.55}$$

In this case $[A^e]$ is the element matrix. which takes the form of a 3x3 matrix for triangular elements in 2D.

$$[A^e] = \begin{bmatrix} A_{11}^e & A_{12}^e & A_{13}^e \\ A_{21}^e & A_{22}^e & A_{23}^e \\ A_{31}^e & A_{32}^e & A_{33}^e \end{bmatrix} \tag{2.56}$$

The matrix elements A_{ij}^e in equation 2.56 represent the coupling with the nodes i and j [28]. Boundary conditions are then applied to the element matrices, i.e the values of the nodes on certain boundaries are changed according to specified boundary conditions. This is followed by an assembly procedure to simplify the problem to allow more sparse matrices.

2.7.1.3 Assembly and solution of matrices

Equation 2.55 represents a set of equations for each element. This can be summed up as it is in 2.52, for the whole model as it is in equation 2.57. The system matrix form is reached in equation 2.58, where $[\varphi]$ is known as the global nodal vector with elements c_i^e

$$\sum_{e=1}^{N_e} [A^e][\varphi^e] = \sum_{e=1}^{N_e} [F^e] \tag{2.57}$$

$$[A][\varphi] = [F] \tag{2.58}$$

The assembly procedure reduces the repetitions due to common local nodes, and ensures the continuity of φ from element to element.

The system matrix is solved afterwards to determine the unknown vector φ of potentials at the nodes. Correct nodal values exist at the point where the functional is stationary with respect to the expansion coefficients c_i. Therefore its derivative with respect to c_i is equalled to zero, which is called the Ritz Method [28]. Or due to the linearity of the assembly process, the derivative of F^e with respect to c_i^e are taken and then summed to zero as in equation 2.59.

$$\frac{\partial F}{\partial c_i} = \sum_{e=1}^{N_e} \frac{\partial F^e}{\partial c^e} = [0] \tag{2.59}$$

The resulting set of simultaneous equations are solved to yield the unknown vector φ. There are many available solvers for FEM that is available. Some solvers options are L-U decomposition, iterative solvers and Pardiso [105] depending on the problem under consideration.

2.8 Methods and model features in comsol

This section describes the model features and their implementation in Comsol. The model features in Comsol were thoroughly analysed and tested prior to designing the devices in this thesis. Detailed analysis of these features can be found in appendix B. Finite Element terminology defines the regions separated by boundaries in the simulation domain, as subdomains. Every subdomain is configurable to represent different materials, and are separated by boundaries which relate different media to each other and express the behaviour of the waves passing through using boundary conditions. When designing the devices considered in this thesis, scripts are built to adjust various properties and geometry of these subdomains and solve the resulting models.

General issues with modelling electromagnetic problems involve concepts such as speed, memory consumption, geometry representation, and truncation of infinite domain, accurate source representation and postprocessing.

The most attractive feature of FEM is its speed and meshing abilities. Comsol implements an automatic mesh generator, which uses triangular elements of multi-scale. When modelling complicated geometries, the meshing is adjusted to give more accuracy in very thin regions.

A common problem in electromagnetics modelling is the truncation of the infinite domain. The truncation issue is common to all numerical modelling methods, however, FDTD method can circumvent this problem for some cases, by reducing the time limits the simulations are run. When plane waves are obstructed abruptly without the appropriate boundary conditions, they will introduce unwanted components in the frequency domain.
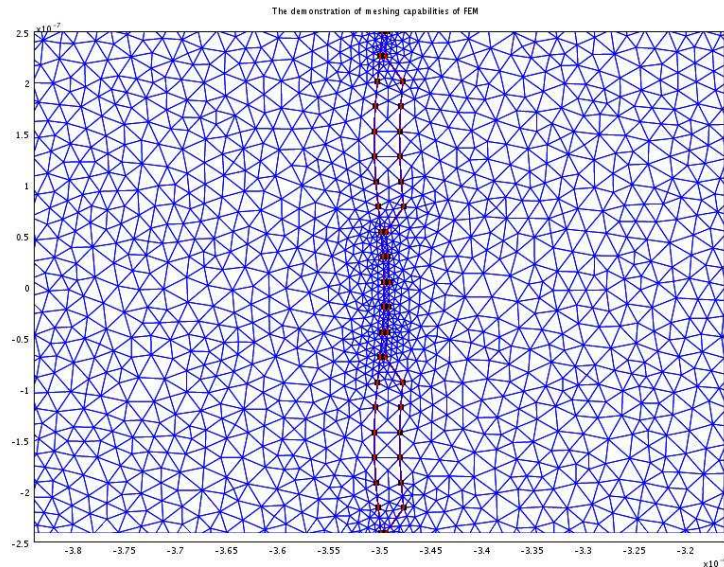
Figure 2.6: The demonstration of advantageous meshing abilities of FEM. As can be seen, the curved lines can be represented efficiently due to triangular elements, and the presence of variable size elements make it possible to represent curved surfaces with high resolution at a reduced cost in terms of computing power.

This decreases the numerical accuracy of the solution and the theoretical accuracy of the model.

Photonic structures are tend to be of repeating nature, especially when their dimensions are smaller than the wavelength of the light being used. Such structures will be represented using periodic boundary conditions, which is explained in section 2.8.2. However, it is necessary and important to remove the periodicity and look at isolated structures, to analyse the devices in more detail. When the devices considered involve dealing with media of semi-infinite geometry, Perfectly Matched Layers (PMLs) are used to truncate the simulation domain.

2.8.1   Truncation of EM models using Perfectly Matched Layers (PMLs)

A generic problem when modelling time-harmonic electromagnetic waves in open space, is making the simulation space finite.

2.8.1.1 Problems with the truncation of open space

Theory of harmonic waves, indicate that the solution for a wave, namely a plane wave, would be valid only at infinite dimensions. This is because the wave fronts of plane waves are planes of infinite size. Trimming parts of this solution so as to fit it into a finite simulation domain would result in diffraction effects, because such a domain would resemble an aperture which would change the shape of the spectrum in the frequency domain. Additionally, in the outer boundaries of the domain, inappropriate boundary conditions results in unwanted reflections since the waves that would normally go undisturbed are obstructed by these outer boundaries [106]. The method used to take care of this problem need to either block and absorb the incoming plane wave, or to pass it through without any reflections.

2.8.1.2 Properties of PMLs

Perfectly Matched Layers (PMLs) are artificial media created to circumvent the problems associated with the truncation of the simulation domain [107].

PMLs are anisotropic artificial media whose parameters are designed for a method of complex coordinate-stretching (for details see appendix B.1), in order to attenuate the waves going through them exponentially [54] through the width of the layer(see figure 2.7). PMLs are ideally "perfectly matched" to the medium they terminate, hence, as far as the incoming field is concerned, there is no boundary between the actual medium and the artificial absorbing layer.

PMLs are generally set to absorb the component of the field that is normal to the boundary between the actual and the artificial domain. After that component is removed, the parallel component is taken care of by the terminating boundary condition. Hence, different formulations are necessary for spherical coordinates or for when the boundary is not straight. PMLs are very useful features in FEM modelling and are easily set up in Comsol, especially for rectangular coordinates.

2.8.1.3   Limitations of PMLs

The presence of PMLs imply more mesh elements, hence they require more CPU time and memory due the complex mathematics involved in their implementation, and the resulting enlargement of the simulation domain, which is a fundamental problem especially in 3D simulations. Also, since they have to be set up with a general sense of absorbing direction, their performance gets worse for the cases when there is a very wide range of angles (see section B.2). Figure 2.7 shows what happens to the norm of the wave as it enters a PML region.
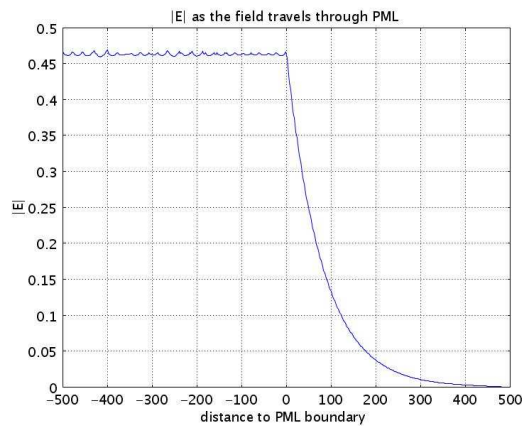
Figure 2.7: The exponential decay of the norm of the E-field upon entering the PML layer

Detailed background in the opeartion of the PMLs and their shortcomings can be found in section B.1. PML absorption decrease with the incidence angle (see section B.2). For very oblique incidence angles (towards grazing incidence), the reflections from the PMLs could become big enough to create standing wave patterns. Figure 2.10(a) show these reflections. Figure 2.10 (b) shows the total field amplitude when passing through a layer of PML with various incidence angles.

PML performance can be improved upto a level by discretising them more finely. However, the existence of PMLs alone increases the computation time, due to increased amount of complex calculations involved. PML settings were automatically optimised by Comsol. Only the geometry and the desired absorbing direction is specified.

2.8.2 Using periodic boundaries in comsol

A model can be truncated using periodic boundary conditions(PBCs), if it can be assumed to be infinitely repeating in one dimension, with a given period. In all the models considered in this thesis, the infinitely repeating assumption was used. Using periodic boundary conditions, it is possible to model any infinite structure given that one unit cell can be modelled and the relationship between each unit cell is known. It is also possible to model some isolated structures using this condition, if the simulation domain can be made big enough to assume that the periodicity k-vector is too small to have an effect on the results.

In this section the use of PBC is described and the implementation of plane waves with these settings is discussed. (refer to appendix B.3 for more detailed explanation of PBCs)

2.8.2.1 Periodic boundary conditions description

The repeating nature, i.e. periodicity, can be realised by deciding what part of the sample is unique, and analysing/implementing the input-output relationship of a unit cell with its neighbouring unit cells.

When the periodicity assumption holds, the fields crossing the boundaries between the individual unit cells, are equal or related, i.e fields leaving one side of the a cell, should be related to the fields coming in from the other side [54]. Figure 2.8 explains this situation in a simple form. Experimentally, this assumption is validated by having a laser spotsize, that is much bigger than a unit cell. This ensures that the intensity is uniform over many periods.

Application of the PBCs to the model will solve the problem due to the diffraction of the incident plane waves, by effectively making the simulation domain infinitely repeating. Since the model can be considered as infinitely repeating, the source boundary can be set-up to be infinite size. This makes the incident excitation (plane waves) in the model mathematically correct.
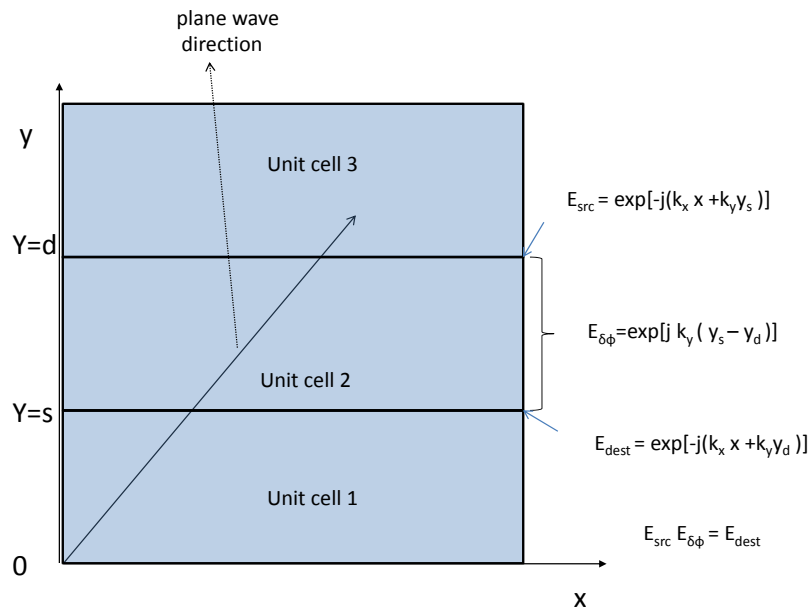
Figure 2.8: A graphical explanation of how the periodic boundary conditions apply. The field leaving the unit cell 1(E_{dest}) is related to the field entering unit cell 2. The field leaving unit cell 2 (E_{src})is related to the field entering unit cell 3. Since all the unit cells are equivalent, it is adequate to model one unit cell and apply periodic boundary conditions so as to realise the input output relationship. At different incidence angles, the y-component of the k-vector is different, requiring an adjustment of the phase of the incoming and outgoing fields to/from the unit cell. When the phase between the source and destination boundaries (bottom and top boundaries of a unit cell) is not 2π ($E_{\delta\phi}$)this phase has to be taken into account.

2.8.3 The effect of periodic boundary conditions and PMLs on the source boundary

The source was modelled by setting up an outer boundary with a plane wave expression. It is worth noting that this type of source boundary can be used accurately only when some sort of infinity assumption, such as the periodic assumption is made in the corresponding direction.

A perfect plane wave source can be used when PBCs are utilised. PBCs will enforce an infinitely repeating condition along the y-direction as a default, hence making the plane wave expressions mathemetically correct. The source boundaries are set to a value, which will force them to equate the field on that boundary to the source amplitude. When there is incoming field to this boundary, it will reflect, yielding multiple reflections if there is a reflective layer in the model opposite to the source boundary. The source boundary is therefore set behind a PML to avoid the possibility of multiple reflections from the models.

The input field expression is written in terms of the coordinate variables and the k-vectors, which were all dependent on this input angle variable. The source boundary is prescribed by a series of variables which are updated every time the model is run. Equation 2.60 shows these variables.

$$
\begin{aligned}
k &= \frac{2\pi n_1}{\lambda_0} \\
k_x &= k\cos(\theta) \\
k_y &= ksin(\theta) \\
H_z &= \exp(-j(k_x x + k_y y))
\end{aligned}
\tag{2.60}
$$

Since the field was entering the domain through the PML layer on the left, the x-component of both the E_x and E_y fields were heavily attenuated before entering the glass subdomain. The performance of the PMLs are known to degrade with increasing the angle of incidence, because the decay constant of the absorbing material is proportional to $k_x = kcos(\theta_{in})$. Also, for very oblique incidence angles (close to grazing incidence), due to the transverse nature of light, the x-component of the field dominates the amplitude.

This effect can be seen as an envelope when the maximum amplitudes from each input
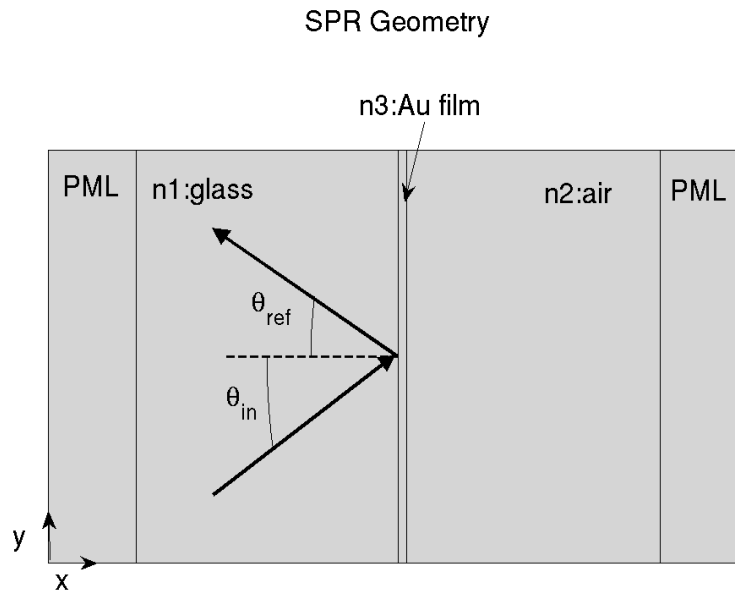
SPR Geometry

Figure 2.9: The general model geometry shown on a simple Comsol Model. The subdomains from left to right are PML, glass(n1) , Au(n3), air(n2), PML. The excitation comes from behind the left PML through the glass layer. The reflected field at various incidence angles, θ_{in}, is observed

angle are collected to form a plot (refer to Figure 2.10(b). Any changes in the reflected field, or the transmitted field will be seen superimposed onto this envelope. This envelope is dealt with by normalising every field value by the incoming field amplitude, before postprocessing.

2.8.4   Farfield extraction

The devices such as self-calibrated plasmon sensor, plasmon resonator and the eCHOT(evanescent CHeap Optical Transducers) structures use the periodicity assumptions described above. In this scenario, there is a distinct reflecting feature present, dividing the medium into a reflected side and a transmitted side. In all cases, the FEM models only incorporate a small, finite domain due to the computing resources. In this section, the extraction of farfield information from such nearfield data is described. Farfield informally means what happens to the fields far away, when the evanescent waves have died off, over large distances. In farfield region, the angular field distribution is independent of the distance from the source.
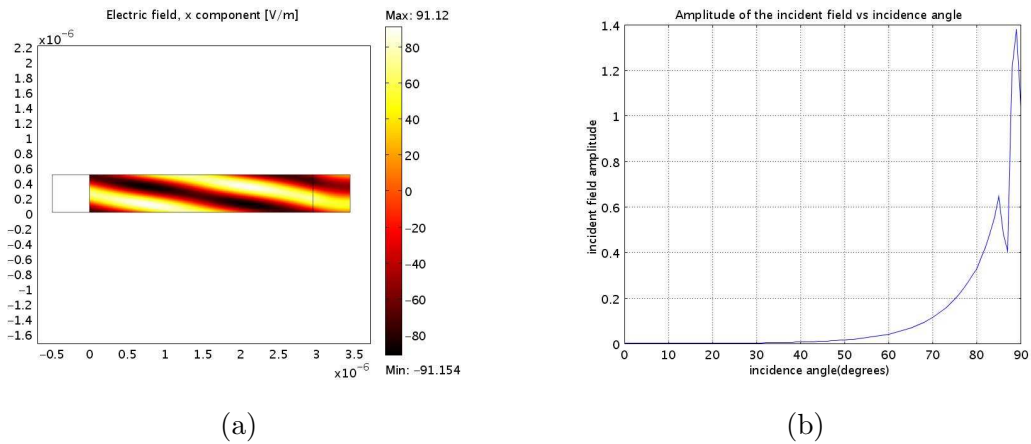
Figure 2.10: **(a)**The x-component of the E-field at 80°. It can be noted that at grazing incidence angle, there seems to be a problem, because the field looks like a very faint standing wave. Also there is an increase in the maximum field amplitude, as the angle of incidence increases. **(b)** The norm of the field amplitude as the input angle changes. As can be seen, the amplitude goes up with the input angle due to the decreased absorbance in the PML layers. The feature around 80° - 90° is attributed to the reflection from rightmost PML

When simple periodic models including only flat interfaces are modelled, data such as reflectivity may be obtained looking at the standing wave patterns in the nearfield data (for more details and the calculation of reflectivity from near field data, refer to section B.5).

When there are more than one reflected components present in the model such as different diffracted orders or scattered fields, the standing wave patterns may get very complicated. Especially in the cases of gratings, where waves with different directions and amplitudes will be seen to superimpose in the simulation domain. In these cases it is useful to be able to resolve these components, so as to be able to analyse a specific one that may be of interest. In order to do this, it is necessary to look at the farfield spectrum of the fields, where different propagating fields will be well separated form each other. .

Farfield information is used in the design the devices presented in this thesis. It is an important tool in studying optical components. An integral equation called Stratton-Chu equations are used to extract the farfield information from the models considered. Implementation of this equation involves obtaining the field information along a closed surface. This boundary/surface is referred to as the source boundary.

2.8.4.1 Farfield extraction using modified implementation of the Stratton-Chu equation

In Comsol 3.4 the farfield is calculated from the near field solution. In order to use the Stratton-Chu formulation, a boundary around which integration is to be carried out is required. This boundary is called the source boundary and it has to be placed in the model prior to meshing. Far-field is then calculated on the specified boundary.

Farfield extraction for the problems in this thesis could not be done using the Stratton-Chu equation given in Comsol 3.4, which was designed for antenna radiation patterns. This implied that it assumed that the fields were in air and there was a net source enclosed by the integral surface. When there was no net source enclosed in the farfield source boundary, the incoming field cancelled out the outgoing field leaving a zero net field. This is referred to as the boundary shadowing problem.

Consequently, a new farfield extraction system was developed using a modified version of the Stratton-Chu equations. This included an angular apodisation function which allows the extraction of the farfield for fields which did not originate inside the source boundary. Scripts (see appendix A.2.2, A.2.4) used the parameters such as refractive index of the medium, the source radius etc. as inputs. Using these constants, the modified Stratton-Chu equation (see equation 2.61) was implemented (script can be found in Appendix A.2.4). The field was extracted around a circle (in 2D) and integrated as shown in equation 2.61.

The farfield source boundary resembles an aperture in functionality. The observation of the farfield is done through this aperture. Therefore the size of this aperture is reciprocally related to the plane wave profile width in the farfield spectrum. The aperture was chosen to be circular so as to maintain a uniform shape for the visible aperture from all observation angles.

The original Stratton-Chu equation (refer to equation B.35) was modified with the necessary constants to work in any medium. (rewritten with the new constants in equation 2.61).

$$E_p = \frac{jk_n}{4\pi} r_0 \times \int [\mathbf{n} \times \mathbf{E_{ap}} - \eta_n \mathbf{r}_0 \times (\mathbf{n} \times \mathbf{H_{ap}})] exp(jk_n \mathbf{r}.\mathbf{r}_0) dS \qquad (2.61)$$

Where,

$\mathbf{E}$ and $\mathbf{H}$: Electric and Magnetic Fields on the

the surface S enclosing the source

$\mathbf{r_0}$: unit vector pointing from origin to the field point p

$\mathbf{n}$: unit normal to surface S

η_n : impedance of the ambient medium

k_n : wave number in the ambient medium

$\mathbf{r}$: radius vector of the surface S.(not a unit vector)

$\mathbf{E_p}$: calculated far field at point p

The importance of the change from $\mathbf{E}, \mathbf{H}$ in the original equations to $\mathbf{E_{ap}}, \mathbf{H_{ap}}$ has to do with the apodising aperture function. Stratton-Chu equation is designed for the situation when there is a scatterer or a source in the middle of an enclosing medium. In order to keep the validity of this assumption, and be able to use this equation for sourceless media, i.e account for the fields that simply pass through the aperture undisturbed, an aperture function was used to apodise the extracted nearfield. This apodising aperture function solved the boundary shadowing problem by breaking the balance of the incoming and the outgoing plane wave in a sourceless medium.

Apodisation was done by multiplying the extracted field with a waveform in a way which nullified the field incoming to the source boundary from outside, while keeping the outgoing field. Therefore the maximum of this apodisation waveform is set to the desired farfield observation angle. The waveform for the apodising aperture function is gaussian. Thus, the field along the source boundary would appear as if it originated from a source, located within the source boundary, radiating in the desired farfield direction. When there is no field propagating in the direction looked at, the result of the apodisation is zero, hence would describe a source not radiating. (The details of the aperture function is shown in section B.7.)

The data obtained from the farfield data corresponding to the incoming waves is also used in the normalisation procedures, associated by the uneven absorption of the left PML layer (see 2.8.3). When the far field is extracted, the total far field is calculated by
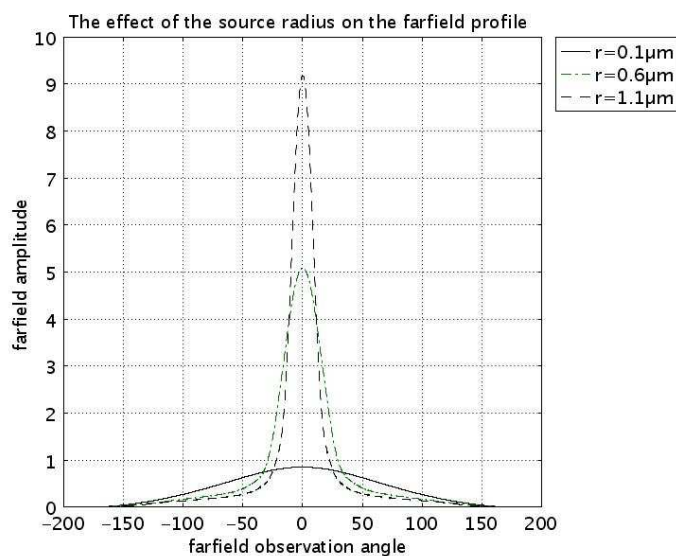
Figure 2.11: Effect of varying the far field source size in a model made up of only glass. The incoming field has thicker far field profile when the radius is smaller and has a narrower profile when the radius is bigger.

Equation 2.62.

$$E_{ff_{total}} = \sqrt{\left|E_{ff_x}\right|^2 + \left|E_{ff_y}\right|^2 + \left|E_{ff_z}\right|^2} \tag{2.62}$$

The incoming field could be extracted in the first medium after going through the PMLS. The normalisation problem due to the angle dependent incident field absorption was thus solved by normalising the whole farfield spectrum by the incident field at each incidence angle(shown in appendix A.2.3).

This could not be done without farfield extraction, since the input field and reflected field together with any unknown diffracted orders (in the case of a grating) would all be combined in the near field.

The farfield extraction is especially useful in the modelling of gratings, since it makes the diffracted orders visible. The ability to extract the farfield information from an arbitrary periodic structure will facilitate the design and analysis of novel and useful structures.

2.9 Summary

This chapter demonstrates a preliminary analysis before taking on the big modelling task. The guidelines for the technique used in design of the devices in this thesis are introduced. Background about electromagnetic waves, surface plasmons and Rayleigh waves were given to help understand the device mechanisms.

This was followed by the mechanisms and basic equations for each modelling tool used in this thesis. These were the analytical solutions to planar stratified media, FEM and FDTD. Background knowledge in the modelling methods and their implementation is important since they are related to the abilities of the modelling methods which dictate which modelling tool is more appropriate in terms of speed, accuracy, and versatility.

After the background for modelling tools, the model features of FEM and Comsol were discussed so as to understand how certain features were implemented and how certain shortcomings of Comsol were circumvented or solved. Issues such as PMLs, periodic boundaries, source boundaries and farfield extraction were discussed. Given this background it is possible to understand the route in the design process and its relationship to the choice of modelling methods used in this thesis.

Chapter 3

Applications and Experimental Background

3.1 Introduction

This chapter will give background on the applications for the devices considered in chapter 4 together with the experimental methods. The discussion in this chapter helps in the design process by clarifying the practical issues and points out the importance of the devices in their application areas.

The discussion of applications of the devices in science and engineering is important since it puts the devices in context in their application areas in the current trend of technology. Also, it helps to understand what techniques have and have not worked in the similar application areas, in the past. Therefore the appropriate experimental techniques are chosen considering the applications.

The experimental methods are necessary in the design of the device, since they determine the set-up, hence what input output metric is to be/can be measured from the fabricated prototype. The design process is carried out in light of the experimental techniques.

First, different techniques to excite surface plasmons are described. This is followed by different uses of surface plasmons and the applications of the devices considered in this thesis to these areas.

Then different methods to generate/detect ultrasound on a surface are described briefly. The main focus remains on the laser generation/detection of ultrasound, more specifically SAW. This is followed by the definition and applications of picosecond ultrasonics in the generation/detection of ultrahigh frequency SAW.

3.2 Optical excitation of surface plasmons

The k-vector of light in glass can be matched to that of SPs hence enabling the excitation of SPs through glass which has a refractive index greater than that of air, thus increasing the k-vector hence the momentum of the incoming photons. The easiest way to excite surface plasmons is through a configuration known as "the Kretschmann Configuration" [90]. This method uses a technique that is called Attenuated Total Internal Reflection (ATR). Another method called the "Otto Configuration" [90] can also be used, but due to the nature of this setup it becomes practically more difficult to observe the plasmons directly. Otto Configuration is shown in Figure 3.2. Another practical way can be achieved if a high NA (Numerical Aperture) objective is available. There are other methods such as generation via the coupling of internally reflected light to SPs in a metal coated fibre optic, which are out of the scope of this thesis.

3.2.1 Kretschmann configuration

Kretschmann Configuration involves light entering into a prism whose backside has been coated with a layer of suitable metal of a particular thickness, eg gold. Part of the incoming wave penetrates the metal. The incoming EM wave has to be plane-polarised (p-polarised) for free electrons in the metal to perfectly couple to the incoming wave, exciting surface plasmons. As the angle of incidence is varied, at a certain angle which is referred as the "resonance angle", the projection of the wave vector that is parallel to the interface matches that of the surface plasmon at a particular frequency of the incoming field. The free electrons of the metal start vibrating resonantly with the frequency of the light that has penetrated into the metal. Light energy is thus transferred into the SP energy. The reverse process also takes place and the SP radiates back into prism [108]. This is due to the leakage radiation, which is a characteristic of surface plasmons on thin films. SPs do

Kretschmann Configuration
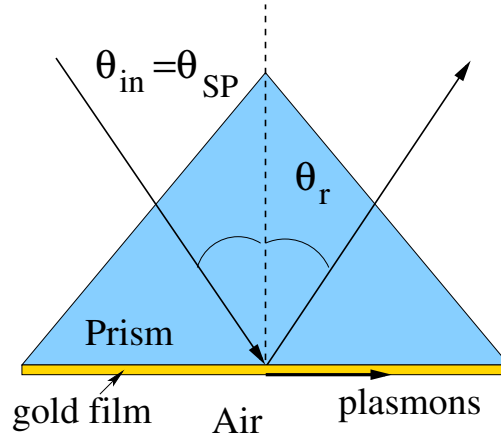
Figure 3.1: Kretschmann Configuration. A minimum in the total internal reflection is observed when the incoming energy is coupled onto the flat metal film. This is referred to as "attenuated total reflection(ATR)"

not radiate into air when flat surfaces are considered. At the resonance angle, the phase difference between the light that is reflected directly off the metal surface and the light re-radiate by the plasmons reaches 180 degrees making them completely out of phase. The two waves being almost equal in amplitude, destructively interfere to give a sharp dip in the reflected light intensity. This particular thickness is approximately 48nm for gold at $\lambda = 633nm$. A common SPR dip for glass-gold-air set up, when 639nm wavelength is considered is around 44 deg. For smooth surfaces, the resonance angle is always greater than the critical angle of the prism/outer dielectric interface, and is called the attenuated total reflection (ATR) angle [90]. The k-vector necessary to excite plasmons will only transfer through the metal film, with the evanescent field observed when there is total internal reflection taking place. Figure 3.1 shows a schematic sketch of observing the ATR using Kretschmann configuration.

3.2.2 Otto configuration

In the Otto configuration, a prism is located very close to the surface of a medium with a negative dielectric constant. As the angle of incidence was increased, beyond the critical angle, when there is total internal reflection, the reflected light intensity was observed.
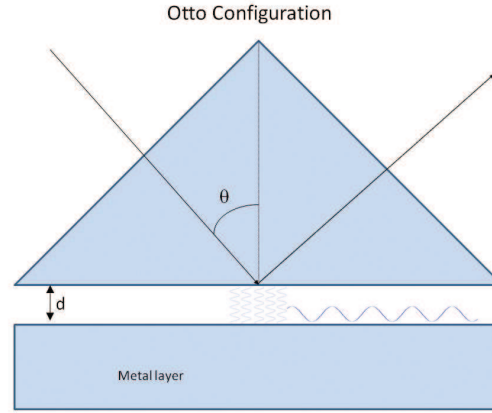
Figure 3.2: Otto Configuration for the excitation of Surface Plasmons. The incidence angle is varied till the k-vector of the evanescent field is equal to the plasmon k-vector. When the k-vector matching takes place, the energy is transferred to the plasmons and a minimum in the total reflected light takes place.

While TIR is taking place, evanescent waves with larger wave-vectors pass through the medium separating the glass prism and the negative-dielectric-constant layer, namely a metal at those frequencies, exciting surface plasmons at its surface. The excitation of the plasmon appears as an attenuation of the total internal reflection in the prism. However, since the separation d in Figure3.2 cannot be precisely measured at all times, and since the plasmons cannot be observed directly, this method frequently is superseded by the Kretschmann configuration.

3.2.3 Grating couplers

Surface plasmons can also be excited by the use of a grating coupler. A grating coupler is a periodic structure which adds the necessary k-vector to the incoming photons to build up the necessary plasmon k-vector.

$$k_{sp} = nk_g + k_{in} \tag{3.1}$$

$$k_{sp} = n\frac{2\pi}{\lambda_g} + \frac{2\pi n}{\lambda_o} \tag{3.2}$$

Where n is an integer, k_{sp} is the plasmon k-vector, k_g is the grating k-vector, λ_o free space optical wavelength, λ_g is the grating period. By adjusting and using such gratings,

plasmons can be launched from many media [109, 110, 111, 112], thereby not limiting their excitation to certain materials, wavelengths and incidence angles.

3.2.4 Coupling with high NA objective

Surface plasmons can also be excited on a metal film deposited on a glass substrate (cover slip), by the use of an oil-immersion high NA objective [113].

The technique involves the use of an objective to illuminate and to subsequently collect information about the reflected spectrum. The range of angles that can be input into the objective from either side depends on the Numerical Aperture(NA). Numerical Aperture of a system is defined as the range of angles over which the system can accept or emit light. (Angles are measured from the optical axis). Therefore, the bigger the NA, the more diverse angles can be focused on the sample, and more diverse scattering can be collected.

$$NA = n\sin(\theta) \tag{3.3}$$

Where, θ: the angle between the optical axis and the maximum incidence angle

n: refractive index of the ambient medium

In this form of SP generation, a rotating diffuser can be used to create a plane of pseudo-point sources, which is imaged on to the back focal plane of the objective. Point sources on the back focal plane correspond to plane waves travelling at various angles at the focal plane, where the sample is located. The incidence angles of these plane waves depend on the location of the point sources. Points further away from the optical axis yield increased incidence angles. Figure 3.3 explains this situation. The reflected parallel plane waves, moving away from the focal plane to the objective, are subsequently focused onto points at the back focal plane, which is imaged back onto a camera (refer to Figure 3.3.

In order to observe[1] the excitation of plasmons, it is necessary to observe the back focal plane of the objective. Figure 3.4 shows the back focal plane of the objective when a coverslip is coated with 50nm of gold. In this figure the plasmon dip is clearly seen on

---

[1]on the reflected side in the farfield
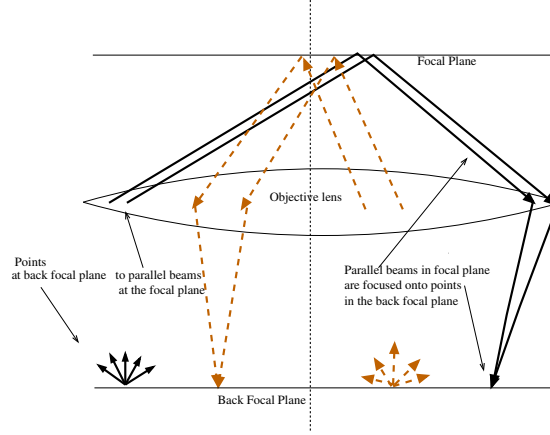
Figure 3.3: Explanation of the operating principles of the objective-based coupling to SPs. The point sources on the back focal plane correspond to parallel plane waves travelling at the conjugate focal plane. The location of the point source dictates the incidence angle and the polarisation of the plane waves at the focal plane. Different locations on the back focal plane illumination corresponds to different incidence angles in the focal plane. Conversely, the parallel plane waves reflected off the sample at the focal plane are focused on to points at the back focal plane. Different reflection angles will correspond to different points on the back focal plane.

both sides of the optical axis. The radial direction in figure 3.4 corresponds to different incidence angles. There seems to be no dip at the radial positions corresponding to SP excitation, at angular positions 90° and 270° (towards the top and bottom of the figure). This is because these locations correspond to s-polarisation, which is normal to the plane of incidence. SP cannot be excited using s-polarisation in this geometry.

This technique is used and improved in section 4.4.5 to verify the operation of the plasmon resonator, where a detailed explanation of a system which can generate/detect SPs is also given.

For the plasmons excited this way, the wavelength of plasmons can be expressed as in equation 3.5.

$$k_p = k_{in}\sin(\theta_p) \tag{3.4}$$

$$\lambda_p = \frac{\lambda_{in}}{n_1\sin(\theta_p)} \tag{3.5}$$
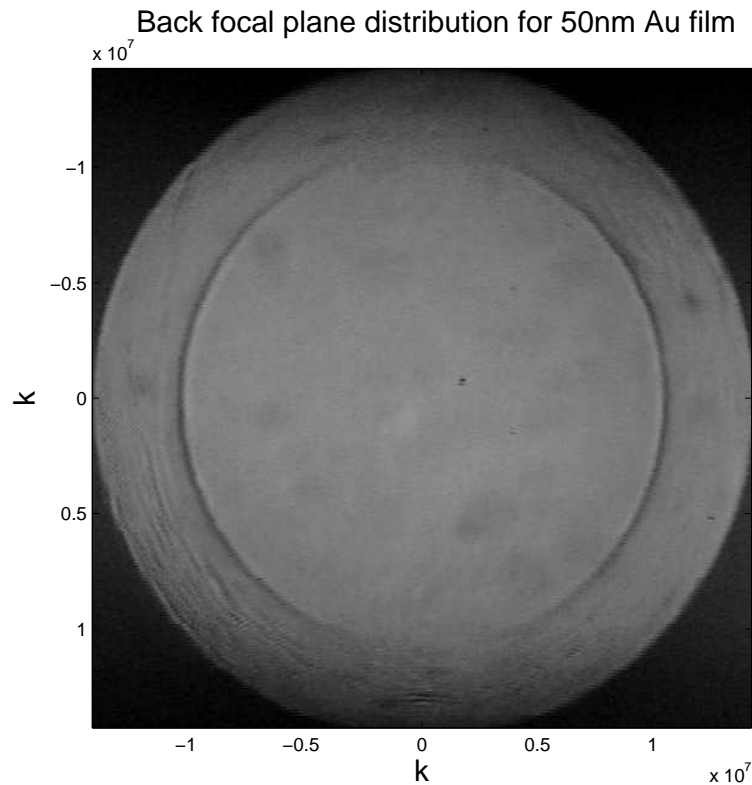
Figure 3.4: Image of the Back Focal Plane of the high NA objective when it is used to couple light to SPs. The sample is a flat gold film the incoming wavelength is 639nm. The arc shaped lines are due to the excitation of plasmons, plasmon dips. It can be seen that the top and the bottom part of the back focal plane does not include a plasmon dip. This is due to the fact that those parts of the plane correspond to s-polarisation, i.e the polarisation is normal to the plane of incidence. Surface plasmons cannot be excited in this configuration using s-polarisation.

3.2.5 Applications of surface plasmons: plasmonics and manipulation of SPs

As the dimensions of the electronics are being reduced, and the speed of data processing is increased, problems relating to high losses of small electronic wires arise. Importance of plasmonics in this aspect has been viewed before (see section 1.3.1).

Plasmons are thought to bring the high frequencies of optics into subwavelength dimensions of electronics. However, before this can be done, many surface plasmon properties have to be investigated and devices analogous to larger optical devices has to be analysed.

3.2.5.1 SP waveguides and potential circuits

It has been shown that surface plasmons can be squeezed passed the diffraction limit [114, 115, 116]. The fact that plasmons can squeeze light beyond diffracted limited dimensions opened up the way to consider them as basic transport mechanisms in very small distances (due to the inherent ohmic losses). Surface plasmon waveguides were demonstrated such as metal-dielectric-metal (MDM) [115, 117, 116], metal-gap-metal(SPGW) [118, 2], arrays of nanoparticles [110, 119], and metal stripes [114, 120]. SPs upon propagating around curvatures was found to be less lossy for sharper curvatures [60].

Plasmon circuits were considered [2] and plasmonic devices analogous to the larger optical ones were to be designed such as SP mirrors, beam splitters [121] and SP lens [122].

SP propagation is observed either by scattered light spectroscopy [123](for propagation on designed features), by near field techniques such as photon scanning tunnelling microscope [120], fluorescence imaging [121], or by leakage radiation [124] inherent to SPs on thin films. SP propagation on metal films, stripes etc. are noted to be lossy mechanisms, which would not go beyond a tens of micrometers [125].

3.2.5.2 SP resonators

In order to make surface plasmon resonators, the reflection mechanism of SPs are very important. SP reflection and transmission from step discontinuities and bars [126, 59, 127, 128] were observed and reflection efficiencies of several structures were addressed in a search for efficient SP mirrors. Also, SP interference resulting from reflectors such as metallic or dielectric obstacles [127, 59, 121] were observed. Surface plasmon reflections from edges and its Fabry-Perot like behaviour in valley-like gaps were shown numerically [59] and the interference pattern was found to depend on $\lambda_{plasmon}/2$. Surface plasmon generation, counter propagation and scattering from superimposed gratings were theoretically analysed by Rochon [129] and standing waves on the gratings were observed.

Additionally, the resonating property of plasmon supporting wires have been analysed both numerically [130, 127] and experimentally using near field techniques [131, 128] and scattered field spectroscopy [123]. The edge reflectivities of such wires were also considered

[123]. Pairs of strips carved from a metal surface via Focused Ion Beam etching were shown to resonate [132], by observation of the standing wave patterns (also modelled using FDTD). Krenn et. al showed via near field optical microscopy that resonators could be obtained by tilting a periodic arrangement of silver wires along the normal to the direction of SP propagation [128]. Finite size waveguides have been shown as plasmon resonators [116], which were adjusted to accommodate deep UV wavelengths (around 10nm). Therefore interference and resonating properties have been considered for cavities in the near field. The Fabry-Perot like properties give rise to an interferometric farfield behaviour of the specular and the retroreflections (see section 4.4).

3.2.5.3 Studies on extending SP propagation length

Most of the plasmonic applications are hindered due to the high loss associated with the plasmon propagation [90].

Different geometries were investigated to excite coupled plasmon modes on very thin metal films, called Long range surface plasmons [133], which could achieve longer propagation distances.

SP behaviour in the presence of gain media has been investigated [9, 8] numerically/theoretically, in an effort to compensate for the loss and provide longer SP propagation distances. Stimulated emission of surface plasmons has been analysed theoretically [134], shown in optically pumped dye solutions [135].

Stimulated emission of SPs have been studied conventionally by measuring the change in reflectance of the analysed structures [135], which offers limited evidence in SP enhancement [136, 135] due to the superposition of the signal onto the directional emission of fluorescent light. In chapter 4, the use of a resonator is proposed, which can be adjusted to cause retroreflection due to reflected plasmons. This structure could find use in the observation of the dynamics of stimulated emission of surface plasmons, by observing the retroreflection rather than the reflection (section 4.4).

3.2.6 The application of surface plasmons in sensing

SPR is effectively a resonant transfer of energy from the optical excitation to the SP wave. The fields associated with SPs decay exponentially away from the metal surface into the

dielectric, therefore confining the energy onto the proximity of the metal surface. Therefore the optical changes right next to the metal surface, in the dielectric medium, namely the transducing medium, modify the propagation constant of the surface plasmon, and change the SPR condition. The change in the SPR condition results in change in the optimum incidence angle or wavelength for the SPR excitation. Consequently, variations in the optical properties of the dielectric medium are observed by interrogating the properties of plasmons. Development of SPR sensors date back to late seventies [125].

Surface Plasmon Resonance offers a new generation in the label-free real-time analysis of bio-molecules, providing information on kinetic processes, concentration and real-time molecule detection [137, 137, 138, 139, 140, 141, 137]. Label-free detection means that the analyte does not need to be labelled by fluorescent molecules. A label-free real-time detection technique offers advantages over an alternative technique, such as fluorescence-based sensors. These methods also provide high sensitivity, however they require either multi-step detection protocols, which limits the real-time measuring ability, or suffer from cross-sensitivity to non-target analyte [137] which reduces accuracy. The first commercial use of SPR in biosensing was in Biacore International AB in 1990 [137]. Biacore [142] is one of the first SPR instruments and market leaders, especially for bigger pharmaceutical companies, with their SPR instruments. Companies such as Nomadics [143] and Reichert [144] also produce, surface plasmon resonance sensors for detecting binding information etc. in real time.

3.2.6.1 Research for application of SPR sensors out of laboratory

When sensing biomolecular interactions using SPR, the sensing area is the area covered by the SP fields. Usually sensing is done by immobilising the analyte on the sensor surface, yielding a thin film of surface bound analyte, which changes the refractive index of the local medium due to the change in concentration. If the binding specific changes happen in a thin film confined to a smaller area than the range of the SPR evanescent field, refractive index changes in the remaining area also affects the sensor behaviour, since it is also probed by the SP field. Such refractive index change could include reduced concantration of the analyte, temperature changes in the ambient medium etc. These effects may shift the resonance in such a way that it may be confused with the binding specific changes. This

issue affects the in field use of the SPR sensors.

A great deal of research has been carried out for the efficient application of SPR sensors for out of the laboratory use [137, 145]. However, in field use requires resistance to the effects of the ambient temperature changes. General approach to work around these issues is to use a reference channel with which the measurements are compared [137, 146, 147]. A way of getting an extra reference channel is the addition of a thin dielectric layer in the illuminated spot, within the sensing area, in order to move the resonance to longer wavelengths in that part of the sensing area [147], yielding a second dip (at longer wavelength) in the spectrum which is not sensitive to the refractive index changes happening due to the analyte.

Another common problem associated with the SPR sensors is calibration [148] between repeated measurements. When the optics move slightly and the angles are shifted one way or another, the collection of reference data is necessary for the calibration of the sensor. In some applications this may be difficult, such as out of laboratory situations, and when microfluidic arrays are involved [148] (since it is difficult to move the analyte out of the sensing area) . This is also a problem when the absolute refractive index is to be measured in environments where temperature changes are inherent, such as food industry and process control [149] or aqueous environments [145].

A sensor arrangement, where the location of the varying plasmon dip could be obtained with respect to a reference point already on the measured signal space, which would be insensitive to all the refractive index changes, would be very helpful in these applications. In [145], this problem was addressed by introducing a reference channel to the fibre optic tip, where the measurements were taken. The concept was similar to that in [147]. This caused the addition of a refractive index insensitive and temperature sensitive SPR dip onto the spectrum, while keeping the refractive index sensitive SPR dip.

In section 4.3, a multilayer structure is suggested that has a refractive index insensitive dip together with a refractive index sensitive dip, which could offer a solution to such problems. The suggested sensor has less sensitivity than the regular SPR sensors, but this may be a price worth paying for the self calibrated nature of the suggested sensor. The structure could be readily adapted for gas sensing which requires a sensing region for refractive index values smaller than 1.1 [148].

3.3 Generation and detection of SAW on surfaces

In NDE of surfaces, transducers are commonly used in two modes, namely the pulse-echo, and pitch-catch techniques. The difference between these two techniques is which signal is detected. If the reflected signal is detected by using the same transducer as the generation, it is called pulse-echo. If it is detected with another transducer, the method is pitch-catch method.

When the surface waves are generated/detected using optical means, i.e lasers, the transducers are called optical transducers. There have been many devices that has been shown in the literature, some of which use purely one method [21, 22, 24], or a combination of methods [150].

3.3.1 Optical generation of ultrasound

There are two physical processes governing the generation of ultrasound with lasers. First is the thermoelastic generation. Thermoelastic generation happens when a laser is incident on the material, and the power density is such that no ablation takes place on the sample [14]. The illumination then causes a local rise of temperature, expansion and resultant stress, momentum conversion, which propagates away from the area. The equations governing the generation will not be discussed in this report. The conceptual explanation of this thermoelastic process is shown in Figure 3.5

The second process in optical ultrasound generation is the ablative regime. In this regime, the optical power density is large and the irradiated area cannot remove the generated heat sufficiently quickly. The instantaneous heat build up causes ablation, therefore evaporation of some material in the inspected volume. This evaporation is transmitted due to momentum conversion to the sample, causing ultrasound, namely longitudinal and surface waves, in the sample. The conceptual explanation can be achieved with the help of Figure 3.6. Note that, even though there are simultaneous thermoelastic forces, these are relatively small [14].

Several techniques have been tried in an effort to generate Rayleigh waves on the surface. The simplest technique is to just focus the light on a line [151] or point, but this is a common reason for damage, which is against the very nature of NDE [14]. If only a

Thermoelastic regime in optical ultrasound generation

Laser irradiation

Thermal expansion

Induced Stress due to
thermal expansion
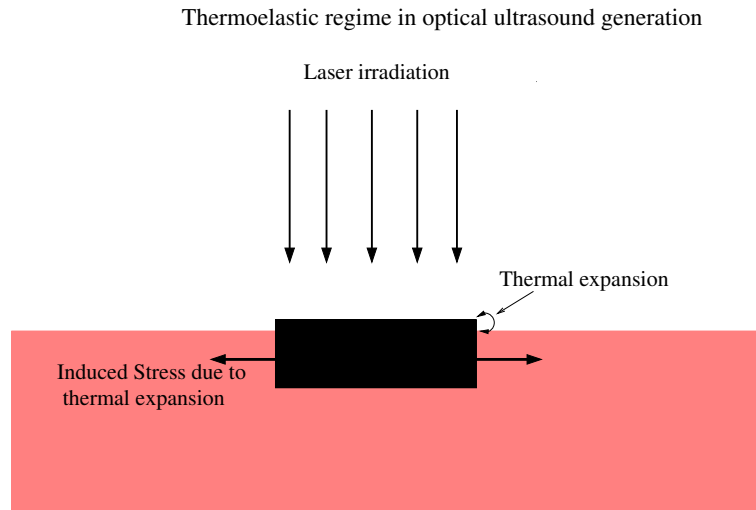
Figure 3.5: The thermoelastic way of generating ultrasound with lasers. Laser irradiates an area, which subsequently goes through thermal expansion, inducing stress in the material

single area is considered, the frequency of the ultrasound generated is related to the pulse duration of the laser being employed. As the pulse duration gets longer, the generation efficiency becomes lower in metals (due to the dissipation of the heat), but the generation frequency decreases as well. In any case, the ultrasonic waves generated are short duration and large-bandwidth, upper bandwith depending on the laser. The signal to noise ratio can be increased by moving to narrow-band regime.

The key point in producing Rayleigh waves of desired frequency, i.e narrowband surface waves, is similar to that in SAW device described in section 1.4.2. It is necessary to create periodic disturbances, in order to generate narrowband Rayleigh waves. This could be thought of as periodically cascading figure 3.5. On metals, there needs to be a periodic contrast in the absorption of optical energy, amount of which will dictate the amplitude of the Rayleigh wave excited. In order to achieve such a contrast, and be able to give directivity to the generated SAW, different techniques have been developed such as imaging a mask to the surface of the metal or shining light through a mask [152], using interferometric techniques to illuminate the sample surface by fringe patterns [153], using lenticular arrays [154], employing diffraction gratings [155] to illuminate the sample surface with different orders of diffraction, using computer generated holograms [156], or using spatial light modulators.

Ablative regime in optical ultrasound generation
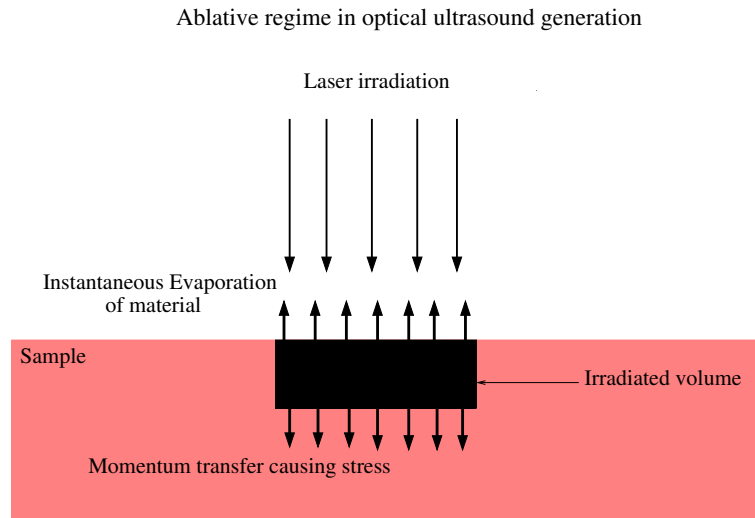
Figure 3.6: Ultrasound generation using the ablative method. The heating of the material causes instant evaporation. The rapid loss of mass is transferred to the bulk substrate by momentum conversion and stress is induced.

Another approach which is of particular importance in this thesis, involves writing the disturbance pattern on a nonabsorbing surface, with an absorbing metal, or vice versa and illuminating with a laser pulse [157, 158, 159]. The frequency of the acoustic wave launched this way depends on the deposited line periodicity, rather than the pulse width of the laser [160]. This implies that the frequency of the acoustic wave generated is not limited by the laser technology, but rather limited by the fabrication technology [158].

This technique is referred to in [157] as generating cheap optical transducer, g-CHOT. A conceptual diagram of g-CHOT is shown in Figure 3.7. g-CHOTs are designed by arranging the finger separations by the desired acoustic wavelength, so as to be able to generate at the desired frequency.

