Novel Electro-Optic Modulator for Silicon Nitride Waveguides

Joshua Simon Male

Doctor of Philosophy

University of York

Physics, Engineering and Technology

June 2023

Abstract

Under the guidance of Moore's Law, the number of transistors on a central processing unit has doubled every two years for over 50 years, but the ever-growing demand for data processing, computational performance, and the general advancement of the capabilities of technology has pushed conventional computing systems, which are limited by the von Neumann bottleneck, to a digital efficiency wall of performance. Optical signal processing and neural network architectures present a possible solution to overcoming the limitations of conventional computing systems, given that computationally heavy Fourier-space computations are trivial in the optical domain. Spatial light modulators play a key role in an all-optical system and are required to perform to a highly-efficient level with the desire to push the dimensions close to that of microelectronic systems.

Electro-optic modulators, a branch of spatial light modulators, operate on the principle that the optical property of a material may be controlled with electrical gating. Indium tin oxide has been widely identified as a suitable material, due to the plasma dispersion effect exhibited in the material. Typically, electro-optic modulators utilise a silicon platform to operate in this wavelength range, but in this work, I propose to use a silicon nitride platform and take advantage of the low-loss operation and high-power throughput of the platform.

In this work, I show that a competitive phase modulator can be realised using a metal-oxide-semiconductor capacitor to modulate the local permittivity in an indium tin oxide layer, the change in permittivity is overlapped with a guided mode in a silicon nitride waveguide platform and Mach-Zehnder Interferometer architecture to produce a phase change. I have demonstrated a half-wave voltage-length between 0.5-1.0Vmm in the visible spectrum, achieved with a dual-mode interaction within the modulator.

The simulations and experiments in this work demonstrate the capabilities of indium tin oxide as a material for electro-optic modulation, and the suitability of silicon nitride as a platform in the wider field of spatial light modulators. I also demonstrate that the electrical thickness of the insulating material in the capacitor may be reduced without affecting the optical behaviour of the modulator, as has been hinted to in the literature.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my sincerest gratitude to the people who have encourage, guided, and supported me over the past four years during my research project. First and foremost, I thank my supervisor, Professor Thomas Krauss for his academic and personal support throughout my project. Over the many years of working together, I have gained experience, knowledge, and skills that I will carry for the rest of my career.

I also thank my examiners Professor Andrea Di Falco and Dr Eugene Avrutin, whose constructive comments, feedback, and corrections have added significant value to this thesis. I extend my gratitude to Professor Kieran Gibson and Dr Laurence Wilson for their role as a thesis advisory panel throughout my research.

I would like to thank Dr Christopher Reardon, for the training, advice, and skills passed down to me and the general project advisory role he has played throughout my project. I also thank Dr Richard John for the general supporting role he plays in the group in all aspects of the lab work undertaken. For his work with ITO and the development of the characterisation technique, I would like to thank Samuel Blair, without whom I would likely have spent months beyond my deadline continuing to characterise ITO. I extend a thank you to Dr Manuel Deckart, Dr George Duffett, Dr Kezheng Li, Dr Donato Conteduca, Dr Alex Drayton, and Rhys Ashton for the teaching, training, or advice extended by them in a variety of fields.

My academic career at the University of York has, amongst other things, been a good-natured and enlightening atmosphere. For this privilege, I thank Professor Steven Johnson, Dr Yue Wang, Dr Casper Kunstmann, Dr Lisa Miller, Dr Matthew Simmons, Dr Giampaolo Pitruzzello, Dr Callum Silver, Dr Augusto Martins, Kalum Thurgood-Parkes, and Francisca D'Rozario. I have had the privilege to work alongside many others in the Photonics and Bio-Inspired Technologies research groups throughout my project and to those group members, I extend my thanks.

I would like to thank my friends and family, far and wide, for their never-ending support throughout my research and their constant encouragement to achieve my goals. I can't thank you all individually but know that I am forever grateful.

I'd like to thank my son, Sebastian, who has done his very best to distract me in the final part of my project and entertain me through the difficult days. Finally, I'd like to thank my wife, Tasha, who has supported me through three degrees now and has tolerated the stress I put on myself to achieve my goals.