3.3.2   Optical detection of ultrasound

Common optical detection regimes are usually grouped into interferometric and non-interferometric. Non-interferometric methods include knife-edge, surface-grating, reflectivity methods [18]. Knife-edge detection uses a beam smaller than the wavelength of the acoustic wave. The displacement due to the acoustic wave causes a change in the beam location, which is detected using a knife-edge or a position sensitive detector. The beam
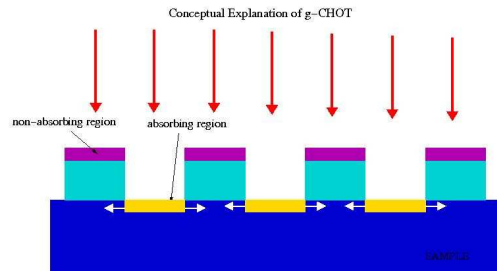
Figure 3.7: Conceptual explanation of the g-CHOT. In this case, non-absorbing lines are drawn on an absorbing sample, and the whole structure is irradiated by a pulsed laser. The difference in the absorption will induce periodic stress on the sample causing the generation of the Rayleigh wave. The use of a pulsed laser is important, since a uniform laser would result in an equilibrium throughout the generation area, resulting in no generation.

will move at the rate the ultrasonic wave perturbs the illuminated area. Surface roughness of the sample will reduce the light going in the direction of specular reflection due to scattering, hence it will reduce the sensitivity of this technique. Consequently, better results are obtained for samples whose surfaces are polished.

Interferometric systems can have high sensitivity, high bandwidth, and can work on rough surfaces and moving components. However, it is not possible to find an interferometer which will gather all these features under one name. Therefore, the optimal type of detection is tightly problem bound. Interferometric techniques include optical heterodyning, differential interferometry, velocity or time-delay interferometry [14, 18]. Details of these methods can be found in [14, 18]. A common interferometer used for the detection of ultrasound is the Michelson interferometer shown in Figure 3.8. In practice, simple interferometry is hard to use and very sensitive, and it will have added complexity in order to deal with problems such as strong sample vibration, air currents and laser pointing instability [161]. The knife-edge and interferometric techniques both suffer from vibration effects and are of similar sensitivities.

A mechanism called d-CHOT has been suggested in [157] which will require no interferometric apparatus. Instead of the interferometric system, metal fingers of height $\lambda_o/8$ (λ_o is the optical wavelength) are printed on the sample and the path difference resulting from the specular reflections from the different parts are detected as changes in the

Michelson Interfometer Probe
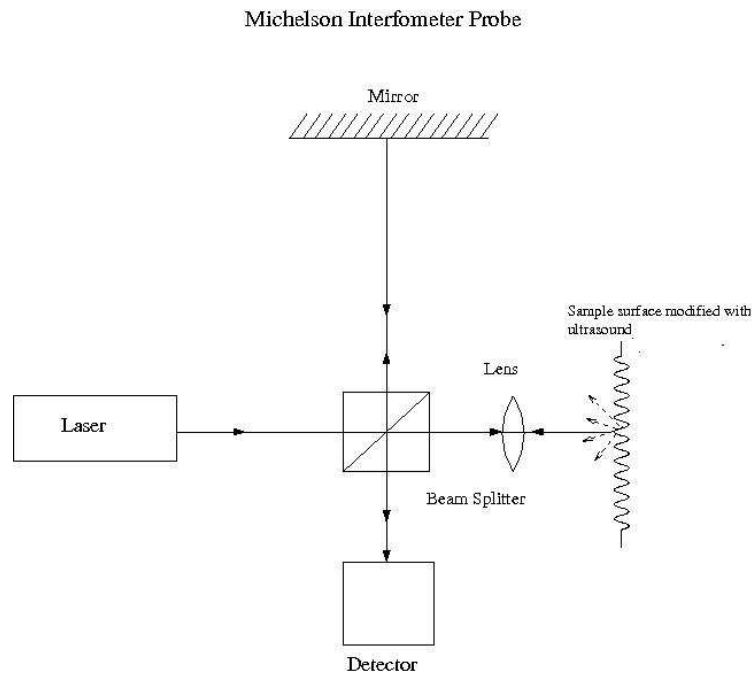
Figure 3.8: The use of Michelson interferometer for the detection of ultrasound.

reflectivity. The operation point, hence the finger height is designed to be located on the part of the resulting sinusoidal interference curve with the steepest slope (which corresponds to path difference of $\lambda_o/4$). The device acts as a reflective diffraction grating, with SAW sensitive diffracted orders. SAW interacts with the device by modulating the fingers, therefore changing the path difference from $\lambda_o/4$. This causes a change in the intensity of the reflection orders, and this change is facilitated by change in energy of the different diffracted orders. Detection can be done by observing any of the diffracted orders. The CHOT is further explained in section 4.5 with appropriate models and simulations. Figure 3.9 explains the operation of the d-CHOT mechanism.

d-CHOT can be conveniently used as a counterpart to the g-CHOT mentioned in the previous part. Together, they form a couple which can generate and detect narrowband ultrasonic waves on metal surfaces. d-CHOT is effectively an in situ common path interferometer, with path difference around $\lambda/4$ which makes it relatively stable to ambient changes, sample vibrations etc [157]. This reduces the need to use external stabilisers, as is the case with some interferometric detection techniques [161]. On the other hand, the knife-edge detectors and interferometric set-ups can be readily used as point detectors,

Conceptual Explanation of g–CHOT
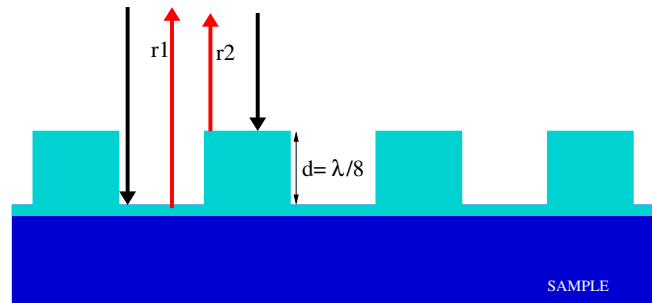
Figure 3.9: A graphical demonstration of the CHOT detection mechanism. The path difference of the reflected light from different parts (r1 and r2) of the metal structure will create an interference effect which is sensitive to the optical path difference, (path difference=$2d = \lambda_o/4$)), hence the finger height d is set to $\lambda_o/8$. An acoustic surface wave will modulate these fingers when it travels through this structure as shown. The modulated fingers will give a rise to a reflectivity difference due to the change in the optical path length and the resultant interferometric effect. This change is facilitated by the shift of the power between the diffracted orders.

while this is not possible with the d-CHOTs.

3.3.3 Ultrahigh frequency SAW

Surface acoustic waves have many applications in microelectronics and NDT&E. Devices based on SAW excitation and detection such as delay lines, resonators, convolvers and filters are widely used elements in microelectronics and telecommunications.

The penetration depth of acoustic surface waves are proportional to the acoustic wavelength, hence for the characterisation of very thin films, or very small defects located less then a micrometer away from the surface, very high frequency (very small wavelength) surface waves are required [162]. Ultrahigh frequency surface acoustic waves generated and detected using picosecond ultrasonics are considered to be very useful in microelectronics and semiconductor industries in the fabrication and testing of components [163]. Spatially distinct generation and detection sites offered by picosecond ultrasonics [163] are very attractive in situations where a surface feature is to be analysed in an approximate region rather than a precise location [164, 165].

3.3.3.1   Picosecond ultrasonics

The the frequency content of the generated ultrasound depends on the pulse duration of the laser. This implies that picosecond lasers could be used to generate very high frequency acoustic waves. This is one of the underlying concepts of a widely used technique called picosecond ultrasonics [14]. In picosecond ultrasonics, a picosecond pulsed laser irradiates the sample with a very short pulse of light, which causes temporal heating, rapid thermal expansion and cooling, resulting in the excitation of broadband acoustic waves [14, 166] in all directions. However, directivity and frequency selectivity can be achieved by using the methods described in 3.3.1.

The acoustic wave is detected using a time delayed "probe" beam at various time delays so as to temporally resolve generated acoustic waves [166]. The early uses of picosecond ultrasonics included launching of broadband acoustic pulses into the bulk of the material, for thin film characterisation [166].

3.3.3.2   Narrowband SAW using picosecond ultrasonics

The concept of spatially modulating the irradiation pattern also applies to ultra-high frequency narrowband SAW generation using picosecond ultrasonics. Ultrahigh frequency surface waves can also be generated optically using different techniques in conjunction with picosecond ultrasonics [160, 159]. However, ultrahigh frequencies, for a given material corresponds to short wavelengths, and for a given optical wavelength the highest frequency of the SAW generated is diffraction limited if conventional methods mentioned above are to be used to create the spatial contrast necessary. A common approach to bypass this limit is to deposit periodic array of thin films of metal onto the sample surface [160, 159] using lithographic methods.

When illuminated, a surface acoustic wave is launched whose wavefront shape depends on the line deposition. The frequency depends on the period of the lines, therefore it is limited by lithographic techniques [158]

3.3.3.3 Laser ultrasonics for ultrahigh frequency SAW

In laser ultrasonics, the improvements in both the lithographic techniques for the generation and interferometric techniques in the detection allow higher frequency SAW to be used in the laboratories.

Generation of SAW at frequencies up to 5GHz using Aluminum stripes [160] on silicon samples has been reported. The detection was done by observing the change in the reflected light due to the acoustooptic effect, i.e the change in the refractive index due to the acoustic effects, in the generation area and a few micrometers away from the generation area [160]. SAW with suboptical wavelengths were generated and detected in [159](upto 23GHz acoustic frequency) using periodic Aluminum stripes on the same patch [159] and on different patches $7.5\mu m$ apart [163], where a sagnac interferometer was used [163].

An issue with dealing with ultrahigh frequency surface acoustic waves is the detection mechanisms are known to be less sensitive [14]. Interferometric methods can be used to detect very low level signals due to acoustooptic effect [160], polarisation sensing [159].

As the frequency of the SAW increases, the detection becomes harder. This is mainly due to the reflected optical field not being able to diffract from gratings designed for very high frequency signals, due to their suboptical wavelength periods. Various interferometric detection techniques are reviewed and their sensitivity concepts are described in [161, 18]. In some cases, harmonics of generating optical wavelength were used to detect in deep UV wavelengths, for acoustic frequencies around 50GHz on sapphire substrates[2], where diffraction from the detection grating is still possible [158].

For lower frequency regimes, where the saw wavelength is such that there are diffracted orders from the detection gratings, devices have been shown to work with simple detection mechanisms [157]. In section 4.5, the design of an optical transducer is discussed, which can be used to generate and detect ultrahigh frequency surface acoustic waves.

3.3.3.4 Numerical modelling in NDT&E

In NDE, generation of ultrasound has been numerically studied thoroughly [167, 168]. It is known that the numerical analysis of laser generated ultrasound in complicated geometries

---

[2]sapphire supports higher SAW velocity 6000m/s than conventional glass 3000m/s

are difficult to analyse [168]. The most common numerical techniques used to analyse this type of situation are BEM and FEM [168]. Laser generation of ultrasound has been modeled using BEM in [167] and FEM [169] where the pulsed laser was modelled as a transient heat source. The propagation of ultrasound as different modes in planar and annular structures have been addressed numerically [170] using FEM. However, for very small features compared to the simulation domain, such as very small cracks, the methods lead to very large problems [168]. These models are thermoelastic models, where heating and displacement are coupled to simulate laser generation of ultrasound and are mostly lower (MHz) frequency range acoustic waves. While the optical generation in MHz range has been modeled numerous times in the literature [167, 168, 169], the ultrahigh frequency acoustic waves has not received much attention so far. In section 4.5, the deformation due to the SAW displacement is coupled into an electromagnetic model to design a transducer which can generate and detect ultrahigh frequency SAW.

3.4 Summary

This chapter has introduced the experimental methods that can be used to generate surface plasmons and ultrasonic waves on surface, while giving background on the application areas of these topics.

Some common surface plasmon generation techniques have been discussed. These were Kretschmann Configuration, Otto configuration, grating coupling and high NA objective coupling. The latter was used in the experiments performed on the fabricated devices in section 4.4.

Then common SAW generation/detection schemes were discussed. Laser ultrasonics were viewed and picosecond ultrasonics were discussed for the generation of high frequency SAW. CHOTs [157] were also discussed, which are used in the generation/detection of the lower frequency (MHz) SAW.

These experimental methods and applications are considered in this chapter are useful while modelling the considered devices in this thesis.

Chapter 4

The Design of Photonic Devices using FEM

4.1 Introduction

As the interest in nano-photonic devices keeps growing; it is becoming crucial to accurately design, optimise and fabricate devices with dimensions of the order of nanometres. Small devices are difficult to fabricate and theoretically analyse. Numerical analysis is important since it can approximate the exact solutions. Fabrication induces some uncertainties which may take the devices away from their operating regions. When the experimental data are obtained from these devices, numerical modelling helps in the explanation of possible discrepancies between the theory and the practice. Especially, when theoretical analysis is impossible.

However, the power of numerical modelling is very dependent on the computational power available. Small devices which involve arbitrary geometries usually result in large problems. As the problem size grows, the computational power needed to solve such problems increase. It becomes impossible to carry out numerical analysis on some devices, due to the lack of computing resources. When designing very small devices with arbitrary geometries, modelling has to be carried out with care.

In this chapter, the implementation of the developed design and optimisation technique is demonstrated (see section 2.1.1) combined with physical insight and inductive reasoning

as an efficient bridge between electromagnetic theory and fabrication of practical nano-photonic/plasmonic devices.

In section 4.2, three software using different modelling techniques (FDTD, FEM and analytical solutions for planar stratified media) are compared on three simple problems. For the implementation of FDTD and FEM, commercial software were used, EM-Explorer and Comsol Multiphysics, respectively. The comparison of different techniques serves as a demonstration of the practicality of the modelling tools common to the literature, on some problems similar to but simpler than the devices considered in following sections. The problems range in difficulty from the modelling aspect and reflect the basic mechanisms for the devices considered in sections 4.3, 4.4 and 4.5. The common properties of these problems are that they all involve resonance(s) and evanescent waves and reflections (more than one in two of the problems). The problems are listed below:

- The first problem considered is a simple SPR case, which is well known in the literature.

- The second problem is a system of planar stratified media, consisting of dielectric and metal films, which incorporate multiple reflections and resonances. This problem is more complicated than the simple SPR problem, since it has many reflected and transmitted components and resonances within its planar stratified geometry, yet it still has an analytical solution.

- The third problem can be classified as a plasmonic grating, which is similar from the modelling aspects to the resonator and the optical transducer later considered. For this problem, EM-Explorer (FDTD) and Comsol (FEM), are verified with each other since there is no simple analytical solution.

In section 4.3, a multi-film scenario is considered to demonstrate and design a novel self-calibrated SPR sensor. A self-calibrated plasmon sensor involves a refractive index(RI) sensitive SPR dip and an RI insensitive constant dip. This adds another dip onto the conventional SPR sensor spectrum and aids the calibration by making a reference point. Therefore between repeated measurements, if any of the optics move, there is no need to recalibrate using reference data. The measurement of the angular separation of the two

dips is necessary to calibrate the sensor.

In section 4.4, a surface plasmon resonator is analysed in detail. This structure involves trapping surface plasmons laterally between two thick metallic reflectors, analogous to a Fabry-Perot cavity. Surface plasmons, consecutively reflecting from such reflectors back and forth forms interference patterns, which depend on the separation between the reflectors. Such structures could be useful in the measurement of surface plasmon propagation length in the presence of gain media and it can find use in lasing applications.

The design process mentioned in 2.1.1 is applied to this problem; the operation of this device and the effect of design parameters on the output parameters are examined in detail. The fabrication of these resonators are described, focusing on the fabrication limitations and their effects on the operation of the devices. Afterwards, the resonating characteristic of the fabricated prototypes is verified experimentally. The similarities and differences between the experimental results and the model predictions are analysed and the effect of the fabrication induced uncertainties on the device operation are explained. Even though a similar model is reported in the literature [59], the key features of this device have not been analysed, such as the observation of the interference patterns by looking at the specular and retroreflected light. Additionally, this device has not been fabricated and experimentally verified before to author's knowledge. this device is fabricated using photolithography and measure the features using AFM. The operation of the resonators are experimentally verified by building and using a system which uses SP excitation via a high NA oil-immersion objective. The agreement of the experimental results with the numerical results are shown, and the effects of fabrication and measurement on the experimental results are explained. Therefore, the connection between the numerical models and the practical issues such as fabrication, measurement and experimental testing are built, by comparing the simulation and the experimental results. Experimental results are seen as the verification of the modelling results, therefore increasing the confidence in the methods. In section 4.5, the design and optimisation of a novel optical ultrasonic transducer, namely eCHOT (Evanescent Cheap Optical Transducer) is demonstrated, using FEM as a tool in the design and optimisation process. This device is an extension of the CHOTs suggested in [157], to subwavelength dimensions and multi-GHz acoustic frequency range. A simplified model, which can be solved using FDTD is constructed initially. It is shown that certain

assumptions have to be made in order to model these structures using FDTD. Then it is shown that these assumptions break when the device has very small dimensions. Even though the geometry and the appearance of this device is very similar to that of CHOTs [157], the operating principle is seen to be fundementally different. A signal enhancement for such transducers is pointed out for a wide range of parameters, at a region where the conventional diffraction theory would predict low signals. This device is designed taking the fabrication limitations into account. These devices are then fabricated using e-beam lithography and images are taken using SEM. Additionally AFM is used to measure these samples. This section demonstrates the application of the design and optimisation technique by successfully analysing a problem which would be impractical to model, design or optimise any other way.

Therefore, the bridge between the theory and the practical applications is established through successfull design and optimisation of three nano-photonic/plasmonic devices, namely a self-calibrated surface plasmon sensor, a plasmon resonator, and a novel optical transducer for ultrahigh frequency surface acoustic waves.

4.1.1 The problem definition and modelling approach

In this part, a brief overview is given on the common issues associated with devices considered in this thesis, from the modelling aspect. Definition of model features and the problem type will assist the determination of the best modelling approach to take, hence the most appropriate modelling tool to use.

The technique to design and optimise nanophotonic devices should be able to deal with devices with arbitrary geometries and different scales, therefore versatility in the tools is very useful. In this thesis, the devices are generally thin metallic devices, which can be assumed to be infinitely repeating along their lengths, separating two semi-infinite media. The devices are illuminated from one side, and related reflection and transmission data are analysed. The first device can be considered as a multi-metal-film stack. The second and the third devices are practically gratings.

Key areas in the applications considered in this thesis are:

- The devices have features or layers which are small with respect to the wavelength.

- The devices are of arbitrary geometry.

- The devices involve evanescent fields. The device/feature size is below the wavelength of light used.

- The devices involve TM waves in monochromatic, i.e time-harmonic situations .

4.2 Comparison of FEM and FDTD software with the analytical solution

This section improves the level of confidence in the modelling tools used in this thesis by verifying and comparing them on simple problems. It also addresses the practicality of the modelling tools by addressing the time required for each modelling tool to solve a given simple problem. The problems considered range in difficulty, and represent the devices demonstrated in the following sections.

In this section, the modelling tools, namely Comsol (FEM), EM-Explorer (FDTD) and the analytical solutions, are compared. EM-Explorer utilised rectangular, orthogonal grids with uniform size. Comsol has different choices for the element shape. The order of the elements (approximation order) is an internal parameter which affects the power of the approximation of the solution in FEM. This order was chosen as quadratic, and it affects the solution times and the solution accuracy. Linear elements was found to yield faster solutions but with insufficient accuracy. Three increasingly difficult and complex problems are considered. The first two have analytical solutions, which are known to be exact. These are planar stratified media involving reflection/transmission, resulting in one total reflected field and one total transmitted field. Background for the analytical solution was given in section 2.5. The last problem is a grating, which may have more than one total reflected and/or transmitted field components.

1. Three methods for a simple case of SPR (geometry shown in figure 4.1) is compared, and the results are shown to agree with each other.

2. A more complicated problem is considered by using a multi-metal-film scenario (the geometry is shown in figure 4.3) and the agreement of the three software is verified.

3. A grating is considered for which there is no simple analytical solution. FEM and FDTD packages are considered for this problem.

In the first two problems, which incorporate only flat interfaces, the reflected farfield spectrum is observed. The near-field transmission is observed for the third problem. Reflectivity information in simple models such as planar stratified media (shown in Figure

4.3) and SPR (model geometry shown in Figure 4.1), could be simply obtained from the near field data (see derivation in Equation B.30).

When there are structures/particles on the metal film, the reflectivity calculation gets more complicated. In these cases, far-field extraction could be used to obtain the reflected spectrum, which was explained in Section 2.8.4. For this problem, the near field transmission is observed as a comparison metric, since it is readily available from the models without the need for any further postprocessing, and it is important to verify the presence of evanescent fields in the models.

4.2.1 Comparison of software on surface plasmon resonance

Surface plasmon resonance is modelled on a flat gold film of thickness 50nm, when the excitation wavelength is 639nm, using all three software in Kretschmann configuration (see figure 3.1). Figure 4.1 shows the model geometry used in Comsol. Figure 4.2 shows the match of EM Explorer, Comsol and analytical results. (The script needed to simulate SPR using EM-Explorer is listed in Appendix A.5.2.)

It can be seen from Figure 4.2 that all three modelling techniques show agreement. It can be seen that Comsol disagrees slightly with the rest, at incidence angles around the critical angle. This is considered as a numerical artifact.

4.2.1.1 Computational performance

For this problem, 2D models of equal sizes were used in both Comsol and EM-Explorer. For all simulations, a PC with Intel duo 2GHz CPU and 8Gb RAM has been used. For a model with 800nm width,$4\mu m$ length 25nm gridsize, EM Explorer took 28 minutes to simulate incidence angles from $40°$ to $47.5°$ with $0.1°$ increments. This corresponds to 76 simulations.

For the same problem in Comsol, when the maximum element size is set to $42.6nm$ which is approximately equal to $\lambda_{n1}/10$ in glass ($n_1 = 1.512$), the solution took around 60 seconds. Analytical solutions were obtained in less than 5 seconds for this problem.

Comsol was able to simulate the range of incidence angles as a parametric simulation, which reduced the simulation times greatly, since meshing had to be done only once. Analytical solutions were almost instantaneous. Therefore, for problems involving planar
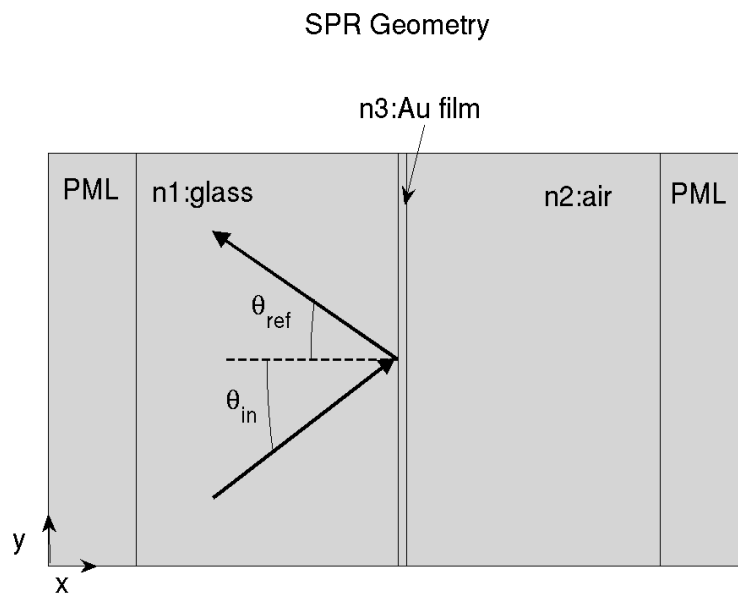
SPR Geometry

Figure 4.1: The general SPR geometry shown on a simple Comsol model. The subdomains from left to right are PML (500nm thickness), glass(n1)($\approx 1.5\mu m$) , gold Au(n3)(50nm thick), air(n2 ($\approx 1.5\mu m$)), PML (500nm). The total simulation size was $\approx 4\mu m \times 0.8\mu m$. The excitation comes from behind the left PML through the glass layer. The reflected field at various incidence angles, θ_{in}, is observed
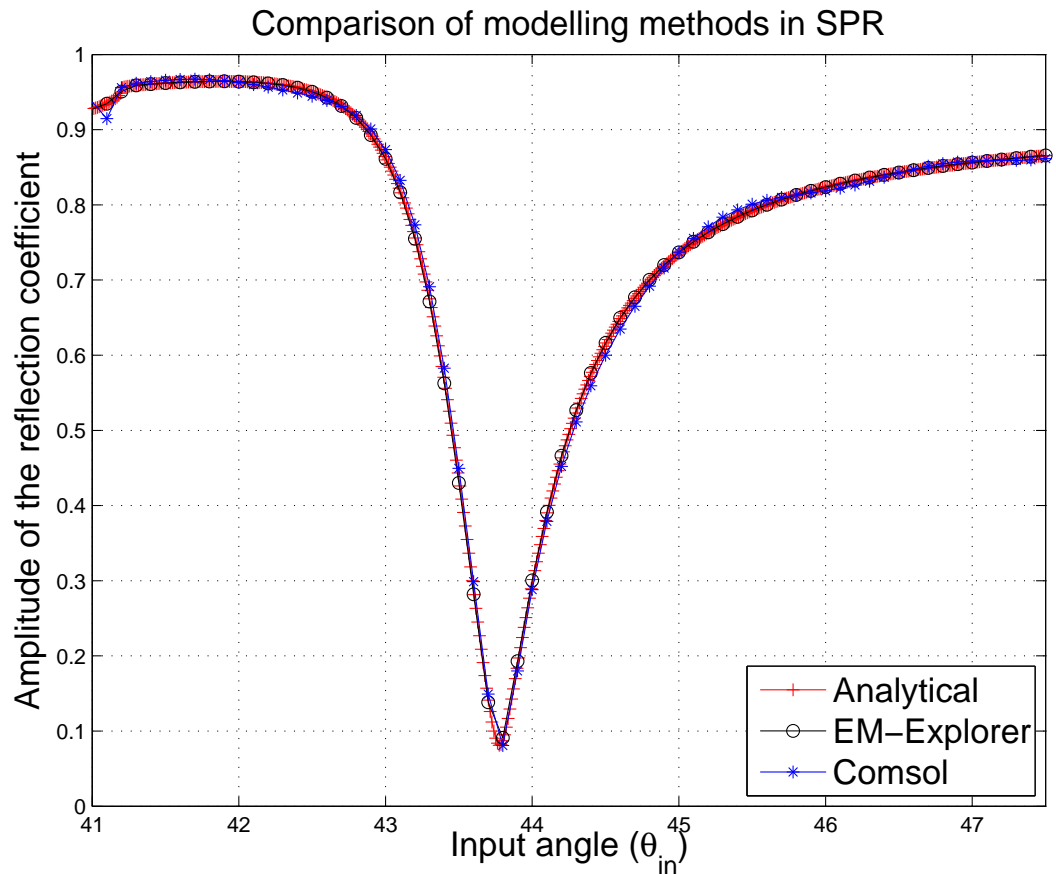
Figure 4.2: Comparison of the analytical results with solutions from EM Explorer and Comsol Multiphysics. The agreement is almost exact with a few numerical errors in Comsol curve around the critical angle. These small numerical errors result when the direction of the transmitted field is perpendicular to the periodic boundaries (along the x-axis in figure 4.1). The geometry and the model details used to obtain this figure is shown in figure 4.1.($n_1 = 1.512$(glass), $n_2 = 1$(air), $n_3 = 0.2 - i3.38$(gold))
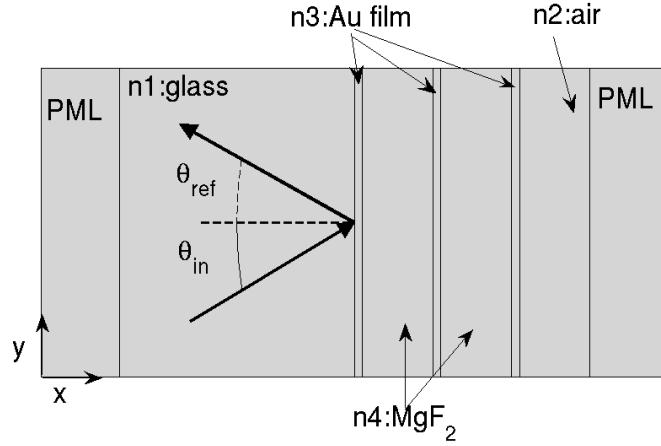
Figure 4.3: Multifilm Geometry used in Comsol. The subdomains from left to right is PML (thickness=500nm), glass (thickness $\approx 1.5\mu m$, $n_1 = 1.512$), Au (50nm thick, $n_3 = 0.2 + i3.38$), MgF_2 (250nm thick, $n_4 \approx 1.38$), Au(50nm thick, $n_3 = 0.2 + i3.38$), MgF_2 (250nm thick, $n_4 \approx 1.38$), Au (50nm thick, $n_3 = 0.2 + i3.38$), air ($\approx 500nm$ thick, $n_2 = 1$), PML (500nm thick). The light is incident from behind the left PML, through the glass layer. The specular reflection is observed in the farfield for various incidence angles θ_{in}. The model size is $\approx 4\mu m \times 0.8\mu m$.

stratified media, the analytical solution must be the method of choice, since it can yield exact solutions much faster than Comsol and EX-Explorer.

4.2.2 Comparison of the software on multi-film geometry (1D-photonic crystal)

In this part, a stratified planar structure is considered as a comparison for the different modelling methods considered. This structure can also be visualised as a 1D photonic crystal, which is an alternative name that is used throughout this section. This structure represents a problem more difficult than the well-known SPR, since it involves multiple reflections and resonances between different layers, yet it still solvable using the analytical solutions. For this problem, EM-Explorer, Comsol and analytical solutions are compared for accuracy and performance.

The geometry of the structure under consideration is shown in figure 4.3. The problem
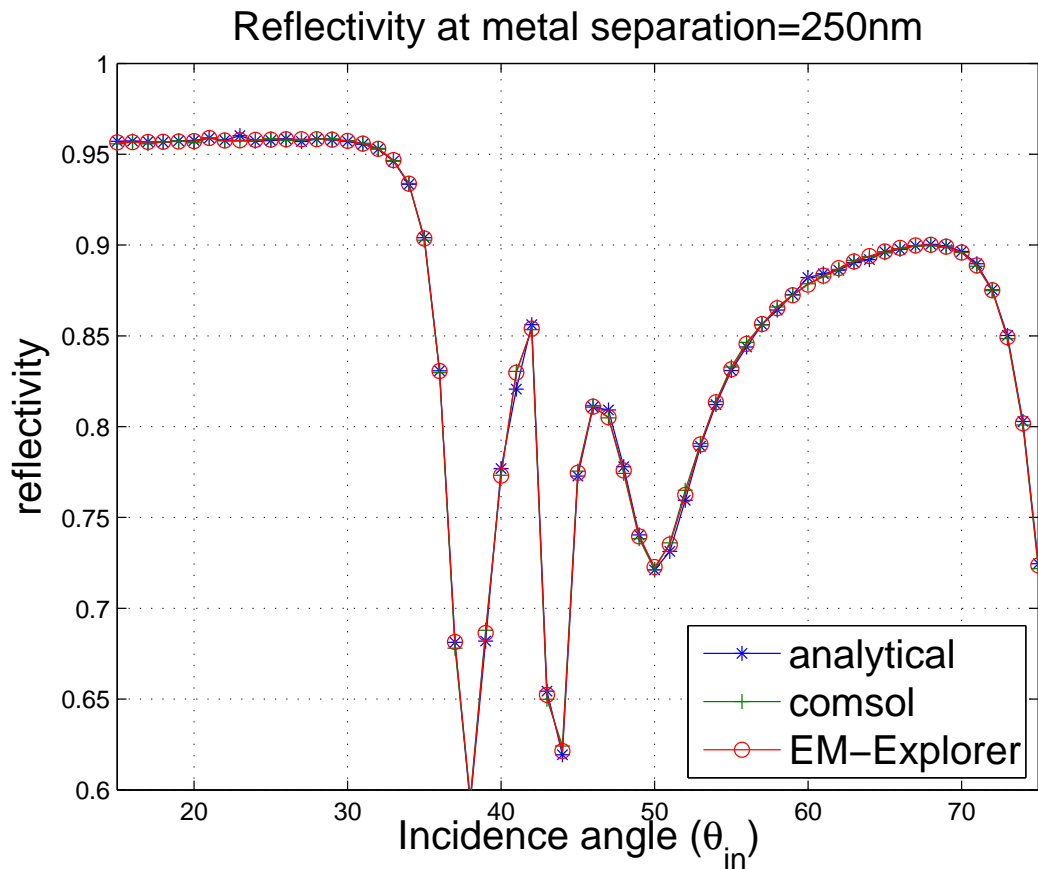
Figure 4.4: The amplitude of the specular reflection coefficient at various incidence angles for the multifilm model. In this model, the metal film separation (or MgF_2 thickness) is 250nm. The results of all the software is seen to agree. The geometry of this model can be seen in Figure 4.3.

involves three gold films, separated by an intermediate layer of refractive index 1.38, which corresponds to either teflon or MgF_2 (n_4 in figure 4.3). The model is periodic, representing an infinite width system. The excitation ($\lambda_{in} = 639nm$) comes from behind the left PML and through the glass layer. The subdomain labeled as n_2 corresponds to air, followed by the associated PML. Figure 4.4 shows the specular reflectivity of such a model for various incidence angles when the gold film thickness is $50nm$ and metal film separation is $250nm$. EM-Explorer, Comsol and the analytical solutions agree.

4.2.2.1 Computational performance

Model sizes used in both Comsol and EM-Explorer were equal for this problem. For a model size $\approx 4\mu m \times 0.8\mu m$ and $25nm$ gridsize, EM Explorer took around 100 minutes to simulate incidence angles from $10°$ to $80°$ with $0.5°$ increments, which corresponds to 140 simulations. This is twice the number of simulations performed previously in SPR problem. The simulation domain size is comparable to that of SPR problem, however more computational time is necessary for bigger structures (thicker intermediate layers). The convergence also depends on the time taken for the resonances to stabilise and the multiple reflections to settle, since the simulation has to be run at many time steps waiting for the steady state to be reached. For a bigger model, this would increase the number of computations, hence the time taken.

For maximum element size $\lambda_{n1}/10$ in glass ($n_1 = 1.512$), the solution of this problem took Comsol around 5 minutes. Analytical solution again were obtained in less than 5 seconds for this problem.

4.2.3 Comparison of Comsol and EM-Explorer on a plasmonic grating

The next comparison problem is a grating which is modelled using both Comsol and EM-Explorer. A grating which supports surface plasmon resonance involves surface plasmon multiple reflections and scattering. This is a more complicated model than the SPR on flat metal. The multiple reflections and the resonant nature imply that the FDTD code has to be run for a longer time, letting them settle. Simulations are run, where the incidence angles are varied and the near field data 100nm above the grating surface is collected. The vertical line was chosen close enough to the grating, so it could capture the evanescent fields but far enough to be not affected by any possible numerical errors around the metal-air boundary. The model geometry for the grating is shown in figure 4.5. It involves a periodic 50nm thick gold obstacle deposited on a 50nm thick gold film, very similar to the resonator model. Here the period of the grating was chosen to be 800nm (along y-axis) and the obstacle size was 400nm (along the y-axis). The model size is $4\mu m \times 0.8\mu m$.

The results for these models are shown in figure 4.6. As can be seen, the results mostly match. However there are some differences as far as colour scaling is concerned.
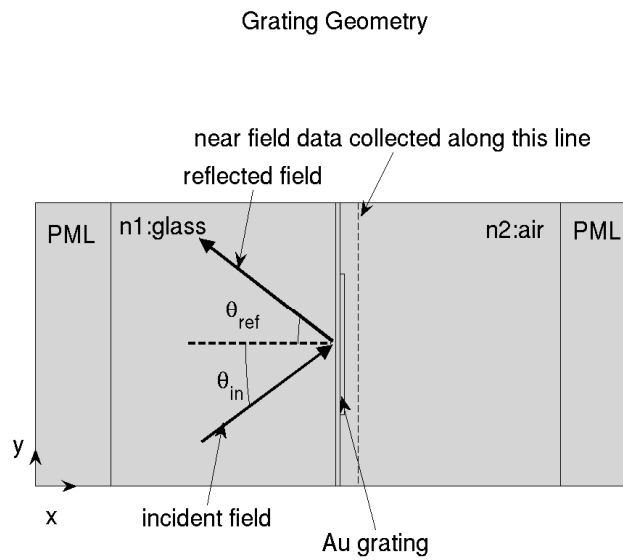
Grating Geometry

Figure 4.5: The geometry of the grating model. The materials are from left to right, PML (500nm thickness), glass ($\approx 1.5\mu m$ thickness, $n_1 = 1.512$) , gold grating (50nm thick film and 50nm thick obstacle, $n_3 = 0.2 + i3.38$) , air($\approx 1.5\mu m$ thickness, $n_2 = 1$) and PML(500nm thickness). Width along the y-axis is the period of the grating, which is 800nm. Field amplitude values for figure 4.6 are collected along the vertical line 100nm above the grating in air. The vertical line was chosen close enough to the grating, so it could capture the evanescent fields but far enough to be not affected by any possible numerical errors around the metal-air boundary. This figure represents a grating that is periodic in the y-direction. Therefore, the grating is constant in the z-direction (into the plane). The model size is $4\mu m \times 0.8\mu m$
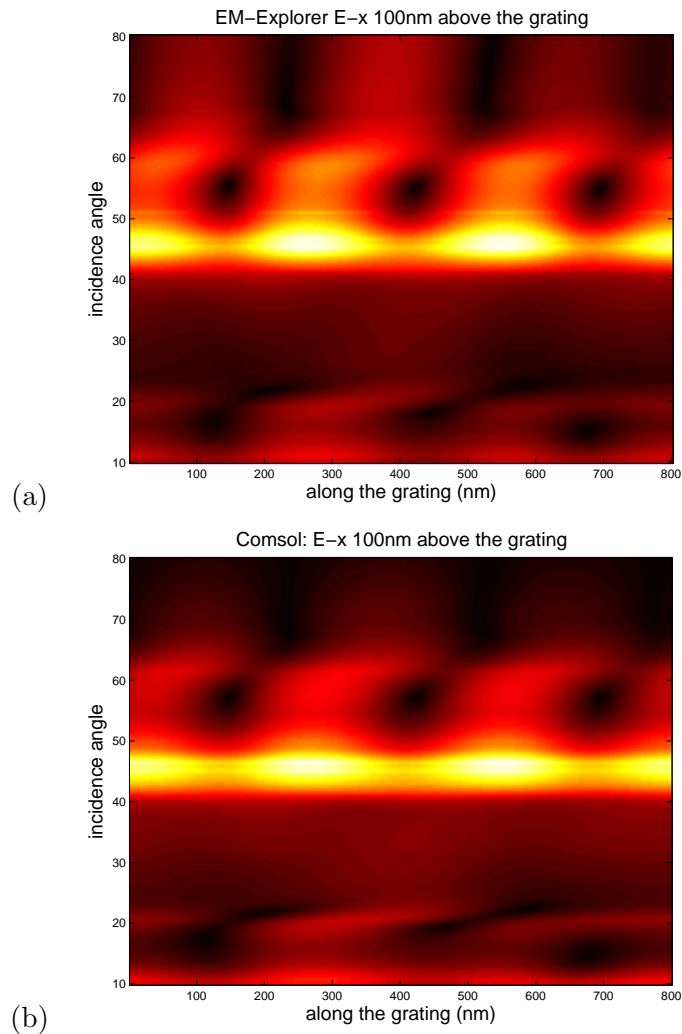
Figure 4.6: Comparison of Comsol and EM-Explorer for the simulation of the grating in figure 4.5. For every incidence angle (y-axis) the x-component of the E-field is extracted from the line 100m above the grating and stored to build these images (the extraction line is shown in figure 4.5, it is along the periodicity of the grating). In both figures a field enhancement is seen due to the SP excitation between $42 - 47°$. This enhancement fades away slowly between $47 - 52°$. **(a)** The EM-Explorer results. **(b)** Comsol results. The results from the two software differ slightly in colorscale at some points. (b) fades away faster than (a) because the EM-Explorer simulation was run for a limited amount of time, which may not have been sufficient for the plasmonic reflections to settle.

These are mainly due to the fact that the model was only run for a limited convergence time. Therefore, even if the simulations were run much longer than Comsol, due to the resonant nature and the multiple reflections, the model could not reach steady state in the FDTD simulation using EM-Explorer. Therefore the small components due to multiple reflections are not accounted for. However, the functional shape of the excitations and maxima correspond to similar angles of incidence, therefore the main characteristics of the figures match.

4.2.3.1 Computational performance

For this problem, 2D models of equal sizes were used in both Comsol and EM-Explorer. For all simulations, a PC with Intel duo 2GHz CPU and 8Gb RAM has been used. For the grating model size is $4\mu m \times 0.8\mu m$, 10nm gridsize, EM Explorer was run for 275 minutes to simulate incidence angles from $10°$ to $80°$ with $0.5°$ increments. FDTD simulations have to be run for a longer time, for the multiple plasmon reflections to settle and the convergence to take place. The combination of the large simulation domain, the increased convergence time and the number of time steps for the steady state, made this simulation very long.

For the same problem Comsol took around 7 minutes when the average element size in the metal grating was 10nm. Therefore Comsol was found to be much faster than EM-Explorer for the simpler cases.

4.2.4 The choice of modelling method

The statements in section 4.1.1 immediately suggest FEM as a suitable candidate for the modelling method of choice. FEM is known to be generally faster than FDTD [28]. EM-Explorer, which utilises FDTD, is found to be slower than Comsol for simulations involving thin films and gratings (see table 4.1). In the range of the problems considered in this thesis, such as the eCHOT structure mentioned in section 4.5, the smallest feature considered could go down to 1-10nm while the domain size can go up to $10\mu m \times 15\mu m$. For such geometries, structural, regular gridding of FDTD (EM-Explorer) would yield very big models, which would be very time consuming to solve.

		SPR	**Multi-film**	**Grating**
EM-Explorer	time	$28min$	$100min$	$275min$
	model size	$4\mu m \times 0.8\mu m$	$4\mu m \times 0.8\mu m$	$4\mu m \times 0.8\mu m$
	grid size	$25nm$	$25nm$	$10nm$
Comsol	time	$\approx 1min$	$\approx 5min$	$\approx 7min$
	model size	$4\mu m \times 0.8\mu m$	$4\mu m \times 0.8\mu m$	$4\mu m \times 0.8\mu m$
	grid size	$25nm$	$25nm$	$10nm$
Analytical		$< 5sec$	$< 5sec$	N/A

Table 4.1: The comparison of the modelling techniques on the simple problems.

For gratings, analytical methods such as RCWA (see section 1.6.1) could also be used. However, RCWA is known for slow convergence for TM waves and it works by representing curved geometries via staircase approximations. Also, it is not as flexible as FEM and would not be appropriate for arbitrary geometries.

Surface plasmons have a rapid exponentially decaying field in one dimension (see section 2.3). Modelling such rapidly varying phenomena involving subwavelength features, in corrugated surfaces requires very fine grids. Especially in corrugated surfaces, since the SP field is maximum on the surface, FDTD fails to characterise the fields efficiently yielding large numerical errors [171], hence taking longer to converge. Convergence of EM-Explorer is controlled via a "convergence monitor"(see section A.5.2) . By selecting appropriate parameters, which are documented for the simple cases considered here [172], convergence issues have been avoided in the implementation of FDTD. FEM's advantage over FDTD stems from this situation.

All the devices considered in this thesis involve resonant effects. The reflections, and field variations in resonant structures take a long time to settle. Time-domain methods such as FDTD requires waiting for these variations to reach steady-state, which may result in very long simulation times. When optimising such structures, this poses a problem, due to the repetitive nature of optimisation. Therefore, Finite Element Method is the most suitable modelling tool for those problems considered in this thesis, which do not have analytical solutions, since it can have mesh with better discretisation, multiresolution and

it is faster for the problems considered in this thesis.

Other modelling methods were considered and ruled out. Method of moments, despite its speed [28] doesn't scale too well with the frequency or size [25]. Besides, it requires the presence of a Green's function. Hence it would be inappropriate to use for these applications. MRTD analysis is not as well known as FEM and FDTD, it is in time-domain [77] and has localisation problems. Therefore, this method would also be inappropriate for this application. TLM is similar to FDTD and would suffer from discretisation problems, and increased simulation times.

4.3   The self-calibrated plasmon sensor

In section 4.2, the agreement of Comsol, EM-Explorer and the analytical solution in the multifilm scenario was verified. Surface plasmons in sensing applications were discussed in section 3.2.6. Surface plasmon resonance sensors offer a new generation in the label-free real-time analysis of bio-molecules, providing information on kinetic processes, concentration and real-time molecule detection [137, 138, 139, 140, 141].

Surface plasmon sensors can have detection mechanisms based on wavelength, amplitude or angular interrogation. In this section the angular interrogation is considered. Angular interrogation looks for the angular position of the SPR dip as the sensor interacts with the analyte. A common problem in surface plasmon based sensors is the calibration of the sensor. This problem limits the use of these sensors in out-of-laboratory environments. Calibration is necessary to differentiate between the dip movements due to the binding of the specific analyte and effects such as non-specific binding and temperature changes. Additionally, calibration [148] between repeated measurements is necessary but difficult. When the optics move slightly between repeated measurements, the angles may shift. Calibration then requires the collection of reference data, and reseting of the plasmon dip, which may be cumbersome in out-of laboratory environments.

Figure 4.7 shows the operating region for a conventional SPR sensor. The angular position (along x-axis) of the dip is seen to change as the refractive index(RI) changes. If there is a move in the angular position of the dip, due to the a slight move in the optics, then there is no way of distinguishing such a change from a analyte-specific dip movement. General approach to work around this issue is to use a reference channel with which the measurements are compared [137, 146, 147] (a more detailed discussion is in 3.2.6). Design and analysis of this structure is carried out using the analytical solutions, since they yield exact solutions faster than FEM. In this section, the application of a multi-film structure as a self-calibrated surface plasmon sensor is demonstrated as a means to solve the calibration problem. In the proposed sensor, there is an additional refractive-index insensitive dip, which could be used as a reference, for the SPR dip calibration.
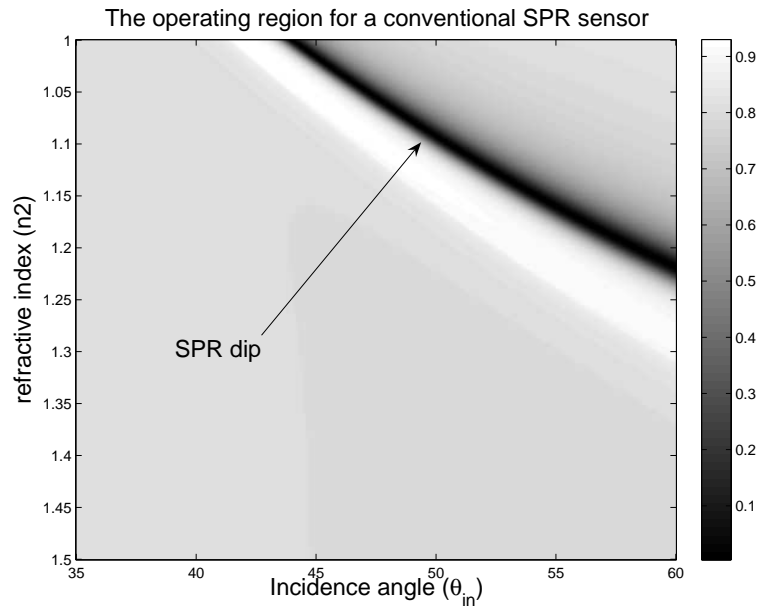
Figure 4.7: The operation region for the SPR sensor. It can be seen that the angular dip position (along x-axis) is sensitive to the refractive index of the analyte (y-axis). When this dip is detected at a different angle, there is no way of differentiating whether this is due to the analyte, or simple change in optics. The model geometry can be seen in figure 4.1. The gold film thickness is 50nm and the optical wavelength is 639nm

4.3.1 The design objective

The self-calibrated surface plasmon sensor is designed so as to result in a refractive index insensitive dip and a RI sensitive dip on the same signal space. The RI insensitive dip should not move if there is a change in the refractive index of the sensing region (see figure 4.8). The desired angular distance between the two dips is small, so as to allow easy detection. The RI sensitive dip needs to be as deep (close to zero) and as thin (small width of the dip) as possible, as this indicates the sensitivity of the refractive index sensitive process in the sensor which is related to change in reflectivity with respect to the angle of incidence, slope of the reflectivity curve. Temperature changes could be accounted for if the two dips moved equally due to the RI changes in the media due to a given temperature change.

4.3.2 The self-calibration mechanism

In this part the operation mechanism of the suggested sensor is analysed. Inductive reasoning is used to lead the modelling and isolate the effects of key features, so as to understand their contribution to the device mechanism.

The operation mechanism and the geometry of the suggested self-calibrated SPR sensor is shown in figure 4.9. There are two dips involved, one of them due to the multiple reflections in the intermediate layer, within the sensor. The other dip is due to the multiple reflections and SP excitations on the metal films. Therefore, the former is not dependent on the analyte changes, but rather on the intermediate layer properties and sensor dimensions.

Figure 4.8 shows the response of the sensor when the design parameters are optimised. It can be seen that when the refractive index of the analyte (n2) is varied, there is a dip with constant angular position and a RI sensitive dip, with different angular locations for each refractive index.

The refractive index insensitive dip is the calibration mechanism, since it serves as a reference that doesn't move with a specific binding that the sensor is designed to detect (here modelled as a change in refractive index n2).

When horizontal lines from figure 4.8 are observed corresponding to different refractive indices (values along y-axis), the movement of the dips (along the x-axis) can be seen more clearly. Figure 4.10(b) shows a cross section through this figure at n2=1, n2=1.01 and n2=1.02. It can be clearly seen that the first dip is constant and broad where as the second dip changes with the refractive index (n2). Figure 4.10(a) shows the same scenario for the traditional (non-calibrated) SPR sensor on a flat gold film (geometry shown in figure 4.1). When this traditional surface plasmon resonance sensor is considered, the changes in refractive index n2 causes a change in the location of the dip, which corresponds to 0.64° change in angular location for a refractive index change of 0.01. The same 0.01 change in refractive index will cause 0.57° change in the angular dip position in the calibrated sensor. Therefore the self calibrated sensor has comparable dip movement for a given change in n2 with the non-calibrated, conventional SPR sensor. On the other hand, both of the dips in the self-calibrated sensor are wide. Additionally , the refractive index sensitive dip is seen to be both wider and higher than the SPR dip in the conventional SPR sensor. This
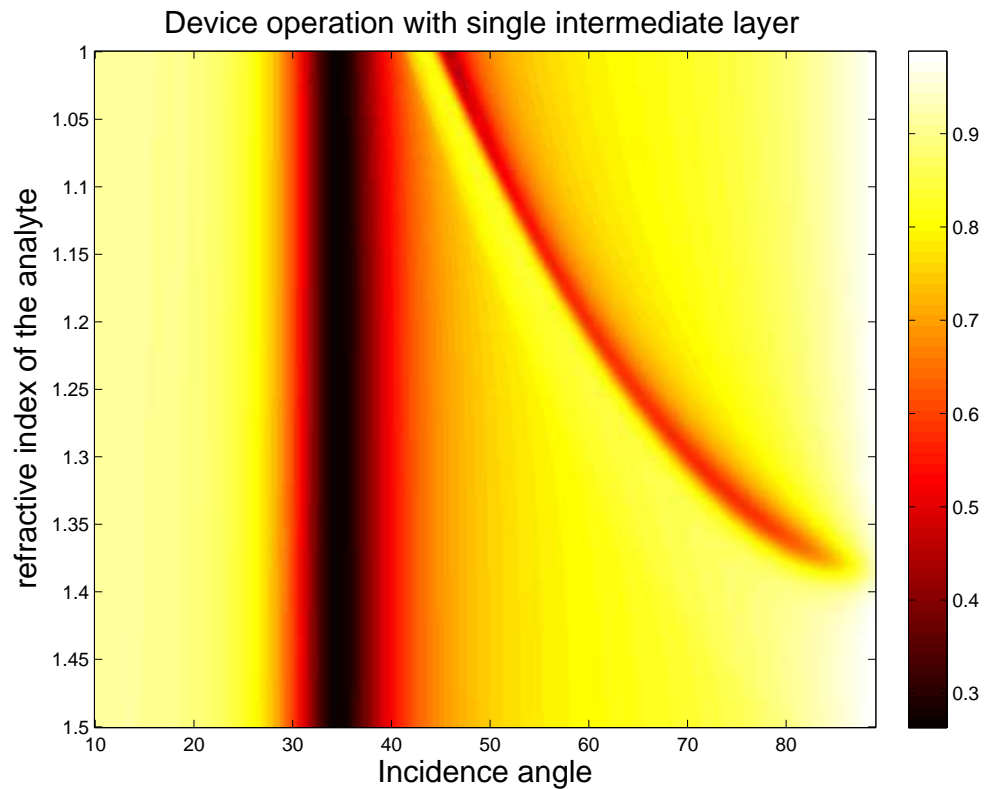
Figure 4.8: The reflectivity at various incidence angles for a simple self-calibrated SP sensor, when the refractive index of air (the subdomain with refractive index n2 in figure 4.9) changes. The thickness of the intermediate (MgF_2) layer is 200nm. Reflected field is observed at various incidence angles(x-axis), at each value of n2(y-axis), which is varied between 1 and 1.5. It can be seen that there is a dip with constant angular location(along x-axis) and a dip that moves along x-axis when the analyte refractive index is varied. This property suggest the use of this structure as a calibrated refractive index sensor.