Declaration of Authorship

I, Joshua Male, declare that this thesis title Novel Electro-Optic Modulator for Silicon Nitride Waveguides is a presentation of original work, and I am the sole author.

- This work was completed wholly while in candidature for a research degree at this University (the University of York). This work was originally submitted in June 2023 for examination and resubmitted with corrections in October 2023.

- This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as reference.

- I have acknowledged all main sources of help.

- Published work included within this thesis is indicated in the list below.

Joshua Male

Publications

- S. F. J. Blair, J. S. Male, S. A. Cavill, C. P. Reardon, and T. F. Krauss, 'Photonic Characterisation of Indium Tin Oxide as a Function of Deposition Conditions', *Nanomaterials,* vol. 13, np. 13, Art. No. 13, Jan. 2023 [1]

Conferences

- Joshua S. Male*, George Duffett, Christopher P. Reardon, Thomas F. Krauss, "Low-Power Photonic Modulators for Neuromorphic Computing", oral presentation at SIOE, virtually presented during the Covid-19 pandemic, June 2021.

- Joshua S. Male*, Samuel F.J. Blair, Christopher P. Reardon, Thomas F. Krauss, "Highly Efficient Dual-Mode Waveguide Modulators in Epsilon-Near-Zero Materials", oral presentation at SIOE, Cardiff, April 2023.

Dedicated to the memory of Richard Holland and Brian Male

1. Introduction and Scope

Chapter Synopsis

This chapter presents an original review of the literature and introduces the research project, inspired by the end of conventional computing systems and the advancements in optical processors. Advancements in science are rarely unprompted and are instead often the result of a need or desire to find the solution to an existing problem. This means that scientific researchers require a keen eye for observation, a compassionate understanding of the fundamental physical processes, and the creative imagination to design and realise a solution that benefits society. The research contained in this thesis is no different. Society's insatiable demands for high-speed data processing and power will already have grown substantially in the time between when I wrote this paragraph and when it will be read. Largely conventional computing has been guided by Moore's Law, though scaling and efficiency limitations are highlighting the need for new computational architectures and hardware. Neuromorphic computing represents a new paradigm for computational systems, specifically speech recognition, natural language processing, and machine learning systems, where the brain-like architecture is highly suited. For machine learning systems, an optical processor is highly desirable. Such a system requires the design and implementation of spatial light modulators, and it is here where I will focus my research.

1.1.   The End of Conventional Computing

I will begin with the observation that drives scientific research, the motivation. For this research, the motivation is the ever-growing demand for data processing, computational performance, and general advancement of the capabilities of technology. It is estimated that approximately 63% of people use the internet, with 5.32 billion people worldwide accessing it through mobile phones. Google receives 99,000 search requests per second, totalling 8.5 billion requests per day. They have invested heavily in machine learning algorithms, to increase the efficiency of its translation tool to 85%. Artificial intelligence (AI) is a key industry, expected to reach a growth rate of 42% by 2027, with 37% of global organisations employing AI in some capacity. While AI as a business tool is on the rise, the largest market is currently the fully- and semi-automatic car industry, estimated to be worth $26 billion by 2030. The demand for data processing is so high because, currently, as a population we generate $2.5x10^{18}$ bytes of data daily [1]. While the computer may have had humble beginnings, the rate of demand for data has drastically increased over the rate of performance.

1.1.1.   Moore's Law

In 1965 Gordon Moore published a prediction that the number of transistors of a central processing unit (CPU) would double every year, a trend dubbed Moore's law [2]. His prediction was corrected in 1975 to state that the number of transistors, i.e., central processing unit (CPU) scaling, would double every two years, but that the cost of computers would halve at the same time. Moore predicted that a more powerful computing architecture and more machines with integrated electronic circuits would lead to increased manufacturing capabilities, a higher demand for semiconductor materials, and an inevitable decrease in production cost.

Moore was specifically referring to integrated electronic circuits with the desire to miniaturise and increase the capabilities of electronic devices in society. While Moore did predict that the advancements in integrated electronics could lead to home computers, automated controls for vehicles, personal portable communications equipment, and a digital telephone communications network, it is not possible he could have predicted that his ideas would have driven the advancement of semiconductor electronics for over 50 years.