The self−calibrated SPR sensor
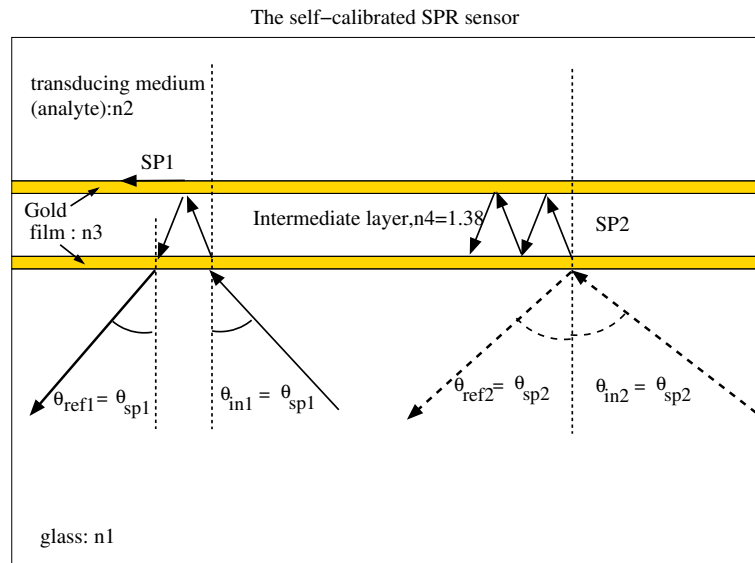
Figure 4.9: The geometry for the self-calibrated plasmon sensor. There are SP related dips contributing to the operation of the sensor in addition to multiple reflections. The first dip is mostly due to the multiple reflections in the intermediate layer. The second dip is due to the SP excitations on MgF_2-gold-sensing layer interface and multiple reflections.

means that the sensitivity of the self-calibrated sensor is less than the conventional SPR sensor. However, the self-calibrated SP sensor has an advantage over the conventional SPR sensor, since it does not need calibration between repeated measurements. It should be noted that the analysis in this section does not consider the affects of the temperature, which is assumed to be constant (The possible effects of temperature are addressed in section 4.3.3.2).

4.3.2.1   Multi-reflections in the intermediate layers

In order to analyse the operating mechanism further, the intermediate layer thickness is reduced to become a very thin film. Figure 4.11(a) shows this situation. As can be seen, there is only one dip in the spectrum, similar to that in figure 4.7. This indicates that when the intermediate layer is made very small, the only dip that remains is the one due to SP excitation on the gold-analyte interface. The SPR dip for a MgF_2-gold-air system corresponds $\theta_{in} = 49°$ (dip starts at around $43°$ at n2=1 due to Snell's law). The SPR dip for a glass-gold-intermediate layer (n4=1.38, which corresponds to RI of MgF_2)
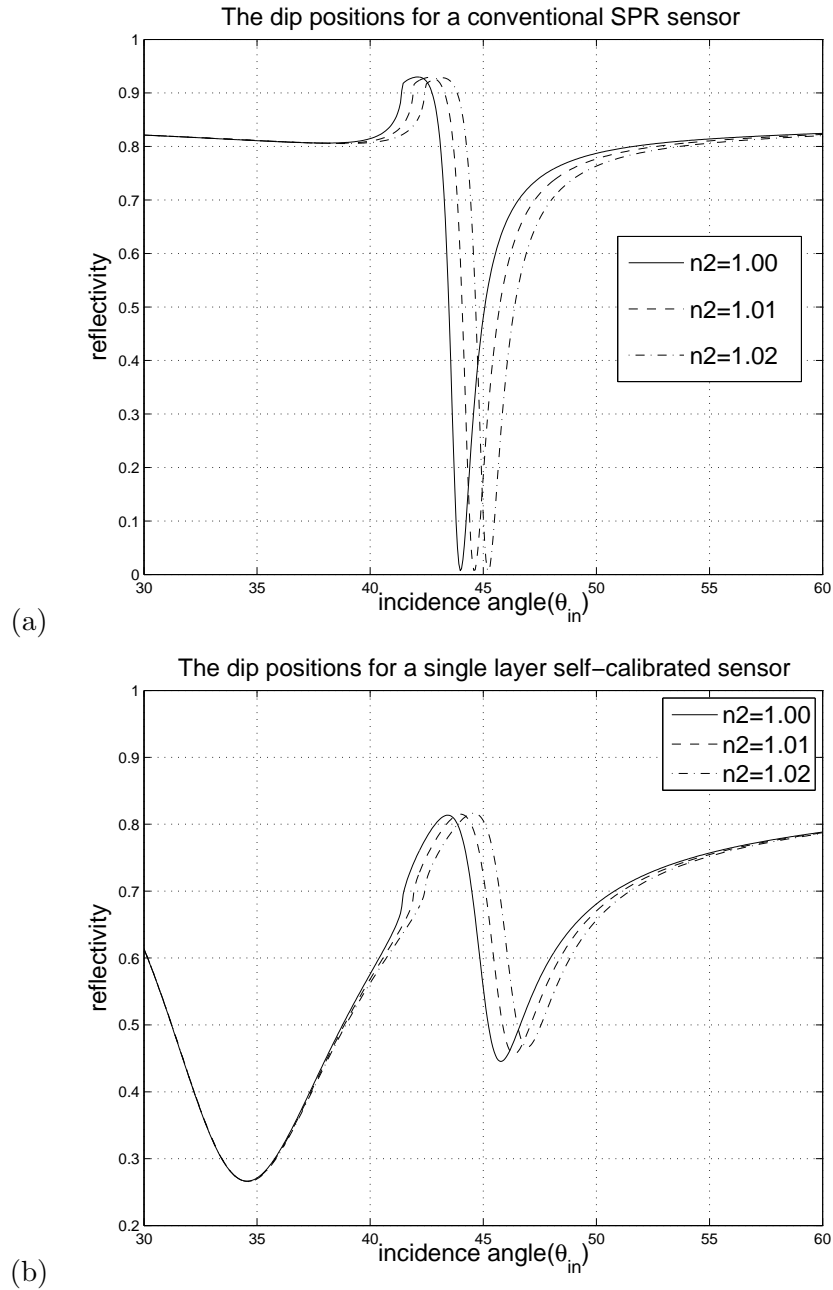
Figure 4.10: **(a)** The demonstration of the movement of the SPR dip in the conventional SPR sensor, as the refractive index changes. The dip movement is 0.64° per 0.01 increase in refractive index. **(b)** The response of the self-calibrated SPR sensor. There is an analyte insensitive and an analyte sensitive dip in the spectrum. The dip movement is calculated to be 0.57° per 0.01 change in analyte refractive index. The dip movement is comparable to the conventional SPR sensor.
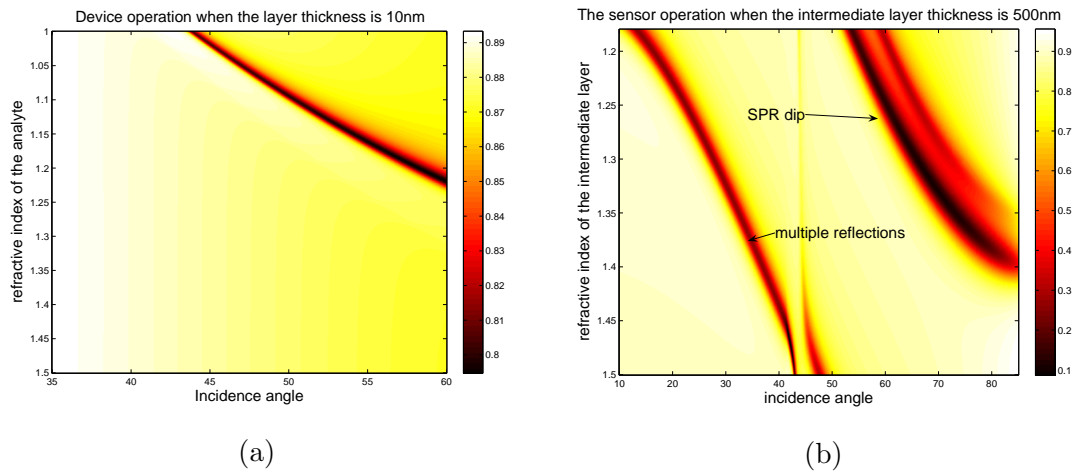
Figure 4.11: **(a)** The operation of the sensor when the thickness of the intermediate layer is 10nm. This completely removes the RI-insensitive dip. The figure resembles figure 4.7. **(b)** The operation of the sensor when the thickness of the intermediate layer is 500nm. In this figure the refractive index of the intermediate layer is varied(along y-axis) so as to demonstrate its SP properties. As can be seen, there are still effects of multireflections etc. since it is not possible to decouple the layers completely.

interface corresponds to $\theta_{in} = 80°$ approximately. This is shown in figure 4.11(b) when the thickness of the intermediate layer is made large enough for the decoupling of the multiple reflections from the SPR dip to partially occur. When the refractive index corresponds to 1.38-($n_{MgF_2} = 1.38$) (y-axis), then there is a dip around 80°. It can be seen that there are still other dips due to interference effects of the multiple reflections. This is because it is not possible to completely decouple the two gold layers.

4.3.3   Design parameters and the optimisation of the sensor

The optimisation technique is applied by varying ambient refractive index and solving the same problem with different incidence angles at different design parameter values. The design parameters for this structure involved the thicknesses and the refractive indices of the layers. A practical range of these parameters are simulated and several images such as the one shown in figure 4.11 are thus obtained.

4.3.3.1 The thickness of the intermediate layer

In order to optimise this device, the effect of the intermediate layer on the device operation has to be found. Then the optimal thickness of the intermediate layer can be determined. The thickness of the intermediate layer was found to affect the operation of the sensor, and the location of the analyte insensitive dip. The analyte insensitive dip is dependent on both the multiple reflections/ interference effects in the intermediate layer and the SP resonance on the glass-gold-MgF_2 system. Therefore the thickness of the layer affects the operation by changing the effective optical path length. Figure 4.12 shows this situation. For low thickness values (figure 4.11(a), figure 4.12 (a)), the angular position of the RI-insensitive dip corresponds to very low angles. This conceptually means that for very low values, the sensor is behaving like a conventional SPR sensor. Figure 4.12(b) and (c) show that as the thickness is further increased, the angular position of the dip approaches and goes past the optimal position shown in figure 4.8, and the two dips start interfering. After this point, this structure cannot be operated as a RI sensor (shown in figure 4.12(d)).

The thickness of the intermediate layers can be adjusted to have an an optimal angular position for the RI-insensitive dips. For practical purposes, the angular locations were chosen to be as close as possible, so as to be able to measure the angular distance between them. The optimal thickness chosen for this sensor was 200nm (see figure 4.8).

4.3.3.2 The refractive index of the intermediate layers

The refractive index of the intermediate layer is important since it dictates which material can be used in the fabrication of this sensor. Besides, difference in temperature, changes the refractive index of both the intermediate layer and the analyte. In conventional non-calibrated sensors, this could be confused with a change in the refractive index due to a binding effect on the sensor surface.

The proposed sensor can detect temperature changes, which could be incorporated into calculations, if the intermediate layer material is chosen as a material with very high low coefficient of refractive index, or the device dimensions were made larger. since one of the dips is only affected by the changes in the intermediate layer and not the analyte.

A change in the refractive index of the intermediate layer due to a change in the tem-
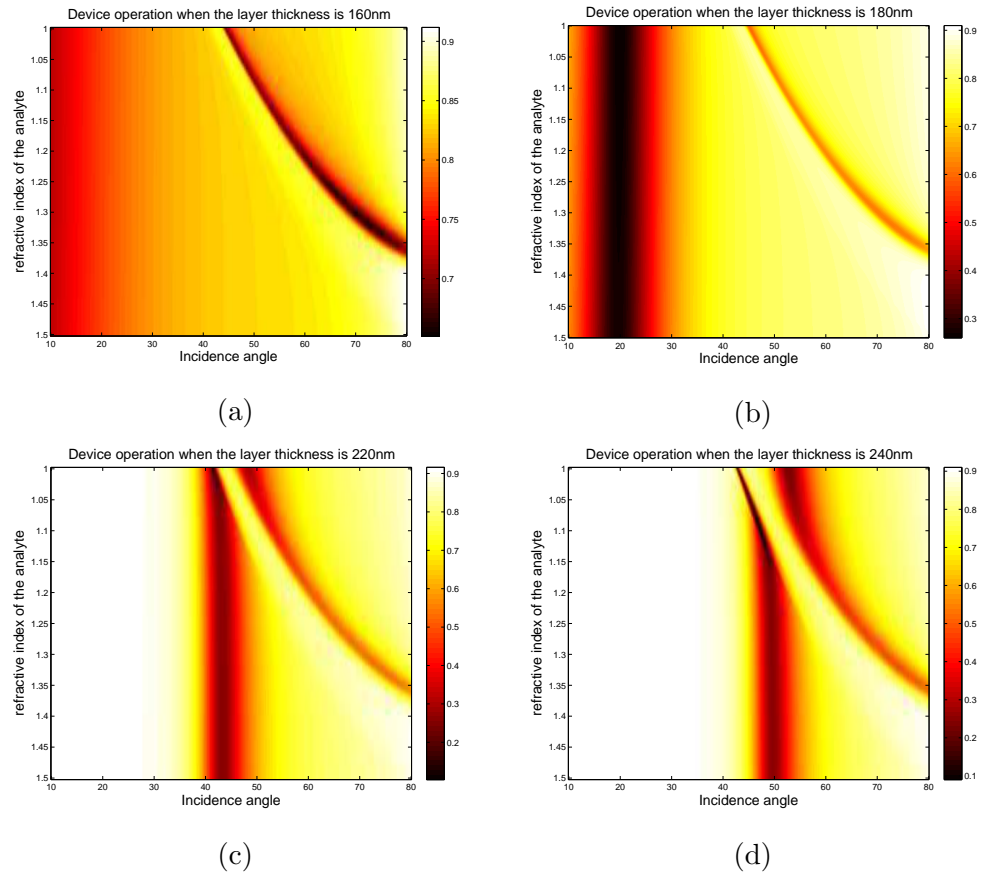
Figure 4.12: The effect of intermediate layer thickness on sensor operation. **(a)** For thickness values smaller than 160, the analyte insensitive dip is below $\theta_{in} = 10°$. **(b)** For intermediate layer thickness around 180nm, the dip with constant angular position is seen around 20°, moving towards the SPR dip. **(c)** The analyte insensitive dip goes past its optimal point as shown in figure 4.8(when thickness is 200nm). It starts interfering with the plasmon dip. **(d)** The dips occupy same angular positions, hence the sensor operation is hindered.
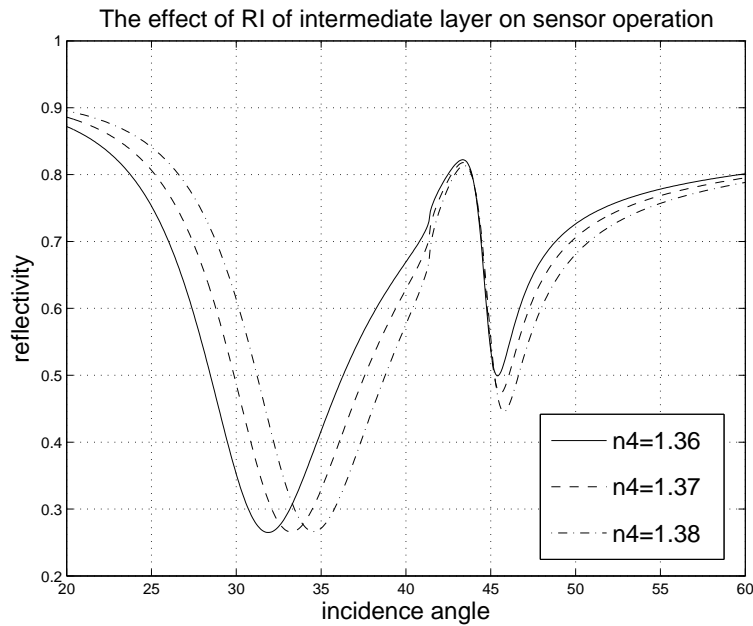
Figure 4.13: The sensor behaviour when the intermediate refractive index changes. Reflectivity is plotted at various incidence angles(x-axis). As can be seen, the angular position of the first dip is dependent only on the refractive index of the intermediate material, hence the temperature. Second dip also moves, which could be compensated by the detection of a temperature difference.

perature, is seen to shift the RI insensitive dip (see figure 4.13). Within a measurement, a change in the angular location of the analyte insensitive dip, would indicate a change in the temperature. Therefore, the effects of such changes could be isolated from binding specific SPR dip measurements. The effect of the intermediate layer refractive index is shown in figure 4.13. As can be seen, the refractive index (n2) insensitive dip, corresponding to lower incidence angles, is moving. A smaller movement of the RI-sensitive dip angular position is also observed.

4.3.3.3 The metal film thickness

The metal film carries the SP fields and affects the multiple reflections. Therefore it affects both of the mechanisms employed in this device. The material for the metal films were chosen as gold due to its resistance to oxidation. Figure 4.14 shows the optimisation process for the metal film thickness. In figure 4.14(a) the operation of the sensor is seen when the metal film thickness is 25nm. As can be expected, the SP dip is wide for this

thickness. In figure 4.14(b), the metal film thickness is increased to 45nm. This is an acceptable thickness for this device. However a thinner dip was seen when the metal film was chosen as 50nm. When the metal film thickness is increased beyond this value, the SPR dip starts fading away, as in figure 4.14(c) and completely disappears in (d).

4.3.4 Optimum Parameters

The optimal parameters for this device are listed in table 4.2 (for $\lambda_o = 639nm$). Metal film was chosen as gold due to its resistance to oxidation. The metal film thickness was chosen to be 50nm, however, $\pm 5nm$ can be tolerated for a decrease in sensitivity. The film thickness affects the sensor operation by adjusting the SP damping term, and the leakage radiation due to the SPs on thin films. The sharpest SP dip, hence the most sensitive sensor will be obtained when the coupling mechanism is optimised. This happens for an optimum film thickness [90]. Too thin metal films, are seen to result in less sensitivity of the dip position and wider dips (refer to 4.14 (a)), while too thick metal films will result in very high dips (4.14 (c)), again resulting in a loss of sensitivity. Therefore the film thickness is chosen to be in the middle, corresponding to the theoretical optimal thickness for the SPR dip [90].

The intermediate layer properties could be chosen so as to yield different responses to the changes in temperature. For example, different thermal coefficients for refractive index could yield designs insensitive to temperature changes, yielding in more robust sensors. MgF_2 was chosen for this case since it is easier to obtain and fabricate. The layer thickness was observed to affect the position of the calibration dip. For some values, this dip was seen to interfere with the SPR dip. Therfore the layer thickness can be chosen so as to yield in a design, that is both flexible to the fabrication tolerances in the deposition of the intermediate layer, and result in a calibration dip that is close enough to the SPR dip to measure easily.

These sensors could be fabricated by successively evaporating metal films and MgF_2 layer. MgF_2 can be deposited or grown on surfaces by electron beam evaporation, or atomic layer deposition techniques [173, 174].
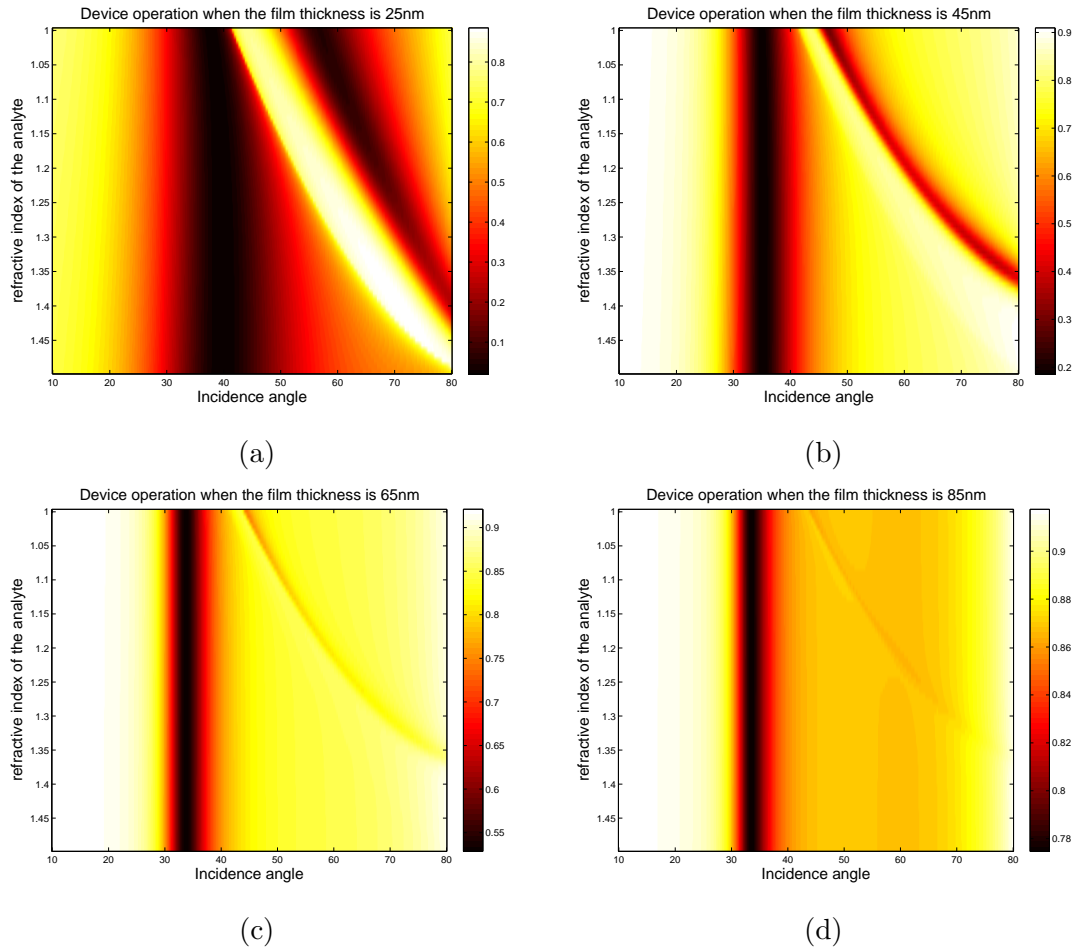
Figure 4.14: The effect of metal film thickness on the sensor operation. **(a)** When the metal film thickness is 25nm. As can be seen the SPR dip is wider than optimal. **(b)** Metal film thickness is 45nm. The SPR dip is seen to be well established at 45nm. The optimal thickness was found to be at 50nm(shown in figure 4.8). **(c)** The sensor operation when the metal film thickness is increased to 65nm. As can be seen, the SPR dip starts fading after the optimal thickness(50nm). **(d)**The SPR dip disappears when the metal thickness is increased to 85nm. The analyte insensitive dip is still observed.

Parameter Name	Parameter Value	Explanation
Metal film material	Gold	Resistance to oxidation
Metal film thickness, t_m	$45nm < t_m < 55nm$	optimum thickness for SPR dip at $48nm$
Intermediate Layer material	Magnesium Fluoride (MgF_2)n=1.38	easy to obtain
Intermediate Layer thickness,T	$170nm < T < 220nm$	chosen to be close to SPR dip

Table 4.2: The optimal parameters for the self-calibrated SPR sensor for 639nm optical wavelength.

4.4 Plasmon resonator

In this section FEM is used to design a surface plasmon resonator. Surface plasmons and their properties have been mentioned before in section 2.3. The interaction of surface plasmons such as reflection [120, 114], refraction [127] or scattering [175] have been studied in an effort to analyse their possible uses as nanophotonic circuit elements. These properties have been modelled in [59, 61, 60]. Surface plasmons have found their use in microscopy [176, 177], sensing [138, 139, 140, 141, 137], and show interesting characteristics to yet be explored.

Plasmons can be excited in a variety of ways. In this thesis, the experimental realisation of SP excitation will involve the use of an high NA oil-immersed objective, in order to use a range of incidence angles, rather than one at a time (described in section 4.4.5). Detection of plasmons, in this thesis, is achieved by observing the leakage radiation [114, 124].

When the plasmons are excited on a thin film between two efficient reflectors, plasmons reflect back and forth and reflected components interfere with each other. This interference can be seen as a modulation of specular and retro reflection in the farfield. The these structures and some similar structures have been suggested in [59, 128]. However, detailed analysis of these structures has not been performed in the literature, such as optimal design for practical applications, fabrication issues, efficiency and experimental verification. These structures have not been fabricated and experimentally verified to the author's knowledge. These structures have been optimised via FEM modelling, fabricated using photolithography, and the resonant behaviour was experimentally verified, which

are explained in that order. Resonators can also be used as a basis of a lasing structure , in the presence of an active layer.

4.4.1 Operating Principle

A block of metal deposited on a thin flat film, can act as a reflector for plasmons [59]. Other types of mirrors are also investigated numerically, such as holes and dielectric reflectors. Reflection efficiency of big metal blocks were found to be superior to these reflectors. Refelectors such as periodic holes in the metal film [123], dielectric blocks and smaller gratings [127, 121, 129] (see equation 3.2 for grating couplers) were investigated in the literature.

Plasmons can be launched on a thin film, in a set-up analogous to Kretschmann Configuration. When there are reflectors present, the reflected plasmons travel in the opposite direction as the launched plasmons. These components will have leakage radiation parallel to but opposite in direction to the incoming field. This will result in a retroreflection. Without the reflector there is only one leakage radiation process in the media, since the plasmon propagation is only on the launched direction. In this case, leakage radiation is in the direction of the specular reflection 180° out of phase with it yielding a dip in the reflectivity of flat metal film on resonance (see section 2.3).

4.4.2 Design objectives

Surface plasmon resonator has been designed so that plasmons can be made to resonate in a cavity. Figure 4.15 shows a conceptual diagram of what happens in such a structure. The detailed operation of the resonators is described in the following section. SPs are launched in the cavity on a thin film and are subsequently reflected back and forth. These components interfere depending on the cavity width, and this interference is observed by looking at the reflected spectrum.

The strength of this interference pattern is dependent on the number of round trips in the cavity. Therefore the propagation length of SPs on the metal film becomes an important issue. The two damping processes inherent to SPs on metal films, namely the internal damping and the radiation damping play important parts in the operation of
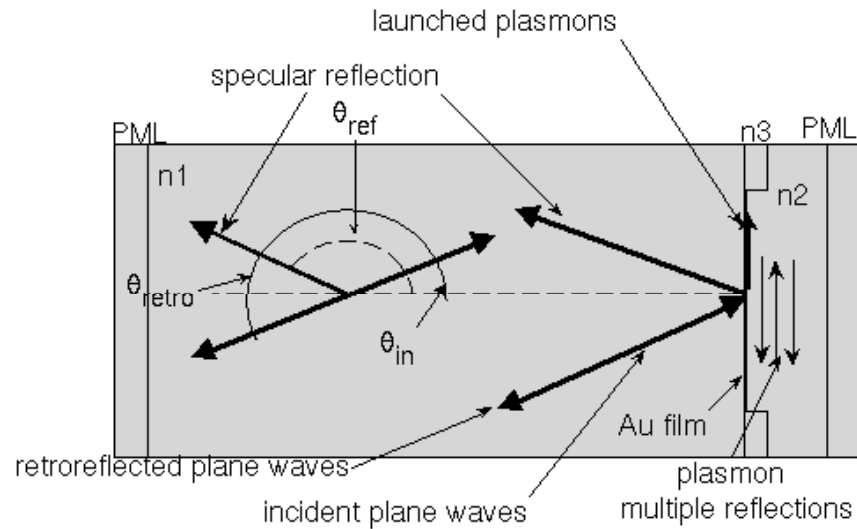
Figure 4.15: The basic model geometry for the plasmon resonator and the demonstration of the leakage radiation components. The subdomains from left to right are, PML, a high RI dielectric(glass), SP supporting metal structure (Ag), lower RI dielectric (air), PML. Plane waves ($\lambda_o = 639nm$) come into the model from behind the PML layer on the left, and pass through the glass subdomain before hitting the Ag structure. Reflected spectrum is collected in the glass subdomain, and the related incidence, reflection and retroreflection components are extracted. There are two groups of leakage radiation due to two components: plasmons travelling in the forward (upwards in the figure) and the backward directions (downwards in the figure). The incident field launches surface plasmons, travelling in the forward direction, which are reflected from the reflectors. As the reflected plasmons propagate backwards, they contribute to the leakage radiation towards the retrodirection. When the back-travelling plasmons are reflected, they propagate in the same direction with the launched plasmons, causing interference, and contributing to the leakage radiation in the direction of specular reflection. All the farfield observation angles are measured from the positive x-axis, anti-clockwise, as shown.

such a device [90], considering the very lossy nature of SPs. Therefore the thickness of the metal film and the material of the metal film should be adjusted to optimise the damping process. Metal film thickness should be optimised with the given conditions, to reduce the internal damping. Radiation damping should not be minimised, since it is the output of interest, and is proportional to the coupling mechanism.

There will also be some additional loss in the cavity due to the reflectors. Therefore they have to be optimised with respect to physical and fabrication limitations. While the material the reflectors are can be made out of may be seen as a physical limitation (how much transmission vs reflection for each material), their dimensions are a fabrication related issue. The light will be coupled in from the substrate and will be collected at the reflected side in farfield, therefore the ultimate aim is to design a resonator with the best coupling efficiency (hence leakage radiation), and the least amount of internal loss. The expectation from this design therefore is not a high quality resonator, due to the very high internal losses.

Design and optimisation of the resonators were done considering the fabrication limitations and the issues with the experimental verification of these devices. Then a base model was created, and scripts were written (listed in Appendix A.2) to investigate several features of the resonators and to extract necessary information. Once the operation is verified, then optimisation of these structures were considered in light of fabrication limitations and experimental details.

The resonating structures can be built by E-beam lithography or photolithography. Photolithography was chosen since it will yield bigger structures/reflectors. Bigger reflectors will be able to trap plasmons better. Fabrication tolerances impose uncertainties and variations on the produced devices. Experimental details consist of issues such as the minimum spot size, optical wavelength, the range of incidence angles that can be used.

The detailed photolithography process is described in section 4.4.4. Here photolithography is briefly mentioned (shown in Figure 4.16) as it applies to the fabrication and the design. The fabrication of the resonators can be summarised as below:

1. A substrate, eg a coverslip, is cleaned and coated with some UV sensitive photoresist material.

1) clean glass substrates
2) spin on the photoresist

Glass substrate

3) Mask Aligner places the mask
on the sample

Mask   pattern to be transferred

4) UV exposure: Tranfers patterns
on the resist

5) Development removes the
exposed resist away yielding voids

6) Metallization process fills out the voids

7) Lift−off process dissolves and removes
photoresist, exposing
the substrate and the metal

8) Second evaporation deposits the metal film

Figure 4.16: Steps of Photolithography as it applies to the fabrication of the resonators. Steps include substrate cleaning, photoresist spinning, mask aligner/UV exposure, development, metallisation, lift-off, second metallisation and second metallisation

2. A mask is prepeared with the desired features/patterns printed on it. These features correspond to the reflectors on the sample. The feature separation is called the cavity width. *Cavity width is a design parameter.*

3. The features on the mask are effectively written on the photo-resist via illumination with UV light through the mask features.

4. The resist is developed with the appropriate developing solution. The pattern is removed exposing the locations where in this case, the reflectors are to be placed.

5. Metallisation(thermal evaporation) fills in the voids left by the development with the desired thickness of metal, to fabricate the reflectors. *The height and width and material of the reflectors are design parameters.*

6. A process called lift-off is implemented to remove (lift-off) the remaining resist from the substrate, leaving only the reflectors on the substrate.

7. A second metallisation deposits the SP supporting metal film. *The thickness and the material of this film are design parameters*

For this wavelength, surface plasmons has a propagation length of about $22\mu m$ on silver film and around $10\mu m$ on gold. Since SPs are dissipative, silver was chosen to support longer propagation distances. Therefore the design parameters are the structure material, the reflector width and height, film thickness and cavity width.

The experimental set-up to verify the operation of these resonators is described in Section 4.4.5. It uses an oil-immersion objective, which allows the excitation of plasmons by supplying a wide range of incidence angles.

The retroreflection and the specular reflection can be simultaneously observed as a measure of efficiency of the resonators. This experimental set-up forms the basis of the numerical model, as it determines how the SPs are generated and detected. In this case SPs are detected in the farfield.

As an initial step, the operation of the resonator, and the resonating property has to be investigated before any optimisation can be done. The analysis of the resonating property will involve changing the cavity width and investigating the interference patterns

by observing the reflected farfield spectrum. Parameters are then optimised according to fabrication and experimental limitations.

The base model is shown in Figure 4.15. As with the previously introduced models, this model also builds on the foundational concepts introduced in Chapter 2. The model is periodic in the y-direction (vertical axis)(refer to section 2.8.2 for periodic boundary conditions). The plane waves come from behind the leftmost PML into the glass subdomain. Farfield spectrum is extracted in the glass subdomain, via the methods described in Section 2.8.4. The scripts to produce models and extract the necessary information is listed in Appendix A.2.2.

4.4.2.1 Basic concepts in the operation of the resonator

The resonator uses two reflectors to trap surface plasmons in a valley-like structure [59]. When there is a pair of reflectors, and the surface plasmons are launched in between them, plasmons are reflected back and forth from each reflector, hence getting trapped to be extinct by the absorption losses. Therefore, the number of reflections depend on the propagation length of SPs on the metal film of interest.

The plasmons travelling in the same direction, due to multiple reflections, yield interference patterns which depend on the cavity width. There is a very small phase shift due to the reflection from the metal walls which modifies the effective cavity width. This adjustment was found to be small for the cavity widths considered. Due to the mask production facilities, integer values of cavity widths rounded to the nearest 100nm values were chosen. Therefore, interference pattern has a period $\frac{\lambda_{plasmon}}{2}$ [59].

The detection of this interference pattern could be done in the near field and was suggested by [178, 59, 128] or could be done in the farfield by using the leakage radiation suggested by [124, 129]. In the presented models, it was possible to look at both, but it was found more appropriate to extract the farfield information so as to be able to easily observe the occurrence experimentally using simpler farfield methods. Near field data is immediately available from the models, as can be seen in Figure 4.17. The farfield spectrum is obtained by applying the methods mentioned in section 2.8.4. An example farfield spectrum is displayed for an optimised resonator in Figure 4.18. Figure 4.18(a) and (b) shows the farfield spectrum of these structures on and off resonance respectively. The
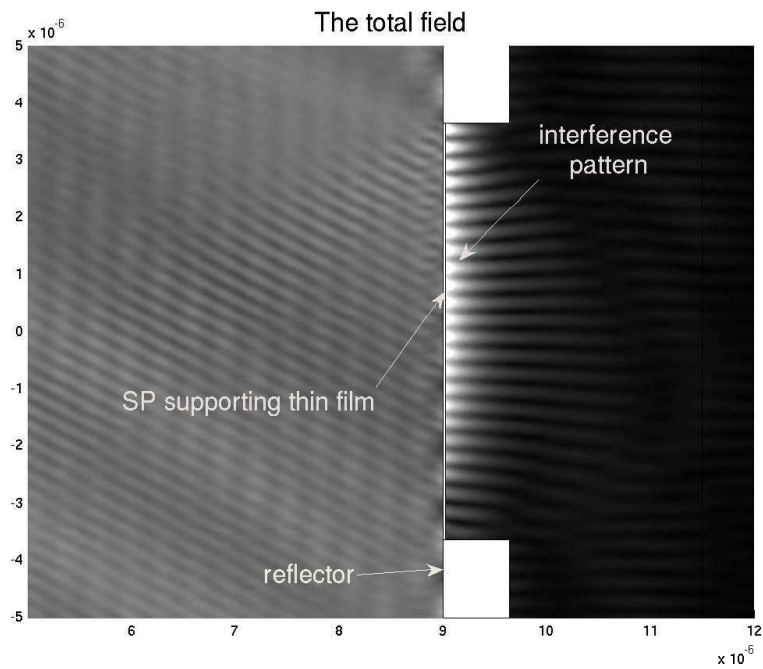
Figure 4.17: The demonstration of the near field data on the resonator. The colour scale has been adjusted so that the contrast of the image is optimised and the visibility of the structure is increased. Also labeled is the interference pattern of the plasmons between the reflectors. As can be seen the contrast of the pattern is less in the lower end of the film. This is simply because there are more plasmons travelling in the middle of the cavity than there is immediately after the metal film. The amount of plasmon coupling is proportional to the length of film considered.

specular and the retroreflections correspond to certain farfield angles, which are explained in figure 4.15. Looking at Figure 4.18 a., there is a field enhancement at farfield angles $220° - 225°$, which corresponds to the retrodirection for an incidence angle between $40° - 45° + 180°$.

Figure 4.15 gives a conceptual explanation of the resonating capability of these structures, and the contributing parts to the reflected spectrum. In this scheme, the dip in the specularly reflected light, will have contributors from the $2m$ reflections (and the generated plasmons), whereas the retroreflection will have contributors from $2m - 1$ reflections, number m depending on the ratio between propagation length of SPs and the separation of the reflectors.

(a)   (b)
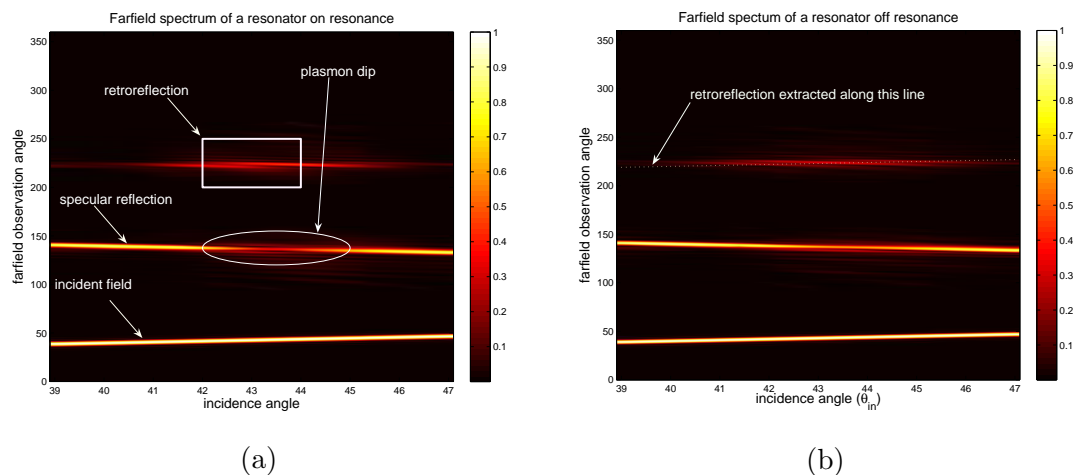
Figure 4.18: The farfield spectrum of the plasmon resonator(normalised with the incident field intensity). The retroreflection can be seen between farfield angles $220° - 225°$ when the incidence angle is $40° - 45°$ (refer to figure 4.15 for the explanation of the conventions for the farfield angles). **(a)** The resonator is on resonance. The dip in the specular reflection is low, and the retroreflection is high compared to (b). Also the SPR dip is wider than the SPR dip on flat films. This broadening is due to the width of the cavity on which plasmons can be excited. **(b)** The farfield spectrum of the resonator close to off-resonance. It can be seen that retroreflection is lower than (a) and the dip in the specular reflection is not as deep. The rising of the dip value shows the destructive interference of the launched plasmons with the reflected ones. Also shown is the dotted lines corresponding to the retroreflection.

4.4.2.2 The retroreflection efficiency of the resonator

The resonator retroreflection efficiency can be measured by observing the retro and the specularly reflected light. More efficient resonators will have increased retroreflection and decreased specular reflection at the plasmon angle. This means that more light is coupled in the structure and more of it constructively interferes in the backward and forward directions. As the cavity width changes, the forward reflected plasmons start to constructively interfere with the launched plasmons, and the backward reflected plasmons interfere with other back reflected plasmons. The presence of the retroreflected light and the plasmon dip verifies the excitation and reflection of the plasmons respectively, as it was done in [124]. The strength of the retroreflected light in the presence of a specular plasmon dip, verifies the efficiency of the device, by demonstrating the strength of the interference pattern.

(a) (b)
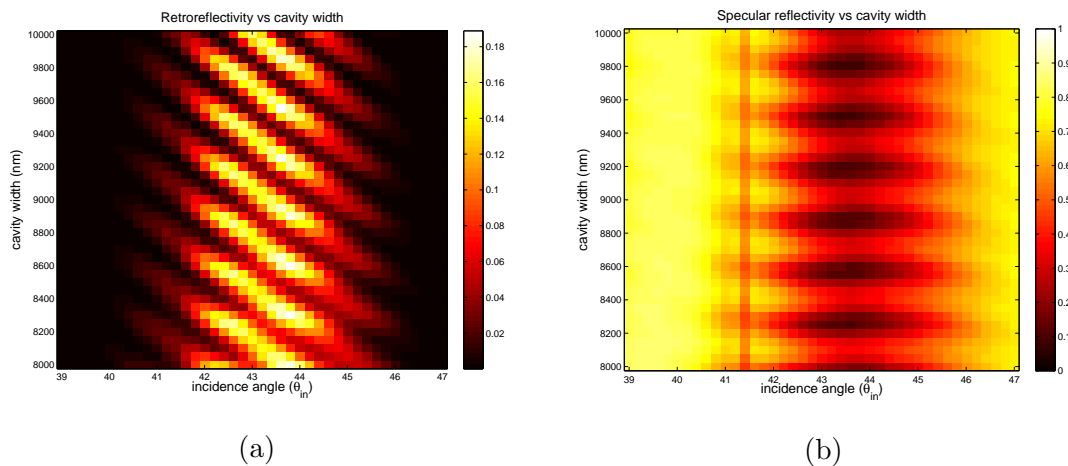
Figure 4.19: Demonstration of cavity width as a condition for resonance. These models had a constant reflector size and film thickness.**(a)** The retroreflection intensity changes with varying cavity width. The modulation of the retroreflection is superimposed on a grating order resulting in the shape shown. **(b)** The plasmon dip as the cavity with changes. The effects of plasmon interference are seen as a sinusoidal modulation pattern in the dip. The broadening of the SPR dip is seen as a result of the limited SP excitation due to the cavity width and the reflectors.

Figure 4.19 shows the simulation results when the cavity width was varied keeping reflector sizes the same. Figure 4.19(a) shows the retroreflectivity at different incidence angles as the cavity width is varied. At a constant incidence angle, if different cavity

widths are considered (this corresponds to a vertical cross section of (a)), a sinusoidal variation is observed. The plasmon dip in figure 4.19(b) goes through a sinusoidal change, due to the interference of the forward travelling plasmons.

However, the dip value is seen to be constraint from above and below. It does not go above a certain value, and does not drop below a certain value. The dip does not rise above a certain value, because the forward reflected plasmons, even when destructively interfering with the launched plasmons, are strongly attenuated due to the losses in the system. Hence they are not able to cause a complete destructive interference to smear out the dip. Additionally, it can be seen that the dip value does not go all the way to zero. The specular reflections from the reflectors form a background on which the interference patterns are superimposed. The dip value can never go to zero, because in the model, not all the light couples to surface plasmons.

The retroreflection angle seems to be modulated/shifted towards lower angles as the cavity width is increased. This is because a periodic model is employed to investigate the resonator. The resonator is effectively a strong grating. The slanted lines are merely high-order diffracted reflections from the grating multiplied with the retroreflection for different angles.

Figure 4.20 describes what happens to the specular and the retroreflectivity at the plasmon related angles (around 43.6°) as the cavity width is varied. Here, figure 4.19 is integrated around the plasmon angles to get rid of the diffraction effects (numerically, integration is done as a summation around the plasmon angles). The physical meaning of this summation is related to the experimental details[1]. A following observations are made from figure 4.20:

- Both of the plots vary sinusoidally, with period around 300nm, which corresponds to approximately $\lambda_{plasmon}/2$ (for silver at 639nm wavelength)(see section 3.2.4 for the calculation of $\lambda_{plasmon}$). This agrees with the results presented in [59].

- The oscillation amplitude of both curves decrease, as the cavity width increases.

- The mean value of the retroreflection curve decreases while cavity width is increased.

---

[1]Experimentally, the retroreflection values and the dip values are extracted from the images, as an integration around the plasmon angles, in order to increase the signal to noise ratio
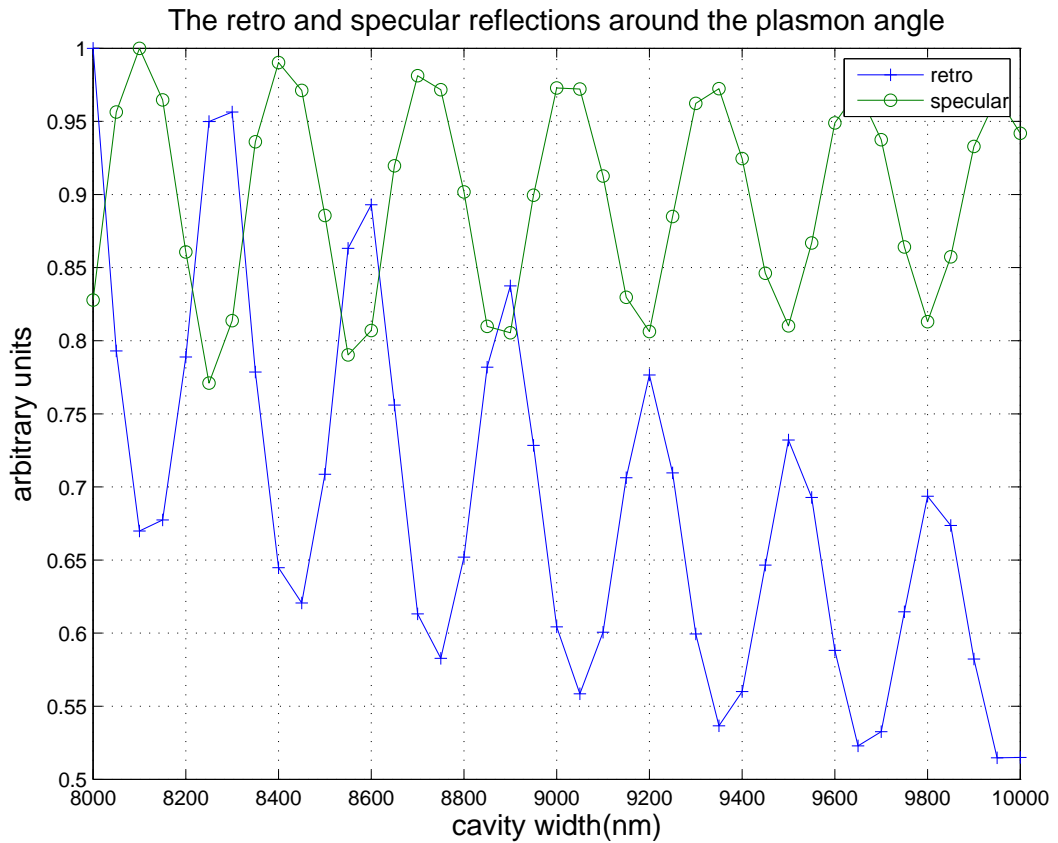
Figure 4.20: The retro and the specular reflections as the cavity width changes around plasmon angles. These correspond to the integration around the plasmon angles (horizontal sum of the images in figure 4.19). The plots are approximately sinusoidal and the two curves are 180° out of phase with each other, which means that a minimum in specular reflection occurs when there is a maximum in the retroreflection. Retroreflection at various angles is detected using the experimental set-up described in section 4.4.5, and it is integrated around the SP dip angles so as to maximise the signal to noise ratio. The two curves are scaled to fit into the same figure so as to demonstrate their out of phase nature.

- The mean value of the specular reflection curve stays around the same values as the cavity width is increased.

The propagation losses in the system reduces the amplitudes of the successive interfering components. As the cavity width is increased, the amplitude of the interfering components decrease due to the propagation losses and the number of effective interfering components decrease. For very high cavity widths, the interfering component amplitudes are so small that the oscillation amplitude approaches to zero and the structure approaches to a simple reflector with constant retroreflection and specular reflection. In this case, the sinusoidal behaviour of the dip will disappear.

The interfering components never completely cancel the launched plasmons or the first SP reflection. Hence there is always an average value over which any interference is superimposed. As the cavity width is increased, the plasmons reach a steady state value. The energy transferred to plasmons stays proportional to the incoming field, which grows with the model size, hence the cavity width. Therefore the specular reflectivity (reflected intensity normalised with the incident intensity) stays constant when the cavity width is increased. The reflected plasmon energy however, is proportional to the steady state value reached by the SPs, which is constant. Therefore as the cavity width is increased, reflection due to this constant field is normalised by an increasing incident field. This causes a decrease in the average value of the retroreflectivity.

4.4.2.3 Analysis of the diffracted orders at the retroreflection angle

When analysed using a periodic model, a change in cavity width corresponds to a change in the periodicity of the model. A change in the periodicity of a grating means that the distribution of the diffracted orders change. The magnitude of the diffracted orders are greater than the magnitude of the effects due to the retroreflection interference. The interference patterns of the retroreflection curves depend on the plasmon wavelength(610nm) which are superimposed on the distribution of the grating orders, which depend on the incoming wavelength(639nm).

The concept of varying diffracted orders is also demonstrated in figure 4.21. Figure 4.21(a) shows the spectrum when the reflector width is varied while the cavity width is
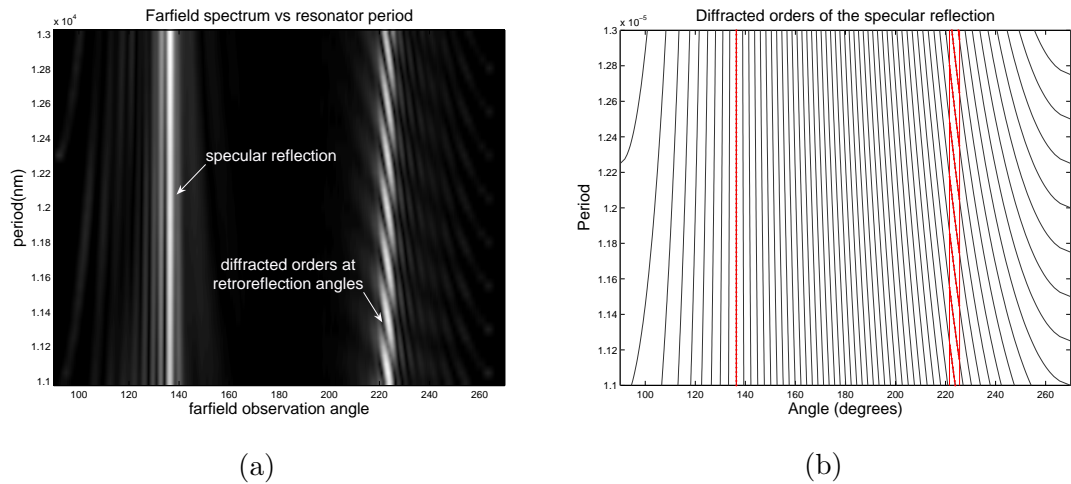
Figure 4.21: Demonstration of the higher grating orders convolved with retroreflection at 43.6°. **(a)** The demonstration of the multiplication of the retroreflection with the diffracted orders when the cavity width is kept constant and the reflector width is varied. In this case the cavity, hence the interference fringes are constant, and the only thing that is varied is the period of the resonator **(b)** The orders of a grating of equivalent period to (a) plotted using the grating equation.

kept constant. This implies changing the period, while the interference pattern is the same. The match with the theoretical grating equation in figure 4.21(b) can be seen. For different incidence angles, different parts of these orders are taken resulting in Figure 4.19(a).

The complication due to the diffracted orders is not present when an isolated resonator is considered with a model using PMLs on all sides (model shown in figure 4.22) as shown in figure 4.23(a). The retroreflectivity here is shown for various cavity widths and incidence angles. A change in the cavity width modulates the retro-reflectivity. In Figure 4.23(b) the retroreflectivity and specular reflectivity at only the plasmon angle is shown vs the cavity widths. It is seen that both curves are modulated and out of phase, indicating that the successful plasmon excitation only happens when there is constructive interference in the cavity, hence retroreflection. In this figure, the retroreflectivity had to be scaled down in order for the modulations of the both curves to be visible. This is due to the fact that the extracted specular reflectivity was only weakly modulated. This results from the effect of the big reflectors in the model dominating the field in the direction of the specular reflection, around the region of farfield extraction. The retroreflection however, does not

Geometry for the isolated resonator
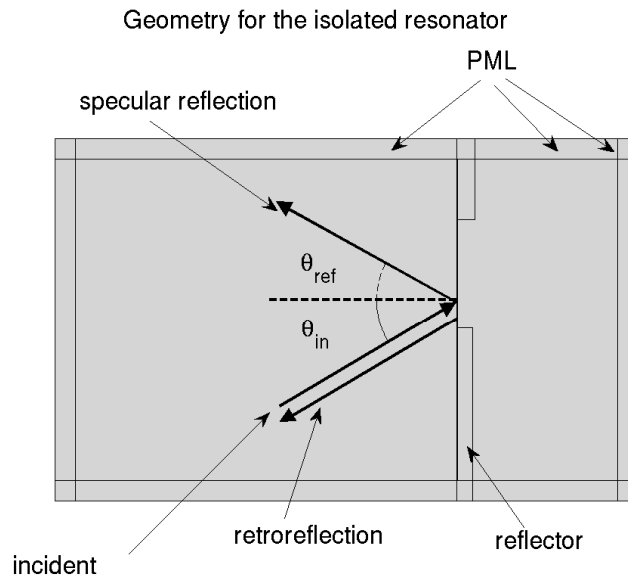
Figure 4.22: The geometry for the isolated resonator model. This model simulates an isolated resonator by replacing periodic boundary conditions with PMLs. It is not as flexible as the periodic model that was considered, however, it doesn't suffer from the complications due to periodicity and the diffracted orders.

suffer from this effect.