While global events, such as the Covid-19 pandemic and the increasing popularity of cryptocurrencies, have driven the price of semiconductors higher and thus removed the

predicted halving of costs from Moore's law, for the most part the original predictions have stood the test of time. However, all technological development eventually reaches an end point [3].

1.1.2.   The End of Moore's Law

From smart devices that comprise the internet of things (IoT) and harvest small volumes of data, to facial recognition software that analyses enormous volumes of data at high speeds, the growth and variety of data processing applications in everyday life has dramatically increased over the last 20 years. Manufacturers have, for the most part, dealt with increasing demand for performance by adding transistors and miniaturising their size, and making minor architectural enhancements such as increasing cache size and complex dynamic instruction level parallelism (ILP).

Original data up to the year 2010 collected by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten. **Year**   Data for 2010-2019 collected by K. Rupp.
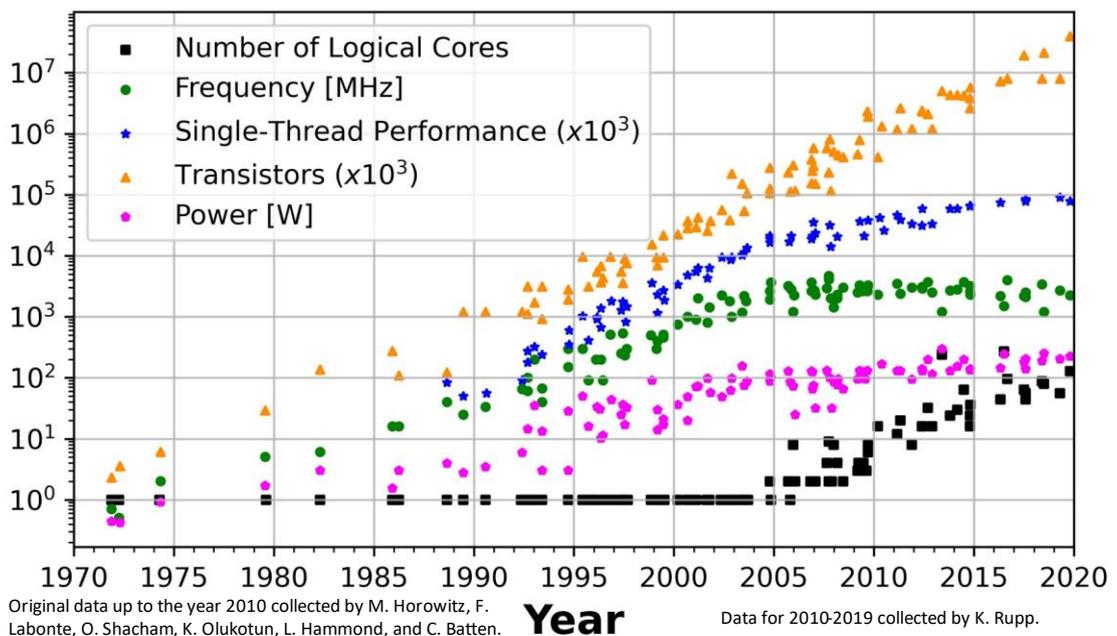
Figure 1.1: 48 years of microprocessor trend data comparing transistors, single-thread performance, clock frequency, thermal design power, and number of logical cores from 1871-2019. Data acquired from K. Rupp [4] and replotted in accordance with CC-BY 4.0.

For the most part, miniaturising the transistor and increasing packing density yielded the performance gains manufacturers desired because of Dennard scaling, whereby the power demand of the transistor scaled with size [5]. As nanometric transistors became more prevalent, CPU speeds stabilised around the 4 GHz mark [6], and Dennard scaling began to breakdown.

CPU clock speed is a measure of the number of cycles executed by a CPU per second, hence the frequency units, and as such a 4 GHz CPU can execute 4 billion cycles per second, where multiple instructions may be handled per cycle, or a string of instructions may take multiple cycles to handle. The performance of the CPU is determined as a combination of its design and speed, which I will discuss in the next section.