It is also necessary to compare the cavity and reflector width induced period changes with simulations involving constant period, in order to relate the two effects. In order to see one order clearly, a constant incidence angle is used. Figures 4.24, 4.25 and 4.26 show these results. All the images were obtained for an incidence angle of 43.6°. Figure 4.24 was obtained by changing the reflector width and keeping the cavity width constant. As can be seen, there is no rippling in the specularly reflected light, this shows that the forward reflected plasmons have the same interference pattern.

Figure 4.25 shows what happens when the period is kept constant and the cavity width is varied. This implies a corresponding variation in the in the reflector width, to keep the period the same. The specular reflection is modulated out of phase with the intensity of the retroreflection. However, it can be seen that the retroreflection ripples are superimposed on a diffracted order, hence their contrast is decreased in this figure. Everything stays on one diffracted order since the period of the resonator is constant.
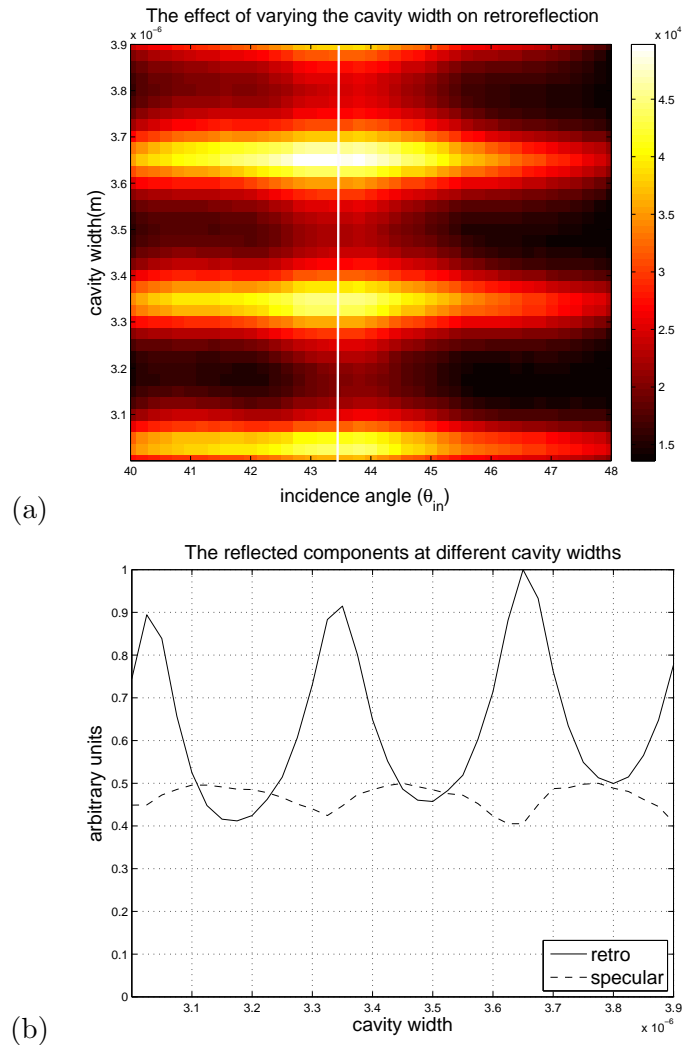
Figure 4.23: **(a)** The retroreflection from an isolated resonator (geometry shown in figure 4.22). As can be seen, the retroreflection does not suffer from the slanted lines (shown in figure 4.19) once the periodicity is removed. **(b)** The retroreflection at around $43.6°$ plotted on the same scale as the specular reflection at the same angle. It can be seen that the curves are out of phase as expected, since the minimum of the dip in the specular reflection occurs on resonance, where there is retroreflection due to constructive interference. The curve of retroreflection was scaled down to fit on the same axes with the specular reflection.
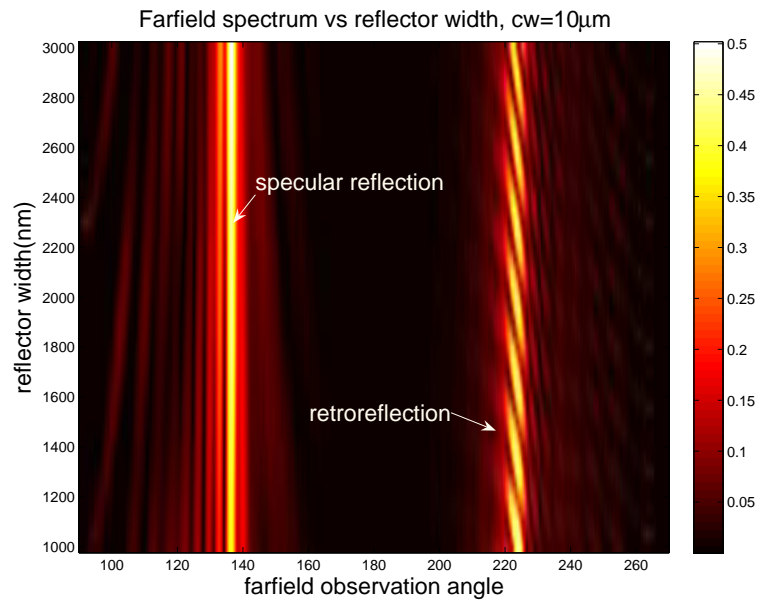
Figure 4.24: The effect of changing the reflector width, while the cavity width is kept constant on the reflected spectrum. A constant incidence angle (43.6°) is considered. In this case the diffraction pattern changes the retroreflection. The specular reflection stays constant, since this depends on only the cavity width, not the period of the resonator.
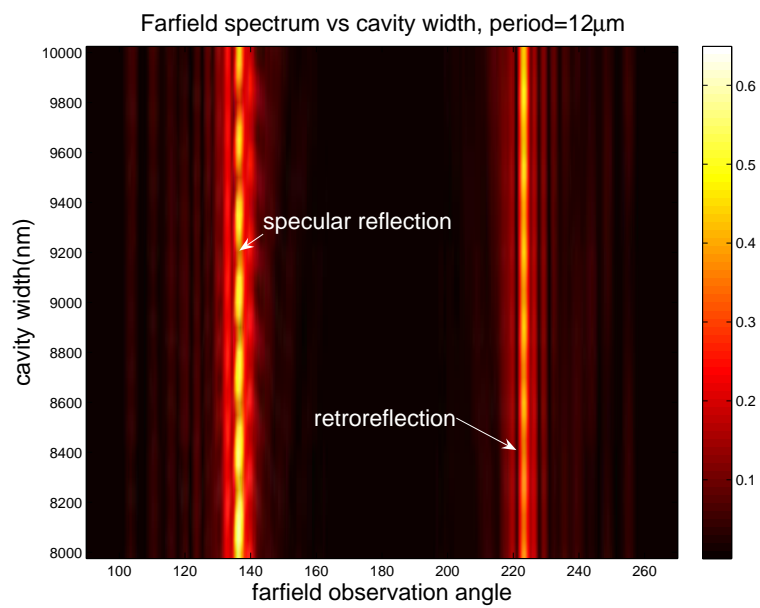
Figure 4.25: The effect of changing the cavity width while keeping the period constant (which means reducing the reflector width). A constant incidence angle (43.6°) is considered. It can be seen that the specular reflection shows the resonating effect, by allowing or banning the propagation of plasmons. In the retro direction, it can be seen that the diffracted orders stay the same as far as distribution, with a mild modulation on a constant background (on the same grating order) due to the interference pattern.
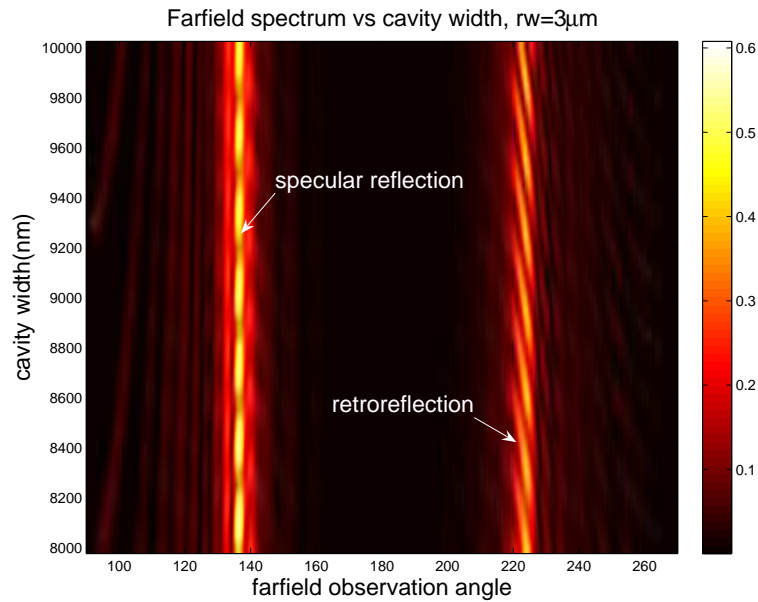
Figure 4.26: The effect of varying cavity width and constant reflector width on the reflected spectrum. A constant incidence angle (43.6°) is considered. If the cavity width is varied when the reflector width is kept constant, the specular reflection shows the modulation due to the resonating property. The retrodirection is the interference pattern superimposed on the diffracted orders. In this case, the diffracted orders will dominate the retroreflected interference pattern

Figure 4.26 shows the effect of changing the cavity width while keeping the reflector width constant. This imples a change in the resonator period. In this figure, rippling of both the specular and the retroreflections indicate that there is interference in both the forward and backward reflected plasmons. It can be seen that the diffracted orders are varying at the retroreflection angles.

The consequences of this concept are seen throughout this section, but it should be noted that the main plasmon dip occurs around 43.6°, where the minimum of the plasmon dip is achieved for 639nm wavelength. However, in the experimental verification, these orders are integrated to yield plots similar to Figure 4.20.

4.4.3 Analysis and optimisation of design parameters

The next step in the design process is to optimise the parameters to get the most practical resonator in light of the fabrication criteria. As thefabrication tolerances were investigated,

it was noted that the available photolithography resources could produce errors on each side of the reflector around $\pm 1\mu m$(along the vertical direction in figure 4.17). Therefore the feature size for the reflectors was chosen to be around $3\mu m$ so as to accommodate for the worst case ($-1\mu m$ error on each side).

The optimisation parameters relating to the reflectors include the reflector width, the side-wall angle, the reflector height and the reflector material. The optimisation parameter for the SP supporting film is the film thickness. The following simulations test these parameters in order to come up with the best combination, which can be easily fabricated.

4.4.3.1 Analysing the reflector width

The reflector width partly determines the efficiency of the resonator by directly affecting the coupling strength, when a repeating structure is considered. Figure 4.27 shows the effect of reflector width on retroreflectivity and specular reflectivity. The width of the reflector, effectively changes the fraction of the incident light that hits the metal film to launch SPs. For example, if the reflector-cavity ratio is 1:1 then 50% of the incident energy can be coupled to plasmons. However if the reflectors are made wider, and this ratio changes to 2:1, only the 30% of the incident energy will launch plasmons, resulting in a decrease in the fractional SP coupling. Therefore, the retroreflection efficiency does not necessarily decrease, but for a given area of illumination, the amount of SP coupling, hence SP reflection and strength of the SPs which are involved decrease.

Figure 4.27 shows the effect of reflector width on the resonating characteristics. A similar analysis was done in figure 4.24. In figure 4.27 the retroreflection and the specular reflection are recorded for various incidence angles. In figure 4.27(a) the diffracted orders are seen. Figure 4.27(b) shows the specular reflection. When the reflector width is increased, the dip value seems to rise, since the fraction of light coupling to SPs is decreasing. Also shown on the figures are the corresponding lines from which the curves in figure 4.28(a) was obtained.

Figure 4.28(a) shows the vertical cross-sections of figures 4.27(a) and (b). Figure 4.28(b) shows the integrated values along the SP angles in figure 4.27 (a) and (b). From figure 4.28(a) it can be seen that the specular reflection remains flat and the retroreflection varies sinusoidally. On the other hand, in figure 4.28(b) there is no sinusoidal variation
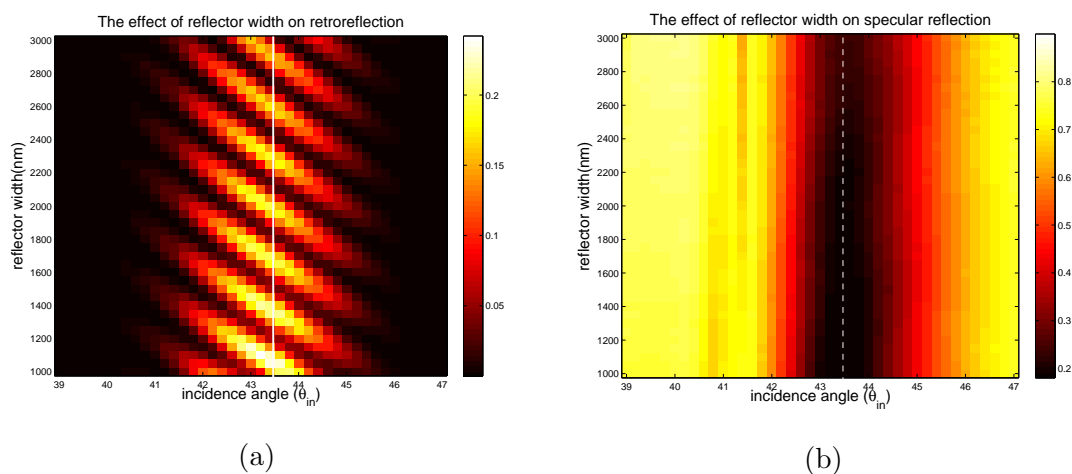
Figure 4.27:   Analysis of the reflector width in the performance of the SP Resonator. **(a)** The retroreflection around the plasmon angle. The retroreflection and the diffracted orders are multiplied in the Fourier domain to yield a modulated sinusoid. **(b)** The reflectivity in the specularly reflected direction. No rippling is observed in the specular direction since the cavity width is the same for this simulation.

in the retroreflected component. This shows that the sinusoidal changes in figure 4.28(a) are due to diffracted orders. Also from these curves, the above mentioned decrease in the fractional coupling efficiency can be seen. The retroreflection is decreasing due to decreased coupling, and the specular reflection is increasing due to the increased reflector-cavity ratio.

Figure 4.29 shows what happens as the reflector width approaches to values below 100nm. The reflecting property decreases, since the metal reflector starts to become transparent. When the reflectors get more transparent, the structure approaches flat metal. This can be seen in figure 4.29(a), where the reflectivity starts looking like surface plasmon resonance on flat smooth films. Therefore as long as the reflector width is grater than around 100nm it does not have a significant effect on the structure, since it is sufficiently reflective. However, large reflector widths reduce the overall efficiency of periodic devices, by decreasing the fraction of the incident energy that couples to SPs.

This analysis of the reflector width gave information about the coupling strength vs reflectivity. Even though the reflector width is mostly a fabrication dictated parameter, due to limitations of photolithography, obtained information is useful when interpreting the
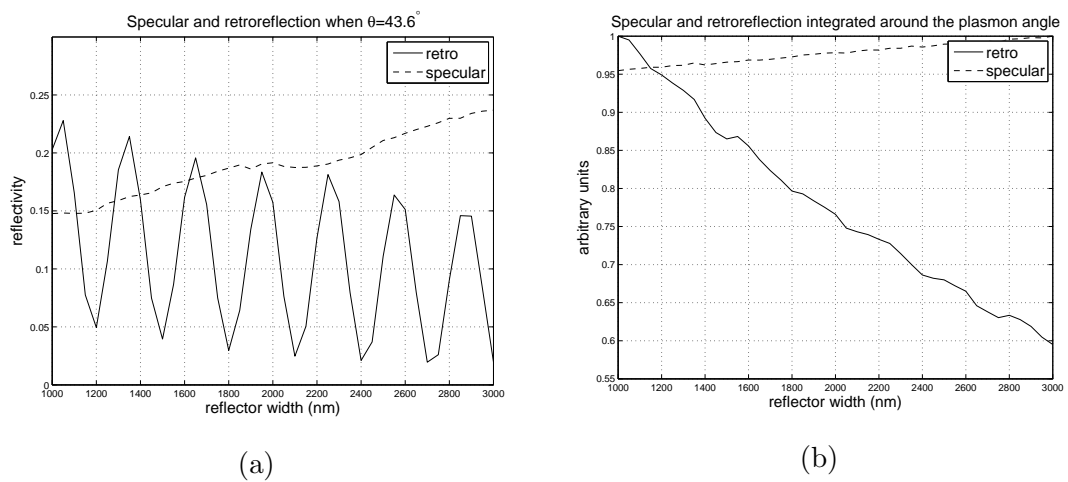
Figure 4.28: **(a)**The specular and the retroreflections at constant incidence angle 43.6°. This corresponds to vertical cross sections of images in figure 4.27(a) and (b). As can be seen, there is modulation on the retroreflection, due to the changing period and the diffracted orders of the resonator. The value of the SP dip is seen to rise, since the fraction of incident light coupling to SPs decreases as the reflector size is increased. **(b)** When the images in figure 4.27 are integrated along the SP angles (the simulated angles of incidence), an increase in specular reflection and a decrease in the retroreflection average value is obtained. The decreasing retroreflection is simply due to the decreased SP coupling,which results in the same interference pattern to happen with less intensity.
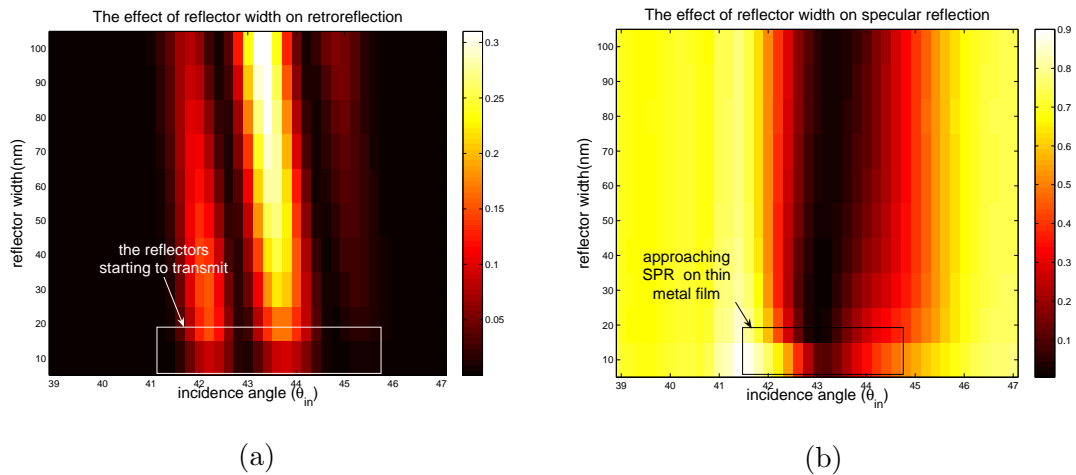
Figure 4.29: Analysis of small reflector widths. **(a)** The effect of having small reflector widths on retroreflectivity. As the reflector width is decreased, the retro-reflection decreases, since the wall reflectivity decreases **(b)** The effect of very small reflector widths on reflectivity. Very small reflector widths will cause decrease in wall reflectivity, hence they will start transmitting more. As can be seen, the dip in the reflectivity is becoming sharper, since there is more coupling due to the increased ratio of film to reflector width. However, there is very little retroreflection, indicating that the plasmons are not reflecting, but rather, being transmitted. The whole structure starts resembling a flat metal film supporting surface plasmons.

fabricated samples. In the following simulations, reflector widths are taken as a nominal $3\mu m$.

4.4.3.2 Analysis of the reflector height

The reflector height is another parameter to optimise in the design process. As the reflector height increases the reflectivity of the metal blocks increase by trapping the surface plasmons better. This parameter has a value above which there is no apparent increase in retroreflectivity. For small reflectors, the efficiency is low, since plasmons will "jump" over smaller obstacles. This is an important parameter which needs to be minimised so as to make the structure easier to fabricate. The fabrication problem for thick reflectors relate to the photosensitive material that was used. The height of this material allows evaporation of only a few hundred nanometres of metal. Higher reflectors could be obtained if two different resist materials were used on top of each other. Simulations were

run to determine an optimum value for the reflector heights according to the fabrication limitations. The results can be seen in Figure 4.30(a). The efficiency increases with the reflector height. Below 400nm there is severe reduction in the efficiency. The reflector height was chosen to be around 250nm due to the fabrication limitations.

4.4.3.3 The effect of film material and film thickness

The efficiency of the resonators is mostly dominated by the propagation length of the SPs on the metal used, hence the absorption of the specific metal at the wavelength of interest. Silver was chosen for the film, since silver has less absorption than gold for 639nm. Gold could be considered, if longer wavelengths were to be used where it is less absorbing. Film thickness simulations were run in order to investigate the optimal coupling thickness for the situation under consideration. Reflector height used in the simulations to model optimum conditions were 750nm and the optimum film thickness was chosen to be 30nm (refer to figure 4.30(b)).
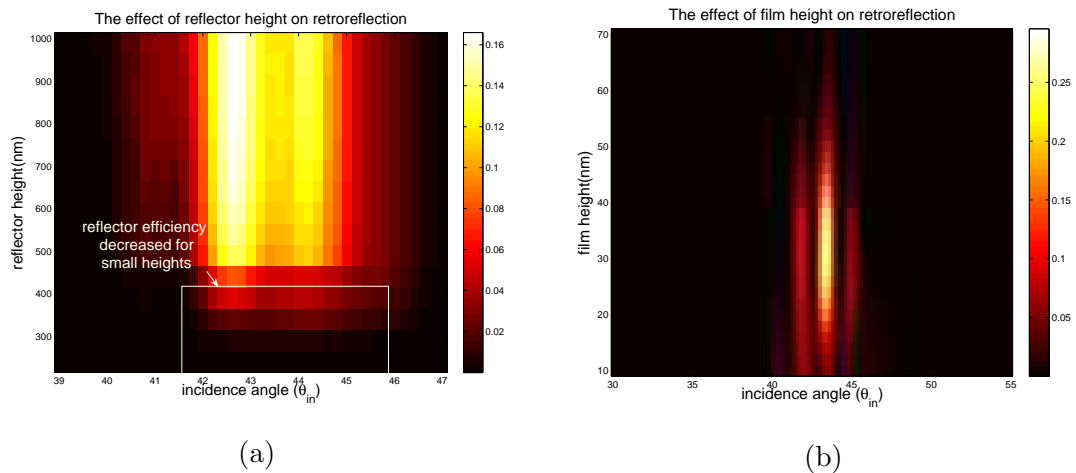
Figure 4.30: **(a)** The effect of the reflector height in the resonator efficiency. As can be seen, for small reflector heights there is less retroreflection. This is due to the fact that SPs can go over such small obstacles. **(b)** The effect of metal film thickness in the resonator efficiency. Film thickness directly related to the coupling to plasmons. The optimal film thickness was found to be around 30nm for silver for $\lambda_{in} = 639nm$

.
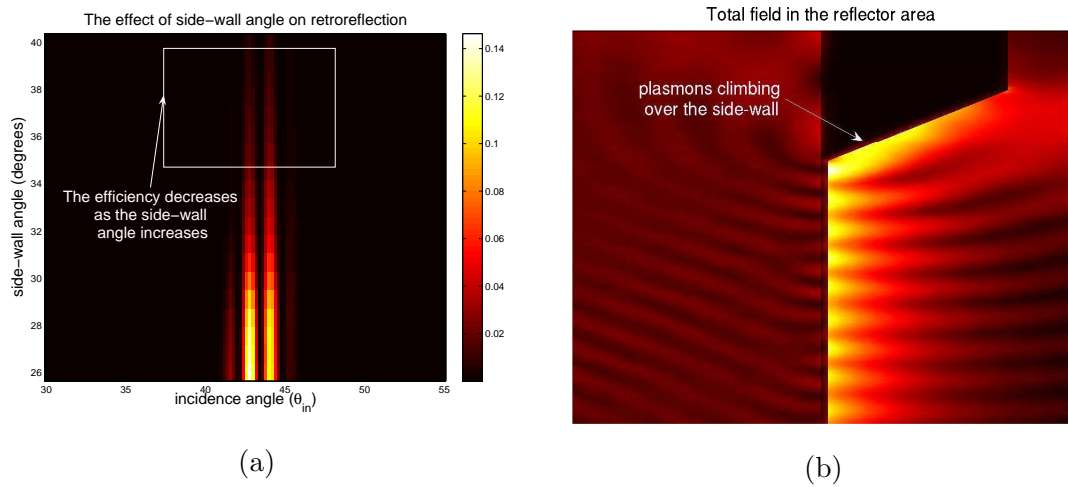
Figure 4.31: The investigation of the side-wall angle. **(a)** The effect of side-wall angle on retroreflectivity. As can be seen, the retroreflectivity is decreasing as the angle is increased. This is because plasmons are able to climb over such obstacles. This decrease is also affected by the change in the effective cavity width, as the sidewall angle is increased. **(b)** The geometrical demonstration of the side-wall angles. The outcut profile is common in over-developed resists. It can be seen that plasmons climb over this obstacle.

4.4.3.4 Side-wall angle

Metal features produced by photolithography will not have exactly vertical side-walls. Depending on the photoresist used and the photolithograpy process, they could have either an undercut profile or an over-cut profile (see figure 4.32). The effect of the side-wall angle is investigated by modifying the reflectors in the models (see figure 4.31(b)) and resolving the problem for several side-wall angle values. The effect of sidewall angle on retroreflection is shown for a resonator in figure 4.31. An increased side-wall angle will not only change the effective optical path between reflectors [59], but also affect the reflectivity of the walls, since the plasmons can "climb over" such reflectors (see figure 4.31(b)). This will move the models away from the ideal situation. Similar effects on wedges has been studied in the literature [59]. It was found that a slope of 70° was found sufficient to confine plasmons. The trapping effect reduced to zero at 45°. This effect has been verified in figure 4.31.

Parameter Name	Parameter Value	Explanation
Metal film material	Silver	low internal loss at $\lambda_o = 639nm$
Metal film thickness, t_m	$20nm < t_m < 40nm$	optimum thickness is $30nm$
Reflector height R_h	$200nm < R_h < 700nm$	high values are desirable, fabrication limited
Reflector Width, R_w	$R_w > 100nm$	$3\mu m$ due to fabrication limitations
Reflector side-wall , angle, R_{sw}	$-30° < R_{sw} < 30°$	vertical walls are preferable, could be overcut or undercut

Table 4.3: The optimal parameters for the plasmon resonator for 639nm optical wavelength.

4.4.3.5 Optimal parameters in the resonator

Several design parameters have been analysed to be able to pick the optimum combination. Effectively, the practical values of all these parameters are fabrication limited. Table 4.3 shows a summary of these parameters

4.4.4 Fabrication

In this section the fabrication process of plasmon resonators are described. These structures can be fabricated using photolithography. The steps involved in the photolithographic process within this context involve substrate cleaning, photoresist application, soft baking, mask alignment, exposure, development and hard-baking. After these stages, a process of metallisation, is followed by lift-off and a secondary metallisation.

The devices devices were fabricated by successively increasing cavity widths to experimentally verify the resonance. Fabricating the resonators similar to those in the curve shown in Figure 4.20 is sufficient to verify the resonating property of these structures. The fabricated devices were to have as high as possible reflectors with $3\mu m$ widths. Resonators had reflectors $3\mu m$ wide so as to accommodate for $\pm 1\mu m$ error on any feature

(reflector) due to photolithography. The maximum possible reflector height depended on the photolithography process and was chosen to be 250nm.

4.4.4.1 Prototype/sample preparation

The first stage of the fabrication process is the mask design. A mask was designed for photolithography process for several devices with different cavity widths and constant reflector sizes. The design mask had features etched away from an opaque chrome layer, and was covered with glass. Each feature is in the form of a line of length $2.5mm$ and width $3\mu m$ (corresponding to the chosen reflector width). Each device was formed by arranging these lines parallel to each other to approximately form a $2.5mm \times 2.5mm$ square patch(see figure 4.33). The separation of the lines within a patch correspond to a certain cavity width for each device. There are several such patches/devices in the mask, cavity widths ranging from $7\mu m$ to $14\mu m$. Adjacent patches/devices are aligned to be next to each other, so as to have similar fabrication errors.

The cavity width is then incremented by 100nm steps. This step size is dictated by the mask production facilities. The mask was then produced by Toppan Photomasks Inc. There is a $\pm150nm$ error on the mask features. This adds to the error in the fabricated prototypes (see section 4.4.6 for the discussion of errors on the fabricated prototypes).

Substrates are cleaned in order to remove organic, inorganic, ionic or metallic impurities together with any debris from the surface, improving adhesion of the photoresist to the substrate. The residues will result in defects in the final resist pattern [179]. Substrates are dipped into various solvents and left in the ultrasonic bath for 10 minutes per solvent. Solvents used are ethyl lactate, acetone, methanol and IPA respectively. After the cleaning process, an oxygen plasma cleaner is used in order to further etch any residues and roughen up the surface, so as to improve adhesion properties. After the cleaning process, the substrate is spincoated at 3800RPM for 30 seconds with photoresist, BPRS-150. This deposits a resist layer of approximately $1\mu m$. The spin time and speed are important, since they dictate the thickness of the deposited resist, hence the reflector height of the fabricated resonators. After the spincoating, the samples are left on a hot plate at $70°$ for 2 minutes in order to let any solvents evaporate and the resist to harden. After this process, the samples are left to rest for 10 minutes, afterwhich the mask aligner is used to

transfer the features from the mask onto the sample.

Mask aligner is set up to a setting called "proximity". In this process, the sample to be patterned is brought to the proximity of the mask on barely contact and a high intensity UV light is used to illuminate the photoresist at the openings of the mask. The photoresist is sensitive to the UV, and the parts that are illuminated will change in composition to become more readily affected by the developer. After the mask aligner stage, the development stage follows.

Development stage is particularly important, since it dictates whether or not the features on the mask is replicated sufficiently. The samples are soaked in the 1:3 BPRS developer water solution for 45 seconds allowing the photoresist to be removed. The ratio of the developer to the water in the solution is critical since it affects the development time. Over-developed samples tend to have features which are irregular and wider than intended, whereas the underdeveloped samples may still have some resist connected to the substrate. The time for development is thus very critical, since it can determine whether or not the features will be properly placed on to the substrate.

A common feature in the samples produced is called "wings" [180], which is an artifact of both the resist type and development stage. Wings appear when the resist profile is not overcut enough, and metallisation results in covering of the edges of the resist as well as the substrate. Wings occur randomly on the samples, at random places on the features. Figure 4.32(a) and (b) shows two ways wings can develop on the samples. In figure 4.32(c), the overcut resist profile can be seen as opposed to an undercut resist profile in (a) and (b). Wings can be overcome by changing the resist used, and by soaking the sample in toluene before development. The soaking hardens the top of the resist and makes it harder to develop. Hence, by the time the top of the film develops to desired feature size, the bottom of the resist layer has developed wider, thereby improving the overcut property. This is a process that needs to be optimised, since over-exposure of the film to toluene will slow down the development of the top of the resist too much and the development may become irregular.

The next stage is to build the reflectors by evaporating 250nm of silver on the sample. This is followed by the lift-off process. The lift-off process involves soaking the sample in acetone in order to remove the resist layer. After the resist was removed, metal stripes can

(a)   (b)   (c)
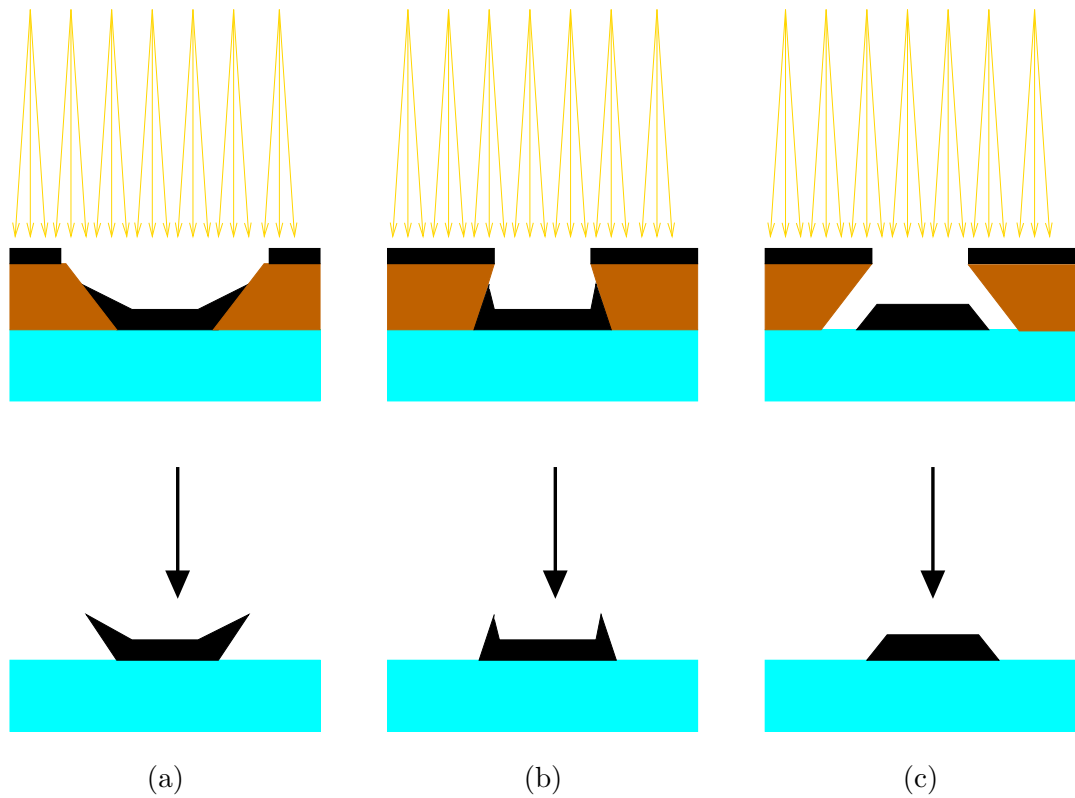
Figure 4.32: Demonstration of the formation of wings during the photolithography process. Demonstration of **(a)** wing formations due to an undercut resist profile **(b)** wing formation due to an insufficient amount of overcut resist profile. **(c)** structure formation without wings, due to sufficient resist overcut profile
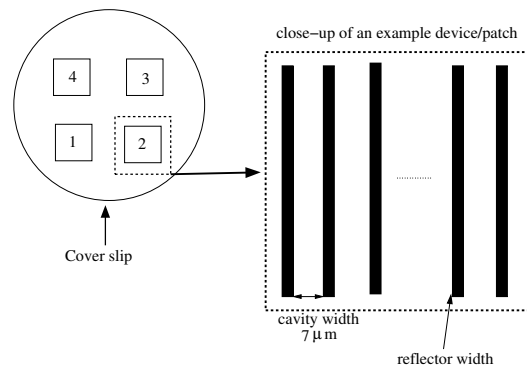
Figure 4.33: The diagram of a sample after photolithography. Every patch corresponds to a specific device. The devices with adjacent attributes are placed next to each other.

be seen on the sample. This stage is followed by a secondary evaporation of 30nm silver, which will result in the metal film formation. Also, 5nm of gold was deposited afterwards, in order to act as a protective layer against oxidation. The samples each contained four different resonators, in the form of patches, corresponding to consecutive cavity widths (see 4.33). The structures are measured using Atomic Force Microscopy(AFM).

4.4.4.2   Measurement of the prototypes

Atomic Force Microscopy is a technique that is capable of resolving and measuring very fine features. It is superior in non-biological applications to most optical methods since it is capable of giving depth information. AFM works by measuring the atomic forces between a very small finite size probe tip (around the range of 50-150nm depending on the tip used) and the sample. The AFM sensor probe tip size that was used had width $= 10\mu m$, length $= 20\mu m$, which corresponds to $25°$ of tip angle (see figure 4.34). There are some errors and uncertainties in the AFM measurements that affect the experimental results. These are discussed in section 4.4.6. Several measurements were taken from the each device so as to be able to average out some of the uncertainties.

.

The cavity widths were measured at approximately $250nm$ above the metal film as shown in figure 4.35. The wings can be seen in figure 4.35. In fig 4.35(b) it is seen that the wings occur at random places along the reflectors within a patch, with random sizes
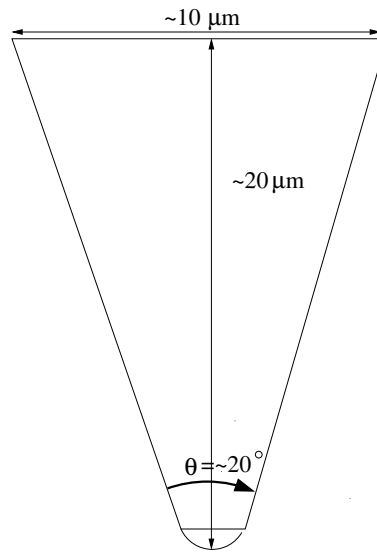
Figure 4.34: The dimensions of the AFM tip. It can be seen that the tip has 20° angle, $10\mu m$ width and $20\mu m$ length.
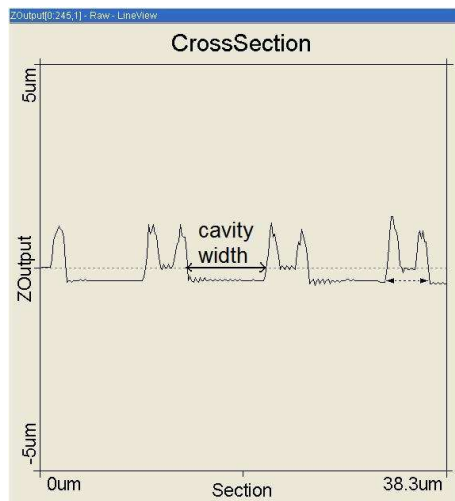
(a)
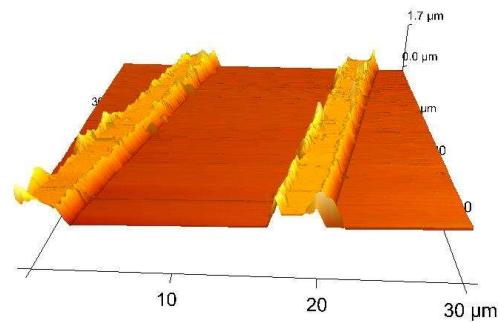
(b)

Figure 4.35: AFM images of the resonator **(a)** Demonstration of the wings using the BPRS-150 resist. **(b)** 3-Dimensional view of the stripes. As can be seen the features have very high wings, which occur at random heights, along the features. The presence of the wings introduce uncertainty to the experimental measurements.

within a certain limit.

4.4.5 Experimental results vs FEM predictions

In order to verify the predictions of the FEM model, it is necessary to test the fabricated prototypes experimentally. In order to do this optical system was set-up and used which could excite SPs and detect the specular and the retroreflections simultaneously. This section focuses on the details of the experimental set-up that was used to verify the predictions made by the simulations in the section 4.4.2. The operation of the system is described followed by the data obtained using this set-up from the fabricated samples.
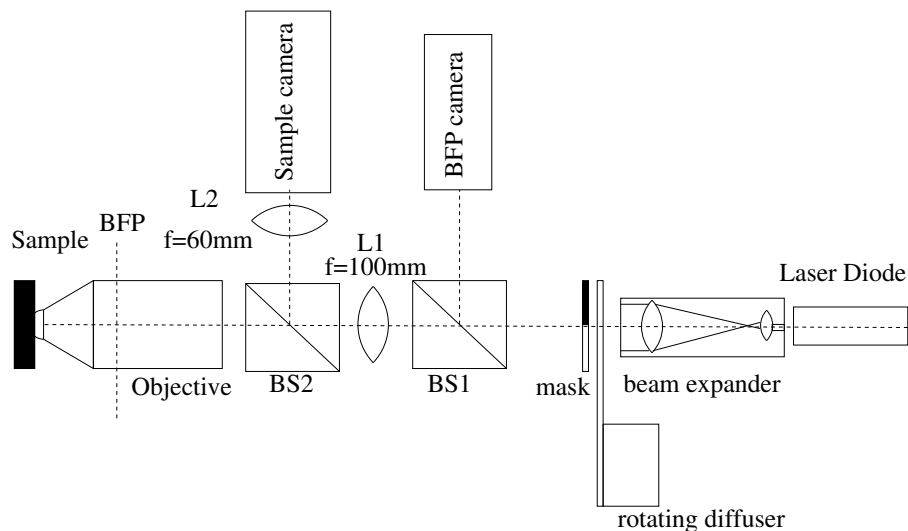
Figure 4.36: The objective based experimental set-up. In this set-up a mask is imaged onto the back focal plane of the objective via lenses L1. The same lens is used to collect light from the back focal plane. Also, an image of the sample is taken using the objective and L2. The use of the mask is to block one half of the back focal plane, so as to be able to collect retroreflection information.

The system involves the use of an objective to illuminate and to subsequently collect information about the reflected spectrum (see section 3.2.4). The range of angles that can be input into the objective from either side depends on the Numerical Aperture(NA).

The schematic of the experimental set-up is shown in Figure 4.36. The system uses a red laser diode, of 639nm wavelength. An oil-immersion (NA=1.45) objective lens is used

Focal Plane

Specular reflection

Optical Axis

Objective lens

Only illumination is from this side
Only the retroreflection
is seen on this side

Mask prevents illumination
to this side of the optical axis

Only the specular reflection
can be seen from this side

retroreflection
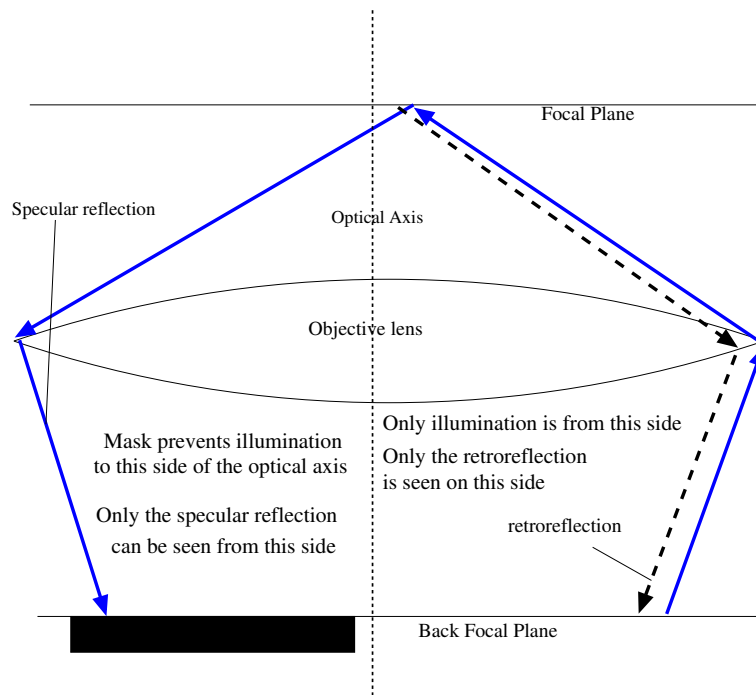
Back Focal Plane

Figure 4.37: Explanation of the operating principles of the masked BFP, objective-based system. A mask is used to block one half of the back focal plane illumination. This will result in the placement of the specular reflection on the side where the illumination is blocked, and the placement of the retroreflection (since there is no specular reflection coming to that side) onto the side that is unblocked. Hence images could be obtained with retroreflection on one side and the specular reflection on the other.

to focus the incoming light onto the sample and to subsequently collect the scattered light in the reflection side of the sample. The immersion oil has a refractive index of 1.518, which allows incidence and reflection angles upto 73° (see equation 3.3). Two cameras are used; one for viewing the sample, and the other to view the back focal plane of the objective(labelled as "Sample Cam" and "BFP cam" respectively). The illumination of the sample is achieved by imaging a plane of pseudo-point sources onto the back focal plane of the objective. This is achieved by the lense L1 in Figure 4.36.

Imaging of the BFP onto the back focal plane camera is done by the same lens L1. The different k-components corresponding to different reflected/scattered directions, emerging from the focal plane will correspond to points of associated intensity in the back focal plane (conjugate plane), which can subsequently be imaged onto the camera. Specifically, different k-vectors will correspond to different radial positions in the back focal plane of the objective (radial position $\propto |k|$).

Figure 4.37 explains what happens at the back focal plane and the focal plane (sample plane). The reflected parallel plane waves, moving away from the focal plane to the objective, are subsequently focused onto points at the back focal plane, which is imaged back onto the BFP camera (refer to Figure 4.37. Subsequently, the sample is imaged using the objective and the L2 onto the sample camera. When the mask is removed from the system, both sides of the objective back focal plane is illuminated with the points sources. In this case the BFP will look like Figure 4.38 for a thin metal film (50nm gold)(see section 3.2.4 for a more detailed explanation).

The oil-immersion (NA=1.45) objective used was a Zeiss Achroplan. The two cameras are 576x768 pixel CCD cameras. A mask was imaged on the BFP when the direction of a propagating SP component is to be detected. A mask is placed at the mask plane in figure 4.36. The mask plane is where the point sources are formed right after passing through the rotating diffuser, and is imaged onto the back focal plane of the objective via L1. Since the locations of these point sources determine the angles of incidence being used, the mask plane is of practical importance.

Figure 4.37 demonstrates the concept used in the detection of the retroreflection. The mask modulates the plane imaged onto the back focal plane, by blocking a half of it. In this case, the detection of retroreflection is of interest. In order to achieve this, illumination
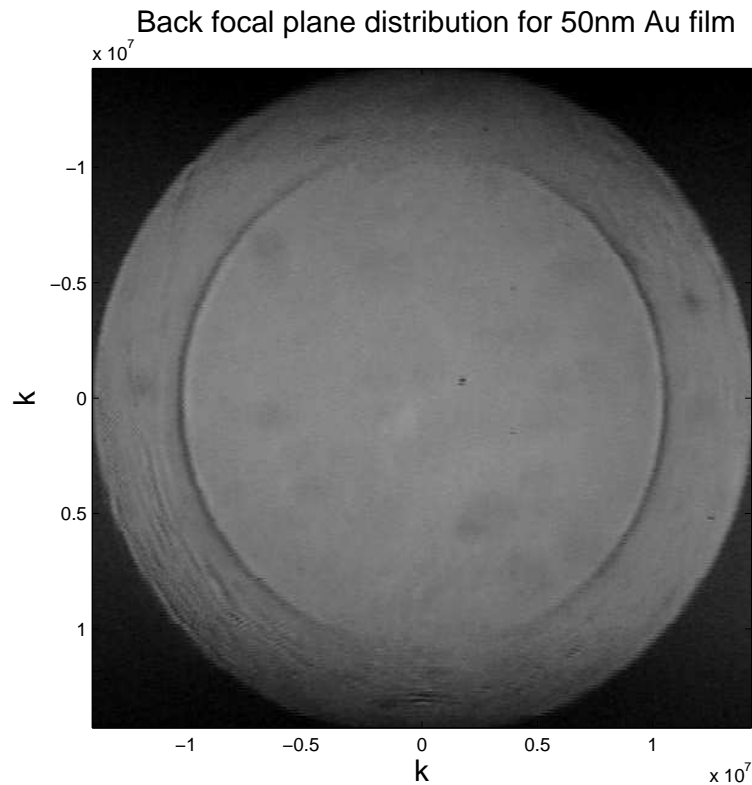
Figure 4.38: Image of the back focal plane of the objective when the sample is a flat gold film(50nm thick). The arc shaped lines are due to the excitation of plasmons. These lines are called the plasmon dip.

to one half of the back focal plane is blocked starting from the optical axis (see figure 4.39 for masked back focal plane). Only the unblocked half of the mask plane illuminates the back focal plane. These points sources correspond to plane waves travelling towards the blocked side, after specular reflection. The specular reflection is focused onto the blocked side which can be imaged as the specular reflection and the plasmon dip. The unblocked side does not receive any specular reflection. The only component that can reflect and be focused onto the unblocked side is the retroreflection. Therefore when dealing with flat gold films, that side will be completely dark. This is demonstrated graphically in in Figure 4.37. Figure 4.39(a) and (b) show images of the back focal plane in the absence and presence of retroreflection, respectively. These images have been adjusted to enhance contrast. It can be seen in figure 4.39(b) that there is a retroreflection where the plasmon dip would be.

x 10⁷ Back focal plane distribution off resonance

x 10⁷ Back focal plane distribution on resonance

(a)   (b)
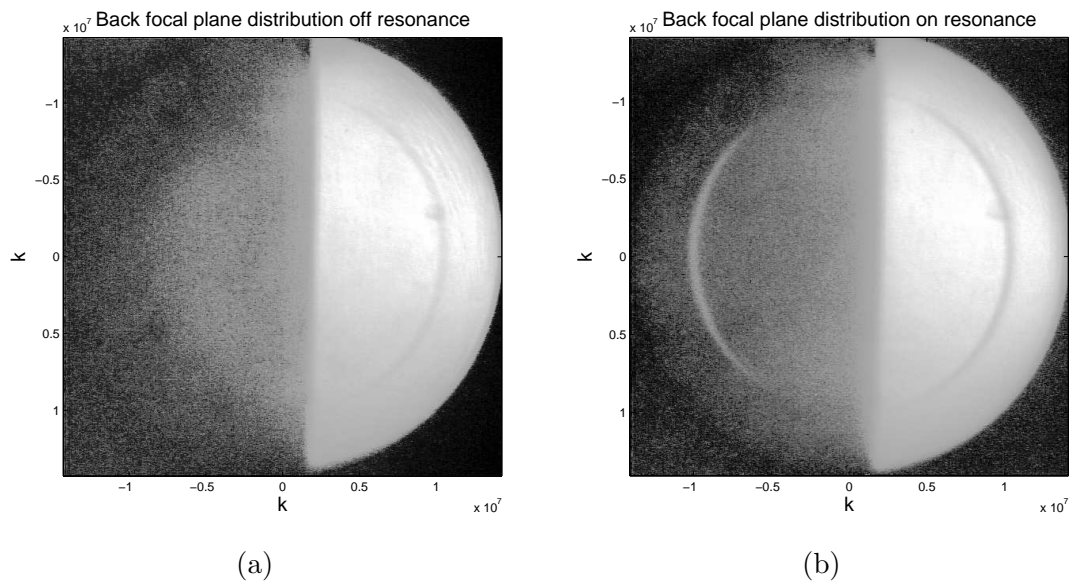
Figure 4.39: The back focal plane distribution of the objective, when a mask is used to observe retroreflection. Both the pictures are used with maximum contrast, so as to be able to emphasise the presence or absence of the retroreflection. **(a)** The back focal plane modified with the mask when there is no retroreflection-cavity off resonance **(b)** The back focal plane of the objective when there is retroreflection, on resonance. The retroreflection is on the left side of the optical axis(retro reflectivity is ≈ 0.02), at the radial location corresponding to the plasmon k-vector. Plasmon dip value (reflectivity) for this figure is 0.7.
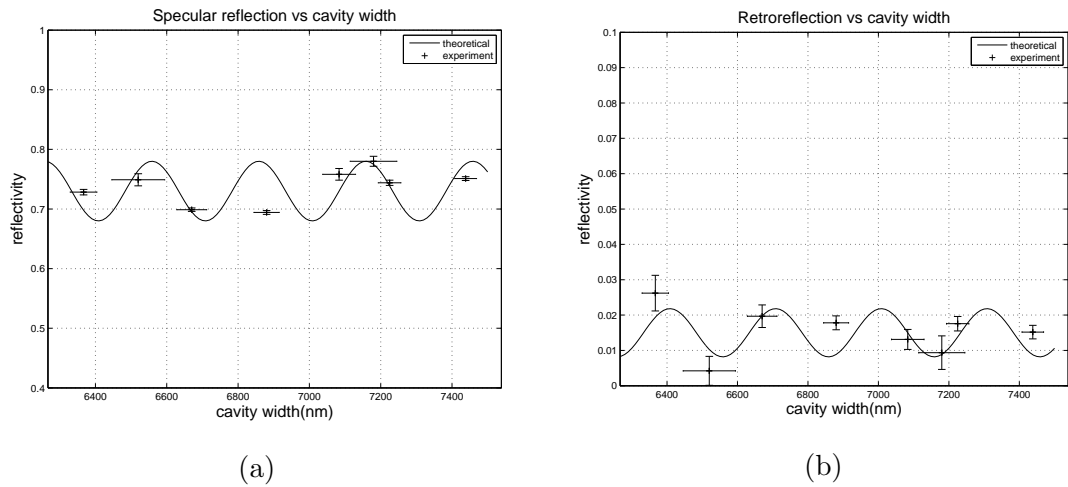
Figure 4.40: **(a)**Comparison of the experimental data and the predictions of the model for the specular reflection. The plots agree within a level of uncertainty at the majority of the time. **(b)** Comparison of the experimental data and the predictions of the FEM model for retro-reflection. The plots and the data points agree for the majority of the time. In both figures the disagreement is emphasised by the undersampled nature of the curve. This is due to the mask fabrication limitations, which enabled 100nm grid size. Therefore, cavity widths are incremented with 100nm steps.

Back focal plane images were taken from each patch, at similar areas to those of AFM measurements. Six images were taken from each patch, so as to be able to assess a measure of variation. The retroreflection data was extracted from the images by using a mask to extract the light only from the corresponding pixels. Such results were averaged for the retroreflection plot. Figure 4.40 shows the experimental results obtained from the fabricated prototypes, as dots with errorbars superimposed on the theoretical curve.

4.4.6 Discussion of results

The experimental results are shown in figure 4.40(a) and (b). The experimental points are plotted on a sinusoid representing the curves in figure 4.20. The theoretical curve has 300nm period.

An initial observation is the the undersampled nature of the experimental points. This is due to the mask fabrication facility, which require the successive cavity width increments to be at least 100nm. The errors due to mask fabrication, photolithography and AFM

measurements move these points away from the theoretical sinusoid. Hence three points for a period is seen to be insufficient to represent the curves clearly. The error from these processes is shown as the horizontal error bars in figure 4.40.

The error/uncertainty due to the mask fabrication is specified by Toppan Photomasks to be $\pm 150nm$. The photolithography process will directly introduce errors on fabricated features. The photolithography process will also indirectly introduce errors in the measurement of the fabricated samples using AFM. The design parameters which are affected by the fabrication include reflector dimensions such as width, side-wall angle (which indirectly effects the cavity width) and reflector height (affects wall reflectivity). There affect the retroreflecting efficiency of this resonator by changing the effective cavity width and the wall reflectivity. The vertical error bars relate to the error on the experimental results obtained from such non-optimal devices.

In section 4.4.3, the individual effects of such design parameters on the operation of the resonators was discussed. In section 4.4.4, the possibility of a common artifact, namely wings,was mentioned due to insufficient cut profile of the resist while doing photolithography.

The presence of the wings directly effect the device dimensions by introducing some side-wall angle, some extra width and height to the features. These parameters have been simulated within the fabrication tolerance limits, and can be seen in figures 4.31, 4.27 and 4.30 respectively. Wings appear when the photo-resist profile is not overcut, enough and the metallisation results in the covering of edges of the resist. Wings occur at unpredictable locations (see figure 4.35(b)) with unpredictable shapes, depending on the evaporation chamber vacuum quality, uneven nature of the resist profile and the consistency. Therefore, they add a level of uncertainty to the actual devices directly.

The uncertainty due to the wings which indirectly affects the accuracy of the AFM measurements include the two mechanisms listed below.

1. The wings represent a sudden and large increase in feature size (refer to figure 4.32(a) and (b)). Sharp features in the sample will be rounded with the tip size (shown in figure 4.35(b)), therefore, the sharpness of the produced image will be limited by the size of the tip [181]. This effect is described in figure 4.41. As can be seen, the route

followed by the tip is not exactly the same as the feature shape, due to the tip size and fetaure size.

2. When there are wings present at the edges of the features, the very sudden change in height will cause an overresponse of the AFM, where the tip will effectively go through a high amplitude vibration. This vibration is dependent on the wing height and will induce uncertainties in the measured feature (see figure 4.42). It affects the cavity width from both sides, corresponding to at least $100nm$ uncertainty in the cavity width (as measured). Figure 4.42(a) shows this vibration. The AFM settings that will yield low vibration on the edge of a feature does not necessarily remain the same on a neighbouring feature due to the randomness of the wings. Hence, it was not possible to get rid of this vibration in all cases.

When many resonators are produced side by side, a change in the reflector width, also affects the cavity width. Whether the features are actually bigger due to fabrication, or if they were measured with uncertainty in the AFM, this has a direct consequence on the cavity width for resonators that are built side by side. Therefore the vertical errorbars does not represent only the experimental errors, but also the imperfections of the fabrication and measurement.

This section serves as a connection between the modelling and the practical aspects of the design and manufacture of photonic/plasmonic devices, where the FEM predictions for a complicated model was shown to match experimental results. The developed design and optimisation technique was implemented and demonstrated. Finite element modelling was used as a tool for the design and optimisation of these devices. The developed technique is used to explore the details of the operating mechanism. The devices are fabricated using photolithography. It is seen that photolithography process has to be carefully tweaked in order to optimise the tolerances. Through the optimisation process, it was possible to see the effects of all the design parameters, hence it was possible to explain the effects of fabrication errors and limitations, such as the wings, even if they were not explicitly modelled via FEM modelling (wings are considered as an increased reflector height and an side-wall angle). It was seen that wings introduced certain amount of error to the experimental results by affecting the features and the AFM measurements.
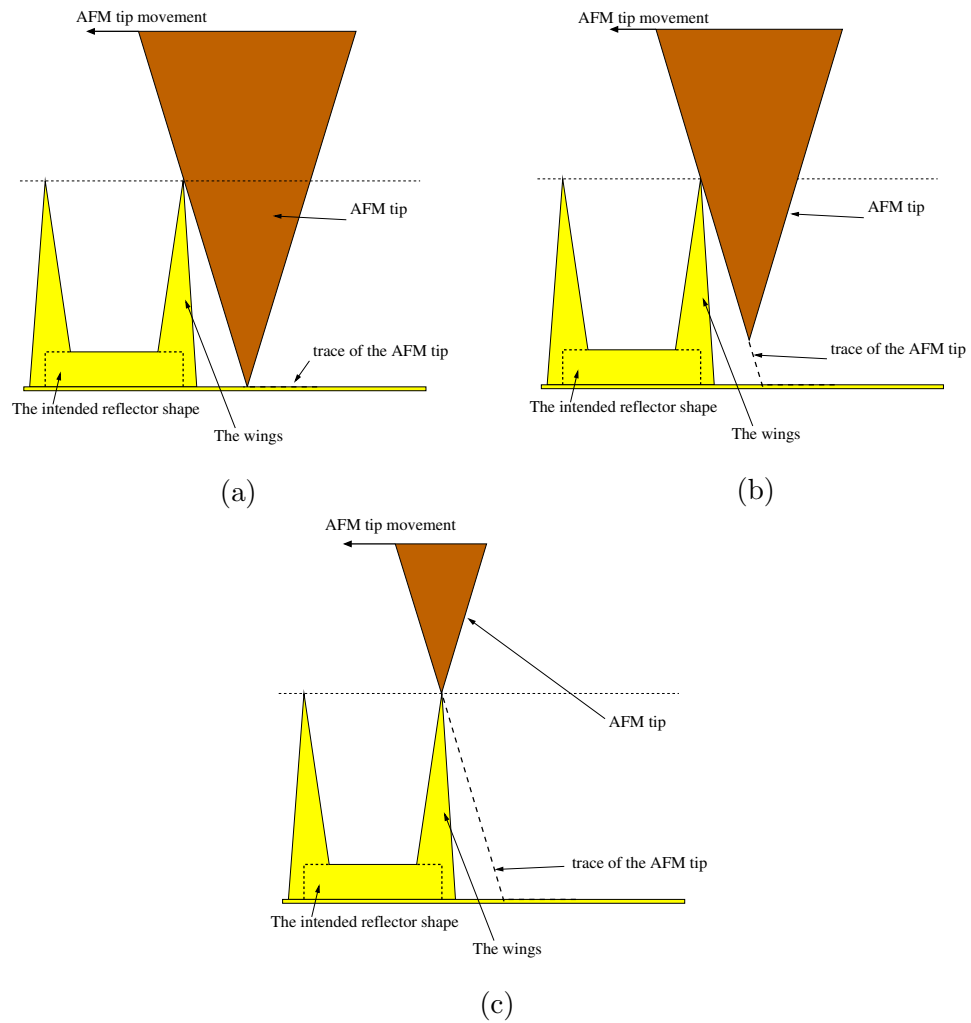
Figure 4.41: The diagram of the AFM tip in the measurements. It can be seen that as the AFM tip moves along from (a) to (c), the very sudden transition in height caused by the wings introduces an error which is dependent on the AFM tip size and the wing height.

(a)   (b)
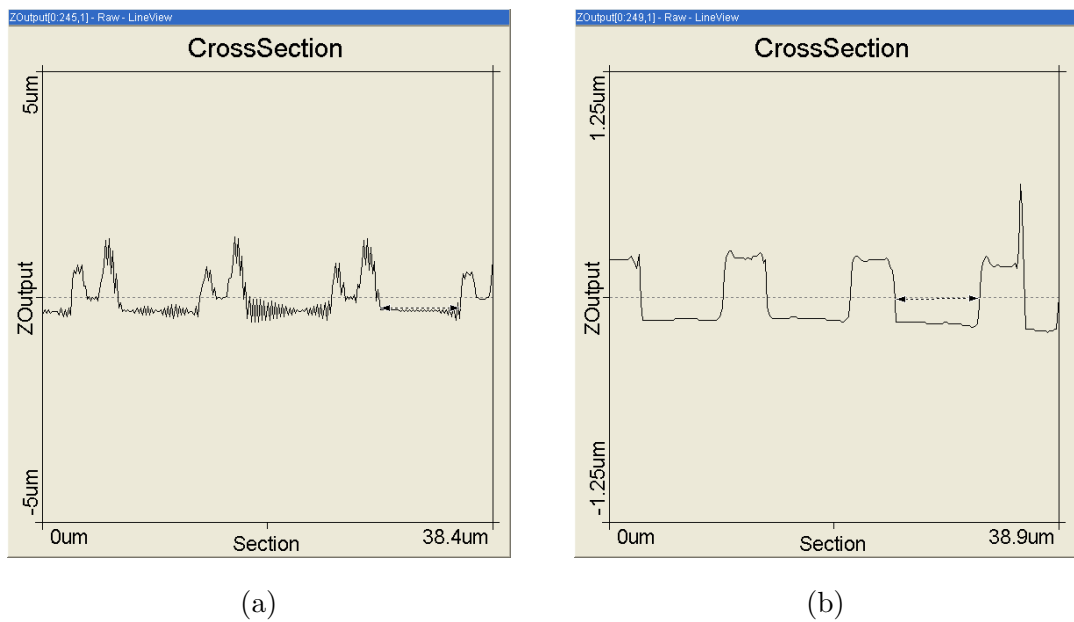
Figure 4.42: The overresponse of the AFM measurements due to the wings.**(a)** The over-response of the AFM measurements. As can be seen the tip is going through a high amplitude vibration as it interacts with the wings **(b)** A different photoresist was used to remove some of the wings. As can be seen, there are no wings present on the features. However, the photolithography process has to be optimised for this resist since only a small portion of the features stayed on the sample.
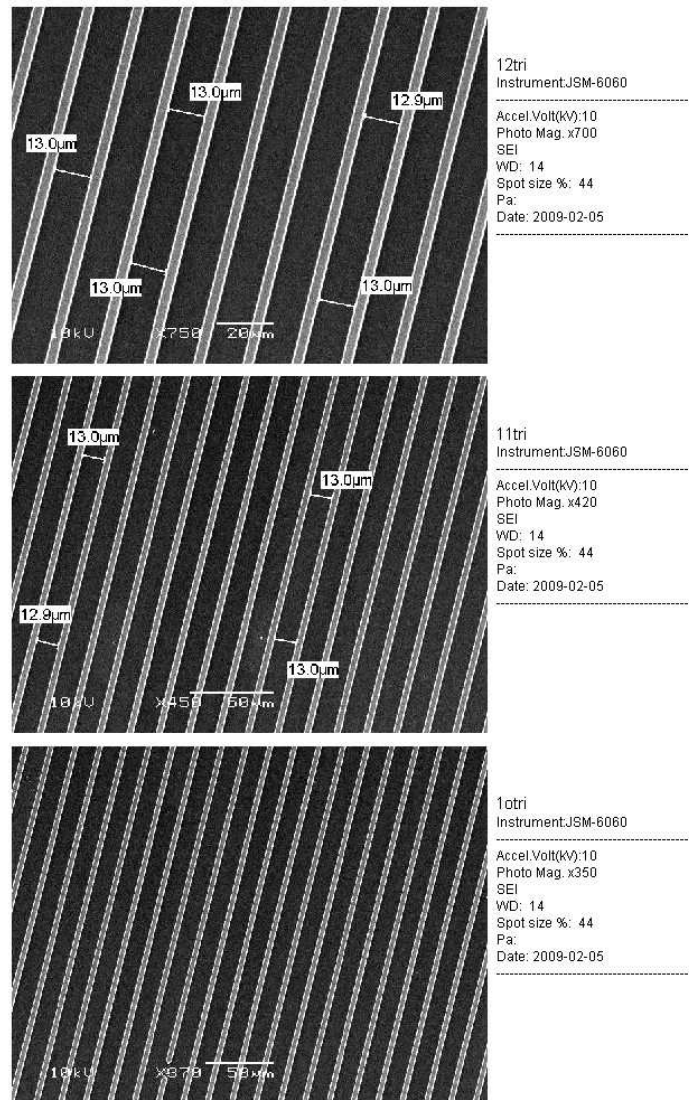
Figure 4.43: The SEM pictures of the samples. As can be seen the wings are still visible on the edges of the samples.

4.5 eCHOT

Ultrasound is a widely used method for nondestructive evaluation [182]. Generation and detection of ultrasound can be done remotely or in contact with the medium of interest. Disadvantages of the contact methods were discussed in section 3.3.

The generation and detection of ultrasound can be performed optically using lasers in a variety of ways [162, 183, 184] including the CHOTs technique [157]. For ultrahigh frequencies, grating fingers on substrates have been considered [159, 160] which were considered in section 3.3.3. The small sub-optical-wavelength gratings required for the detection of SAWs of this frequency range, have not been analysed optically. The exploitation of the resonances and evanescent fields in this scale has not been considered for ultra high frequency SAW detection.

In this section, a novel device, namely eCHOTs, is designed, which can be used in nanoscale ultrasonics to generate and detect ultrahigh frequency SAWs. eCHOTs, rely on a novel operating mechanism, which utilises the interaction between the deformation of the transducer due to the SAW, the optical evanescent fields / resonances involved in sub-optical-wavelength gratings and the thin background film. The interaction between the mechanical deformation, optical resonances and the thin background film forms the key part in this novel operating mechanism, and it yields enhanced signals, where the conventional optical theory predicts lower signals, for a wide range of parameters. The resulting devices are very thin, which makes them very useful in nanoscale ultrasonics. The same transducer geometry can generate and detect these waves, which enables the use of this transducer in pulse-echo techniques as well as pitch-catch.

4.5.1 Definition of CHOTs and eCHOTs

A simple CHOT consists of a simple pattern made on the sample [157]. The pattern modulates the absorption and reflection of the incident light to aid the generation and detection of ultrasound. It is seen throughout this thesis that different transducers are named by using a system of prefixes. The basic idea is a CHOT, which is seen as a transducer attached to a sample, to generate and detect ultrasound. The improved transducers for ultra high frequency ultrasound are named evanescent CHOTs (prefixed with the corresponding

"e", eCHOTs). Each transducer also has a prefix system to determine which part of the transducer system is being referred to. For example, generation part is prefixed with a "g" and the detection part is prefixed with a "d", hence the names, d-eCHOT, d-CHOT, g-eCHOT and g-CHOT.

CHOTs for SAWs are effectively gratings and their operation depends on the energy trade between the 0-order and ± 1 order reflections as the SAW modulates the transducer (see figure 4.46). The presence of the ± 1 diffracted orders depend on the grating period, which is designed to be equal to the SAW wavelength. As this wavelength is reduced, so as to obtain higher frequency acoustic waves, the diffracted orders open up to become evanescent. At this point, the CHOT cannot shift the energy to the diffracted orders , and any shift in energy has to take place through transmission and/or absorption.

The eCHOTs (evanescent CHOTs) does this energy transfer mostly through absorption (see figure 4.57). Figure 4.44 shows the simple operation mechanism of the CHOT and the eCHOT.

4.5.2 Design objectives

The generation mechanism of the eCHOTs and CHOTs are conceptually similar, while the detection mechanisms are fundamentally different from each other. Hence, this section deals mostly with the design issues of the d-CHOT and d-eCHOT. Both of these devices(eCHOTs and CHOTs) are optical gratings of different dimensions. The optical wavelength, λ_o, of interest is 800nm, which corresponds to the wavelength of the laser intended for the experimental verification. CHOTs in the context of this thesis, are diffractive gratings, which have bigger periods than the optical wavelength. CHOTs consist of grating fingers of certain height attached on a 25nm thick background film. They are optimised to be diffractive and work as a localised interferometer.

eCHOTs are smaller, non-diffractive gratings. The interferometric mechanism of the d-CHOTs do not apply to gratings of this size. A novel operating mechanism is thus developed considering the link between the deformation of the transducer as it interacts with the SAW and the consequent modulation of the optical resonances within the transducer. This mechanism requires the presence of a very thin, highly resistive background film ($\approx 2nm$), as opposed to the 25nm thick film in the d-CHOT. The interaction of the

(a)

(b)
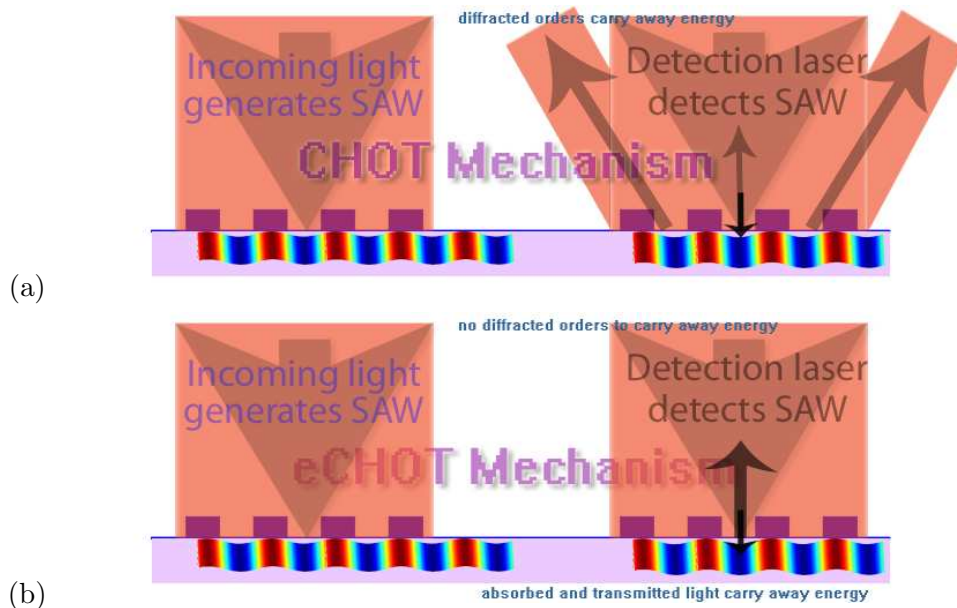
Figure 4.44: **(a)** Demonstration of the CHOT operation diffracted orders carry away the energy, causing a signal, as the SAW interacts with the transducer **(b)** Demonstration of the eCHOT operation. There are no diffracted orders in the eCHOT, therefore the energy must be carried away through transmission and/or absorption, in order to cause a difference in the 0-order reflection. The designed eCHOT does this through absorption.

SAW with this transducer modulates the evanescent fields/resonances, and it moves this thin film in and out of an evanescent field, therefore modulating the absorption/resistive heating in the structure.

Therefore the main design parameter in the eCHOTs is the choice of this thin film thickness, finger height and the transducer material. The resistive heating, transmission and reflection are the only energy mechanisms in the system, therefore these mechanisms need to be optimised to yield a transducer whose deformation yields the best modulation in resistive heating or transmission (this thesis will deal with the optimisation of absorption). The optical wavelength to be used in eCHOTs is also 800nm, and the substrate of choice is glass for the sample fabrication, due to the possible role of transmission of light in the operation of transducers of this size.

4.5.3 Overview of the section

In this section the challenges in the design and optimisation of the eCHOT devices using FEM modelling is described. The eCHOTs are designed by linking the physical displacements caused by SAW to an electromagnetic model, followed by numerous analyses to determine the operation mechanism. The eCHOTs are optimised by varying the design parameters and considering the fabrication limitations. The challenges in the prototype fabrication, which was done by e-beam lithography are mentioned. Experimental verification of such devices could be done with a picosecond ultrasonic system. The fabricated devices could not be measured since the ultrafast laser systems were not available at the time of submission of this work.

The development of the eCHOTs begin with the demonstration of the operation mechanism and shortcomings of the CHOTs. In section 4.5.5, the similarities and the differences between CHOTs and eCHOTs are described. Even though eCHOTs look like a smaller replica of the CHOTs, due to their sub-optical-wavelength size; the optimum parameters, device behaviour and fabrication issues are different. In section 4.5.6, the challenges in the design and optimisation of the d-eCHOT is demonstrated.

The analysis of the d-eCHOTs is followed by the discussion of the generation part of the eCHOT (g-eCHOT). The design issues of the g-eCHOT are briefly explained and metarial choices are demonstrated. However, the FEM modelling of the g-eCHOT mechanism is not covered in detail in this thesis, since the generation is a thermoelastic process and requires the optical model to be coupled with a heat and structural mechanics model.

In section 4.5.8, the fabrication issues of the eCHOT are discussed and the optimum parameters are listed for the fabrication of such a device.

4.5.4 CHOTs

CHOTs combine some ideas from the theory of laser ultrasonics for the generation part [14]. The detection part uses ideas from interferometry in optics, to perform as an in situ interferometer. They are both relatively simple techniques which make the idea very attractive in applications such as flaw detection. CHOTs have been demonstrated in acoustic frequencies around 100MHz for generation and detection [157].

The generation mechanism of the g-CHOT depends on the heating of the sample surface in a spatially varying pattern. The important parameters include the absorption characteristics of the materials considered for the transducer at the desired optical wavelength and the speed of the acoustic wave in the material.

CHOTs have a grating like appearance, and the wavelength of the acoustic signal (λ_a) generated and detected in the CHOTs, is dependent on the period of this grating. In d-CHOTs, the grating fingers act as an in situ interferometric mechanism that introduces a path difference between the parts of the beam reflecting off the fingers and the background film. The passage of the SAW modulates the height of these grating fingers, thereby changing this path difference. This causes the phase of the 0-order reflected components (components reflected from the fingers and the background film) to be different, and their superposition causes a change in the reflected light intensity. The change in the reflected 0-order intensity is balanced by the changes in the intensity of the other diffracted orders. The ideal path difference (where the change in the 0-order intensity is most sensitive to the perturbation in the height of the fingers) is $\lambda_o/4$ [157](λ_0 is the optical wavelength). The finger height in a CHOT is set to be half this path difference, $\lambda_o/8$.

The modelling of the d-CHOTs was done by linking the mechanical SAW displacement to an EM model. This link was achieved by superimposing a sinusoid to the grating structure and running the model for each phase of that sinusoid to simulate a travelling SAW. The reflected field was obtained in the farfield and recorded for each phase of the sinusoid. Figure 4.45 shows the angle conventions used when extracting and displaying the reflected field in the farfield. Figure 4.46 shows the operation principle of the d-CHOT obtained this way. In Figure 4.46 (a) the reflected farfield spectrum is plotted at various sinusoid phases, to see the change in the distribution of energy in the reflected components as the SAW propagates through the d-CHOT. The d-CHOT in this figure was designed for an acoustic wavelength of $1.5\mu m$, when the optical wavelength is 800nm, and the displacement amplitude is 5nm (due to the SAW). It can be seen that there are diffracted orders at around farfield observation angles 147° and 213°. These orders are seen to change in intensity, out of phase with the 0-order reflection (located on the farfield observation angle 180°) as the sinusoid phase changes. The change in 0-order reflected intensity is accounted for by the change in intensity of the higher diffracted orders.
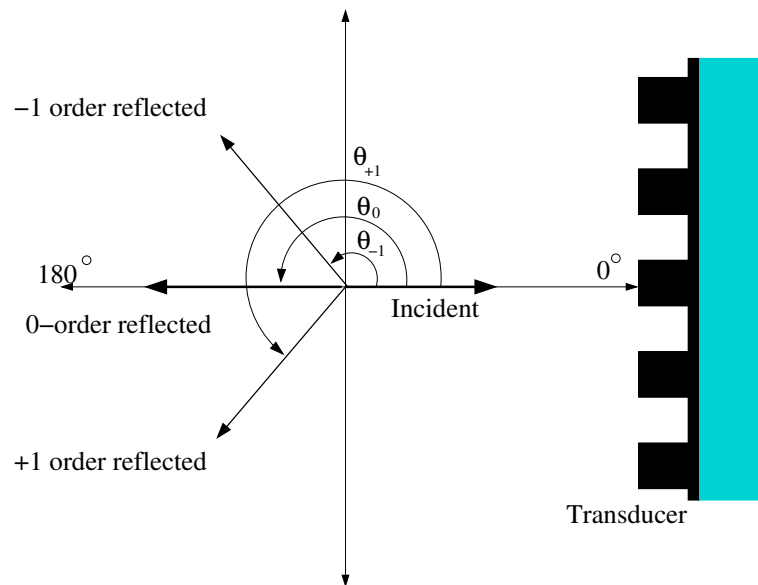
Figure 4.45: The explanation of the conventions about the farfield observation angles in CHOT and eCHOT simulations. The convention is the same in every CHOT or eCHOT model in this thesis. The incident field is considered as a reference at 0° (normal to the surface). Fields travelling in the same direction, eg the transmitted field is also considered at 0° (on the other side of the sample). All the angles are measured anti-clockwise from this reference direction.

(a)   (b)
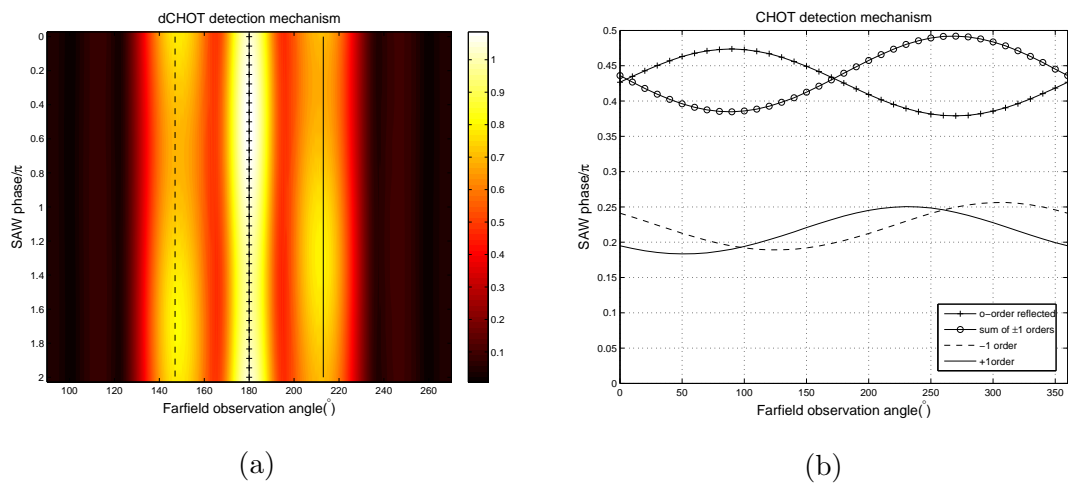
Figure 4.46: The operation characteristics of the d-CHOT. **(a)** The reflected farfield spectrum of the d-CHOT mechanism. The diffracted orders are seen at farfield observation angles, $147°$ and $213°$. The 0-order reflection is also seen at $180°$ **(b)** The intensity of different propagating components vs the SAW phase. This plot shows what happens to the intensity of the different field components as the SAW interacts with the d-CHOT. None of the diffracted orders are individually $180°$ out of phase with the 0-order reflected, however, their sum is completely out of phase with the 0-order reflected field, which suggests that the lost intensity in the 0-order reflection is compensated equally by the diffracted orders.

The signal amplitude in the d-CHOT and d-eCHOT models is defined as half the difference in the 0-order reflected light intensity maximum and minimum as the SAW interacts with the transducer. Therefore, it is simply half the difference between the minimum and the maximum of the reflected field intensity in figure 4.46(b). In both the d-CHOT and the d-eCHOT the minimum value in the 0-order reflected light happens when the sinusoid and the grating are in phase (when the crest of the sinusoid coincides with the middle of a grating finger). Conversely, the maximum value in the 0-order reflected light happens when the sinusoid and the grating are out of phase. With this knowledge, only two phases were run when optimising the transducers.

The d-CHOT was optimised by simulating these two phases of the sinusoid and resolving the model for different combinations of finger heights and the background film heights. For each combination, the phases were simulated and the difference in the reflected field was recorded. Figure 4.47 shows the effect of these parameters on the expected signal from the d-CHOT (made of aluminium, for an acoustic wavelength of $1.5\mu m$ and optical wavelength 800nm). The optimum parameters can then be picked from figure 4.47. For this d-CHOT designed for acoustic wavelength (1500nm), the optimum finger height was 100nm and the optimum background film thickness was $25nm$. The d-CHOTs can be designed using the same finger height and background film thickness for different acoustic wavelengths (for acoustic wavelength grater than the optical wavelength), by simply adjusting the grating period to be equal to the acoustic wavelength of interest.

4.5.5 CHOTs vs eCHOTs

As the desired acoustic frequency increases, the acoustic wavelength hence the grating period decreases. As the grating period gets smaller approaching the optical wavelength, the diffracted orders open towards 0° and 90°. When the acoustic wavelength becomes smaller than the optical wavelength, the diffracted orders become evanescent and cannot propagate. Therefore the interaction of the SAW with the transducer no longer causes the shift of energy from the 0-order reflection to ± 1 order reflections. In this case, the CHOTs need to be re-optimised to cause a shift of energy through either transmission or absorption. eCHOTs are designed and optimised to provide an energy trade that occurs between the 0-order reflected light and absorption in the device.
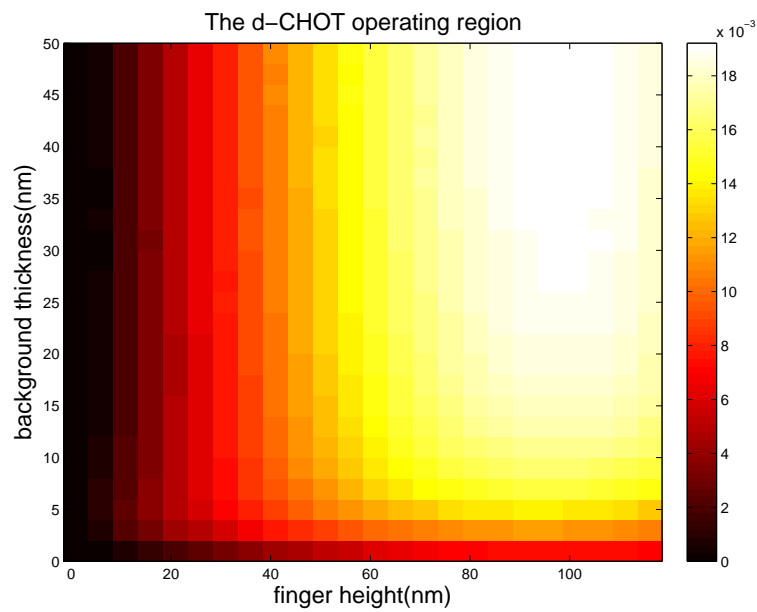
Figure 4.47: The expected signal amplitude from d-CHOT for various design parameters. We define signal amplitude to be half the difference between the maximum reflectivity and the minimum reflectivity as the sinusoid completes one cycle. In this figure y-axis represents the background film thickness and the x-axis represents the finger height. The biggest signal is received when the finger height is around 100nm, which corresponds to $\lambda_{in}/8$, where $\lambda_{in} = 800nm$ is the incident optical wavelength. The displacement amplitude of the SAW was set as 2nm.

eCHOTs can be thought of as an extension of CHOTs into the multi GHz frequency range. The two devices share a similar appearance as they are both grating-like structures. The generation parts of both devices are similar in concept, as they both depend on the spatial contrast in resistive heating applied to the sample (refer to section 3.3.1). However, the detection mechanisms of the two devices are fundamentally different from each other.

d-eCHOTs does not have diffracted orders, the distribution of energy occurs through another mechanism, which is resistive heating and transmission related. The optimum parameters for the d-eCHOTs will be different from that of CHOTs. This will be studied in detail in section 4.5.6, where the eCHOT operation is discussed in detail. The graphical demonstration of its operation can be seen in figure 4.57.

4.5.6   d-eCHOT

The optical set up for the d-eCHOT is simpler than the interferometric techniques that has been mentioned in the literature [18, 185]. The d-eCHOT is designed to operate in subwavelength dimensions.

The general geometry of the eCHOTs can be seen in figure 4.50 (the appearance is similar for CHOTs). The laser illuminates the g-eCHOT and the acoustic surface wave is launched to either side. The acoustic wave travels into the d-eCHOT, and interacts with the grating. A laser illuminates the d-eCHOT and depending on the detection parameters, the energy distribution between reflection, transmission or absorption (resistive heating) changes as the SAW propagates through the transducer. This happens since the displacement of the SAW moves the thin, highly-resistive background film of the d-eCHOT in and out of a localised evanescent field (see figure 4.48). This results in changes in the absorption (resistive heating) (see figure 4.49), transmission and hence the reflected component. The changes in the 0-order reflected component is observed in the farfield and the SAW is detected.

(a) (b)
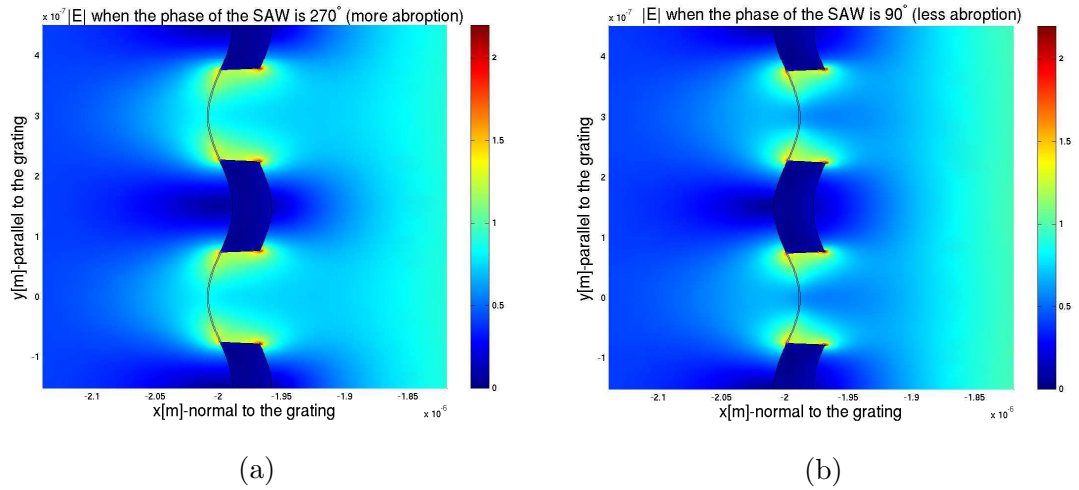
Figure 4.48: The demonstration of the effect of evanescent fields in the eCHOT. **(a)**When the phase of the SAW is such that the evanescent field covers the thin film. The film is thus moved into the evanescent field, which increases the absorption. **(b)** When the phase of the SAW is such that the evanescent field does not cover the thin film. The thin film is thus moved out of the evanescent field, in which case the absorption is less.

(a) (b)
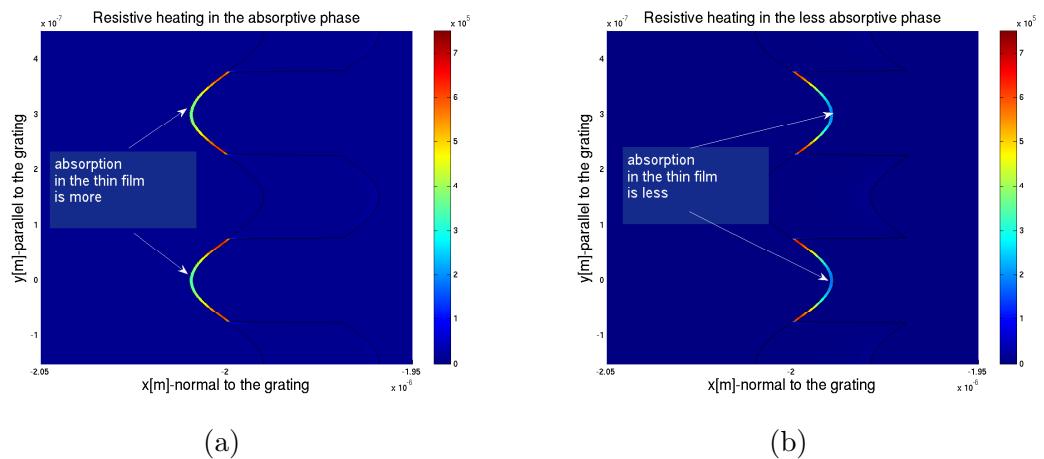
Figure 4.49: The demonstration of the modulation of the absorption in the thin film in eCHOT. **(a)** Resistive heating in the thin film in the absorptive phase. (When the phase of the SAW is such that the thin film moves into an evanescent field increasing the absorption.) **(b)** Resistive heating in the thin film, in the less absorptive phase. (When the phase of the SAW is such that the thin film is moved out of the evanescent field yielding less absorption)
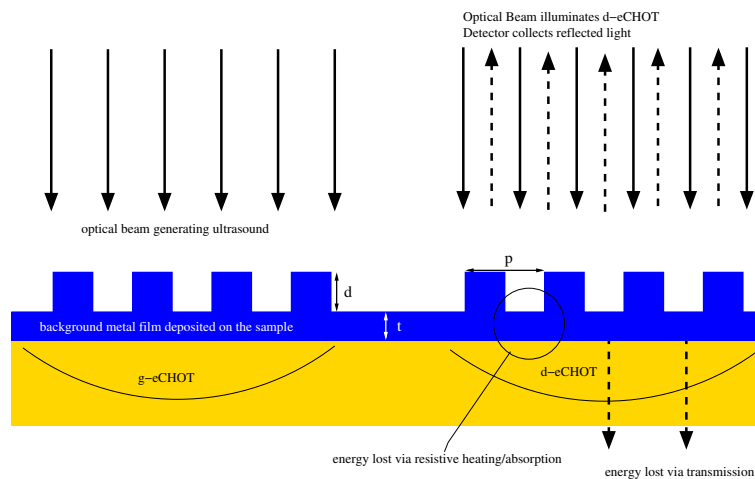
Figure 4.50: The demonstration of the g-eCHOT and the d-eCHOT dimensions on the same sample. Also shown is the basic operation principle of the d-eCHOTs. eCHOT is deposited on the sample as stripes on a thin film background. The g-eCHOT is illuminated, and depending on the parameters, SAW is launched, travelling in both directions. The period (p) determines the wavelength of the SAW generated or detected, respectively. d and t determines the sensitivity of the d-eCHOT. t is the background film thickness which plays a crucial role as a resistive thin film in the operation of this transducer. d is the finger height, which plays an important part in the setting up of the evanescent waves and the resonances key to the operation of the transducer. The acoustic wave modulates the d-eCHOT. When it is illuminated, this modulation causes a change in the intensity in the reflected light as the SAW interacts with the d-eCHOT, hence the acoustic wave is detected. The change in the reflected light is assisted by transmission and absorption. Absorption takes place in the thin background film.

4.5.6.1 Definition of problem and the modelling approach in light of experimental and fabrication limitations

In this thesis, the eCHOTs are designed to generate and detect surface acoustic waves on the same sample [2]. The two parts, the generation and the detection, involve different processes, but some parameters are cross-dependent between the generation and the detection parts when the two devices are to be used on the same sample. When modelling the eCHOTs, this fabrication limitation has to be taken in to consideration. The same transducer geometry needs to both generate and detect the same SAW.

The detection part can be modelled as a coupled mechanical-electromagnetic problem. However, in this thesis, certain assumptions allow it to be modelled as an electromagnetic model. The detection parameters include thickness of the metal fingers and the background metal film together with the material which will do most of the absorption. The main focus of this section will stay on the FEM modelling and optimisation of the detection part.

The eCHOTs are designed for use with a 800nm laser. In practice, a pulsed laser will be used, however from the modelling point of view, time harmonic simulations can be used. This is because the movement of the surface due to the acoustic displacement is much slower than the pulse. Time harmonic simulations are much faster in FEM. The period of the structures is too small compared to the laser spot size, hence a periodic model assumption is valid, where a spatially constant laser intensity is implied.

For a given acoustic wavelength, the optimisation parameters are the finger height, background film thickness, and the eCHOT material. When aluminium transducers are considered, thicker background films have low transmission and low absorption (compared with the maximum, see figure 4.63). The only available components that can carry away the energy of the 0-order reflection are transmission and absorption. Therefore, when aluminium transducers are considered, eCHOTs with very thin background films are expected to perform better than eCHOTs with very thick background films.

Rayleigh wave amplitude decays exponentially away from the surface it is generated. The displacement amplitude is conventionally assumed to dissappear around an acoustic

---

[2]although they could be used separately
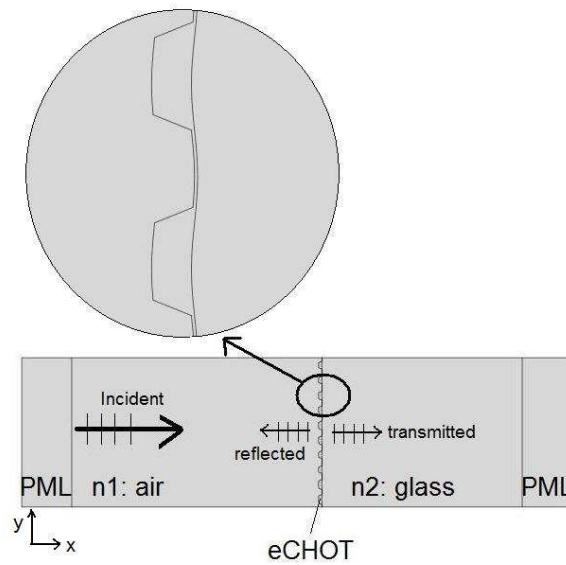
Figure 4.51: The geometry representing the eCHOT. The sinusoidal modulation of the eCHOT due to the acoustic wave is also shown. In this circumstance, where the diffracted orders are evanescent, the incident power is divided between reflection, transmission and absorption (resistive heating). The subdomains from left to right are PML, air, eCHOT(aluminium) glass and PML. The incident plane waves come from behind the left PML and the reflected and transmitted fields are captured in the farfield. The conventions for the farfield observation angles are shown in figure 4.45

wavelength(λ_a) away from the generation surface. In the conventional d-CHOT, this is not a problem since the optimal finger height($\lambda_o/8$) is much smaller than the acoustic wavelength. However, when very high frequency Rayleigh waves are considered, the finger height becomes acoustically significant.

A 50GHz SAW requires a period of 60nm (acoustic wavelength=60nm) to detect Rayleigh waves on glass (v= 3000m/s in glass). If a device was to be designed to operate like d-CHOT does, 100nm finger height would required for the detection of such waves. However, the SAW displacement disappears at this height (away from the generation surface), creating a practical problem. Therefore, eCHOTs are expected to perform worse with very thick fingers. This property however will not be apparent in the FEM simulations, since the displacement will be assumed to be the same all through the structures. This indicates another fabrication metric. The low finger height, combined with thick background film makes it possible to use e-Beam lithography for the fabrication of such devices.

4.5.6.2   The definition of the base model

The reasons behind choosing FEM as the modelling method for nanophotonic/plasmonic devices were discussed before (4.2.4). The basics of the FEM model used for the d-eCHOTs are explained in Chapter 2. They are be briefly mentioned here. The model is periodic in the y-direction, which corresponds to the vertical axis in Figure 4.51. This is due to the assumption that the field has a uniform distribution over the features, and covers all of the features.

Field comes from the left hand side of the metal structure, behind the leftmost PML layer, and modifications to the metal layer in the middle are made through Comsol Script, so as to simulate the passage of an acoustic wave. The modified model is remeshed and rerun. Once the base model is set-up, the optimisation is done by simulating the same situation with different design parameters. The passage of an acoustic wave through such a structure is complicated to model, since it will involve sinusoidal geometry and hence unconventional meshing. Initially, a simple model which would be possible to do with conventional FDTD grids is demonstrated.

4.5.6.3 Assumptions in the transducer models

In all d-eCHOT models, the displacement amplitude is assumed to be the same along the normal to the gratings (x-direction). Therefore, no matter what height the metal fingers are set to, they are acted on by the same displacement amplitude. This assumption is sensible when the thickness of the device considered is smaller than an acoustic wavelength, which is around 300nm for this case. Although, practical reasons make thicknesses around 100nm acoustically significant, due to the increased reflection at thicker fingers.

An issue with modelling a high frequency surface acoustic wave, is that the amplitude of such displacements is in the range of picometers [159]. These are actually extremely small compared to the wavelength of the optical illumination considered. The signal for very small amplitudes will be very noisy(numerically), as the displacement amplitude gets very small compared to the simulation domain. In order to show this the same model is run with different displacement amplitudes. This implies that the maximum and the minimum elements sizes were constant throughout the the models. The results are seen in figure 4.52. As can be seen, smaller displacement amplitudes results in signal with more noise for a given number of mesh elements (which was constant for all displacement amplitudes). Therefore, an assumption is made that the displacement amplitude and the received signal were linearly related for all displacement amplitudes (see figure 4.62). This means that smaller amplitudes will give linearly smaller signal amplitudes and vice versa. Therefore the simulations were run with higher displacement amplitudes than expected in practice, and the signal levels were linearly related to the actual values. Using this assumption, it was possible to reduce the number of mesh elements for a given transducer and decrease the simulation times. This assumption is validated by simulating a range of displacement amplitudes and measuring the difference in reflectivity (twice the signal amplitude) in figure 4.62.

4.5.6.4 The simple transducer models

When a surface wave is travelling through the transducers (d-eCHOT and CHOT), it will cause sinusoidal displacement in the background material and the grating structure, namely the "grating fingers". When the background film and the grating fingers are thick,
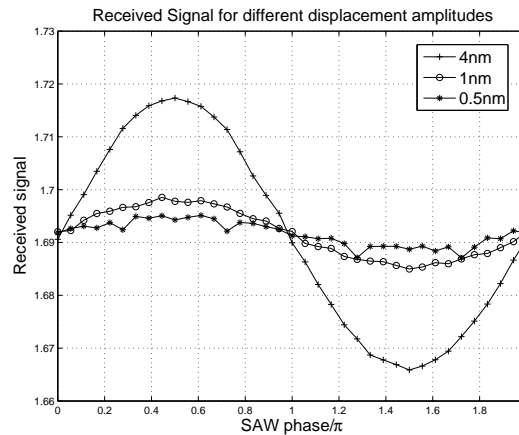
Figure 4.52: The demonstration of the effect of numerical noise on the signal for small displacement amplitudes. The same model with the same number of elements are run for different displacement amplitudes. It can be seen that as the displacement amplitude is reduced, the obtained signal amplitude (reflectivity) decreases and becomes more noisy.In these models, a transducer with 100nm finger thickness and 25nm background film thickness is used. The maximum mesh element size was set to 53nm($\lambda_o/15$) while the minimum element size was around 8nm (in the background film).

the interaction of the SAW with the transducer can be approximated by simply changing the height of the fingers. This means that the SAW is assumed to cause an effective displacement on the finger height, and the displacement in the background material is negligible. In these models, a grating of a certain finger height and background film thickness is modelled. The reflectivity is extracted and stored from this model. The passage of the acoustic wave is then simulated by an effective increase in the finger height. Therefore, another model is built with finger heights increased by the SAW amplitude. The extracted reflectivity from this model is subtracted from the stored value (from the pervious model) in order to obtain the change in the reflectivity associated with the passage of the SAW. This could be done more concisely by modelling transducers with the finger heights and background film thicknesses of interest, and extracting the reflectivity from all models. This will result in a table such as figure 4.54(a). Then this table is differentiated along the finger height dimensions to obtain the change in reflectivity with respect to an effective increase in finger height(due to the passage of the SAW). This is shown in figure 4.54(b).
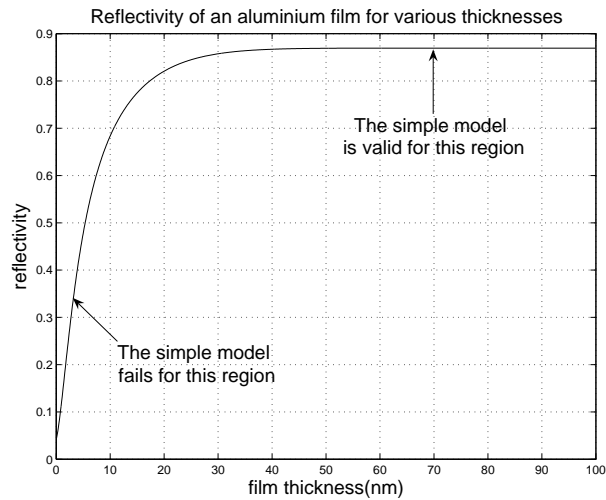
Figure 4.53: The reflectivity of an aluminium film for various thicknesses. The simple model fails for the region where a small increase in the amount of metal causes a large change in the reflectivity. The simple model is valid when the reflectivity can be considered as constant for a small change the amount of metal. This curve was obtained by modelling various thin films using Comsol

When the background film and the grating fingers are very thin, the same amount of increase in the finger height causes a significant increase (percentage) in the amount of metal in the model, hence the reflectivity of the transducer goes up significantly. This is shown in figure 4.53, where the reflectivity of an aluminium film is plotted for various thickness values. The simple model is valid only around the regions where a small change in the amount of metal does not change the reflectivity significantly. Therefore approximation of the interaction of the SAW with the transducers as an effective change in grating finger thickness is unrealistic and yields incorrect results for thin films. For thick films, this simple approximation is valid.

A script was written to model the effects of approximating the interaction of a SAW with the transducer (script listed on appendix A.3). Using this simple model, the finger and the background thicknesses of the transducers were varied, and the reflectivity of the transducers were recorded. In order to see the effects of the parameters, the finger heights are changed for a set of background thickness values and the reflected farfield spectrum are calculated for each combination of background thickness and finger height. For each background thickness, several finger height simulations are made and a table is built (see

figure 4.54(a)). At every background value, the difference is taken along the dimension of the finger height (see figure 4.54), so as to see the amount of change in reflectivity, with respect to a change in the finger height. Therefore, a metric of reflectivity change is obtained at each finger thickness, for all values of background film thickness. This corresponds to a measure of the expected signal amplitude at the associated grating finger and background film thickness values.

The figure 4.54 shows the expected signal amplitude from a simple transducer model (designed for 300nm acoustic wavelength and 800nm optical wavelength) for a range of background film and finger thickness values. Looking at figure 4.54(b) it can be seen that there is a region where there is very high levels of signal amplitude, when background thickness and the finger heights are very small, labeled as the "unrealistic region". This area corresponds to the region where the interaction of the SAW with the transducers cannot be approximated by an effective change in the finger heights. When the background film and the grating fingers are very thin, the interaction of the background film with the SAW has to be considered as well as the interaction of the fingers. A high signal amplitude, when the background and finger thickness values approach to zero, indicates that better devices with higher signal amplitudes can be obtained if no device is there. This does not reflect the realistic situation.

This approximation is useful since it makes the models much simpler, even though it yields unrealistic results for very small device dimensions. The model resulting from such an approximation involves rectangular geometry and can be discretised using regular rectangular grids. Modelling techniques such as FDTD would be able utilise this approximate model. In order to keep the model accurate, however, thin background film and grating fingers would have to be avoided. Therefore a more complex, "improved" model is required which additionally considers the deformation of the background film.