This frequency limit, of approximately 4 GHz, was reached because of power considerations; driving transistors faster and shipping the data around the circuit takes disproportionally more power, so the frequency was a compromise between performance and cooling, also known as the power wall [7]. The power wall caused the rate of development in single thread performance to be reduced considerably [4]. Single-thread performance is the amount of work completed by a program that runs as a single stream of instructions and the time taken to complete these instructions is then the product of computational speed and number of instructions. Single-thread performance, indicated by the blue dots in Figure 1.1, highlights the trend toward the asymptotic end of Moore's law for digital electrons in single-thread performance.

The semiconductor industry overcame the diminishing returns in performance scaling by adopting the integration of multiple cores and processors on the same chip, which came into effect around 2005. Each core can exploit ILP and thread-parallelism to increase processing speeds. The breakdown of Dennard scaling limited the number of simultaneously powered cores with a fixed power and heat extraction rate. The knock-on effect is a phenomenon known as "dark silicon", referring to the fact that not all of the circuit can be powered-up to the operating power due to thermal overload [8].

The onset of the dark silicon phenomenon, and the trends observed in Figure 1.1, are indicative of the fact that digital electronic scaling, and by extension Moore's law, is reaching its inevitable end. Fundamentally, the end of the scaling can be traced to three bottlenecks in conventional computing. The first is the power-wall discussed here, where scaling a system that no longer obeys Dennard scaling requires exponentially more expensive and power demanding cooling systems to prevent system damage. The increased demand on the components of the system causes a shortened life cycle for components and an increase in computational faults with operation time [9]. The other bottlenecks are the limit of electronic interconnects and the energy consumption of digital switching, which I will discuss next [10].

1.1.3.   The Von Neumann Bottleneck

The scaling of CPU architectures and innovations in performance are not always matched by faster memory technologies and memory access techniques. There are two major design philosophies in conventional computing architectures when discussing how the CPU accesses a memory unit: the Von Neumann and Harvard architectures.
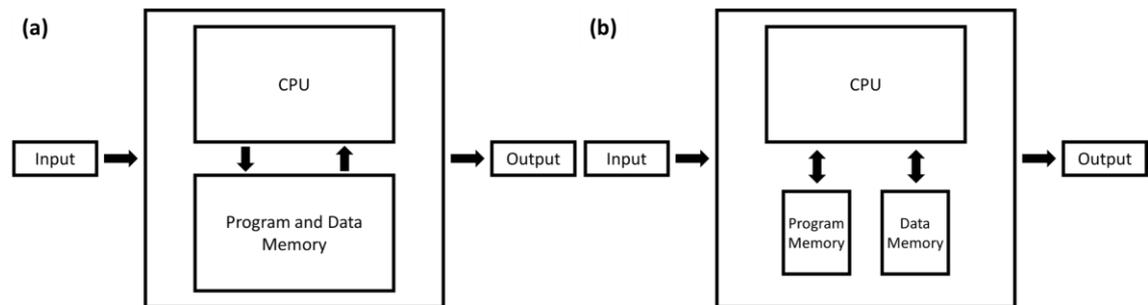
Figure 1.2: (a) Von Neumann and (b) Harvard computing architecture.

The Von Neumann architecture (Figure 1.2(a)) comprises a CPU and a memory unit responsible for all program elements, data, and instructions; whereas the Harvard architecture (Figure 1.2(b)) has the memory unit divided into a program memory unit for instructions and a data memory unit. As the memory unit can only be accessed once per clock cycle, the decision of which architecture to use is dependent on the application. Von Neumann architecture require 2 clock cycles per instruction but has the distinct advantage of having program and data instructions in the same unit [11]. The Von Neumann architecture is typically used for personal computing due to the simple and inexpensive design, where the Harvard architecture is more suited to signal processing and microcontrollers because of the increase in processing speed.

Both the Von Neumann and Harvard architectures suffer the same flaw, which is that the memory access is reliant on CPU-to-memory speed, and inevitably this leads to a performance bottleneck commonly known as the Von Neumann bottleneck [12], or memory-wall. The Von Neumann bottleneck becomes more of an issue with performance and memory capacity scaling as the system is still fundamentally limited by access speed between the two units. The bottleneck between CPU and memory leads not only to the reduced single-thread performance, but also to a power density problem in conventional computing systems [13].