The simple transducer model was improved by taking into account the actual deformation of the transducers, rather than approximating it as an effective change. This was done by actually modelling a sinusoidal geometry. The sinusoidal geometry would not be efficiently resolved within practical computational limits using regular rectangular grids. This can be seen as a demonstration of the fact that such devices would be impractical to design and optimise using methods such as FDTD. FEM is able to resolve thin films,
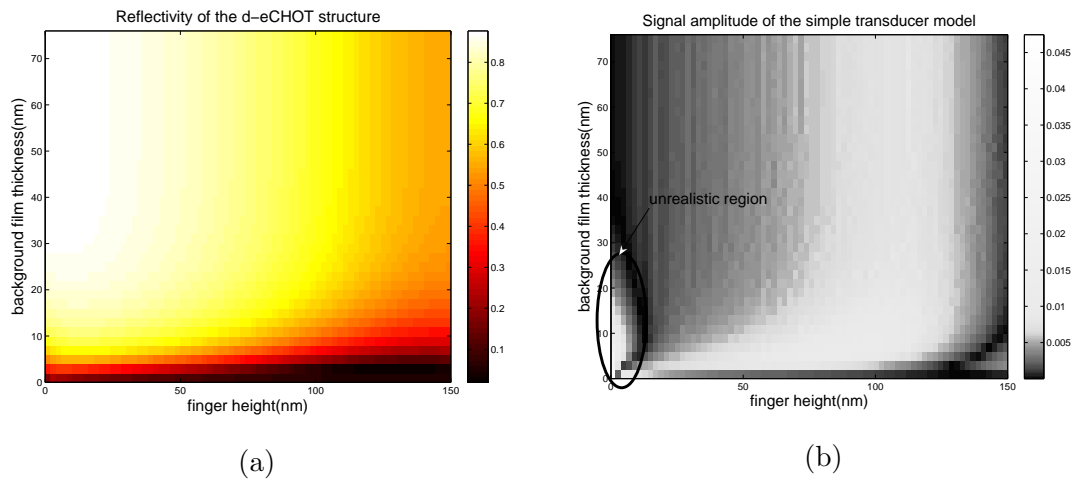
Figure 4.54: **(a)** The observed reflectivity at different values of background film and finger thickness from the simple transducer model. The model is designed for 300nm acoustic wavelength and 800nm optical wavelength. **(b)** The expected signal amplitude from the simple transducer model at different finger and background film thickness values. The colourmap of this image has been changed so as to increase the contrast. This image has been obtained by taking the difference of the consecutive reflectivity values along the horizontal direction in (a). This corresponds to taking the reflectivity difference of two devices with different finger thickness(difference of 2nm in this simulation). The average modulation due to sinusoidal displacement is thus represented by increments in finger thickness. There is a region of very high signal amplitude which corresponds to devices with thin background film and grating fingers. In this region, the SAW interaction with the transducer can no longer be approximated as an effective change in finger thickness.

and complex geometries involving sinusoidal deformation, while keeping a moderate simulation size. This reduces the simulation times and makes the optimisation of such small transducers possible.

4.5.6.5 The improved transducer model

In order to verify what happens for the unrealistic values shown in figure 4.54, a more realistic model was created where a sinusoid was superimposed on the structure. Resulting model looked like Figure 4.51. In these models, the propagation of the Rayleigh wave through the structure is approximated by a sinusoid with changing phases. As the phase of the sinusoid is varied, the reflectivity is observed (refer to Figure 4.57). The difference between the maximum and the minimum values of the reflectivity are recorded for one cycle of sinusoid. The signal changed sinusoidally with the phase of the sinusoid, hence the difference between the maximum and the minimum of the reflectivity corresponds to twice the signal amplitude.

Figure 4.55(a) and (b) show the results of the simulations for the simple and the improved models for the region where the assumptions for the simple model is valid, i.e. the displacement can be approximated by simply changing finger heights. It can be seen that the two methods yield similar signal amplitudes especially for high background and finger thicknesses. However, for small values of finger and background film thicknesses, the models start to disagree.

Thin background films and grating fingers (the region marked in figure 4.55(b)) are investigated further and in more detail for the possibility of more sensitive transducers, using the improved model. The result is shown in figure 4.56. In this region it can be seen that the device exhibits high sensitivity in a region where the simple diffraction theory describing the behaviour of the regular CHOT [157] indicates low sensitivity. This is an attractive region not only due to the increased signal amplitudes but also because thin background film and grating fingers are less perturbing to the acoustic waves.

Looking at figure 4.55(b) and 4.56, two distinct regions are recognised where the devices seem to operate. First one is the region corresponding to thin background films and grating fingers and it yields high signal amplitudes. The second region, even though it yields smaller signal amplitude, is the analogue to the d-CHOT, where the finger heights are
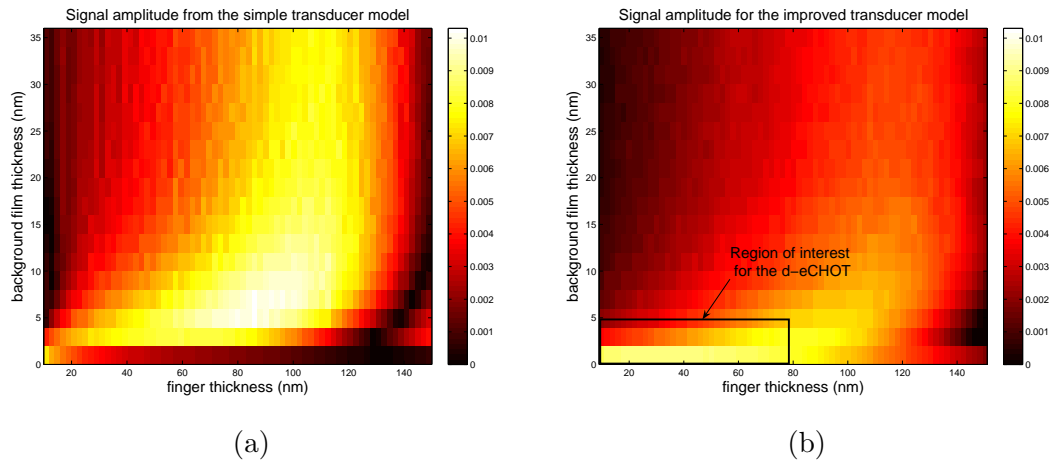
Figure 4.55: **(a)** Signal amplitude from the simple transducer model. **(b)** Signal amplitude obtained from the improved transducer model. For both figures, the finger height is taken from 10nm. After this value the simple model yields similar results to that of the realistic model. It can still be seen that, for small background film thickness values, there is a significant mismatch between the simple and the improved models. Also shown in (b) is the interesting region that is further investigated in figure 4.56

arranged so as to create an interferometric effect. This region incorporates finger heights around $\lambda_o/8$.

In order to determine the mechanism involved in this region, a device is picked from the eCHOT region in figure 4.56. Additionally another device is picked from figure 4.55(b) with high background and finger thickness values, specifications matching to that of CHOTs (a graphical demonstration of the device operation regions are also shown in figure 4.61). One device from each region (eCHOT and CHOT) will be considered, and their operation will be discussed in detail, in terms of the field components involved.

The reflected farfield spectrum of the eCHOT, is shown in figure 4.57(a). The 0-order reflection is seen to change in intensity as the SAW propagates. Figure 4.57(b) shows the related field components, as the SAW interacts with the eCHOT. The field components are obtained by extraction from spectra such as the one in figure 4.57(a). The incident, transmitted and reflected fields are readily available from these images, where the absorption is calculated as equation 4.1, where I is a measure of intensity[3]

---

[3]calculated using the appropriate normalisation considering the refractive index of the media
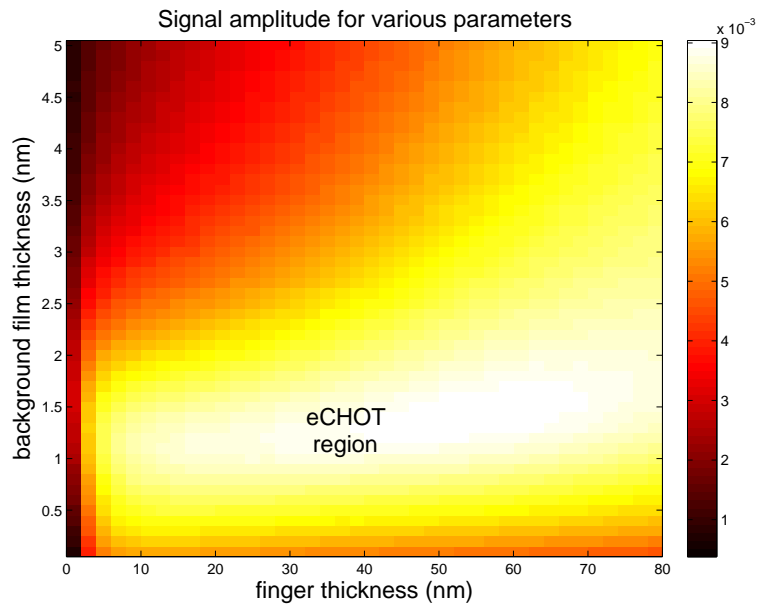
Figure 4.56: Signal amplitude from the transducers at low background film thickness values, produced with the improved transducer model, at parameters for which the simple transducer model is not valid. The displacement amplitude used was 2nm. The transducers are designed for 300nm acoustic wavelength (grating period is 300nm) and 800nm optical wavelength. It is seen that for a small range of background film thickness values, there is enhanced signal amplitude. This is region referred to as the "eCHOT region". d-eCHOTs can be fabricated with the parameters in this region to yield higher signal amplitudes, for the acoustic and optical wavelength considered. It is seen that if there is no background film, then the signal amplitude drops. This suggests that conventional methods of depositing metal fingers on a substrate (without the background film) for the detection of SAW yields smaller signal amplitudes than the eCHOTs designed from this region.

$(|E|^2)$.

$$I_{absorbed} = I_{incident} - I_{reflected} - I_{transmitted} \tag{4.1}$$

In this figure it can be seen that the reflected energy and the absorbed energy is out of phase with each other. The shape of the transmitted energy curve is not very clear in this figure, since the modulation on the transmission is very small relative to the other field components (shown in table 4.4), but is it plotted in figure 4.58.

The detailed analysis of the field components are shown in table 4.4. Each device is tested for two different displacement amplitudes, 2nm and 5nm, in order to show the effect of displacement amplitude on the signal. The field components are retrieved from the reflected and the transmitted farfield spectra. At the desired farfield observation angle (corresponding to 180° according to figure 4.45), the field is observed as the phase of the imposed sinusoid changes.

The square of the total field is taken to obtain measure of intensity. This is followed by the normalisation of all extracted components with the incident intensity. The difference between the maximum and the minimum of the components are noted as "the range" and the average value is noted as "average". The reflected intensity is noted as the signal, therefore the range of the reflected intensity corresponds to twice the signal amplitude. The range of the transmitted intensity and the absorption is dependent on the displacement amplitude for both devices. However, the average values of the field components depend on the device parameters rather than the SAW properties.

In eCHOT, the thin background film yields higher average values for transmission and absorption. The change in the reflected intensity is mostly compensated for by the absorption, rather than the transmission. The transmitted intensity remains with small range, noisy at small displacement amplitudes, therefore even though it is modulated by the SAW, and is affected by the displacement amplitude, it plays a small role in the energy trade mechanism.

Observing table 4.4 it can be seen that the displacement amplitude only increases the range of the field components. In the d-CHOT, the background film is thicker. The transmission, is nearly zero, but is in phase with the absorbed energy and out of phase
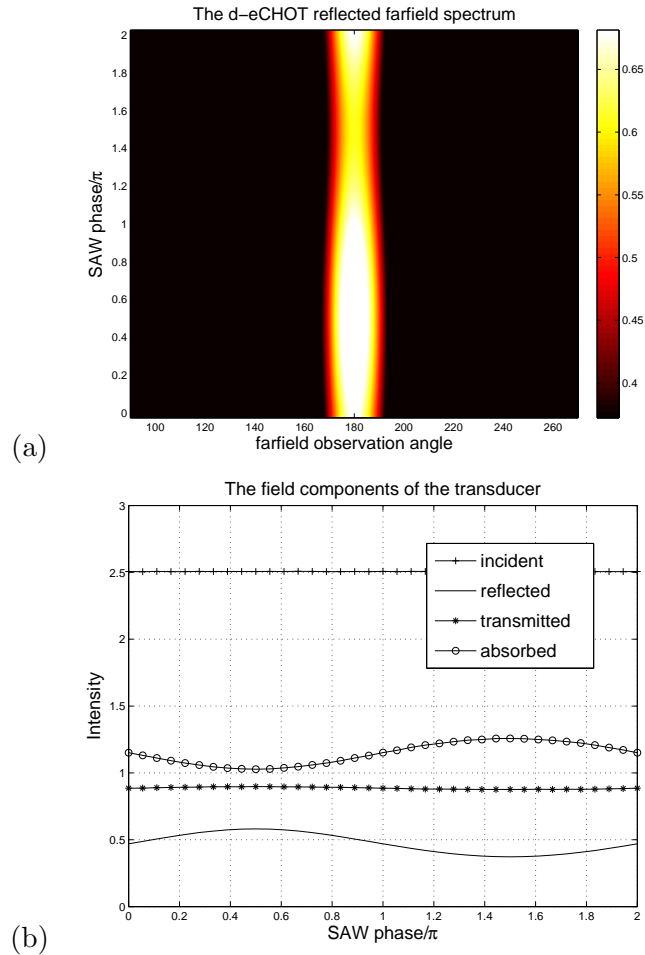
Figure 4.57: **(a)** The reflected far-field spectrum of the d-eCHOT with background film thickness 2nm and finger height 40nm. (Designed for 300nm acoustic wavelength and 800nm optical wavelength).The displacement amplitude used is 10nm. It can be seen that there is no other major component carrying energy away from the d-eCHOT in the reflected direction, other than the one in 180°, i.e. specular reflection(0-order). **(b)** The components of light involved in the operation of the d-eCHOT. These curves are obtained by extracting the field amplitudes from the farfield spectra, followed by the necessary normalisations. The absorption is calculated by subtracting the reflected and transmitted intensity from the incident. It can be seen that as the SAW propagates, the incident intensity is constant, whereas the reflected intensity (the signal) and the absorption are 180° out of phase with each other. The transmitted intensity also has a sinusoidal nature out of phase with the absorption, however, the amplitude of the sinusoid is too small to be visible on this axis. This plot shows that the energy is taken out of the specular reflection via absorption.
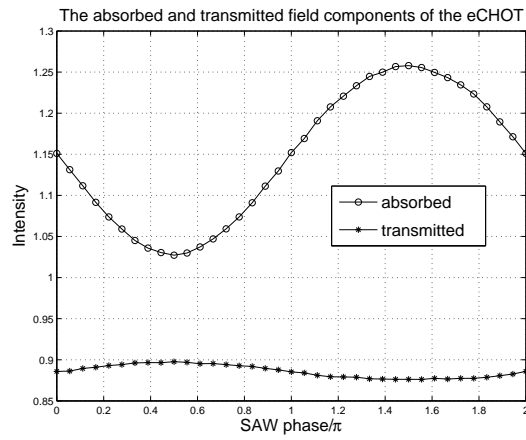
Figure 4.58: The transmitted and the absorbed field components of the d-eCHOT when the background film is 2nm, finger thickness is 40nm, displacement amplitude is 10nm, optical wavelength 800nm and acoustic wavelength 300nm. As can be seen, the fields are 180° out of phase with each other. All of the field components for this device is shown in figure 4.57

	Displacement amplitude= 2nm				Displacement amplitude= 5nm		
		range	average			range	average
eCHOT	**absorption**	0.0177	0.4550	**absorption**	0.0445	0.4555
	transmission	0.0022	0.3289	**transmission**	0.0041	0.3294
	reflection	0.0161	0.2162	**reflection**	0.0404	0.2150
		range	average			range	average
CHOT	**absorption**	0.0115	0.3154	**absorption**	0.0273	0.3163
	transmission	2.5e-4	0.0140	**transmission**	5.8e-4	0.0142
	reflection	0.0117	0.6706	**reflection**	0.0279	0.6696

Table 4.4: The contribution of the field components to the operation of the d-eCHOT and the d-CHOT. The involved field components in the two devices considered. eCHOT device considered has a background thickness of 2nm and finger height 40nm. The CHOT device has a background thickness 17nm and finger height 100nm. The field components considered are absorption, transmission and reflection. These values are obtained by extracting the incident, reflected and transmitted intensities from the corresponding farfield spectra (reflected spectrum shown for the eCHOT device in figure 4.57)

with the incident. The distribution of energy in the reflected component is assisted by both the transmission and the absorption.

The average values of both the absorption and the transmission is less than they are in eCHOT. This is as expected outcome since the thick background film (17nm) transmits and absorbs less than a thin background film (1nm) (see figure 4.59).

The d-CHOT mechanism is as expected and is similar to that of nonevanescent d-CHOTs, with a difference that instead of diffracted orders, transmission and absorption handle the energy gain/loss from the reflection. For thick background films, when the absorption is minimal and transmission is zero, the d-CHOT is not expected to perform well. Because, an interaction with the SAW does not modulate the absorption in the thick background film as well to cause a signal.

In the d-eCHOT, the background film is thin. This allows the device to transmit and absorb more. The transmitted intensity curve is 180° out of phase with the absorption (see figure 4.58), and in phase with the incoming energy. This shows that the device behaviour is dominated by absorption, and transmission does not play a part in the mechanism. Therefore the SAW merely moves the resistive thin film in and out of an evanescent field, causing a change in the absorption. When the thin film moves into the field, there is more absorption and the transmission decreases as well as the reflection, vice versa.

Therefore, despite their similar appearance, the operation of the eCHOT is fundamentally different from CHOT. It does not depend on any interferometric effect due to the path difference between the reflected field components. Rather, the operation depends on the displacement of a highly resistive thin film in and out of an evanescent field, hence increasing and decreasing the resistive losses. The effect can be seen from table 4.4, when the absorption range increases with displacement amplitude. It can also be shown by extracting the time-averaged resistive heating from the models (see figure 4.60). Figure 4.60(a) is obtained by extracting the resistive heating in the metal structure, and taking the average of the resistive heating for each phase of the sinusoid. The resistive heating is calculated in Comsol using equation 4.2 [186].

$$Q_{av} = \int_V \mathbf{J} \cdot \mathbf{E} dV \tag{4.2}$$
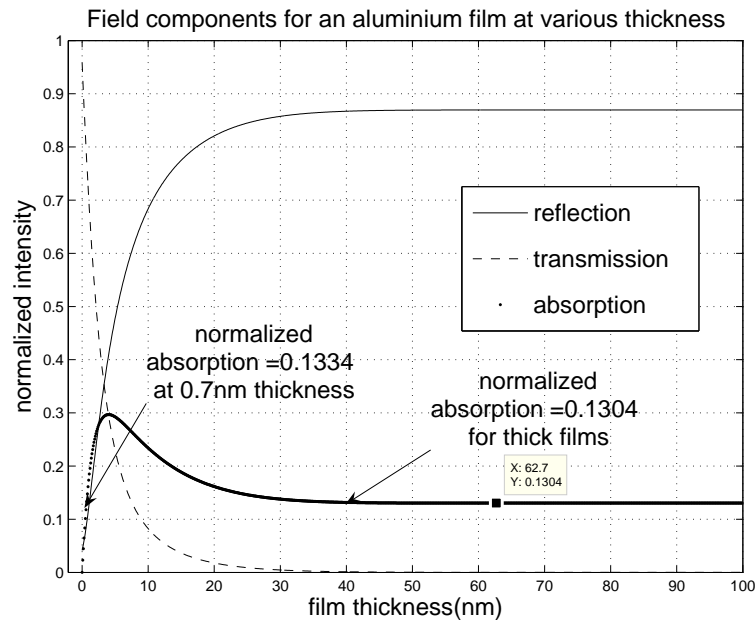
Figure 4.59: The reflected, transmitted and absorbed field components for aluminium films of various thicknesses. It can be seen that, for film thicknesses greater than 0.7nm the film absorbs more and transmits more than very thick films.

Where, **J**: Current density

E: Electric Field

Then the difference in resistive heating is calculated as a percentage of the resistive heating when the displacement amplitude is zero. It can be seen that high displacement amplitudes, result in a bigger range in resistive heating. Figure 4.60(b) shows the difference in extracted reflected and transmitted intensity as a percentage of the incoming intensity. Also shown is the calculated absorption. It can be seen that the increase in absorption has a bigger slope than the increase in reflection. This is a result of the transmission and absorption being out of phase with each other.

Figure 4.61 shows the operating regions of the transducers designed for acoustic wavelength 300nm and optical wavelength 800nm, with 2nm displacement amplitude. The regions corresponding to d-eCHOTs or d-CHOTs are labeled. It is seen that eCHOTs offer higher signal amplitudes than the CHOTs.

Figure 4.62 compares the two devices considered in various displacement amplitudes. As can be seen, the d-eCHOT has more signal amplitude for all displacement amplitudes

Figure 4.60: **(a)**The mean time-averaged resistive heating in the d-eCHOT models (extracted from the background film). It can be seen that the resistive heating is increasing with the displacement amplitude, as a percentage of the resistive heating when there is no SAW(when the displacement amplitude is zero). It can be seen that the increase is linear with the displacement amplitude. This demonstrates the contribution of the resistive heating to the device operation **(b)** The difference in the reflected and the transmitted field intensity of an eCHOT as a percentage of the incident intensity at different displacement amplitudes. The reflected component is essentially twice the signal amplitude, which grows linearly. The transmitted component has a smaller slope. It can be seen that the amount of change in transmission is not big enough to compensate for the increase in the reflected component. This further demonstrates that the absorption/resistive heating is the main operating mechanism in d-eCHOT. Also shown is the calculated absorption from the model. It has a bigger slope than the difference in reflection, since the transmission is out of phase with the absorption.

considered. This figure was obtained by varying the displacement amplitude and recording the signal amplitude for d-eCHOT and d-CHOT. Also shown is the linear rise of the signal amplitude with the displacement amplitude, which shows that the sensitivity of the devices (sensitivity corresponds to the slope of the curves in figure 4.62) stays the same for all displacement amplitudes considered. This also verifies that the assumptions used for the linear scaling of the displacement amplitudes. It is seen that d-eCHOTs have higher signal amplitude than d-CHOTs for all displacement amplitudes, for the acoustic wavelength considered. Besides, d-eCHOTs are thinner structures, which correspond to a smaller perturbation to the acoustic waves.

4.5.7   g-eCHOT design

The generation part can be modelled as a coupled optical-thermal-mechanical problem [187], where a structure is heated to cause displacement. The foundation on which the generation part is built on is explained in detail within the literature [159, 162]. The heating is provided optically and the acoustic wave is observed as a displacement.

The parameters that need to be analysed in the generation part consist of grating finger and background film thicknesses and materials.

The FEM modelling of the g-eCHOT is not covered in this thesis. Instead, the main focus will be the simulations of the absorption characteristics of metals for the optical wavelength considered. As far as the design of the g-eCHOT is concerned, the most crucial thing to consider is the material absorption characteristics at different thicknesses.

The main principle behind the g-eCHOT is that consecutive absorbing and non-absorbing surfaces are illuminated with the same irradiance, and the contrast between the rapid expansion and cooling processes at these regions generates a Rayleigh wave, of a certain wavelength, that travels through and along the surface of the background metal film.

The wavelength of the surface acoustic wave generated depends on the period of the features in the generation part. The period of the grating is set for a given acoustic wavelength and the thickness of the background layer depends mostly on the d-eCHOT design, which was discussed in the previous section. The generation part needed to have fingers sufficiently thick to not melt when generating a signal powerful enough to be
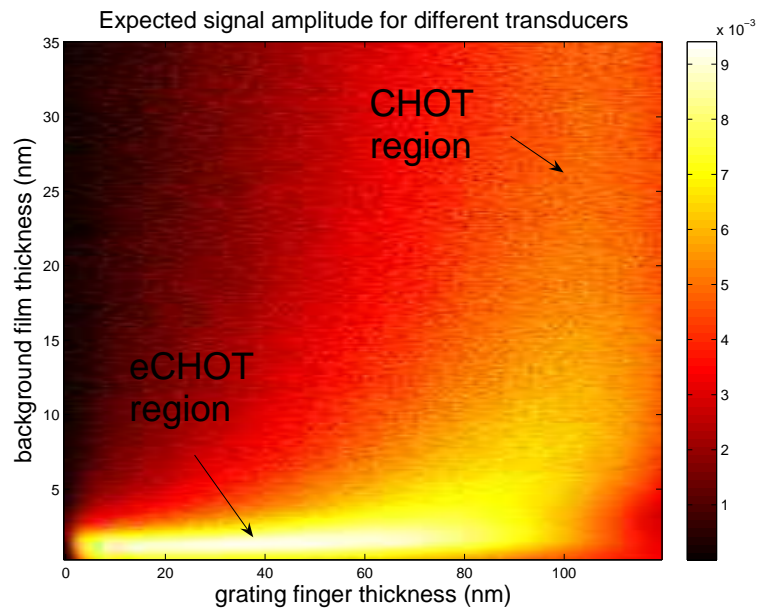
Figure 4.61: The demonstration of the operating region for the CHOT and the eCHOT. The signal amplitude from the transducers at various finger heights and background thicknesses. It can be seen that for high background thicknesses and finger heights, the signal amplitude is smaller than that of the eCHOT region. When the finger thickness is 100nm, which is $\lambda_o/8$, the signal amplitude is bigger than it is for other finger thicknesses (assuming same background film thickness). This operating mechanism is analogous the d-CHOT, with the energy compensation required for the interferometric effect taking place in the absorbed field rather than diffracted. The operation region for the d-eCHOT is seen at low background and grating finger thickness. In the eCHOT region, where the background film thickness is around 1nm-3nm, the path difference relationship for the conventional d-CHOT does not hold. However, there is increased signal amplitudes. This is due to the thin background film acting as a highly resistive layer. As the SAW propagates, the deformation moves this thin film in and out of a field, causing changes in the resistive heating (absorption), hence the changes in the reflectivity.

CHAPTER 4. THE DESIGN OF PHOTONIC DEVICES USING FEM
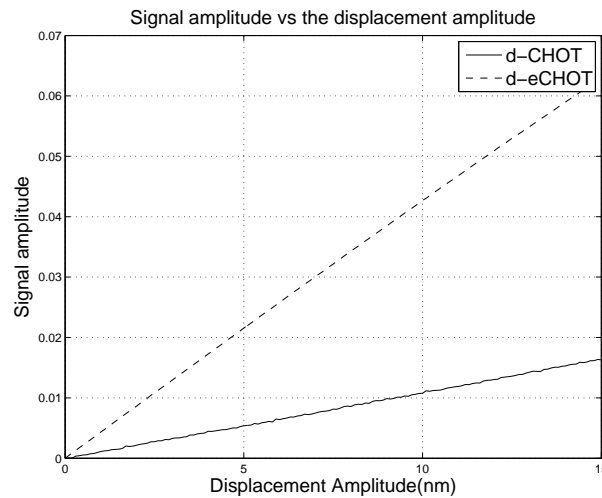
Figure 4.62: The validation of the assumption that the displacement amplitude and the signal amplitude can be scaled down linearly. Hence it is not necessary to simulate with very small, realistic displacement amplitudes. On this figure, the signal amplitude from a transducer with the same characteristics as a d-CHOT, where the finger heights are arranged to be around $\lambda/8$ and background around $20nm$ are plotted on the same axis with a d-eCHOT where the FEM simulations have labeled more sensitive, (background$= 2nm$ and finger height$=40nm$). It is shown that the latter devices will have higher signal amplitude for a given displacement amplitude than a transducer with d-CHOT like parameters.

detected by d-eCHOT. Also the fingers need to be thin enough to present only a small perturbation to the acoustic waves.

The spatial absorption contrast is generated in the g-eCHOT by the selection of the finger heights and the background thickness in figure 4.50. The device is designed to work with optical wavelength 800nm. The selection of the particular metal is based on both the absorption characteristics and mechanical characteristics. It is crucial for this metal to have similar acoustic/mechanical properties to that of glass and possibility to have contrasting absorption at different thicknesses. From figure 4.63 it is seen that for aluminium the absorption at 1nm is more than absorption at 40nm.

The absorption characteristics are shown in Figure 4.63. Among these metals, aluminium and chromium have similar acoustic impedances to that of glass (the substrate material), which will enable acoustic waves to propagate at close speeds within the materials. Therefore, aluminium was chosen as the material for the eCHOTs.
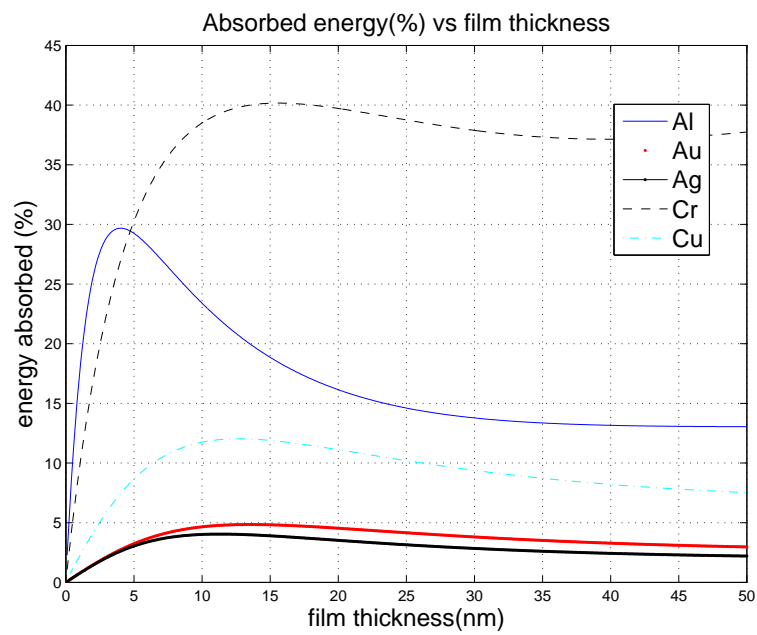
Figure 4.63: Comparison of the percentage of incident energy that is absorbed by different thicknesses of metal films. Different metals absorb different amounts at various film thicknesses. The metals with higher absorptions can be chosen with more confidence, since the d-eCHOT operation is also based on resistive heating.

Parameter Name	Parameter Value	Explanation
Transducer material	Aluminum	High loss, thickness dependent absorption
Background film thickness, t	$0.5nm < t < 3nm$	optimum thickness is $1.25nm$, to be specified as 4nm above optimal due to oxidation
finger height, d	$30nm < d < 70nm$	low values are more desirable
grating period, p	$300nm$	designed for $\lambda_a = 300nm$

Table 4.5: The optimal parameters for the eCHOT (both d-eCHOT and g-eCHOT)

4.5.8 Fabrication-design issues

In this section the main focus will remain on a specific design that was fabricated using e-beam lithography. Other designs are possible, which involve different finger heights for generation and detection, which can be fabricated using methods such as Focused Ion Beam (FIB) etching. FIB etching does not have simultaneous depth information, hence experimentation is necessary to get the finger heights, especially the very thin background film within tolerance limits. This was be done by etching of test samples where several FIB settings are tried, followed by the AFM measurement of the produced features. These measurements can be used as feedback to relate different settings to different depths. However, this method is prone to uncertainty due to AFM measurements, hence will only be approximate. When dealing with tight tolerance limits such as 1-3nm of background film thickness, this method was found to be inappropriate. Photolithography was considered, especially since the eCHOTs are very thin. However, the errors involved in photolithography were too high for reliable fabrication of the eCHOTs(see section 4.4.4for photolithography process and the associated errors).

The optimal design parameters for eCHOTs are shown in table 4.5. Both d-eCHOT and the g-eCHOT requires $2nm \pm 1nm$ background film thickness and $40nm \pm 10nm$

finger thickness. Different grating pitches were to be fabricated for different acoustic wavelengths. E-beam lithography was used to fabricate the prototypes. The devices were fabricated by Eulitha [188]. An issue about the fabrication of the eCHOTs was the oxidation of Aluminum. Device dimensions were specified as 4nm thicker than intended to allow a very thin layer of oxide to immediately form on these layers [189]. Simulations were run to show that a 4nm thick Al_2O_3(refractive index of 1.7) did not effect the operation of the eCHOT. The SEM pictures of the grating with 300nm period is shown in figure 4.66. The AFM picture of an 400nm grating is shown in figure 4.67.

E-beam lithography requires a multi-step process, and involves the use of similar patches for detection and generation. It involves marking a desired area on a resist using an electron beam(e-beam). Afterwards, either the marked area or everywhere else on the sample is removed depending on the type of resist used. For this application, the use of positive resist is considered as an example. When using a resist of the "positive" kind, the parts exposed to the e-beam loosen and become more sensitive to the developing solution. When the resist is developed, the exposed areas are removed. After a stage of metallisation, the remaining resist can be removed, during a process called "lift-off", yielding the metal fingers of the drawn shape on the substrate. The background film is then deposited on the substrate via a secondary evaporation.

In the g-eCHOT the generating mechanism can either depend on the absorption difference between a 2nm film and a 40nm finger, or if the high power laser causes ablation of the very thin metal film, it could genrate due to the absorbing and non-absorbing nature of metal fingers and the substrate.

4.5.9 Experimental Issues

The experimental verification could be done using an ultrafast laser in a pump-probe arrangement. The suggested experimental set-up would utilise an ultrafast laser, with $\approx 100fs$ pulse length. This system is conceptually represented in figure 4.64. The laser beam is split into two, the pump and the probe beams using a polarising beam splitter. Pump beam is used to irradiate the g-eCHOT. The probe beam beam is passed through a variable optical delay line to temporally resolve the measurements. Then a motorised mirror is used to change its angle of incidence. Therefore, after being rejoined with the
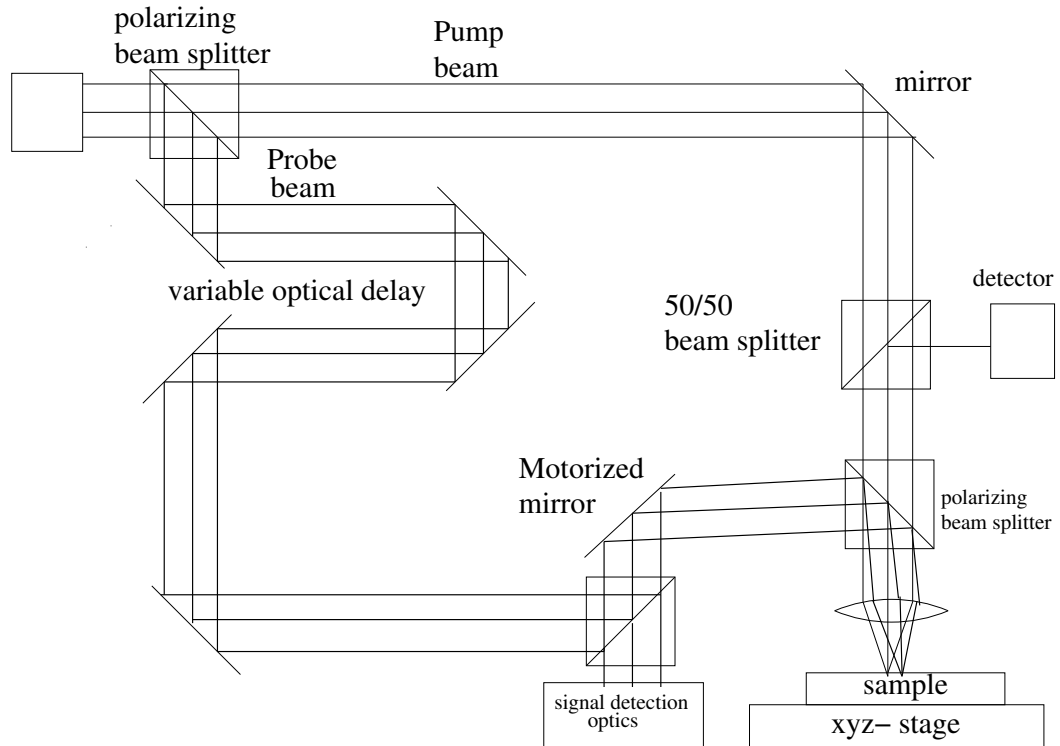
Figure 4.64: The conceptual explanation of the experimental set-up that can be used to test the eCHOTs. The system involves an ultrafast laser. The beam is separated into two, called the pump and probe. The probe beam is passed through a variable optical delay, which is used to temporally resolve the measurements. The g-eCHOT is irradiated by the pump beam and the d-eCHOT by the probe beam. The reflected light from the d-eCHOTs is then directed to the detectors.

probe via a polarising beam splitter, the probe beam enters the objective with a certain angle. Pump beam enters the objective with normal incidence hence the two beams are focused on different locations on the sample. The probe beam follows the same path back to the mirror and is directed to the detector optics.

Alignment can be done by blocking the probe beam, and moving the sample, and using the detector optics to place them beam on the g-eCHOT. Then the sample is fixed and the pump beam is blocked. The motorised mirror is then moved and the probe beam is placed on the d-eCHOT. After this alignment, measurements are taken by moving the variable optical delay mirrors with a known speed and taking data at known times, followed by averaging to produce the necessary signals.

The experimental verification could not be done before the thesis submission, since

For e−beam lithography

G−eCHOT

D−eCHOT

p/2   p

d

Aluminum   Glass

|t
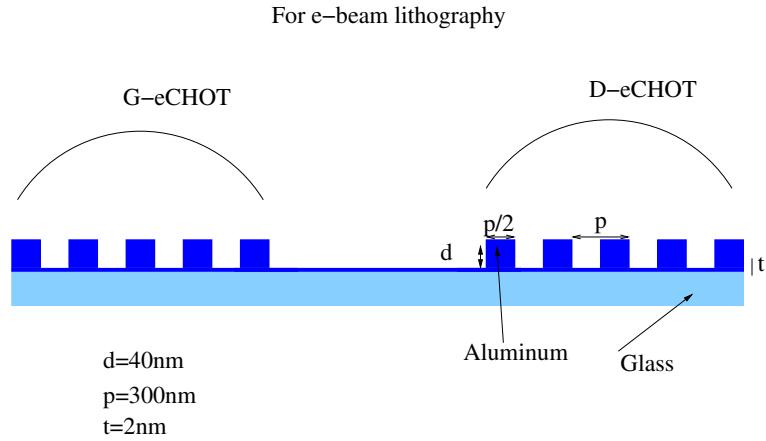
d=40nm
p=300nm
t=2nm

Figure 4.65: The eCHOT design and the specification for the e-beam lithography. The specifications for the device is $t = 2nm \pm 1nm$ and $d = 40nm \pm 10nm$ $p = 250, 300, 350, 450, 550nm$. This design involves using the gratings of equal dimensions for generation and detection mechanisms. Many periods for the grating was fabricated to test the eCHOT for different acoustic wavelengths and the separation between the g-eCHOT and d-eCHOT were varied to test the SAW velocity. If these devices were to be fabricated with FIB, an optimal device would involve a thin background film in the d-eCHOT and no background film in the g-eCHOT.

V 1 = 154.5 nm

300nm   Mag = 114.06 K X   EHT = 2.00 kV   Signal A = InLens   WD = 6 mm   High Current = Off   Date :23 Jul 2009
              Stage at T = -2.5 °   Stage at X = 77.958 mm   Stage at Y = 50.376 mm   User Name = DAIS   Time :9:40:45
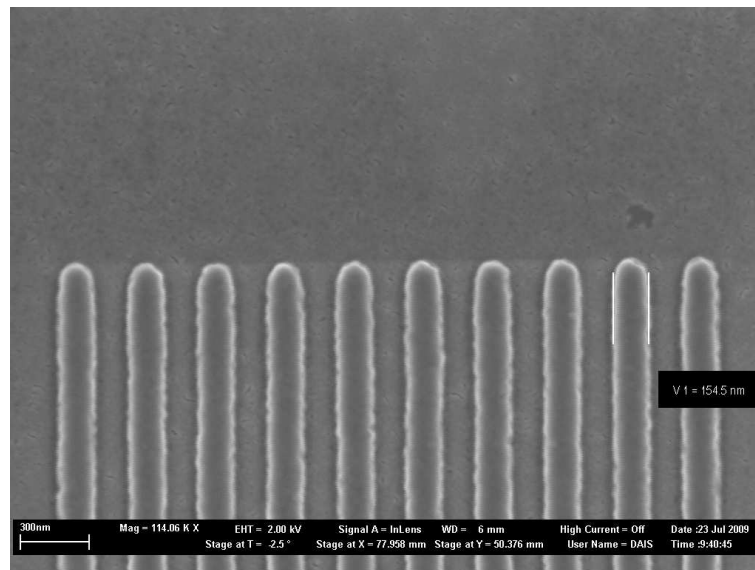
Figure 4.66: The SEM images of the fabricated eCHOT samples with 300nm period.
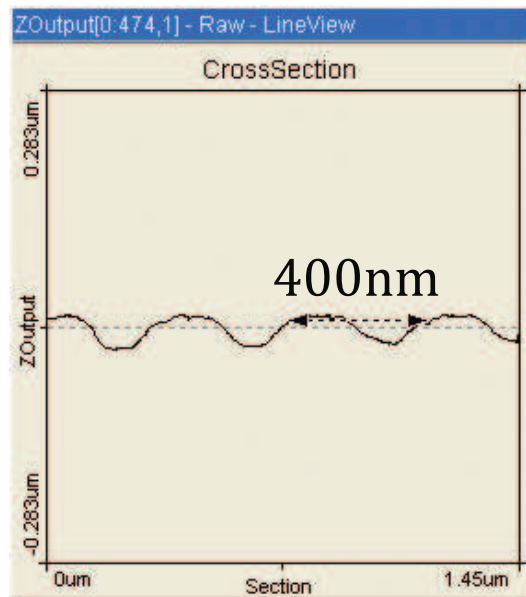
Figure 4.67: The cross section of the fabricated gratings taken with AFM. The finger heights are seen to be around 40nm. The period is 400nm.

the ultrafast laser system present in the department was unavailable due to the technical problems at the time of this work. The fabricated samples were put in an air tight container to wait for the laser system to become available again. The results are intended for publication after this thesis.

4.6   Summary

In this chapter, the development and implementation of a technique was demonstrated, that combines numerical modelling with physical insight to analyse and optimise the operating mechanism of novel nanophotonic devices that can fabricated and experimentally tested. It was found that knowledge and experience was necessary in electromagnetics theory, computational electromagnetics(FEM specifically), applied optics, experimental methods and micro/nano fabrication techniques.

The design process and the determination of the optimisation parameters was led by inductive reasoning, and numerical modelling was used heavily to test the device operation in different scenarios so as to gain physical understanding of the processes that may not have analytical solutions. The optimisation of the devices have been automated, although much knowledge about the workings of the numerical modelling methods was required to be able to model different scenarios. Given the appropriate optimisation parameters, the scripts were able to carry out the numerical modelling to adjust the model properties and loop through the optimisation parameters while looking at the reflected/transmitted light, to pick the parameters that could be used in the optimal fabrication of the devices.

In section 4.1, the general problem definition in terms of the numerical modelling is briefly discussed. The common features of the devices considered in this chapter that affect the choice of modelling methods/features are mentioned.

In the section 4.2 the FEM , analytical solutions, and FDTD was compared in three increasingly difficult problems in terms of speed and agreement with each other. These problems were simpler versions of the devices designed in this chapter. The modelling tools were compared first using a simple geometry involving surface plasmon resonance on a thin flat metal film. Agreement was observed among all three modelling tools. Afterwards, a system involving three planar metal films, stacked on top of each other was modelled. All three modelling tools were shown to agree for this structure, which was more complicated than the simple surface plasmon resonance case, since it involved multiple reflections and resonances. Afterwards, a grating-like structure was considered, for which there was no analytical solution. For this structure which could also support plasmons, the FEM simulations agreed with FDTD simulations. The analytical solution

was found to be faster than FEM and FDTD when it could be applied. FEM was found to be more appropriate than FDTD for the purpose of designing and optimising nanophotonic devices in a monochromatic environment. This is due to FEM's speed and ability to model arbitrary structures of different scales, which provides a substantial advantage in terms of time and accuracy within the repetitive nature of optimisation.

In section 4.3, a novel self-calibrated surface plasmon refractive index sensor is designed and optimised using the analytical solutions. The operation of this device and the effect of every feature in the operation of the device are investigated in detail. It is shown that the multireflections and surface plasmon excitations within the multifilm structure result in two dips. One of these dips is sensitive to the refractive index changes of the transducing medium whereas the other one is not. The use of this mechanism as a self-calibrated plasmon sensor is demonstrated. The sensitivity of this sensor is compared to that of conventional SPR sensors, and it was concluded that even though the suggested sensor has reduced sensitivity, there may be a scenarios where this is a price worth paying, for the possible movement of this sensor to in the field usage. A fabrication method was suggested that could be used to fabricate such sensors.

In section 4.4, a surface plasmon resonator was designed and optimised using FEM. It is shown that the multiple reflections of surface plasmons in the resonator cause an interference pattern, which could be detected in the farfield as a retroreflection whose intensity changes periodically with the cavity width. After the detailed analysis, design and optimisation of the operating mechanism, the fabrication of these devices were described and experiments were set-up in order to verify their operation. The fabricated prototypes were then measured using AFM, and experimentally verified using the optical system built. The optical system used a high NA objective to illuminate the sample and detect the reflected light at various angles. Using this system, the presence or absence of the retroreflection was detected from the resonators. The experimental results were discussed, and it was shown that they matched the predictions made using the FEM model with finite uncertainty. The origins of this uncertainty is discussed and related to the inefficiency of the fabrication and the measurement methods. There are fabrication related artifacts which directly cause a certain level of uncertainty in the device dimensions and indirectly cause uncertainty in the measurement of the devices. Some of the modelling results re-

lating to the operation of the resonators match a similar model found in the literature [59]. However the fabrication and experimental testing of a practical resonator has not been done. This device, while being unique and important in its application area, serves another purpose in this chapter. It serves as a problem which does not have analytical solutions readily available, but could be modelled with FEM and FDTD, and fabricated and verified experimentally. Additionally, this problem points out the challenges in design of a practical photonic device, whose operation is sensitive to device dimensions, considering the fabrication and experimental issues. Therfore, this problem ties the numerical modelling to practical fabrication and experimental issues.

In the final section, a novel nanophotonic device (eCHOTs) was considered, which could generate and detect ultrahigh frequency SAWs. The operation of this device was analysed in detail, its differences from that of a more conventional optical transducer designed for lower acoustic frequencies (d-CHOT [157]), was described. While the detection mechanism of the CHOTs is based on an in situ interferometer whose operation depends on an energy gain/loss mechanism between the diffracted orders, the eCHOT detection is based on a an optimised energy gain loss mechanism depending on resistive losses. At ultrahigh acoustic frequencies, where the SAW wavelength is smaller than the optical wavelength, the diffracted orders are not present to carry away any energy, hence the sensitivity of the CHOTs suffer. It was shown that once the orders are evanescent, a more sensitive design could be obtained by abandoning the d-CHOT criteria, and subsequently changing the operation mechanism fundementally. It was also pointed out that the eCHOTs operate at a region that is not predicted by the conventional diffraction theory, to yield enhanced signal amplitude for a wide range of parameters at thickness values that do not pose large perturbations to the acoustic waves. The eCHOTs were designed to optimise the resistive heating dependant mechanism in the structures, while keeping in mind the fabrication and experimental limitations. The physical displacement caused by the SAW was linked to the electromagnetic model by deforming the transducer geometry sinusoidally with a changing phase, so as to simulate the propagation of a SAW. FEM was chosen as the modelling tool in this circumstance rather than the more conventional optical modelling method FDTD, due to its superior meshing abilities, and its inherent speed. eCHOTs were fabricated using e-beam lithography and measured using SEM and AFM.

This section served as a statement that the optimisation method that was developed, when used with sensible judgement and insight, is capable to design and optimise devices that would otherwise be impractical to understand, model or design. Additionally, a novel device and an operating mechanism was introduced at an operating region that would normally be predicted with low sensitivity using the conventional diffraction theory.

This chapter showed that by carefully leading the numerical simulations, using knowledge, experience and insight, it is possible to design devices that cannot be predicted or designed using other methods. This chapter tied the developed design and optimisation environment to practical applications and showed that FEM modelling is a valuable tool in the design and optimisation of arbitrary shaped subwavelength photonic devices.

Chapter 5

Conclusion and future work

The aim of the work in this thesis was to develop a technique for the design and optimi-
sation of nanophotonic-plasmonic devices. Nanophotonic devices are difficult to model,
design and optimise, due to their small size, arbitrary geometry and variable scales.

This thesis demonstrated and reached the following goals:

- Nanophotonic devices involving thin films, evanescent waves, small arbitrary ge-
  ometries and large scales can be/were designed and optimised using an appropriate
  numerical modelling tool.

- A technique was developed employing the appropriate tool, to analyse the effects of
  the design parameters. The range of values for the optimisation parameters were
  determined by careful examination of the fabrication and experimental limitations.

  - The analysis was done by parameterising the design metrics and resolving the
    problem for each combination, simultaneously recording an output metric of
    interest.

  - The optimisation plots were produced, and the appropriate design parameters
    and combinations were picked.

- The technique was applied to simple problems with analytical solutions.

- The technique was applied to problems without analytical solutions and the applica-
  tion of the technique was demonstrated considering the fabrication and experimental

201

limitations.

- The technique was applied to design a novel structure which was impractical to model otherwise. Therefore demonstrating the competence of the technique.

In this thesis, three photonic/plasmonic devices were modelled, designed and optimised. These devices demonstrate different challenges in the modelling and design while being of considerable interest in their own right as nanostructures for sensing and measurement. The work extended from the design of simpler micrometre sized devices, to the design of novel devices with suboptical wavelength dimensions and complicated geometries.

In section 4.2, FEM, FDTD and the analytical solutions were compared in three different problems, which represent simple versions of the devices considered.

It was concluded that FEM, FDTD and the analytical solutions agreed in the results. The analytical solution was found to be faster than FEM and FDTD for problems involving planar stratified media, while FEM was found to be much faster than FDTD for the problems where there was no analytical solution. This was attributed to the superior meshing abilities of FEM resulting in much smaller models, and the general technique of FEM being more memory rahter than CPU time dependent. Given the serial nature of the optimisation technique FEM was seen as a more advantageous method in modelling small photonic/plasmonic devices with arbitrary geometries.

In section 4.3, the developed technique was demonstrated on a multifilm model, using the analytical solutions to design and optimise a self-calibrated surface plasmon sensor.

In section 4.4, FEM was used to design and optimise a surface plasmon resonator. The operating principle of this device had been demonstrated using FDTD in the literature [59]. The operating principle of the FEM designed device and the demonstrated device was found to be the same. This device was designed and optimised with the fabrication limitations in mind, and fabricated them according to the designs using photolithography. The fabricated devices were then measured using AFM and experimentally verified. This problem tied the connection between the design and the practice. This device could to be used to study surface plasmon propagation length and different amplification media via the observation of the retroreflection experimentally.

In section 4.5, an alternative, novel operating mechanism for the remote laser-based

detection of ultrahigh frequency surface acoustic waves is demonstrated, which is based on resistive heating. The designs presented in section 4.5 transduce with increased signal levels for a wider range of parameters using this mechanism. This mechanism was based on the ultrasonic wave to move a very thin resistive metal film in and out of a localised evanescent field. This device involved very thin films, evanescent waves and sinusoidal geometry, which would be impractical to model using other design tools. However, in this section it was shown that the design and optimisation of this device can be done using careful choice of the design parameters with physical insight to lead the developed technique and numerical modelling. This demonstrated the ability of the technique in the design and optimisation of the devices whose operation cannot be predicted efficiently using only optical theory, and that cannot be designed and analysed any other way. eCHOTs were fabricated using e-beam lithography and were measured using AFM. The experimental verification of the eCHOTs is pending.

The future work involves research towards using modelling as an exclusive tool for design and analysis of practical and complicated nanophotonic devices. In the ultrasonics applications the future work can be summarised as follows:

- Experimentally verify the operation of fabricated eCHOT samples. This can be seen as the next step in the improvement of confidence levels for the design technique.

- Use Finite Element Method to further optimise the mechanism presented in 4.5.6 to design new, better ultrasonic transducers, that can be applied to other environments such as biological samples. Consider nanoparticles for biological samples.

- Investigate the interaction of ultrasonics with coated nanoparticles in 3D.

Although only 2D models are considered in this thesis, a drawback of FEM as a modelling tool is realised when 3D modelling is considered. FDTD may be programmed to do computations in parallel, sharing the work among multiple CPUs or workstations, whereas this is far more complicated in FEM [25]. Relatively simple 3D FEM models force the limits of the current state of the art workstations. It is not possible to model very detailed geometries, with high discretisation in 3D. The advantages of FEM over FDTD may be outruled in some cases when such 3D models are considered. However, with

the increase of computational power, according to Moore's Law [190], 3D models should be within reach on an ordinary desktop workstation in the next few years. Then the optimisation can be done among multiple computers by solving each model in a separate computer and saving the results, which can be connected with another script. This thesis was concerned with 2D models, hence the computer resources, especially memory usage, was not a limiting factor in the models considered.

In the sensing applications, the future work involves fabricating and experimentally testing the self-calibrated plasmon sensor. Besides, new materials, other than MgF_2 may have to be investigated depending on the applications. In plasmonics, the future work involves investigating several amplification media for surface plasmons.