1.1.4.   The Digital Efficiency Wall

Computational efficiency is measured in "multiple and accumulation" (MAC) operation per Joule and has been consistently doubling every 1.5 years with computational speed for around 50 years, an effect referred to as Koomey's law [14]. Just like Moore's law however, the scaling of efficiency has begun to reach an asymptote of around 10 GMAC/J, and has only doubled once in the last decade [10]. With current trends predicting the number of IoT devices to reach 50 billion by 2030, the world to produce 463 Eb of data by 2025 [15], and the power consumption of data centres to reach 33% of the world's energy consumption by 2030 [16], it is now imperative that the power budget of the computational systems we use becomes a top priority.

The limit of the capabilities of conventional computing systems, regarding an increase in performance without the increase of power demand, is allowing a new era of information technology with research shifting to neuromorphic architectures and optical transistors, which are discussed in the next sections. Figure 1.3 shows the speed and efficiency metrics achievable by conventional electronic systems, shown by the digital and microwave electronics regions, and the metrics offered by neuromorphic electronics and photonics platforms. The areas are approximated based on the qualitative trade-offs of each platform by Lima et al. [10].
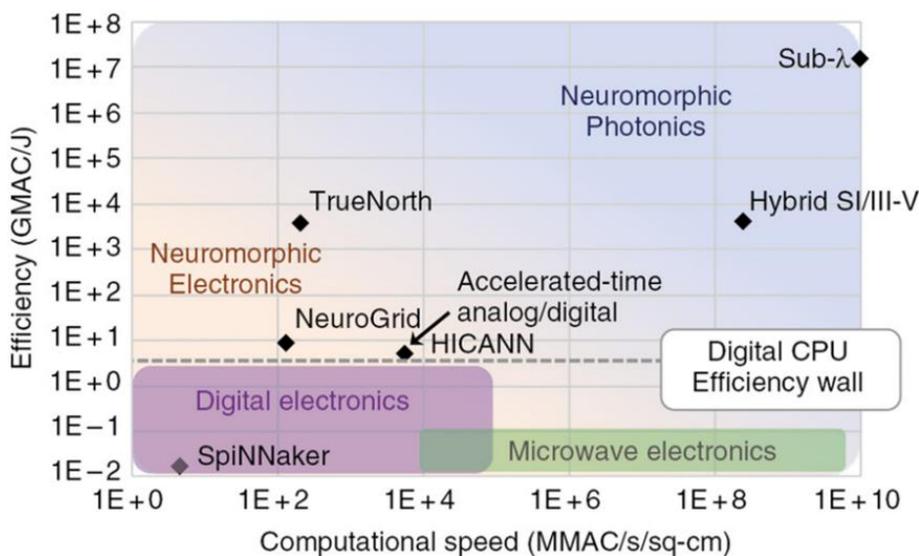
Figure 1.3: Computational efficiency and the digital efficiency wall. Figure taken from neuromorphic photonics [10] under CC-BY. Hybrid Si/III-V is a silicon and group III-V semiconductor simulated neural network PIC, sub-λ refers to subwavelength photonics, and the neuromorphic electronics platforms are discussed in later sections.

The emergence of photonic integrated circuits (PICs), which are commonplace in fast ethernet switches and supercomputers, presents an exciting choice for replacing conventional electrical

interconnects. Photonic systems offer several advantages over electronics in data transmission, such as: high bandwidth, high resolution, and high interconnectivity through very high-frequency operation, natively high spatial resolution, and free space operation, respectively [17]. The key advantage of optical technology is that there is no power penalty in moving data through the system, a major downfall in electronic systems leading to the power wall.

Optical interconnects in high-performance computing systems have been considered since the mid-1980s [18], [19], with a reignition of interest by Benner et al. [20] almost 2 decades ago. Optical cabling replacing electrical wiring in supercomputers was used by IBM in the Roadrunner supercomputer [21], the first petaflop machine, with 5 Gb/s bitrates per channel. VCSELs [22], [23] and silicon photonics [24]–[26] platforms have been reported as exceeding 50 GB/s bitrates, with an silicon photonic modulator system reportedly exceeding a 100 Gb/s rate on chip [27].