Appendix A

Algorithms and Scripts

A.1 The Analytical Model for multifilm and SPR models

In this section the script for the main function is given in pseudocode for the understanding of algorithm. for a detailed script of Fresnel coefficient calculation, refer to section A.5.

```
global r1p r1s r2p r2s rrep1p rrep1s trep2p trep2s
clear all,close all
set parameters % sets the range of ambient
%refractive indices, metal
%film separations, incidence angles,
% number of repeating layers
for all ambient_refractive_index
for all metal_separations
for each incidence angle
    call the coefficient calculator
end
build matrices and append to matrices
end
end
plot matrices
```

A.2 Comsol Script for Plasmon Resonator Models

In this section, the scripts used in the plasmon resonator models are listed in pseudocode for the understanding of the mechanism. A main file was used to build, solve and save models. A farfield extractor file was used to load the models, extract the farfield, build matrices and save the raw data. A final script was used to load the farfield data files one by one and extract the necessary information to appropriate matrices.

A.2.1 Main file

The following script, given in pseudocode, was used to load the basic model and create similar models with different parameters. This model enabled quick analysis of different paremeters in the model, such as film thickness, structure period etc.

```
clear all
set parameters
sae parameters
load base_file % this is the basic fem struct
      % that has a pre-build structure.

define the parameter to change = param
for param values
  retrieve base model
scale geometry according to parameters
fix boundary conditions due to changed geometry
mesh problem
solve problem
assign a name; % to load later for postprocessing
save solution with the assigned name
end
```

A.2.2 Far Field Extraction Scripts

The scripts below, written in pseudocode, were used to extract the far field from the models produced by scripts in appendix A.2.1.

A.2.3 The main Far-field extraction file

The following script was written (listed here in pseudocode) in order load models that had previously been produced, by the scripts in A.2.1 one by one, to extract and save information from them. For a given model, it will set some parameters for farfield extraction and call the function to calculate the far-field. Besides, it will also extract reflectivity and retroreflectivity information and save it to a file to be read into arrays by another script in section A.2.5.

```
clear all
close all
load parameters
%load the parameter ranges
% for which the script is goint to run
param=desired parameter
set up flags
for values of param
build filename from parameters;
load filename;
set farfield extraction parameters according to arguments;
% these are the radius, wavelength, numner of source points,
% observation angles, refractive index,
% and the location of farfield source
call farfield_extraction function
retrieve unnormalized fields.
normalize fields
extract data from farfield plots
build matrices
```

```
save matrices
end
```

A.2.4   Calaculating Far field from the Extracted Field

The following script was used to implement the Stratton-Chu equation using the correct constants for any medium, and circumventing the boundary shadowing problem using the aperture function. The following is written in pseudocode. For a more detailed, functional script, refer to A.5.3

```
function [total field old,  Ex,Ey,unnormalized total field,...
phase info]=calculate_farfield(fem,radius,...
location,number of sourcepoints,...
observation angles,incidence angles,...
wavelength,ambient refractive index,flags)
set constants
describe farfield source circle
%build source points from input arguments such as number of
%sourcepoints, radius and source location
set dS % the unit area for the integration
set normals % for the integration
set unit vectors  %for the integration
aperture=gaussian % calculate the gaussian aperture

extract Ex, Ey, Ez, Hx, Hy, Hz % from the input argument fem
for incidence angles
for farfield observation angles
calculate position vectors
calculate apodization aperture
%apodize extracted fields
Ex_ap=Ex * aperture
```

```
calculate the FF by intergrating around the source
end
end
build and append matrices to send out
end
calculate the output variables
```

A.2.5 The Data Collection Script

The following sript was written so as to load the files produced by script in appendix A.2.3, and form the necessary matrices.

```
clear all
load parameters
for parameter_of_interest
  load related_file_encrypted_with_parameter
store necessary matrices
end
save(['data_noref',arg11,arg4])
```

A.3 The approximate eCHOT Model

This section lists the scripts that is used in the approximate eCHOT model. The main file built, solved and saved the models. The farfield extractor file loaded these files and extracted the farfield into matrices that are saved as workspaces. These workspaces are loaded and data are extracted into matrices via another script.

A.3.1 Generating Script

The script that generates the approximate eCHOT model is shown below in pseudoscript form.

```
clear all
%parameters
```

```
set background_values
set finger_height
set geometry_parameters
save all_parameters
load base_model

for background_values
for finger_height
load base_model
set necessary variables in model
scale geometry_objects
update model
build fingers
build background
change fingers using geometry_parameters
change background using geometry_parameters
update model
work around comsol 3.4 boundaries
update, mesh , solve
save models
end
end
```

A.3.2 Extraction Script

The following script, written here in pseudocode extracts far field information from already produced models.

```
load all_parameters

for all background_values
for all finger_height
```

```
build filename from parameters

load corresponding file

set farfield_parameters

set offset and radius and resolution of farfield source circle

set wavelength and ambient medium

calculate farfield

build unnormalizsed total_field matrix

save results

end

end
```

A.3.3 Loading Script

The following script loads each farfield file and extracts information to necessary files.

```
clear all

clear all

load all_parameters


for all background_values

 for all finger_heights

build filename avccording to all_parameters

load file

build matrices

end

end

save built matrices
```

A.4 Example: The Realistic eCHOT Model

This section will list the scripts necessary to produce the realistic echot models. The farfield calculation script is shown in section A.5.3.

A.4.1   The main file

```
clear all
bckgrn_left_edge=-2e-6;
%the metal structures left edge, it determines where the metal ends
factor=10000;
%factor for multiplication of parameters with decimal points.
% allows easier loading/saving by removing points
background=[0.1:0.1:5];  %background metal thickness
geomsize=3000;  % geometry size along y-axis
xint=25;
%point separation (nm). This will sample the sinusoid produced
rayleigh_amplitude=5; %amplitude of the displacement sinusoid
chrome_height=[1:2:80]; % height of the metal fingers
s_factor=geomsize./3000;
%Scale factor in case a different geometry size is required
plist=[0];
% parameter list of the solver.
%  only normal incidence is considered
phase1=[0 :10:360].*pi./180; %phase of the sinusoid
phase=[phase1(9) phase1(28)];


arg4='_Al_bs';arg3='0360';arg2='.ws';
 %flags and arguments for saving
ycount=1;icount=1;
flload('CHOT_base_sinusoid_deneme_fliplr_bothsides_bigger.mph');
%load the base file
fem_base=fem;
for hcount=1:length(chrome_height)
for zcount=1:length(background)
        for icount=1:length(phase)
```

```
fem=fem_base; %laod base file

fem.const{10}=num2str(geomsize);

 %modify geomsize constant

   arg1=['_geom',num2str(geomsize),...

'_Cr',num2str(chrome_height(hcount)),...

'_phs',num2str(icount),...

'_amp',num2str(rayleigh_amplitude),...

'_bck',num2str(background(zcount))];

temp=fem.draw.s.name;

%  fing and scale every geometry object

for jcount=1:length(fem.draw.s.name)

objname=temp{jcount};

d1=drawgetobj(fem,objname);

if ((size(d1))(:,2)>1)

d1=geomcomp(d1);

end

d1=scale(d1,1,s_factor(ycount),0,0);

fem=drawsetobj(fem,objname,d1);

end

%build sinusoid

[x,y,err,y_back]=sin_maker(phase(icount),...

geomsize(ycount).*1e-9,...

xint.*1e-9,...

rayleigh_amplitude.*1e-9,...

chrome_height(hcount).*1e-9);

if err==1

disp('error')

end

   fem=geomanalyze(fem); %update model

CO1=drawgetobj(fem,'CO1'); % get the metal

%change the geometryof the metal to add the sinusoid on
```

```
y_geom=y+bckgrn_left_edge+background(zcount).*1e-9;

y_geom_bck=y_back+bckgrn_left_edge;

c=geomedit(CO1);

f=x;

d=line1(y_geom,f);

e=line1(y_geom_bck,f);

dd=geomedit(d);

ee=geomedit(e);

dd1=c;

dd1_2=geomedit(line1([y_geom_bck(1),y_geom(1)],...

[-geomsize(ycount)./2.*1e-9,-geomsize(ycount)./2.*1e-9]));

dd1_122=geomedit(line1([y_geom_bck(length(y_geom_bck)),...

y_geom(length(y_geom))],...

[geomsize(ycount)./2.*1e-9,...

geomsize(ycount)./2.*1e-9]));

dd1=[dd(1) dd1_2 dd(2:length(dd)) dd1_122];

c(1:length(dd1))=dd1;

c(length(dd1)+1:length(c))=ee;

cobj=geomcsg(c,'solidify','on');

 %solidify to remove any potential boundary problems

fem=drawsetobj(fem,'CO1',cobj); temp=fem.appl{1}.bnd;

[fem,map]=geomanalyze(fem); %update geometry

fem=multiphysics(fem);

%update equations and boundary

% conditions according to the new geometry

temp1=fem.appl{1}.bnd;  pec=fem.appl{1}.bnd.ind(1);

pec_loc=find(fem.appl{1}.bnd.ind==pec);

% find the location of pec boundary condidtions.

% All the outside boundaries except source

src_loc=length(fem.appl{1}.bnd.ind);

%find the source
```

```
src=fem.appl{1}.bnd.ind(src_loc);

fem.appl{1}.bnd.ind(:)=2;

%make everything 'cont'

fem.appl{1}.bnd.ind(pec_loc)=pec;

 %put back your original pecs

fem.appl{1}.bnd.ind(src_loc)=src; %put back your sources

fem.appl{1}.bnd.ind(pec_loc(2:length(pec_loc)))=5;

fem.appl{1}.equ.ind=[1 2 3 4 5];

   fem=multiphysics(fem);

fem.mesh = meshinit(fem,'hmax',800e-9/30);

 fem.xmesh = meshextend(fem);

o_f_rayleigh(zcount).phase(icount).filename=['g_chot',...

num2str(geomsize(ycount)),arg1,arg4];

%used when separate models need to

%  be run on different computers in the

%optimization to save time.

%these files are saved with

% different filenames and joined later on.

fprintf('solving file named : %s \n',...

o_f_rayleigh(zcount).phase(icount).filename);

fem.sol=femlin(fem,'pname','tet',...

'plist',plist,'Conjugate','on'); %solve model

%extract your farfield and save your first matrix.

offset=[0.5e-6,0]; %offest to go to the air side

[tfcount(icount,:)]=CHOT_ff_extractor_single(fem,arg1,...

arg2,arg3,arg4,ycount,plist,...

geomsize,offset); %extract farfield from air

offset=[-4e-6,0]; %offset to go to the glass side

[tf_trans(icount,:)]=CHOT_ff_extractor_single(fem,arg1,...

arg2,arg3,arg4,ycount,plist,...

geomsize,offset); %extract farfield from glass
```

```
        end
%  build necessary matrices
ref(hcount,zcount,:)=tfcount(:,1);
 incom(hcount,zcount,:)=tfcount(:,2);
trans(hcount,zcount,:)=tf_trans(:,2);
 pml_ref(hcount,zcount,:)=tf_trans(:,1);
sensit_curve(hcount,zcount,:)=
(ref(hcount,zcount,:).^2)./(incom(hcount,zcount,:).^2);
trans_curve(hcount,zcount,:)=
(trans(hcount,zcount,:).^2)./(incom(hcount,zcount,:).^2);
end
end
%build sensitivity matrices
d_max=sensit_curve(:,:,1);
d_min=sensit_curve(:,:,2); sen=d_max-d_min;
imagesc([chrome_height(1) chrome_height(length(chrome_height))],...
[background(1) background(length(background))],flipud(sen'))
axis([chrome_height(1) chrome_height(length(chrome_height))
 background(1) background(length(background))])
colorbar on
 xlabel('Finger height(nm)')
ylabel('Background thickness (nm)')
title('The Sensitivity of the CHOT structure for various
 background thickness and finger heights (realistic model)')
save(['data_Al_latest_bck',num2str(background(1)),...
     '_',num2str(background(length(background))),...
     '_Cr',num2str(chrome_height(1)),'_',...
     num2str(chrome_height(length(chrome_height))),...
     '_amp',num2str(rayleigh_amplitude)])
trans1=trans.^2./(1.512);
ref1=ref.^2; incom1=incom.^2;
```

```
absorption=squeeze(incom1(:,:,1))-squeeze(trans1(:,:,1)+ref1(:,:,1));
subplot(2,1,1)
plot(chrome_height,squeeze(incom1(:,:,1)),'b',...
chrome_height,squeeze(ref1(:,:,1)),'r',...
chrome_height,squeeze(pml_ref(:,:,1).^2),'g+-',...
chrome_height,squeeze(trans1(:,:,1)),'c-o',...
chrome_height,absorption,'y',...
chrome_height,squeeze(ref1(:,:,1)+trans1(:,:,1)),'k')
legend('input','reflected',...
'reflection from pml','transmitted/\epsilon_g',...
'transmitted+reflected','absorbed')
 xlabel('finger height')
ylabel('Energy')
title('Energy levels at 0 background')
grid on
subplot(2,1,2)
plot(chrome_height,sen)
grid on
 xlabel('finger height')
ylabel('sensitivity')
title('Sensitivity at various finger  heights')




function [x,y,err,...
y_back]=sin_maker(phase,geomsize,xint,...
amplitude,ridge_height)
%phase=phase of the sinosoid
%geomsize: geometry size in nm
%xint: interval of points in nanometers, determines
%  how well the sin is sampled.
```

```
% has to be changed in the base model first. -->
%  first change it in flat metal,
%  then get the geometry and build another base model
% amplitude=amplitude iof the perturbation


x=[-geomsize/2:xint:geomsize/2];
y=ridge_height.*(sin(2*pi*x./300e-9+3*pi/2)>0)+
amplitude.*sin(2*pi*x./300e-9+phase);
y_back=amplitude.*sin(2*pi*x./300e-9+phase);
y1=y(2:length(y));
y2=y(1:length(y)-1);
err=0;
if(~isempty(find(y1==y2)))


err=1;
exit
end
end;
```

A.5   Important Scripts

This section lists some of the important scripts that were used in Matlab, Comsol Script and tcl for EM Explorer.

A.5.1   Fresnel Coefficients

This function is used in the calculation of the Fresnel coefficients in the analytical model. It is called for SPR and multifilm problems.

```
function
[rs,rp,ts,tp]=reftraextr(sthinc0,...
n0,n1,n2,nrep1,nrep2,n5,d1,d2,drep1,...
```

```
    drep2,numrep_layers,lamda)


%This function is written by Mike Somekh.
cthinc0=sqrt(1.-(sthinc0.*sthinc0));
sthinc1=n0*sthinc0/n1;
cthinc1=sqrt(1.-(sthinc1.*sthinc1));
sthinc2=n0*sthinc0/n2;
cthinc2=sqrt(1.-(sthinc2.*sthinc2));
 sthincrep1=n0*sthinc0/nrep1;
cthincrep1=sqrt(1.-(sthincrep1.*sthincrep1));
sthincrep2=n0*sthinc0/nrep2;
cthincrep2=sqrt(1.-(sthincrep2.*sthincrep2));
 sthinc5=n0*sthinc0/n5;
cthinc5=sqrt(1.-(sthinc5.*sthinc5));
delta1=2.*pi*n1*d1*cthinc1/lamda;
 delta2=2.*pi*n2*d2*cthinc2/lamda;
deltarep1=2.*pi*nrep1*drep1*cthincrep1/lamda;
deltarep2=2.*pi*nrep2*drep2*cthincrep2/lamda;
prop1(1,1)=exp(-i*delta1);
 prop1(2,2)=exp(i*delta1); prop1(1,2)=0;
prop1(2,1)=0; prop2(1,1)=exp(-i*delta2);
 prop2(2,2)=exp(i*delta2);
prop2(1,2)=0; prop2(2,1)=0;
proprep1(1,1)=exp(-i*deltarep1);
proprep1(2,2)=exp(i*deltarep1);
proprep1(1,2)=0; proprep1(2,1)=0;
proprep2(1,1)=exp(-i*deltarep2);
proprep2(2,2)=exp(i*deltarep2);
proprep2(1,2)=0; proprep2(2,1)=0;
r1p=((n0*cthinc1)-(n1*cthinc0))
./((n0*cthinc1)+(n1*cthinc0));
```

```
r1s=((n0*cthinc0)-(n1*cthinc1))
./((n0*cthinc0)+(n1*cthinc1));
t1p=(2.*n0*cthinc0)./
((n0*cthinc1)+(n1*cthinc0));
t1s=(2.*n0*cthinc0)./
((n0*cthinc0)+(n1*cthinc1));
t2p=(2.*n1*cthinc1)./
((n1*cthinc2)+(n2*cthinc1));
t2s=(2.*n1*cthinc1)./
((n1*cthinc1)+(n2*cthinc2));
r2p=((n1*cthinc2)-(n2*cthinc1))
./((n1*cthinc2)+(n2*cthinc1));
r2s=((n1*cthinc1)-(n2*cthinc2))
./((n1*cthinc1)+(n2*cthinc2));
trep1p=(2.*n2*cthinc2)./
((n2*cthincrep1)+(nrep1*cthinc2));
trep1s=(2.*n2*cthinc2)./
((n2*cthinc2)+(nrep1*cthincrep1));
rrep1p=((n2*cthincrep1)-(nrep1*cthinc2))
./((n2*cthincrep1)+(nrep1*cthinc2)
);
rrep1s=((n2*cthinc2)-(nrep1*cthincrep1))
./((n2*cthinc2)+(nrep1*cthincrep1)
);
trep2p=(2.*nrep1*cthincrep1)./
((nrep1*cthincrep2)+(nrep2*cthincrep1));
%interface between rep1
trep2s=(2.*nrep1*cthincrep1)./
((nrep1*cthincrep1)+(nrep2*cthincrep2));
rrep2p=((nrep1*cthincrep2)-(nrep2*cthincrep1)
)./((nrep1*cthincrep2)+(nrep2
```

```
*cthincrep1));
rrep2s=((nrep1*cthincrep1)-(nrep2*cthincrep2))
./((nrep1*cthincrep1)+(nrep2
*cthincrep2));
trep21p=(2.*nrep2*cthincrep2)./
((nrep2*cthincrep1)+(nrep1*cthincrep2));
trep21s=(2.*nrep2*cthincrep2)./
((nrep2*cthincrep2)+(nrep1*cthincrep1)); %
and rep1
rrep21p=((nrep2*cthincrep1)-(nrep1*cthincrep2))
./((nrep2*cthincrep1)+(nrep
1*cthincrep2));
rrep21s=((nrep2*cthincrep2)-(nrep1*cthincrep1))
./((nrep2*cthincrep2)+(nrep
1*cthincrep1));
% this one is used when
% numrep_layers>1 so there is a 21 interface
r5p=((nrep2*cthinc5)-(n5*cthincrep2))
./((nrep2*cthinc5)+(n5*cthincrep2));
r5s=((nrep2*cthincrep2)-(n5*cthinc5))
./((nrep2*cthincrep2)+(n5*cthinc5));
t5p=(2.*nrep2*cthincrep2)
./((nrep2*cthinc5)+(n5*cthincrep2));
t5s=(2.*nrep2*cthincrep2)
./((nrep2*cthincrep2)+(n5*cthinc5));
I1p(1,1)=1.; I1p(1,2)=r1p;
I1p(2,2)=1.; I1p(2,1)=r1p; I1p=I1p/t1p;
I2p(1,1)=1.; I2p(1,2)=r2p; I2p(2,2)=1.;
I2p(2,1)=r2p; I2p=I2p/t2p;
Irep1p(1,1)=1.; Irep1p(1,2)=rrep1p;
Irep1p(2,2)=1.; Irep1p(2,1)=rrep1p;
```

```
Irep1p=Irep1p/trep1p; Irep2p(1,1)=1.;

Irep2p(1,2)=rrep2p; Irep2p(2,2)=1.;

Irep2p(2,1)=rrep2p; Irep2p=Irep2p/trep2p;

Irep21p(1,1)=1.;   Irep21p(1,2)=rrep21p;

Irep21p(2,2)=1.; Irep21p(2,1)=rrep21p;

Irep21p=Irep21p/trep21p; I5p(1,1)=1.;

I5p(1,2)=r5p; I5p(2,2)=1.; I5p(2,1)=r5p;

I5p=I5p/t5p; I1s(1,1)=1.; I1s(1,2)=r1s;

I1s(2,2)=1.; I1s(2,1)=r1s; I1s=I1s/t1s;

I2s(1,1)=1.; I2s(1,2)=r2s; I2s(2,2)=1.;

I2s(2,1)=r2s; I2s=I2s/t2s; Irep1s(1,1)=1.;

Irep1s(1,2)=rrep1s; Irep1s(2,2)=1.; Irep1s(2,1)=rrep1s;

Irep1s=Irep1s/trep1s; Irep2s(1,1)=1.;

Irep2s(1,2)=rrep2s; Irep2s(2,2)=1.;

Irep2s(2,1)=rrep2s; Irep2s=Irep2s/trep2s;

Irep21s(1,1)=1.; Irep21s(1,2)=rrep21s;

Irep21s(2,2)=1.; Irep21s(2,1)=rrep21s;

Irep21s=Irep21s/trep21s; I5s(1,1)=1.;

I5s(1,2)=r5s; I5s(2,2)=1.; I5s(2,1)=r5s;

I5s=I5s/t5s; tempp=I5p; temps=I5s;

aa=1; %to convert to a number, not really necessary

for m1=1:numrep_layers

   last=(m1==numrep_layers);

   last=last*aa;

   tempp=((last*Irep1p)+(~last*Irep21p))*

proprep1*Irep2p*proprep2*tempp;

   temps=((last*Irep1s)+(~last*Irep21s))*

proprep1*Irep2s*proprep2*temps;

end

Stotp=I1p*prop1*I2p*prop2*tempp;

Stots=I1s*prop1*I2s*prop2*temps;
```

```
 tp=1./Stotp(1,1); rp=Stotp(2,1)*tp;
ts=1./Stots(1,1); rs=Stots(2,1)*ts;
```

A.5.2   TCL code for EM Explorer (SPR Model)

```
namespace import emxp::* #
Loop over polarization #foreach POLARIZATION {TE TM} { set
POLARIZATION TM # Loop over angle of incidence for {set THETA 40.0}
{$THETA <= 50.0} {set THETA [expr $THETA+0.2]} { set INFO
${POLARIZATION}_${THETA} # Set incident e-field components if
{$POLARIZATION == "TE"} {
  set EFIELD ey; set EX 0.0; set EY 1.0
} else {
  set EFIELD ex; set EX 1.0; set EY 0.0
} set WAVELENGTH  0.639 set N1 1.518; set K1 0 set N2 0.2; set K2
3.38 set N3 1; set K3 0 set GRIDSIZE 0.0025 # Dielectric thickness
set H1 0.3; set H2 0.05; set H3 0.65
#################################### # Setup Grid
####################################
 #set LX $H1;
  set LX $GRIDSIZE;
 # One cell in y direction
  set LY $GRIDSIZE; # One cell in
y direction
set LZ [expr $H1+$H2+$H3]
 set NX [expr int($LX/$GRIDSIZE+0.5)]
 set NY [expr int($LY/$GRIDSIZE+0.5)]
 set NZ [expr int($LZ/$GRIDSIZE+0.5)]
 set SUBNX 10; set SUBNY 10; set SUBNZ 10
 grid lx=$LX ly=$LY lz=$LZ nx=$NX ny=$NY nz=$NZ subnx=$SUBNX
subny=$SUBNY subnz=$SUBNZ n0=$N1 k0=$K1
#################################### # Setup Geometry (i.e.,
```

```
% the 2nd dielectric)

##########################################

zc=$ZC xw=$XW yw=$YW zw=$ZW n=$N2 k=$K2

set ZTHICK1 $H1 set ZTHICK2

$H2 set ZTHICK3 $H3

set zc1 [expr 0+0.5*$ZTHICK1]

set zc2 [expr (2*$ZTHICK1+$ZTHICK2)/2]

set zc3 [expr (2*($zc2+($ZTHICK2/2))+$ZTHICK3)/2]

set xc [expr 0.5*$LX]

set yc [expr 0.5*$LY]

set x1w $LX;

set y1w $LY set n1 1.518; set k1 0 set

n2 0.2; set k2 3.38 set n3 1;

set k3 0 solid shape=box xc=$xc yc=$yc

zc=$zc1 xw=$x1w yw=$y1w zw=$ZTHICK1 n=$n1 k=$k1 solid shape=box

xc=$xc yc=$yc zc=$zc2 xw=$x1w yw=$y1w zw=$ZTHICK2 n=$n2 k=$k2 solid

shape=box xc=$xc yc=$yc zc=$zc3 xw=$x1w yw=$y1w zw=$ZTHICK3 n=$n3

k=$k3


## # Setup Incident Planewave

######################

#set THETA 44 set KW [expr

2.0*3.1415927/$WAVELENGTH]; # Wave number

set KX [expr $N1*$KW*sin($THETA/180.0*3.1415927)]

set KY 0.0

set ZSRC 0.1

planewave wavelength=$WAVELENGTH z_source=$ZSRC kx_inc=$KX

ky_inc=$KY ex_inc=$EX ey_inc=$EY

####### Setup Convergence Monitor

##########################

set i 0; set j 0; set k 2;
```

```
#set k [expr int(($ZSRC+0.1)/$GRIDSIZE+0.5)]

set INTERVAL 100;

set SAMPLE_SIZE 10;

set TOLERANCE 0.01 convergence field=ex i=$i j=$j

k=$k interval=$INTERVAL sample_size=$SAMPLE_SIZE

tolerance=$TOLERANCE


########## # Output Geometry

#############################

output property=geom

file=geom_${INFO}.vtk


########### # Run the Solver

#######################################

set MAX_STEPS 1000000 run

n_steps=$MAX_STEPS output property=$EFIELD attribute=amplitude

file=${INFO}_a.vtk \

        imin=0 imax=$NX jmin=0 jmax=$NY kmin=0 kmax=$NZ

output property=$EFIELD attribute=phase file=${INFO}_p.vtk \

        imin=0 imax=$NX jmin=0 jmax=$NY kmin=0 kmax=$NZ

output property=convergence file=${INFO}_convergence.txt

########################### # Output E-field

#############################

set property ex; set attribute

amplitude set filename ${INFO}_ex_all.vtk

set imin 0; set imax $NX

set jmin 0; set jmax $NY set kmin 0;

set kmax $NZ #output

property=$property

attribute=$attribute file=$filename \

imin=$imin imax=$imax jmin=$jmin
```

```
jmax=$jmax kmin=$kmin kmax=$kmax
###################################

set filename reftry_${INFO}.vtk set K1 0 ;

set K2 $NZ set I1 1 ; set

I2 1 set Y1 [expr $NY/2] ;

 set Y2 $Y1 output property=$property

attribute=$attribute file=$filename \

       imin=$I1 imax=$I2 jmin=$Y1

jmax=$Y2 kmin=$K1 kmax=$K2

output property=ez attribute=amplitude file=

ez_${INFO}_amp.vtk \

       imin=$I1 imax=$I2 jmin=$Y1 jmax=$Y2 kmin=$K1 kmax=$K2


#scat_func property=$EFIELD attribute=amplitude

file=scat_${INFO}_a.vtk k=$NZ #scat_func property=$EFIELD

attribute=phase

file=scat_${INFO}_p.vtk k=$NZ #scat_func

property=$EFIELD attribute=amplitude file=scatft_${INFO}_a.vtk k=$NZ

ft=true #scat_func property=$EFIELD attribute=phase

file=scatft_${INFO}_p.vtk k=$NZ

ft=true #farfield property=$EFIELD

attribute=amplitude file=far_${INFO}_a.vtk

 zmin=100 zmax=110 nz=10

#farfield property=$EFIELD attribute=phase

file=far_${INFO}_p.vtk zmin=100 zmax=110 nz=10 #image

file=image_${INFO}.vtk k=$NZ na=0.8 mag=2.5 defocus1=-0.1

defocus2=0.1 ndefocus=2 #set k [expr

int(($ZSRC+0.1)/$GRIDSIZE+0.5)+2]

 set k 2 gauge action=measure

info=R_$INFO i=0 j=0 k=$k
```

```
k=$k gauge action=output file=gauge.txt reset;
# This is very important
gauge action=output file=
gauge.txt gauge action=reset exit
```

A.5.3 Farfield Extraction Script

The following is the exact script for the farfield extraction and the implementation of the Stratton-Chu equations.

```
%% function [total_field,E_ff_x,...
E_ff_y,E_ff_z]=calculate_farfield(fem,...
radius,offset,N,angles,sols,w,n1)
%% function for calculating the FF
% from a comsol fem structure
%% assuming:
%% radius=radius of the ff source
%% offset=offset of the ff source
%% N=number of points to
% % extract from the model for the source
%% angles=the angles in the farfield that you want
%% sols =the solutions you want to use from the model
%% w=the FREE SPACE wavelength
%% n1=the (uniform) refractive index at / inside the source
%%
%% This model uses a Gaussian wrap around
% % apodisation technique to permit
%% the calculation of the FF
% % for source free areas / volumes
%% This effectively sets the field at
% %  the back of the source to 0 and the
%% field at the front to 1, it has nice
```

```
% % soft edges to reduce ringing and
%% side band effects while maintaining good resolution



function [total_field_old,E_ff_x,E_ff_y,...
field_raw,tot_phase,total_field_new]
=calculate_farfield(fem,radius,...
offset,N,angles,sols,w,n1,arg2)


%% set some constants -
% really we should read these from the model (fem)
eps_0=8.8541878e-12;
mu_0=4*pi*1e-7;


%% assume we are encased in an
% infinite medium refractive index n1
k1=2*pi*n1/w;
mu_r1=1;
eps_r1=n1.*n1/mu_r1; %% n=sqrt(eps_r mu_r)
neta1=sqrt(mu_0./eps_0.*mu_r1./eps_r1);


%% describe circle for the source
ang=linspace(-pi,pi,N);
x=offset(1)+radius*cos(ang);
y=offset(2)+radius*sin(ang);
z=0;
%% we are 2d hence z=0 but we use 3d
% coords because the field may have z components
p=[x;y]; %% convert coords to comsol point array


dang=ang(2)-ang(1);
```

```matlab
dS=dang*radius; %% length constant of integration


%% find the normal, r0 and position vectors
% (since we are center on 0,0,0 then they are the same)
nx=cos(ang); ny=sin(ang);    nz=0;
n(1,:)=nx;  n(2,:)=ny;      n(3,:)=nz;
r(1,:)=x;     r(2,:)=y;       r(3,:)=z;


%% Calculate the Gaussian aperture
%% the angle offset compensates to that the
% aperture angle and the dir of the ff match
angle_offset=0;%N/2
aperture=gaussian(N/8,N);
[Ex,Ey,Ez,Hx,Hy,Hz]=postinterp(fem,'Ex_emwh','Ey_emwh',...
'Ez_emwh','Hx','Hy','Hz',p,'solnum','all');


for sol=1:length(sols)
%% convert to vector form and remove
%  nans (they represent zero components in the field).
E(1,:)=nanto0(Ex(sol,:));
E(2,:)=nanto0(Ey(sol,:));
E(3,:)=nanto0(Ez(sol,:));
H(1,:)=nanto0(Hx(sol,:));
H(2,:)=nanto0(Hy(sol,:));
H(3,:)=nanto0(Hz(sol,:));
%% gather data for a crude normalisation process
average_field=mean(mean(abs(E)));
different_field(sol) = max(sqrt(sum( E.*conj(E))));
for i=1:length(angles);
%% calculate the position vector r0
r0(1,1)=cos(angles(i));
```

```
r0(2,1)=sin(angles(i));

r0(3,1)=0;

r00=repmat(r0,1,N); %% matrix form

%% compute the aperture at the FF angle

ap_shift=mod(round(N/2/pi*angles(i))+angle_offset,N);

ap=repmat(shift_aperture(aperture,ap_shift),3,1);

%% impose the aperture on the fields

if strcmp(arg2,'no_ap')

Ep=E;

Hp=H;

else

Ep=E.*ap;

Hp=H.*ap;

end

%% calculate the FF by intergrating around the source

E_ff(:,i)=sqrt(-1).*k1/4/pi.*...

cross(r0,sum((cross(n,Ep)-neta1.*...

cross(r00,cross(n,Hp))).*...

exp(sqrt(-1).*k1.*repmat(dot(r,r00),...

3,1)),2).*dS);

end

store(:,:,sol)=E_ff./average_field;

store1(:,:,sol)=

(sqrt((sum(E_ff.*conj(E_ff))))).  /different_field(sol);

store_raw(:,:,sol)=E_ff;

end


%% get the total field...

total_field_old=sqrt(squeeze(sum(store.*conj(store))));

%total_field_old=sqrt(sum(store.*conj(store)));

total_field_new=squeeze(store1);
```

```
field_raw=store_raw;
tot_phase=squeeze(unwrap(angle(store)));
E_ff_x=E_ff(1,:);
E_ff_y=E_ff(2,:);
E_ff_z=E_ff(3,:);
store1=store1;
%%%%%%%%%%%%%%%%%%%%%%

function n=nanto0(n)
i=find(isnan(n));
n(i)=zeros(size(i));


%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%
function b=shift_aperture(a,n)
m=length(a);
b((n+1):m)=a(1:(m-n));
b(1:n)=a((1+m-n):m);


%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%5
function g=gaussian(w,N,off,norm)
if(nargin==2)
off=0;norm=0;
end;
if(nargin==3)
norm=0;
end;


w=w*w;
if(length(N)==1)
```

```
h=N/2;
x=-h:(h-1);
else
x=N;
end
x=x+off;
g=exp(-x.*x./w);
if(norm==1)
g=g./sum(g);
end;
```

Appendix B

Appendix B

B.1 Operation of Perfectly Matched Layers

Perfectly matched layers are used in all the models mentioned in this thesis. In some models, their effects or inefficiencies remain invisible, whereas in some other models, their artifacts due to reduced performance is very visible. In order to understand the nature of such artifacts, it is necessary to have some understanding of the workings of the PMLs.

In this section, the concept of complex coordinate stretching, and its application in the implementation of the artificial absorbing layers, PMLs, is described for rectangular coordinates for simplicity.

When truncating an open surface, a coordinate at the interface between the PML and the simulation domain can be expressed as (n, t_1, t_2). Where $\vec{t_1}$ and $\vec{t_2}$ are tangential components and $\vec{n} = \vec{t_1} \times \vec{t_2}$ is the normal component of the unit vector to this interface. PMLs work by the complex stretching of the coordinate n [191, 186]. Comsol does this stretching via Equation B.1. This specific equation involves some terms that works towards the optimisation of the PML parameters according to the PML size and the incoming wavelength. The PML layers have a 3x3 tensor of complex elements as a dielectric constant, and a complex magnetic permeability [192].

$$n' = n\frac{\lambda}{\delta n}(1 - i) \tag{B.1}$$

The coordinate n is replaced by its complex counterpart n'. δn is the thickness of the

PML layer, and λ is conventionally the wavelength of the incoming wave in the matched medium. More generally, this stretching can be expressed as Equation B.2.

$$n' = \int_0^n s_n(\xi)d\xi \tag{B.2}$$

In this formulation s_n is the coordinate variable. After the stretching, the field on the modified coordinate is expressed by a modified set of equations as in Equation B.3 [191].

$$\begin{aligned}
\tilde{\nabla} \times \mathbf{E}^c &= j\omega\mu\mathbf{H}^c \\
\tilde{\nabla} \times \mathbf{H}^c &= j\omega\epsilon\mathbf{E}^c \\
\tilde{\nabla} \cdot \mathbf{D}^c &= 0 \\
\tilde{\nabla} \cdot \mathbf{B}^c &= 0
\end{aligned} \tag{B.3}$$

The term complex stretching is frequently used in this context and the effects of it can be explained as follows. A coordinate is being stretched into a complex number so as to be able to make up an imaginary part which will add a term that will cause absorption. The stretched coordinate will make the PML domain appear as a wavelength long,in terms of the stretched coordinates. The new fields $\mathbf{E}^c$ and $\mathbf{H}^c$ on the stretched coordinates can be related to the original ones by Equation B.5. The tangential components of both set of fields obey the same boundary conditions.

$$\begin{aligned}
E_n &= s_n E_n^c, \\
E_{t_1} &= s_{t_1} E_{t_1}^c \\
E_{t_2} &= s_{t_2} E_{t_2}^c
\end{aligned} \tag{B.4}$$

$$\begin{aligned}
H_n &= s_n H_n^c \\
H_{t_1} &= s_{t_1} H_{t_1}^c \\
H_{t_2} &= s_{t_2} H_{t_2}^c
\end{aligned} \tag{B.5}$$

Perfectly matched layers, in the transformed coordinates, can be viewed in terms of the

general coordinates as a layer with anisotropic permittivity and permeability, $\tilde{\mu}$, $\tilde{\epsilon}$ [107]. These two are represented as complex tensors (see Equation B.6.)

$$\frac{\tilde{\mu}}{\mu} = \frac{\tilde{\epsilon}}{\epsilon} = \begin{bmatrix} \frac{s_{t_1} s_{t_2}}{s_n} & & \\ & \frac{s_n s_{t_2}}{s_{t_1}} & \\ & & \frac{s_{t_1} s_n}{s_{t_2}} \end{bmatrix} \tag{B.6}$$

The vector form of the fields in Cartesian System can be written as in Equation B.7. For TM waves, the z-component will be assumed to be zero. Equation B.7 reflects this assumption together with the fact that we will be dealing with time-harmonic waves (absence of t from equations).

$$\mathbf{E} = \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} \exp\left(-j(k_x x + k_y y + k_z z)\right) \\ \exp(-j(k_x x + k_y y + k_z z)) \\ \exp(-j(k_x x + k_y y + k_z z)) \end{bmatrix}$$

$$E_z = 0 \quad \text{for p-polarisation}$$

$$k_z = 0 \quad \text{no waves propagating in z-direction}$$

$$\therefore \mathbf{E} = \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} \exp\left(-j(k_x x + k_y y + k_z z)\right) \\ \exp(-j(k_x x + k_y y + k_z z)) \\ 0 \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} \exp\left(-j(k_x x + k_y y)\right) \\ \exp\left(-j(k_x x + k_y y)\right) \\ 0 \end{bmatrix} \tag{B.7}$$

In the simple cases, when x is transformed into the complex $\tilde{x} = x + jf(x)$, then the field expressions change as follows.

$$\mathbf{E}^c = s_n \begin{bmatrix} E_x^c \\ E_y^c \\ E_z^c \end{bmatrix} = s_n \begin{bmatrix} \exp(-j(k_x\tilde{x} + k_yy)) \\ \exp(-j(k_x\tilde{x} + k_yy)) \\ 0 \end{bmatrix} \quad \text{(B.8)}$$

$$= s_n \begin{bmatrix} \exp(-j(k_x\mathrm{Re}(\tilde{x}) - jk_x\mathrm{Im}(\tilde{x}) + k_yy)) \\ \exp(-j(k_x\mathrm{Re}(\tilde{x}) - jk_x\mathrm{Im}(\tilde{x}) + k_yy)) \\ 0 \end{bmatrix} \quad \text{(B.9)}$$

The coordinate stretching is done such that $\mathrm{Re}(\tilde{x}) = x$ for this simple case. When the stretching also scales the real part, only the algebra changes. In both cases, an exponentially decaying term appears in the field expressions once the complex coordinate transformation is complete.

$$\mathbf{E}^c = s_n \begin{bmatrix} \exp(-jk_xx - k_x\mathrm{Im}(\tilde{x}) - jk_yy)) \\ \exp(-jk_xx - k_x\mathrm{Im}(\tilde{x}) - jk_yy)) \\ 0 \end{bmatrix}$$

$$= s_n \begin{bmatrix} \exp(-j(k_xx + k_yy))\exp(-k_x\mathrm{Im}(\tilde{x})) \\ \exp(-j(k_xx + k_yy))\exp(-k_x\mathrm{Im}(\tilde{x})) \\ 0 \end{bmatrix}$$

$$= s_n \exp(-k_x\mathrm{Im}(\tilde{x})) \begin{bmatrix} \exp(-j(k_xx + k_yy)) \\ \exp(-j(k_xx + k_yy)) \\ 0 \end{bmatrix} \quad \text{(B.10)}$$

$$= s_n \exp(-k_x\mathrm{Im}(\tilde{x}))\mathbf{E} \quad \text{(B.11)}$$

In the wave equations, the differential operator becomes $\partial\tilde{x} \rightarrow (1 + j\frac{df}{dx})\partial x$. If we replace $\frac{df}{dx} \rightarrow \frac{\sigma_x(x)}{\omega}$, the partial derivative $\frac{\partial}{\partial x}$ becomes as it is in Equation B.12. For when $\sigma_x(x) = 0$ the partial derivatives, hence the field equations is unchanged. When $\sigma_x(x) > 0$ then the field is exponentially decaying according to Equations B.12 and B.11. This derivative has to do with the decay rate.

$$\frac{\partial}{\partial x} \rightarrow \frac{1}{1 + j(\frac{\sigma_x(x)}{\omega})}\frac{\partial}{\partial x} \quad \text{(B.12)}$$

B.2   Problems with Perfectly Matched Layers

Perfectly Matched Layers are very good solutions to the truncation problem. However perfect they are in theory, once the numerical errors come into play, their performance start deteriorating. Problems with the PML layers arise when there aren't enough mesh elements in the model. In this case, the result is spurious reflections. Once the domain, hence the exact wave equations are discretised, there will be numerical differences in the actual values and the matched values from the PML. The boundary separating the actual domain and the PML will thus no longer be transparent(reflectionless). This imperfection will cause numerical reflections, which can be minimised when the exact wave equation is efficiently discretised. These numerical reflections are not due to any physical occurrence, they are simply spurious [107, 193] and need to be minimised. Further minimisation can be achieved by keeping the $\sigma_x(x)$ in equation B.12 small, therefore keeping the exponential decay rate smaller. This will limit sudden numerical transitions. Sudden transitions will correspond to high spatial frequencies and the need to use a finer grid(to be able to represent such frequencies). However, when the rate is too small, the PML layer will need to be much wider, so as to be able to eradicate the incoming waves by the time the layer is terminated by a hard-wall boundary condition such as Perfect Electric Conductor(PEC). This is a disadvantage, because the number of elements will grow with the size of the PML and the computing requirements will therefore increase. An optimal value needs to be found for each model of interest.

Another inefficiency of the PMLs is the angle dependency. It can be seen from Equation B.11 that exponential decay carried out by this layer is dependent on k_x. This term is therefore angle dependent, because $k_x = kcos(\theta)$, where θ is the angle of incidence (conventionally taken as the angle made with the x-axis). Therefore as $\theta \rightarrow 90°$ $k_x\text{Im}(\tilde{x}) \rightarrow 0$, $\exp(-k_x\text{Im}(\tilde{x})) \rightarrow 1$. This means that the decay rate will go to 0 for when grazing incidence is approached, and since the exponential goes to one, the field in PML approaches to the field in the actual simulation domain.

Another case when the PML will fail is when the medium is not x-invariant, i.e it is inhomogeneous in x (or y depending on the orientation). This case however will not be covered since it doesn't relate to the concepts covered in this thesis. Related to this
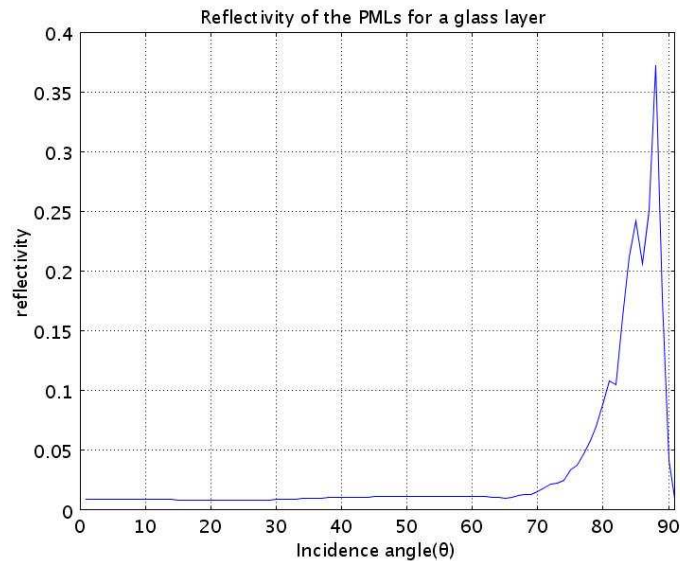
Figure B.1: The reflection coefficient of the right PML when the θ_{in} is varied in the model. The decrease in the performance of the PMLs as the θ_{in} is observed. It should be noted that as 90° incidence angle is reached, the reflectivity is reduced, not due to the performance of the PMLs, but due to the terminating boundary condition, i.e perfect electric conductor.

concept, when there are Bloch waves (periodicity)present in the stretched direction, PMLs will not work.

Comsol 3.4 has an automatic script to optimise the PML parameters, if the coordinate system is chosen to be one of the default settings listed in the software. For example, for cartesian coordinates, it automatically adjusts the decay constant so as to make it optimal. Therefore changing PML layer width slightly not have much effect on their reflectance. However, this automatic setting can be disabled to make the decay constant manual for unconventional case.

Figure B.1 shows the reflection coefficient of a PML (eg. right PML in figure 2.9) with respect to the angle of incidence.

B.3 Periodic Boundary Conditions (PBCs) in Comsol

Comsol implements PBCs using "source" and "destination" boundaries. Simply explained, Comsol will essentially interpolate the variables of interest in the source and destination

boundaries along the boundary elements [194] and establish the necessary boundary relationships. This relationship effectively nests the unit cells together this creating a periodic structure. For Floquet periodicity the following relationship is useful.

$$
\begin{aligned}
E_{src} &= \exp\left[-j(k_x x + k_y y_s)\right] \\
E_{dest} &= \exp\left[-j(k_x x + k_y y_d)\right] \\
y_s &> y_d \\
E_{src} E_{\delta\phi} &= E_{dest} \\
E_{\delta\phi} &= \exp\left[jk_y(y_s - y_d))\right]
\end{aligned}
\tag{B.13}
$$

B.4   Implementation of Periodic Boundary Conditions

Floquet theorem uses a phase term to relate fields at different locations within a periodic structure(see Equation B.14). This is used in the implementation of periodic boundary conditions. The periodic boundaries are stitched via equation B.13 for arbitrary angle of incidence.

$$
E_x(x,y)|_{b1} = E_x(x,y)\exp(j * phase)|_{b2}
\tag{B.14}
$$

It is important to understand the capabilities of this sort of simulations, in order to be able to correctly model structures. The math behind the operation of periodic boundaries is shown below. $phase = k_n * sin(\theta_{in}) * L_y$ is the factor expressing a generalised form of periodicity for input angles.

The phase difference is calculated taking the actual geometry size L_y and the angle of incidence into consideration. At the source boundary, the field is extracted and the predetermined phase difference between the boundaries is enforced before equating the fields at source and destination boundaries. Consider figure B.2, assuming it is positioned on the x axis, so that the bottom boundary corresponds to y=0 , depending on the angle of incidence, the field on the top boundary can be expressed as follows:

$$E = \exp(-j(k_0 sin(\theta_i)y + k_0 cos(\theta_i)x)) \tag{B.15}$$

We can ignore the term with x in it, since x is not periodic and x direction is truncated with PML so that any excessive component will be absorbed. However, at the source boundary, y=d (which is a multiple of the periodicity), the field can be expressed as follows.

$$E = \exp(-jk_0 sin(\theta_i))d) \tag{B.16}$$

At the destination boundary, y=0

$$E = \exp(-jk_0 sin(\theta_i)y) = \exp(k_0 sin(\theta_i)0) = 1 \tag{B.17}$$

Where, d is the period of the model, which corresponds to the length of the geometry in y-direction. Using angle of incidence as a variable in the model, it is possible to calculate the phase term in Equation B.16. The field incoming to a periodic structure, may be scattered at angles related to the period. This situation is very common in diffraction gratings. Structures such as the plasmon resonator and the dCHOT structures, closely resemble this situation. In this case, the scattered field and the incoming field can be related to the periodicity. Following set of equations explain this situation.

$$k_y = k_o sin\theta \tag{B.18}$$

$$k_x = k_o cos\theta \tag{B.19}$$

$$sin\theta_{T_1} = \frac{\lambda}{\lambda_g}m + sin\theta_f \tag{B.20}$$

$$k_o sin\theta_{T_1} = k_o \frac{\lambda}{\lambda_g}m + k_o sin\theta_f \tag{B.21}$$

$$k_o sin\theta_{T_1} = \frac{2\pi}{\lambda}\frac{\lambda}{\lambda_g}m + k_o sin\theta_f \tag{B.22}$$

$$k_g = \frac{2\pi}{\lambda_g} \tag{B.23}$$

$$k_o sin\theta_{T_1} = k_g m + k_o sin\theta_f \tag{B.24}$$
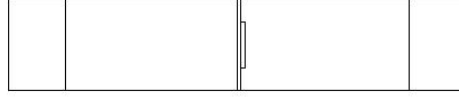
Figure B.2: The geometry used to model gratings and similar periodic structures. The leftmost and the rightmost subdomains are PMLs and the model is periodic. All the scattered fields from this structure are periodic with the geometry size in the vertical direction.

Where, T_1 : First transmitted order from a periodic structure.

k_o : k-vector in air

λ : wavelength

λ_g : grating period

θ_f : incidence angle

k_g : grating wave-vector

Electric field at the top boundary can be expressed as $E = E_o \exp(-jk_y y)$. substituting Equation B.24 into this expression will yield

$$E = E_o \exp(-j(k_o sin(\theta_{T_1})y) \tag{B.25}$$

$$
\begin{aligned}
E &= E_o \exp(-j(k_g m + k_o sin\theta_f)y) \\
&= E_o \exp(-j(k_g m)y) \exp(-j(k_o sin\theta_f)y)
\end{aligned}
\tag{B.26}
$$

In equation B.26, $\exp(-jk_g my)$ is periodic in the direction parallel to the metal surface. Whereas, $\exp(k_0 sin\theta_f y)$ will be taken care of via the Floquet theorem.

The relationship between the incoming light and the scattered light due to the grating enables the adjustment of the periodic boundary conditions so as to accommodate each angle of interest, as long as the length of the geometry is an integer multiple of the periodicity being considered. The boundary conditions at the periodic boundaries are adjusted so as to equate the part of the k_o that fits into the length of the geometry in the

direction of periodicity, then each angle may be accommodated for. The adjustment is done by subtracting part of the k vector that would otherwise exceed the length of the geometry, also called "phase stitching", followed by equating the field at the periodic boundaries. The excess phase of the E field normal to the periodicity is subtracted and equalised to the field at the next periodic boundary so as to accurately implement the physical situation. This is implemented automatically in Comsol, where choices of symmetric, antisymmetric and Floquet periodicity conditions are available.

B.5   Reflectivity Extraction from flat interfaced models

When simple periodic models including only flat interfaces are modelled, data such as reflectivity may be obtained looking at the standing wave patterns in the nearfield data (refer to Figure B.3). This is due to the fact that there is only one spatial frequency component present in the model, therefore there is only one reflected wave direction for each incoming wave direction. This type of reflected spectrum can be seen as a simple standing wave pattern. When the absolute value of H_z is plotted, the maximum and the minimum are related to the incoming and the reflected field values. The troughs of this sinusoid looking curve is the reflected field subtracted from the incoming field, and the peaks correspond to where the incoming wave constructively interferes with the reflected field. As can be seen from FigureB.3b, the $abs(H_z)$ looks like a sinusoidal, if a horizontal section is to be extracted.

The reflectivity in the models where only flat layers were modelled, were obtained by extracting the absolute value of Hz along the x-axis (it is constant along the y-axis) and calculating the minima and the maxima. The relationship may be explained in Equation B.31.
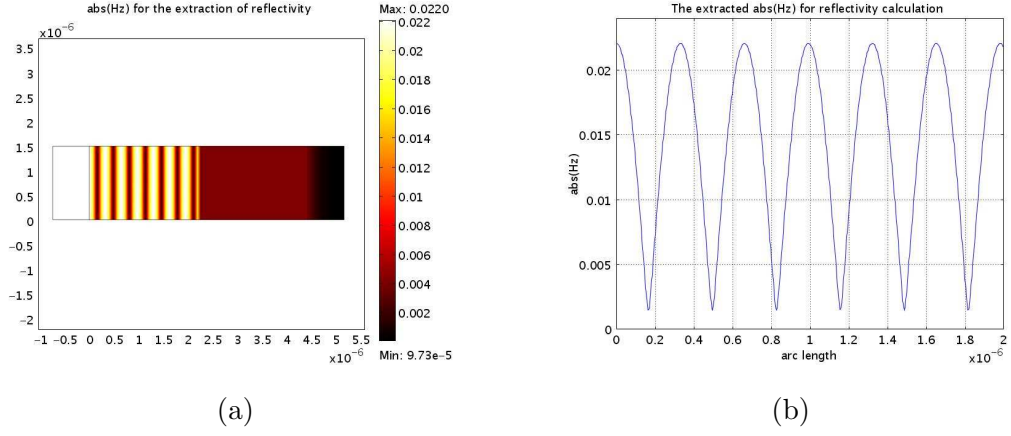
Figure B.3: **(a)** abs(H_z) at SPR angle. It can be seen the along the x-axis, it is like a sinusoid **(b)** The sinusoidal looking nature of abs(Hz) is used in the calculation of reflectivity.

$$r = I_r/I_i \tag{B.27}$$

$$max|Hz| = I_r + I_i, min|Hz| = I_i - I_r \tag{B.28}$$

$$I_i = \frac{max|Hz| + min|Hz|}{2} \tag{B.29}$$

$$I_r = \frac{max|Hz| - min|H_z|}{2} \tag{B.30}$$

$$r = \frac{max|Hz| - min|Hz|}{max|Hz| + min|Hz|} \tag{B.31}$$

B.6 Farfield Calculation using Stratton-Chu Formulation

Stratton-Chu formulation is an integral solution to Maxwell's equations [195]. The formulation was made in an effort to describe diffraction through holes around the size of the incoming wavelength, in order to fix the conceptual errors made in the application of the Kirchoff Diffraction problem [196]. Kirchoff formulation had three main assumptions that formed the basis of his calculations [197]. These could be summarised as follows:

1. The numerical aperture (NA) is reasonably small

2. The incident radiation is unpolarised

3. The Fresnel number ($N = \frac{a^2}{\lambda f}$) is very large, i.e. $N \gg 10$.

Where, λ is the wavelength of the radiation

a is the radius of the aperture

For microwave analysis, these approximations/assumptions became invalid. In an effort to circumvent this problem, Stratton and Chu used a vectorised form of Green's Theorem, to derive integral solutions to Maxwell's equations. It came to be known as vector diffraction theorem [88, 196]. Another approach to the same problem has been done by Fradin, by using the Lorentz reciprocity [195]. The detailed derivation of the formulations are beyond the scope of this section. But the general form and its conceptual derivation is mentioned here. Stratton-Chu equations can be written in general form as equation B.32

$$\mathbf{E}(\mathbf{x}) = \int_S \left[ik(\mathbf{n} \times \mathbf{B})G + (\mathbf{n} \times \mathbf{E}) \times \nabla'G + (\mathbf{n} \cdot \mathbf{E})\nabla'G \right] dS' \qquad \text{(B.32)}$$

Where, $\mathbf{n}$: unit normal vector on the surface of interest

G :is the vector Green's function

The Green's function is a primary method to solve inhomogeneous boundary value problems.[198]. In bounded domains, any representation formula, whether it is in electrodynamics or electrostatics, has to take into consideration the boundary values [199]. Green's function for an operator L, is the function which will yield the a solution. This technique may be used to solve problems in the form of $\nabla^2 u(\mathbf{x}) = f(\mathbf{x})$.

The Green's function takes the form, $\mathbf{G} = \frac{e^{-jkR}}{R}\mathbf{r}$, where R is the distance between the source point and the observation point. This simply describes a spherical wave, also known as Huygen's secondary sources/wavelets.

When the farfield approximation is applied to the equation B.32. Green's function which involves the position vectors will change as follows [200] (refer to equation B.34) and the generalised Stratton-Chu Integral Equation will transform into its far field specialised form as in equation B.35. Note that this is the special case for Stratton-Chu equations for antennas when there is a source enclosed within the integral surface [1]. In this case,

---

[1]derivation outside the scope of this thesis

the charge distribution on the antenna itself is not known/necessary but the fields around such a source are known within a certain accuracy.

$$G(\mathbf{x}, \mathbf{x_0}) \quad = \quad \frac{e^{-jk|\mathbf{r}-\mathbf{r_o}|}}{4\pi|\mathbf{r} - \mathbf{r_o}|} = \frac{e^{-jk_0 r}}{4\pi r} e^{jk_0(\mathbf{r}\mathbf{r_o})} \tag{B.33}$$

$$\nabla' G(\mathbf{x}, \mathbf{x_o}) \quad = \quad j\mathbf{r}k_0 \frac{e^{-jk_0 r}}{4\pi r} e^{jk_0 \mathbf{r}\mathbf{r_o}} \tag{B.34}$$

Therefore, in the FEM simulations, when the near field solution is obtained for the fields, it is integrated along a boundary or a surface and evaluated at a point p at infinity with a predefined angular position [186] in order to get the far field distribution at the limit of infinity. In Comsol's calculations, the radiating source, e.g. the antenna is considered to be in the vicinity of the origin.

$$E_p = \frac{jk_0}{4\pi} r_0 \times \int [\mathbf{n} \times \mathbf{E} - \eta_0 \mathbf{r}_0 \times (\mathbf{n} \times \mathbf{H})] exp(jk_0 \mathbf{r}.\mathbf{r}_0) dS \tag{B.35}$$

Where, $\mathbf{E}$ and $\mathbf{H}=$: Electric and Magnetic Fields on the aperture, ie the surface S enclosing the sou

$\quad$ $\mathbf{r}_0 =$: unit vector pointing from origin to the field point p

$\quad$ $\mathbf{n} =$: unit normal to surface S

$\quad$ $\eta_0 =$: free space impedance $\eta_0 = \sqrt{\mu_0/\epsilon_0}$

$\quad$ k_0 : Free space wave number

$\quad$ $\mathbf{r}$: radius vector of the surface S.(not a unit vector)

$\quad$ $\mathbf{E_p}$: calculated far field at point p

Equation B.35 is originally designed and set to work in free space, since it was meant for antenna design. It is used when there is a source in the enclosed volume, to calculate the far field radiation pattern. It is actually a special case of the Original Formula, which is given in equation B.32. In the simulations considered in this thesis, there is generally no source enclosed in the integration surface. In such cases this formula must return zero, (or numerical noise, refer to figure B.6). However, we will use an apodisation technique (described in section 2.8.4) which will modify the extracted near field, so as to consider only the fields along the farfield observation angle considered.

B.6.1 Far Field Concept in Comsol

Comsol calculates the far electric field using a formulation called Stratton-Chu Formulation [186] [2] set to be used only in air.

Comsol requires an explicit definition of a boundary which encloses a source, a radiating element, located at origin. The farfield is extracted by applying the Stratton-Chu(refer to equation B.35) formula to the enclosing boundary, i.e the circle on the left in the glass medium in Figure B.4. In Comsol by default, the model is meshed and solved first and updated afterwards in order to carry out the integral along the specified far field source boundary. There are some issues that must be addressed when using Comsol's farfield extraction. Some of these issues eventually proved limiting and caused replacement of Comsol's way with a manually constructed script.

In the case of the periodic structures, from a mathematical modelling point of view, an infinitely repeating structure is being represented by formulating a solution for one or two repeating elements with appropriate boundary conditions. The far field will be extracted from the near field. The near field (source) that needs to be integrated will be of finite size since the solution is of finite size. But the theoretical solution is valid for an infinite domain. The source boundary that will be integrated will therefore represent an aperture, hence the formulation will correspond to that of Stratton-Chu. The situation resembles looking into the solution for an infinite size model through a finite aperture. Therefore, the shape and the size of the aperture will have a signature on the results.

In theory, distinct frequency components for infinite structures/sources are to be represented with infinitely thin spikes, i.e delta functions in the Fourier Domain. However, the representation of the infinite domain via PBCs and making the farfield source finite size (equal to the periodicity), will have consequences as described below. The spike in the theory will have to be modified with the spectrum of the aperture. The profile/shape of the aperture in Fourier Domain will depend on the shape of the source boundary as seen from the specific far field angle.

This area was chosen to be circular, so that it would look the same from all farfield

---

[2]Observations and notes on Comsol's farfield extraction were based on the Comsol Multiphysics Version 3.4
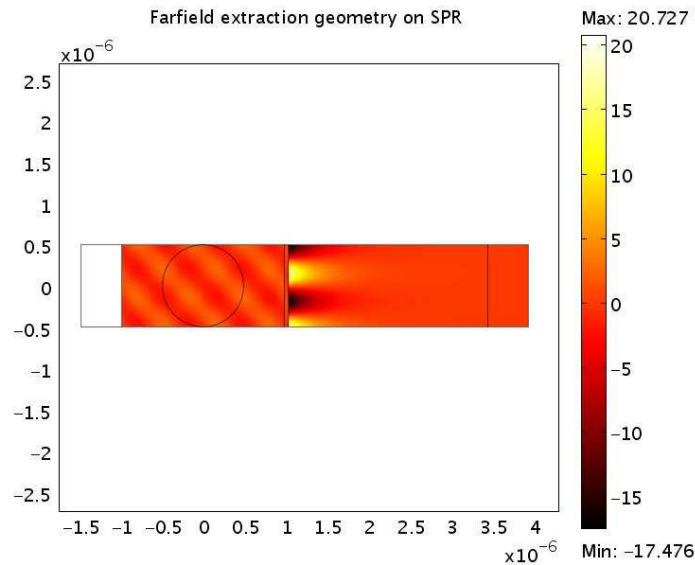
Figure B.4: The geometry for a simple 2D simulation designed for the extraction of farfield. Farfield source is the circle on the left hand side, which corresponds to glass medium.

angles. Had it been any other shape than a circle, it's width would have effectively been different at certain farfield angles, giving rise to nonuniform widths, (and heights, due to the conservation of energy) of aperture profiles in the Fourier domain, with respect to various farfield angles.

In the case where the aperture is circular, as it will be in this thesis, the effect of aperture size can be seen in the farfield distributions. (refer to Figure 2.11).

The shape of the aperture will result in different profile shapes, where the size will result in different profile widths. The width of the profile will affect the resolution of the extracted spectrum. The diameter of the aperture was generally chosen to be as big as possible, so as to increase the farfield observation angle resolution. For periodic structures, it was chosen to be at least one period of the structure. It is seen that for structures with different periodicities, the aperture size is also different. Different aperture sizes will appear in the results with different profile widths. Smaller apertures will yield larger profiles (refer to Figure 2.11).

In models such as the plasmon resonator, 1D photonic crystal and eCHOT structure, we are looking to analyse the farfield distribution at a reflected side. Since there is an infinite length assumption, it is not possible to surround such structures with a farfield
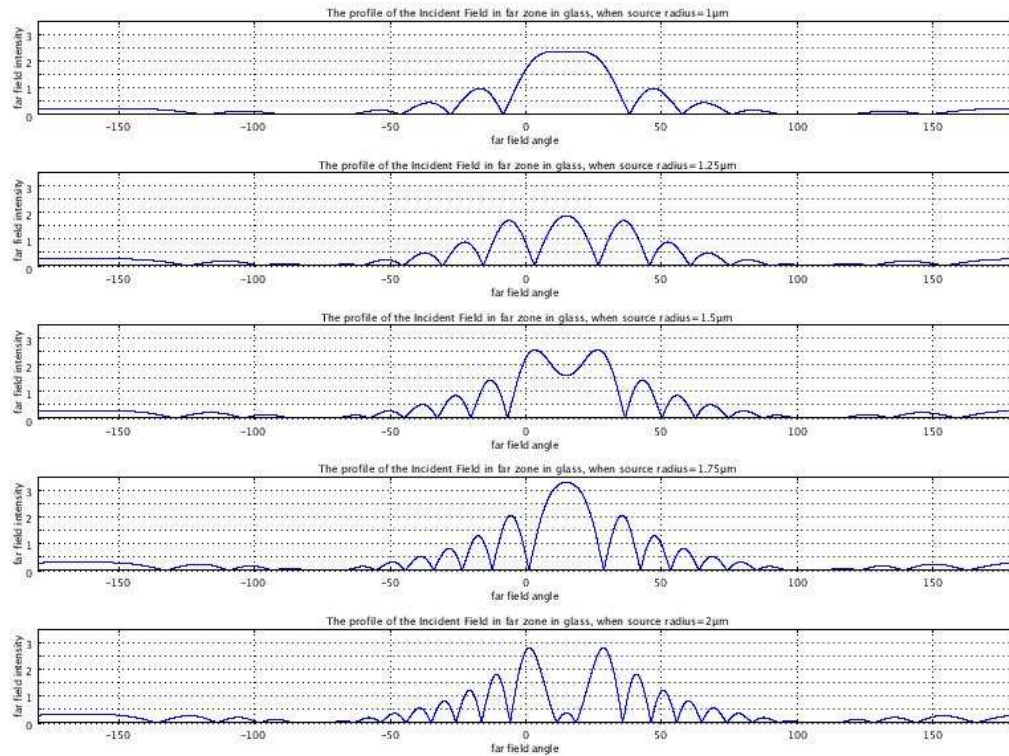
Figure B.5: The far field profiles for a plane wave travelling in glass medium at 15 ° with different source radii (top to bottom: $1\mu m$, $1.25\mu m$, $1.5\mu m$, $1.75\mu m$, $2\mu m$). The profiles are obtained by using Comsol's default far field calculation methods. For some radii the results are misleading, and look correct. However, changing the source size, changes the shape of the profile rather than the width. The main lobe which would correspond to the incoming field, is overtaken by the sidelobes for some source sizes. This shouldn't happen for a plain wave travelling in an isotropic medium. This could be explained as: When the plane wave is going across the aperture, the fields crossing one side of the boundary gets cancelled with fields crossing the opposite side, due to oppositely directed surface normals on the corresponding points along the boundary. Hence the outgoing field will be shadowed by the incoming one.

circle. Therefore it is only possible to look at a reflected/transmitted side, which will yield a farfield source enclosing empty domain. This concept raises an important limitation of Comsol's default farfield extraction method. The Stratton-Chu formulation in Comsol [186] gives the farfield distribution for when there is a net source surrounded by the "source boundary". Otherwise, the normal vectors in the opposite sides of the aperture, will yield in a cancellation, and the results obtained will be merely numerical artifacts. This concept is referred to as *boundary shadowing*. Shadowing happens when there is a plane wave entering and exiting the far field source without change such as reflection etc. When these components are summed up over the source domain, they theoretically cancel each other due to the normal vectors being of opposite sign. In this frame, the boundary through which the plane wave enters the source is said to "shadow" the exit boundary. Figures B.5 show that for some parameters, the mainlobe for some source sizes will dissappear, whereas for some others it looks normal. This concept is important in this thesis, because it resulted in the manual implementation of the Stratton-Chu formula, in order to circumvent problems associated with default means of farfield extraction in Comsol.

The formulation for the farfield extraction in Comsol is meant to be used in free space. The formula was designed for diffraction through apertures [88], and specialised for antenna radiation into free space. Hence for models where the far field is necessary at another medium, the formulation leads to wrong conclusions. Comsol doesn't recognise that the fields are not in free space (as in the models that will be discussed in this report), so it doesn't make necessary adjustments for variables such as the wavelength in a specific medium, and constants such as the permittivity or permeability in the Equation B.35. Comsol 3.4 farfield extraction by default is capable of implementing only this equation whatever the surrounding medium is. Hence, for media other than air, it will always arrive at a wrong result. In order to be able to implement the Stratton-Chu formulation in any other medium, it was necessary to modify the original equation so that the related parameters could be set to be in the medium of our choice.

Comsol's bugs in the implementation of the far field Stratton-Chu equation, together with the above described concept of boundary shadowing yielded erroneous results. Consequently, it was decided that far field extraction was to be achieved by means of manual
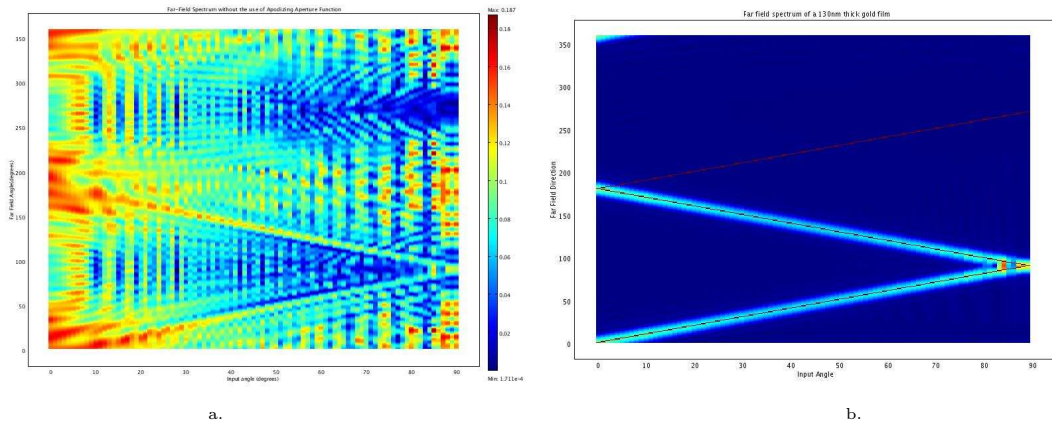
Figure B.6: **a.**The effect of boundary shadowing in the absence of the apodising aperture function. It can be seen that the fields cancel each other mostly. There is a very faint locus that could be interpreted as the incoming and the outgoing field. But this is well overtaken by the numerical noise amplified by the integration. **b.** The same model using the apodising aperture function. The incoming and the outgoing fields are clearly seen.

implementation of the version of the far field Stratton-Chu equation, modified for media other than free space. This way, the complete control of the field values and equations can be assumed, and the boundary shadowing problem could be circumvented.

B.7 The apodising aperture function in the implementation of the modified Stratton-Chu Equations

Figure B.6a shows the effect of the aperture function on the results. As can be seen, the whole spectrum is garbled with numerical noise. The results clear up as soon as the aperture function is used.

The total field was calculated and reflectivity values were obtained by sampling the whole spectrum at corresponding farfield angles. This was a matrix was build that had $360°$ of farfield directions for each input angle. Figure B.6 depicts the far field spectrum of a metal slab which was 130nm thick. Also on the same figure, the locations the spectrum was sampled at can be seen. The incoming and the reflected light is easy to recognise(the red lines that follow the peaks) The third line, which goes $180°$ out of phase with the incoming light is the direction of the retro direction.

Bibliography

[1] H. Diltbacher, J. R. Krenn, G. Schider, A. Leitner, and F. R. Aussenegg, "Two-dimensional optics with surface plasmon polaritons," *Applied Physics Letters*, vol. 81, no. 10, 2002.

[2] K. Tanaka and M. Tanaka, "Simulations of nanometric optical circuits based on surface plasmon polariton gap waveguide," *Applied Physics Letters*, vol. 82, no. 8, 2003.

[3] J. R. Krenn, B. Lamprecht, H. Ditlbacher, G. Schider, M. Salerno, A. Leitner, and F. R. Aussenegg, "Non-diffraction-limited light transport by gold nanowires," *Europhysics Letters*, vol. 60, no. 5, 2002.

[4] E. Ozbay, "Plasmonics: Merging photonics and electronics at nanoscale dimensions," *Science*, vol. 311, no. 189, 2006.

[5] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays," *Nature*, vol. 391, 1998.

[6] U. Schroter and D. Heitmann, "Surface plasmon enhanced transmission through metallic gratings," *Physical Review B*, vol. 58, no. 23, 1998.

[7] E. Cubukcu, N. Yu, E. J. Smythe, L. Diehl, K. B. Crozier, and F. Capasso, "Plasmonic laser antennas and related devices," *Selected topics in Quantum Electronics,IEEE Journal of*, vol. 14, no. 6, 2008.

[8] I. D. Leon and P. Berini, "Theory of surface plasmon-polariton amplification in planar structures incorporating dipolar gain media," *Physical Review B*, vol. 78, 2008.

[9] M. P. Nezhad, K. Tetz, and Y. Fainman, "Gain assisted propagation of surface plasmon polaritons on planar metallic waveguides," *Optics Express*, vol. 12, no. 17, 2004.

[10] J. Tominaga, C. Mihalcea, D. Buchel, H. Fukuda, T. Nakano, N. Atoda, H. Fuji, and T. Kikukawa, "Local plasmon photonic transistor," *Applied Physics Letters*, vol. 78, no. 17, 2001.

[11] L. Rayleigh, "On waves propagating along the plane surface of an elastic solid," *Proceedings of the London Mathematical Society*, vol. 17, 1885.

[12] J. David N. Cheeke, *Fundementals and Appliations of Ultrasonic waves*. CRC Press, 1st ed., 2002.

[13] A. Dhar, "Paul langevin," *Resonance*, vol. 10, 2005.

[14] C. B. Scruby and L. E. Drain, *Laser Ultrasonics: Techniques and Applications*. Adam Hilger, 1990.

[15] S. E. Jacobs and M. J. Thornley, "The lethal action of ultrasonic waves on bacteria suspended in milk and other liquids," *Journal of Applied Microbiology*, vol. 17, 1953.

[16] T. J. Mason, "Sonochemistry: current uses and future prospects in the chemical and processing industries," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 357, pp. 355–369, 1999.

[17] C. J. Hellier, *Handbook of nondestructive evaluation*. McGraw-Hill, 2001.

[18] J.-P. Monchalin, "Optical detection of ultrasound," *IEEE Transactions on ultrasonics, ferroelectrics and frequency control*, vol. UFFC-33, no. 5, 1986.

[19] M. Clark, S. Sharples, and M. Somekh, "Non-contact acoustic microscopy," *Measurement, Science and Technology*, vol. 11, 2000.

[20] M. G. Somekh, M. Liu, H. P. Ho, and C. W. See, "An accurate non-contacting laser based system for surface wave velocity measurement," *Measurement Science and Technology*, vol. 6, 1995.

[21] H. Talaat and E. Burstein , "Direct electromagnetic generation and detection of surface elastic waves on conducting solids. ii. experiment," in *1973 Ultrasonics Symposium*, pp. 569 – 571, 1973.

[22] P. Klimo, "Electromagnetic generation of ultrasonic surface waves," *Journal of Physics E: Scientific Instruments*, vol. 8, 1975.

[23] D. S. Ballentine, R. M. White, S. I. Martin, A. J. Ricco, E. T. Zellers, G. C. Frye, and H. Wohltjen, *Acoustic Wave Sensors: Theory Design and Physico-Chemical Applications.* Academic Press, 1st ed., 1997.

[24] R. M. White, "Surface elastic waves," *Proceedings of the IEEE*, vol. 58, no. 8, 1970.

[25] D. B. Davidson, *Computational Electromagnetics for RF and Microwave Engineering.* Cambridge University Press, 2005.

[26] J. K. Sykulski, A. Krawczyk, and S. Wiak, "Computational electromagnetics in poland : a review," *COMPEL: International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 17, 1998.

[27] C. W. Trowbridge and J. K. Sykulski, "Some key developments in computational electromagnetics and their attribution," *IEEE Transactions on Magnetics*, vol. 42, no. 4, 2006.

[28] R. Garg, *Analytical and Computational Methods in Electromagnetics.* Artech House, 2008.

[29] P. Barber and C. Yeh, "Scattering of electromagnetic waves by arbitrarily shaped dielectric bodies," *Applied Optics*, vol. 14, 1975.

[30] C. Yeh, "Backscattering cross section of a dielectric elliptical cylinder," *Journal of Optical Society of America*, vol. 55, 1965.

[31] J. A. Stratton, *Electromagnetic Theory.* McGraw-Hill, Newyork, 1941.

[32] F. M. Kahnert, "Numerical methods in electromgnetics scattering theory," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 79-80, 2003.

[33] C. Yeh, "Perturbation approach to the diffraction of electromagnetic waves by arbitrarily shaped dielectric obstacles," *Physical Review*, vol. 135, no. 5A, 1964.

[34] G. Pan, "Orthogonal Wavelets with Applications in Electromagnetics," *IEEE Transactions on Magnetics*, vol. 32, no. 3, 1996.

[35] D. J. Griffiths, *Introduction to Electrodynamics*. Pearson : Benjamin Cummings, 1999.

[36] J. Schwinger, "Euclidean quantum electrodynamics," *Physical Review*, vol. 115, no. 3, 1959.

[37] G. H. Geodecke and S. G. O'Brien, "Scattering by irregular inhomogenous particles via the digitized Green's function algorithm," *Applied Optics*, vol. 27, 1988.

[38] R. Schinzinger, P. A, and A. Laura, *Conformal Mapping: Methods and Applications*. Courier Dover Publications, 2nd ed., 2003.

[39] J. W. Dettnman, *Mathematical Methods in Physics and Engineering*. New York: MCGraw-Hill, 1962.

[40] E. Butkov, *Mathematical Physics*. MA: Addison-Wesley, 1968.

[41] M. Neviere and E. Popov, *Light Propagation in Periodic Media: Differential Theory and Design (Optical Engineering)*. CRC, 1st ed., 2002.

[42] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *Journal of Optical Sociaty of America*, vol. 71, no. 7, 1981.

[43] X. D. Hoa, M. Tabrizian, and A. G. Kirk, "Rigorous coupled-wave analysis of surface plasmon enhancement from patterned immobilization on nanogratings," *Journal of sensors*, vol. 2009, 2009.

[44] D. Zhang, P. Wang, X. Jiao, G. Yuan, J. Zhang, C. Chen, H. Ming, and R. Rao, "Investigation of the sensitivity of h-shaped nano-grating surface plasmon resonance biosensors using rigorous coupled wave analysis," *Applied Physics A*, vol. 89, 2007.

[45] P. P. Banerjee and J. Jarem, *Computational Methods for electromagnetic and optical systems*. CRC, 2nd ed., 2000.

[46] M. G. Moharam and T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings," *Journal of Optical Society of America*, vol. 72, no. 10, 1982.

[47] M. M. Ney, "Method of moments as applied to electromagnetic problems," *IEEE Transactions on Microwave Theory Techniques*, vol. 33, no. 10, 1985.

[48] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*. IEEE Press, 1998.

[49] B. T. Draine and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," *Journal of Optical Society of America*, vol. 11, no. 4, 1993.

[50] J. L. Volakis, A. Chatterjee, and L. C. Kempel, *Finite Element Method for Electromagnetics*. IEEE Press, 1998.

[51] T. Eibert and V. Hansen, "Calculation of unbounded field problems in free space by a 3D FEM/BEM-hybrid approach," *Journal of Electromagnetic Waves and Applications*, vol. 10, no. 1, 1996.

[52] K. Umashankar and A. Taflove, *Computational Electromagnetics*. Artech House, 1993.

[53] K. Umashankar, S. Nimmagadda, and A. Taflove, "Application of integral equation and method of moments for electrically very large scatterers using spatial decomposition technique," *IEEE Transactions on Antennas and Propagation*, vol. 40, 1992.

[54] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, 2nd ed., 2000.

[55] Yee Kane S., "Numerical Solution of Initial Boundary Value Problems Involving Maxwells Equations in Isotropic Media," *IEEE T. Antenn. Propag.*, vol. 14, 1966.

[56] A. Taflove, "Application of finite-difference time-domain method to sinusoidal steady-state electromagnetic-penetration porblems," *IEEE Transactions on electromagnetic compatibility*, vol. EMC-22, no. 3, 1980.

[57] "2comu website." `http://www.2comu.com.cn/products/OpenGEMS.html`, 2007.

[58] "The Computational Physicist." `http://www.thecomputationalphysicist.com/`, 2009.

[59] U. Schroter, S. Seider, S. Tode, and D. Heitmann, "Surface plasmon reflection at edges and resonance effects in metal bars," *Ultramicroscopy*, vol. 68, 1997.

[60] K. Hasegawa, J. U. Nockel, and M. Deutsch, "Curvature-induced radiation of surface plasmon polaritons propagating around bends," *Applied Physics Letters*, vol. 84, 2004.

[61] B. Wang and G. P. Wang, "Surface plasmon polariton propagation in nanoscale metal gap waveguidess," *Optics Letters*, vol. 27, no. 17, 2004.

[62] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *Journal of Optical of Society of America*, vol. 71, no. 7, 1981.

[63] E. Popov, M. Neviere, B. Gralak, and G. Tayeb, "Staircase approximation validity for arbitrary-shaped gratings," *Journal of Optical Society of America :A*, vol. 19, no. 1, 2002.

[64] Y. Liu and Sarris C. D., "AMR-FDTD: a dynamically adaptive mesh refinement scheme for the finite-diffrerence time-domin technique," *2005 IEEE Antennas and Propagation Society International Symposium*, vol. 1A, 2005.

[65] M. M. Afande, K. Wu, M. Giroux, and R. G. Bosisio, "A finite-difference frequency-domain method that introduces condensed nodes and image principle," *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 4, 1995.

[66] A. Chabory, B. de Hon, W. Schilders, and A. Tijhuis, "Fast transform based preconditioners for 2d finite-difference frequency-domain - waveguides and periodic structures," *Journal of Computational Physics*, vol. 227, no. 16, pp. 7755 – 7767, 2008.

[67] S. M. Rao, *Time Domain Electromagnetics*. Academic Press, 1999.

[68] P. Sewel, T. M. Benson, C. Christopoulos, D. W. P. Thomas, A. Vukovic, and J. G. Wykes, "Transmission-Linbe Modeling (TLM) Based Upon Unstructured Tetrahedral Meshes," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 6, 2005.

[69] C. Christopoulos, *The transmission-line modeling (TLM) method in electromagnetics.* Piscataway, NJ:IEEE Press, 1995.

[70] M. N. O. Sadiku and C. N. Obiozor, "A Comparison of Time-Domain Finite difference (FDTD) and Transmission-Line Modeling (TLM) Methods," *IEEE Southeastcon*, 2000.

[71] Z. Chen and M. M. Ney and W. J. R. Hoefer, "A new finite-difference time-domain formulation and its equivalence with the TLM symmetrical condensed node," *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, no. 12, 1991.

[72] C. Christopoulos and J. L. Herring, "The application of transmission-line modelling (tlm) to electromagnetic compatibility problems," *IEEE Transactions on Electromagnetic Comaptibility*, vol. 35, no. 2, 1993.

[73] M. G. Madhan and R. Neelakandan, "An improved transmission line laser model for multimode laser diodes incorporating thermal effects," *Optical and Quantum Electronics*, vol. 40, no. 8, 2008.

[74] T. Dogaru and L. Carin, "Scattering Analysis by the Multiresolution Time-Domain Method Using Compactly Supported Wavelet Systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, 2002.

[75] T. Dogaru and L. Carin, "Multiresolution Time-Domain Using CDF Biorthogonal Wavelets," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 5, 2001.

[76] E. M. Tentzeris, A. Cangellaris, L. P. B. Katehi, and J. Harvey, "Multiresolution Time-Domain (MRTD) Adaptive Schemes Using Arbitrary Resolutions of Wavelets," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 2, 2002.

[77] Y. Chen, Q. Cao, and R. Mittra, *Multiresolution Time Domain scheme for Electromagnetic Engineering.* Wiley-Interscience, 2005.

[78] T. Dogaru and L. Carin, "Application of Haar Wavelet-Based Multiresolution Time-Domain Schemes to Electromagnetic Scattering Problems," *IEEE Transactions on Antennas and Microwave Propagation*, vol. 50, no. 6, 2002.

[79] M. Krumpholz and L. P. B. Katehi, "MRTD: New Time-Domain Schemes Based on Multiresolution Analysis," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 4, 1996.

[80] R. Robertson, E. Tentzeris, M. Krumpholz, and L. P. B. Katehi, "Modelling of Dielectric Cavity structures using multiresolution time-domain analysis," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 11, 1998.

[81] R. Letizia and S. S. A Obayya, "Efficient multiresolution time-domain analysis of arbitrarily shaped photonic devices," *Institution of Engineering and Technology: Optoelectronics*, vol. 2, 2008.

[82] Y. W. Cheong, Y. M. Lee, K. H. Ra, J. G. Kang, and C. C. Shin, "Wavelet-Galerkin Scheme of Time-Dependent Inhomogeneous Electromagnetic Problems," *IEEE Microwave and Guided Wave Letters*, vol. 9, no. 8, 1999.

[83] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods.* Springer, 2nd ed., 2000.

[84] R. Courant, "Variational methods for the solution of problems of equilibrium and vibrations," *Bull. Amer. Math. Soc*, vol. 49, no. 1, 1943.

[85] K. K. Gupta and J. L. Meek, "A brief history of the beginning of the finite element method," *International Journal for Numerical Methods in Engineering*, vol. 39, 1996.

[86] K. H. Huebner, D. L. Dewhirst, D. E. Smith, and T. G. Byrom, *The Finite Element Method for Engineers.* Wiley-IEEE, 4th ed., 2001.

[87] E. Hecht, *Optics.* Addison-Wesley Longman, 3rd ed., 1998.

[88] S. J. Orfanidis, *Electromagnetic Waves and Antennas.* `http://www.ece.rutgers.edu/`, 2004.

[89] S. G. Lipson, H. Lipson, and D. S. Tannhauser, *Optical Physics.* Cambridge University Press, 3rd ed., 1995.

[90] H. Raether, "Surface plasmons on smooth and rough surfaces and on gratings," *Springer Tracts in Modern Physics*, vol. 111, 1986.

[91] H. Raether, "Excitation of plasmons and interband transitions by electrons," *Springer Tracts in Modern Physics*, vol. 88, 1980.

[92] D. Hornauer, H. Kapitza, and H. Raether, "The dispersion relation of surface plasmons on rough surfaces," *Journal of Physics D: Applied Physics*, vol. 7, 1974.

[93] B. Raj, V. Rajendran, and P. Palanichamy, *Science and Technology of Ultrasonics.* Alpha Science International, 1st ed., 2004.

[94] C. Minton, "Inspection of metals with ultrasonic surface waves," *Nondestructive Testing*, vol. 12, 1954.

[95] E. G. Cook and H. E. V. Valkenburg, "Surface waves at ultrasonic frequencies," *ASTM Bulletin*, 1954.

[96] G. Alers, M. Tennison, R. Thompson, and B. Tittmann, "Visualization of surface elastic-waves on structural-materials," *Ultrasonics*, vol. 11, no. 4, 1973.

[97] Q. Shan and R. J. Dewhurst, "Surface-breaking fatigue crack detection using laser ultrasound," *Applied Physics Letters*, vol. 62, no. 12, 1993.

[98] D. A. Cook and Y. H. Berthelot, "Detection of small surface-breaking fatigue cracks in steel using scattering of rayleigh waves," *NDTE International*, vol. 34, 2001.

[99] I. Pape, C. W. Lawrence, P. D. W. S. G. Roberts, G. A. D. Briggs, O. V. Kolosov, A. W. Hey, C. F. Paines, and B. K. Tanner, "Evaluation of Polishing Damage in Alumina," *Philosopphical Magazine A*, vol. 80, no. 8, 2000.

[100] I. J. Collison, T. Stratoudaki, M. Clark, and M. G. Somekh, "Measurement of elastic nonlinearity using remote laser ultrasonics and CHeap optical transducers and dual frequency surface acoustic waves," *Ultrsonics*, vol. 48, 2008.

[101] M. Clark, S. D. Sharples, and M. Somekh, "Fast, all-optical rayleigh wave microscope: Imaging on isotropic and anisotropic materials," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency control*, vol. 47, no. 1, 2000.

[102] I. A. Viktorov, *Rayleigh and Lamb Waves: Physical Theory and Applications.* Plenum Press, New York, 1967.

[103] R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized light.* Elsevier, 1986.

[104] J. Jin, *The Finite Element Method in Electromagnetics.* John Wiley and sons, 1st ed., 1993.

[105] C. Multiphysics, *COMSOL 3.4 Multiphysics Users Guide.* Comsol, 2006.

[106] T. Fevens and H. Jiang, "Absorbing boundary conditions for the schrodinger equation," *Journal of Scientific Computations*, vol. 21, no. 1, 1995.

[107] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 114, pp. 185–200, 1994.

[108] X. Yina, L. Hesselink, Z. W. Liu, N. Fang, and X. Zhang, "Large positive and negative optical beam displacements due to surface plasmon resonance," *Applied Physics Letters*, vol. 85, no. 3, p. 372, 2004.

[109] U. Schroeter and D. Heitmann, "Grating couplers for surface plasmons excited on thin metal films in the kretschmann-raether configuration," *Physical Review B*, vol. 60, no. 7, 1999.

[110] W. L. Barnes, A. Dereux, and T. W. Ebbesen, "Surface plasmon subwavelength optics," *Nature*, vol. 424, pp. 824–830, 2003.

[111] W. Barnes, S. Kitson, T. Preist, and J. Sambles, "Photonic surfaces for surface-plasmon polaritons," *Journal of Optical Society of America A*, vol. 14, no. 7, 1997.

[112] S. Kiton, W. Barnes, and J.R.Sambles, "Photonic band gaps in metallic microcavities," *Journal of Applied Physics*, vol. 84, no. 5, 1998.

[113] P. Torok and F.-J. Kao, *Optical Imaging and Microscopy: Techniques and Advanced Systems.* Springer Series in Optical Sciences, 2nd ed., 2007.

[114] J. R. Krenn and J.-C. Weeber, "Surface plasmon polaritons in metal stripes and wires," *Philosophical Transactions of the Royal Society*, vol. 362, no. 1817, 2004.

[115] R. Yang, M. A. G. Abushagur, and Z. Lu, "Efficiently squeezing near infrared light into a 21nm-by-24nm nanospot," *Optics Express*, vol. 16, no. 24, 2008.

[116] H. T. Miyazaki and Y. Kurokawa, "Squeezing visible light waves into a 3-nm-thick and 55-nm-long plasmon cavity," *Physical Review Letters*, vol. 96, 2006.

[117] G. Veronis and S. Fan, "Bends and splitters in metal-dielectric-metal subwavelength plasmonic waveguides," *Applied physics letters*, vol. 87, 2005.

[118] K. Tanaka and M. Tanaka, "Simulations of nanometric optical circuits: Open-type surface plasmon polariton gap waveguides," *Japanese Journal of Applied Physics*, vol. 42, 2003.

[119] S. A. Maier and H. A. Atwater, "Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures," *Journal of Applied Physics*, vol. 98, 2005.

[120] J. Weeber, J. Krenn, A. Dereux, B. Lamprecht, Y. Lacroute, and J. P. Goudonnet, "Near-field observation of surface plasmon polariton propagation on thin metal stripes," *Physical Review B*, vol. 64, 2001.

[121] H. Ditlbacher, J. R. Krenn, G. Schider, A. Lietner, and F. R. Aussenegg, "Two-dimensional optics with surface plasmon polaritons," *Applied Physics Letters*, vol. 81, no. 10, 2002.

[122] Z. Liu, J. M. Steele, W. Srituravanich, Y. Pikus, C. Sung, and X. Zhang, "Focusing surface plasmons with a plasmonic lens," *Nano Letters*, vol. 5, no. 9, 2005.

[123] H. Ditlbacher, A. Hohenau, D. Wagner, U. Kreibig, M. Rogers, F. Hofer, F. R. Aussenegg, and J. R. Krenn, "Silver nanowires as surface plasmon resonators," *Physical Review Letters*, vol. 95, 2005.

[124] A. Drezet, A. Hohenau, D. Koller, A. Stepanov, H. Ditlbacher, B. Steinberger, F. Aussenegg, A. Leitner, and J. Krenn, "Leakage radiation microscopy of surface plasmon polaritons," *Materials Science and Engineering B*, vol. 149, 2008.

[125] J. Homola, S. S. Yee, and G. Gauglitz, "Surface plasmon resonance sensors :review," *Sensors and Actuators*, vol. B54, 1999.

[126] H. A. Jamid and S. J. Al-Bader, "Reflection and transmission of surface plasmon mode at a step discontinuity," *IEEE Photonics Technology Letters*, vol. 9, no. 2, 1997.

[127] T. P. Shen, R. F. Wallis, A. A. Maradudin, and G. I. Stegeman, "Interference phenomena in the refraction of a surface polariton by vertical dielectric barriers," *Optical Society of America*, vol. 23, no. 4, 1984.

[128] J. R. Krenn, R. Wolf, A. Lietner, and F. Aussenegg, "Near-field optical imaging the surface plasmon fields of lithographically designed nanostructures," *Optics Communications*, vol. 137, 1997.

[129] P. L. Rochon and L. Levesque, "Standing wave surface plasmon mediated forward and backward scattering," *Optics Express*, vol. 14, no. 26, 2006.

[130] G. D. Valle, T. Sondergaard, and S. I. Bozhevolnyi, "Plasmon-polariton nano-strip resonators: from visible to infra-red," *Optics Express*, vol. 16, no. 10, 2008.

[131] M. Allione, V. V. Temnov, Y. Fedutik, U. Woggon, and M. V. Artemyev, "Surface plasmon mediated interference phenomena in low-q silver nanowire cavities," *Nano Letters*, vol. 8, no. 1, 2008.

[132] M. K. abd E J R Vesseur and A. Polman, "Fabry-Perot resonators for surface plasmon polaritons probed by cathodoluminescence," *Applied Phyiscs letters*, vol. 94, 2009.

[133] D. Sarid, "Long-range surface-plasma waves on very thin metal films," *Physical Review Letters*, vol. 47, no. 26, 1981.

[134] D. J. Bergman and M. I. Stockman, "Surface plasmon amplification by stimulated emission of radiation: Quantum generation of coherent surface plasmons in nanosystems," *Physical Review Letters*, vol. 90, no. 2, 2003.

[135] J. Seidel, S. Grafstrom, and L. Eng, "Stimulated emission of surface plasmons at the interface between a silver film and an optically pumped dye solution," *Physical Review Letters*, vol. 94, 2005.

[136] M. Ambati, S. H. Nam, E. Ullin-Avila, D. A. Genov, G. Bartal, and X. Zhang, "Observation of stimulated emission of surface plasmon polaritons," *Nano Letters*, vol. 8, no. 11, 2008.

[137] J. Homola, "Present and future of surface plasmon resonance biosensors," *Analytical and Bioanalytical Chemistry*, vol. 377, no. 3, pp. 528–539, 2003.

[138] J. Dostálek and W. Knoll, "Biosensors based on surface plasmon-enhanced fluorescence spectroscopy (review)," *Biointerphases*, vol. 3, no. 3, 2008.

[139] C. R. Yonzon, E. Jeoung, S. Zou, G. C. Schatz, M. Mrksich, and R. P. V. Duyne, "A comparative analysis of localized and propagating surface plasmon resonance sensors: The binding of concanavalin a to a monosaccharide functionalized self-assembled monolayer," *Journal of American Chemical Society Articles*, vol. 126, no. 39, 2004.

[140] S. J. Chen, F. C. Chien, G. Y. Lin, and K. C. Lee, "Enhancement of the resolution of surface plasmon resonance biosensors by control of the size and distribution of nanoparticles," *Optics Letters*, vol. 29, no. 12, pp. 1390–1392, 2004.

[141] J. Homola, S. S. Yee, and G. Gauglitz, "Surface plasmon resonance sensors:review," *Sensors and Actuators B:Chemical*, vol. 54, no. 1-2, pp. 3–15, 1999.

[142] General Electric Company, "Biacore." `http://www.biacore.com/lifesciences/index.html`, 2009.

[143] I. Nomadics, "Sensiq website." `http://www.discoversensiq.com/`, 2009.

[144] Reichert, Inc. , "Reichert Surface Plasmon Resonance (SPR) Systems." `http://reichertspr.com/`, 2009.

[145] Z. Zhang, P. Zhao, F. Sun, G. Xiao, and Y. Wu, "Self-referencing in optical-fiber surface plasmon sensors," *IEEE Photonics Technology Letters*, vol. 19, no. 24, 2007.

[146] G. G. Nenninger, J. B. Clendenning, C. E. Furlong, and S. S. Yee, "Reference-compensated biosensing using a dual-channel surface plasmon resonance sensor system based on a planar lightpipe configuration," *Sensors and Actuators B*, vol. 51, 1998.

[147] J. Homola, H. B. Lu, and S. S. Yee, "Dual-channel surface plasmon resonance sensor with spectral discrimination of sensing channels using dielectric overlayer," *Electronics Letters*, vol. 35, no. 13, 1999.

[148] D. J. Gentleman, L. A. Obando, J.-F. Masson, J. R. Hollaway, and K. S. Booksh, "Calibration of fiber optic based surface plasmon resonance sensors in aqueous systems," *Analytica Chimica Acta*, vol. 515, 2004.

[149] K. S. Johnston and S. S. Yee, "Calibration of surface plasmon resonance refractometers using locally weighted parametric regression," *Analytical Chemistry*, vol. 69, 1997.

[150] D. A. Hutchins, F. Hauser, and T. Goetz, "Surface waves using laser generation and electromagnetic acoustic transducer detection," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. UFFC-33, no. 5, 1986.

[151] A. A. Kolomenskii, A. M. Lomonosov, R. Kuschnereit, and P. Hess, "Laser generation and detection of strongly nonlinear elastic surface pulses," *Physical Review Letters*, vol. 79, no. 7, 1997.

[152] S. R. Nagel and M. A. Saifi, "Generation of surface acoustic waves by means of a c.w laser," *Electronics letters*, vol. 16, no. 12, 1980.

[153] H. Nishino, Y. Tsukahara, Y. Nagata, T. Koda, and K. Yamanaka, "Excitation of high frequency surface acoustic waves by phase velocity scanning of a laser interference fringe," *Applied physics letters*, vol. 62, no. 17, 1993.

[154] S. Kenderian, B. B. Djordjevic, and R. E. G. Jr, "Narrow band laser-generated surface acoustic waves using a formed source in the ablative regime," *Journal of Acoustic Society of America*, vol. 113, no. 1, 2003.

[155] J. Huang, S. Krishnaswamy, and J. D. Achenbach, "Laser generation of narrow-band surface waves," *Journal of Acoustical Society of America*, vol. 92, no. 5, 1992.

[156] M. Clark, F. Linnane, S. D. Sharples, and M. G. Somekh, "Frequency control in laser ultrasound with computer generated holography," *Applied physics letters*, vol. 72, no. 16, 1998.

[157] T. Stratoudaki, J. A. Hernandez, M. Clark, and M. G. Somekh, "Cheap optical transducers(chots) for narrowband ultrasonic applications," *Measurement Science and Technology*, vol. 18, 2007.

[158] M. E. Siemens, Q. Li, M. M. Murnane, H. C. Kapteyn, R. Yang, E. H. Anderson, and K. A. Nelson, "High-frequency surface acoustic wave propagation in nanostructures characterized by coherent extreme ultraviolent beams," *Applied Physics Letters*, vol. 94, 2009.

[159] D. H. Hurley and K. L. Telschow, "Picosecond surface acoustic waves using a sub-optical wavelength absorption grating," *Physical Review B*, vol. 66, 2002.

[160] B. Bonello, A. Ajinou, V. Richard, P. Djemia, and S. M. Cherif, "Surface acoustic waves in the GHz range generated by periodically patterned metallic stripes illuminated by an ultrashort laser pulse," *Journal of Acoustic Society of America*, vol. 110, 2001.

[161] C. Glorieux, J. D. Beers, E. H. Bentefour, K. V. de Rostyne, and K. A. Nelson, "Phase mask based interferometer: Operation principle, performance, and application to thermoelastic phenomena," *Review of Scientific Instruments*, vol. 75, 2004.

[162] S. Krishnaswamy, *Ultrasonic Nondestructive Evaluation: Engineering and Biological Material Characterization.* CRC press, 2004.

[163] D. H. Hurley, "Optical generation and spatially distinct interferometric detection of ultrhigh frequency surface acoustic waves," *Applied Physics Letters*, vol. 88, 2006.

[164] Y. Takagaki, P. V. Santos, E. Wiebicke, O. Brandt, H.-P. Schonherr, and K. H. Ploog, "Superhigh-frequency surface-acoustic-wave transducers using AIN layers grown on SiC substrates," *Applied Physics Letters*, vol. 81, no. 14, 2002.

[165] I. V. Kukushkin, J. H. Smet, L. Hoppel, U. Waizmann, M. Riek, W. Wegscheider, and K. von Klitzing, "Ultrahigh-frequency surface acoustic waves for finite wave-vector spectroscopy of two-dimensional electrons," *Applied Physics Letters*, vol. 85, no. 19, 2004.

[166] C. Thomsen, H. T. Grahn, H. J. Maris, and J. Tauc, "Surface generation and detection of phonons by picosecond light pulses," *Physical Review B*, vol. 34, no. 6, 1986.

[167] J. Guan, Z. Shen, J. Lu, X. Ni, J. Wang, and B. Xu, "Finite element analysis of the scanning laser line source technique," *Japanese Journal of Applied Physics*, vol. 45, no. 6A, 2006.

[168] A. Pantano and D. Cerniglia, "Simulation of laser generated ultrasound with application to defect detection," *Applied Physics A*, vol. 91, 2008.

[169] B. Xu, Z. Shen, X. Ni, and J. Liu, "Numerical simulation of laser-generated ultrasound by the finite element method," *Journal of applied physics*, vol. 95, no. 4, 2004.

[170] F. Moser, L. J. Jacobs, and J. Qu, "Modeling elastic wave propagation in waveguides with the finite element method," *NDTE Intrenational*, vol. 32, 1999.

[171] M. Yan and M. Qui, "Analysis of surface plasmon polariton using anisotropic finite elements," *IEEE Photonics Technology Letters*, vol. 19, no. 22, 2007.

[172] "Em explorer website." `http://www.emexplorer.net/`, 2009.

[173] X. Xu, Z. Tang, J. Shao, and Z. Fan, "The study on the interface adhesion comparison of the MgF_2, Al_2O_3, SiO_2 and Ag thin films," *Applied Surface Science*, vol. 245, 2005.

[174] D. Jacob, F. Peiro, E. Quesnel, and D. Ristau, "Microstructure and composition of mgf_2 optical coatings grown on si subtrate by pvd and ibs processes," *Thin Solid Films*, vol. 360, 2000.

[175] J. A. Sanchez-Gil, "Surface defect scattering of surface plasmon polaritons: Mirrors and light emitters," *Applied Physics Letters*, vol. 73, no. 24, 1998.

[176] X. Luo and T. Ishihara, "Surface plasmon resonant interference nanolithography technique," *Applied Phyiscs Letters*, vol. 84, no. 23, 2004.

[177] B. Huang, F. Yu, and R. N. Zare, "Surface plasmon resonance imaging using a high numerical aperture microscope objective," *Analytical Chemistry*, vol. 79, 2007.

[178] J. Krenn, R. Wolf, A. Leitner, and F. R. Aussenegg, "Near-field optical imaging the surface plasmon fields of lithographically designed nanostructures," *Optics Communications*, vol. 137, 1997.

[179] C. Mack, *Fundemental Principles of Optical Lithography*. John Wiley and Sons, 2007.

[180] G. C. Fiaccabrino and M. Koudelka-Hep, "Thin-film microfabrication of electrochemical transducers," *Electroanalysis*, vol. 10, no. 4, 1998.

[181] C. F. H. Gondran and D. K. Michelson, "Effect of probe tip size on atomic force microscopy roughness values for very smooth samples," *Journal of Vacuum Science and Technology A*, vol. 24, 2006.

[182] J. Achenbach, "Quantitative nondestructive evaluation," *International Journal of Solids and Strucutres*, vol. 37, 2000.

[183] C. Valle and J. W. L. Jr., "Flaw localization using the reassigned spectrogram on laser-generated and detected lamb modes," *Ultrasonics*, vol. 39, 2002.

BIBLIOGRAPHY

BIBLIOGRAPHY

[184] A. Ruiz and P. B. Nagy, "Laser-ultrasonic surface wave dispersion measurements on surface-treated metals," *Ultrasonics*, vol. 42, 2004.

[185] P. A. Mante, J. F. Robillard, and A. Devos, "Complete thin film mechanical characterization using picosecond ultasonics and nanostructured transducers: experimental demosntration on SiO_2," *Applied Physics Letters*, vol. 93, 2008.

[186] C. Multiphysics, *Comsol Multiphysics 3.3 RF User Guide*, 2007.

[187] J. Wang, Z. Shen, B. Xu, and X. Ni, "Numerical simulation of laser-generated ultrasound in nonmetallic materials," *Optics and Laser Technology*, vol. 39, 2007.

[188] Eulitha, "Eulitha website." `http://www.eulitha.com/`, 2008.

[189] T. Campbell, R. K. Kalia, A. Nakano, P. Vashishta, S. Ogata, and S. Rodgers, "Dynamics of oxidation of aluminium nanoclusters using variable charge molecular-dynamics simulations on parallel computers," *Physical Review letters*, vol. 82, no. 24, 1999.

[190] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, 1965.

[191] Y. Xiao and Y. Lu, "Combination of PML and ABC for Scattering Problem," *IEEE Transactions on Magnetics*, vol. 37, no. 5, 2001.

[192] F. L. Teizeira and W. C. Chew, "General closed-form pml constitutive tensors to match arbitrary bianisotropic and dispersive linear media," *IEEE Microwave and Guided Wave Letters*, vol. 8, no. 6, 1998.

[193] J.-P. Berenger, "Three-dimensional perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 127, pp. 363–379, 1996.

[194] Comsol Website, "Comsol knowledge base." `http://www.comsol.com/support/knowledgebase/977.php`, 2006.

[195] E. J. Rothwell and M. J. Cloud, *Electromagnetics*. CRC Press, 2001.

[196] J. A. Stratton and L. J. Chu, "Diffraction theory of electromagnetic waves," *Physical Review*, vol. 56, 1939.

[197] W. Hsu and R. Barakat, "Stratton chu vectorial diffraction of electromagnetic fields by apertures with application to small-fresnel-number systems," *Optical Society of America A*, vol. 11, no. 2, 1994.

[198] A. D. Polyanin and V. F. Zaitsev, *Handbook of Exact Solutions for Ordinary Differential Equations*. Chapman and Hall/CRC, 2nd ed., 2003.

[199] S. Salsa, *Partial Differential Equations in Action from Modelling to Theory*. Springer, 2008.

[200] J. C. Ravey and P. Mazeron, "Light scattering in the physical optics approximation; application to large spheroids," *Journal of Optics*, vol. 13, no. 5, pp. 273–282, 1982.