Integration of optical interconnects with microelectronic systems has shown an increase in efficiency and computation speed, but for optical systems to truly overhaul electronic systems, the issues leading to the digital efficiency and memory walls also needs to be addressed. High interconnectivity, on either platform, has the capabilities to overcome the memory wall when combined with a new, biology-inspired, architecture; but the high bandwidth capabilities of optical systems, not possible with purely electronic systems, will truly overcome the digital efficiency wall. The architectural change required is known as "neuromorphic computing".

1.2. Neuromorphic Computing

Neuromorphic computing systems represent a new paradigm of computing systems, namely their ability to be taught, rather than directly programmed, in a process known as "machine learning", and can be applied to speech recognition, natural language processing, and machine vision applications [28]. In this section I will outline the foundations of neuromorphic computing, and the further limitations of electronic systems.

1.2.1. Neuromorphic Computing

Neuromorphic computing refers to an architecture that resembles the human brain, which can perform tasks beyond supercomputers for a fraction of the power consumption. There are 10^{11} neurons in the brain operating with a total power of 23 W, or 230 pW per neuron. The brain is

estimated to operate at 10^{20} MAC/s, and is therefore estimated to operate with an approximate efficiency of 4.35×10^{18} MAC/s/W [29]. While neural networks tend to use MAC/s as a figure of merit, it is far more common for conventional computing systems to use floating-point operations per second (FLOP), where 1 FLOP/s is equal to 2 MAC/s [30]. It follows that the human brain operates with an efficiency of 8.7×10^{18} FLOP/s/W (8.7 EFLOP/s/W). Hewlett Packard's Frontier supercomputer is rated as the fastest exa-scale supercomputer, achieving 1.102 EFLOP/s while consuming 21 MW of power. Frontier utilises a combined CPU and graphics processing unit (GPU) architecture to operate with an efficiency of 62.68 GFLOP/s/W [31], meaning the efficiency of the human brain is five orders of magnitude beyond that of leading supercomputers. There is clear evidence that distributed processing is a possible solution to the constraints of conventional computing architectures [10].

While the brain is, therefore, a natural standard for information processing, it differs greatly from computers today and instead operates on timed events known as "spikes". Neurons receive the spikes and integrate over time, i.e., the spikes contain no other information other than the time and source of generation, known as "event driven computing" [32]. A neuromorphic process, in contrast to the point-to-point memory-processor communication employed by the Von Neumann architecture, requires many interconnects. The volume of interconnects requires a significant amount of multicasting and introduces capacitive loads and radiative problems with electrical links [29]. Spiked neural networks (SNNs) typically employ the following techniques to overcome the electronic limitations; an event-driven inter-chip encoding technique known as address event representation (AER) to communicate the location and timing information of sparse neural events, crossbar time-division multiplexing (TDM) to transmit and receive signals over a common path through synchronised switches, and packet switching to transfer small pieces of data across various networks. Despite these techniques, SNNs can only achieve biological timescales in the kHz regime [33], [34].

1.2.2. Neuromorphic Photonics

For neuromorphic processing of high-bandwidth data in the GHz-regime, however, a different approach to interconnection is required, and photonics holds the answer. While photonics has, for some time, dominated information transmission, it has been electronics that has monopolised information transformation. If photonics is to fill the role of computation, a new architecture is required. Optical waveguides, unlike complementary metal-oxide-semiconductor

(CMOS) gates that draw energy when called upon, can passively carry multiplexed signals through a network.

The concept of optical systems replacing electronics is not new, the limiting factors have been the physical constraints of optical systems, the limitation of optical-digital interfaces, and the absence of a robust integration platform [35]. A photonic system built around the use of optical interconnects would be able to support nonlinear optoelectronic devices with high energy efficiency and bandwidth, with small latency, thus sidestepping the current electronic interconnect pitfalls.

There has been a shift in the landscape for photonic chips in recent years, and for good reason, Figure 1.3, references [36], PICs [37], [38], and other optical interconnects research is predicting an increase in processing speed by 6 orders of magnitude compared to neuromorphic electronics [39].

1.2.3.   The Artificial Neuron

The three elements to a neural network are: the neuron (a set of nonlinear nodes), the network (configurable interconnects), and the learning (information representation). An artificial neuron is illustrated in Figure 1.4. Inputs (x) enter the system where they are weighted (w) by the learning algorithm and summed with a bias (b). The weighted addition is then fed to an activation function y(x) to produce an output (y).
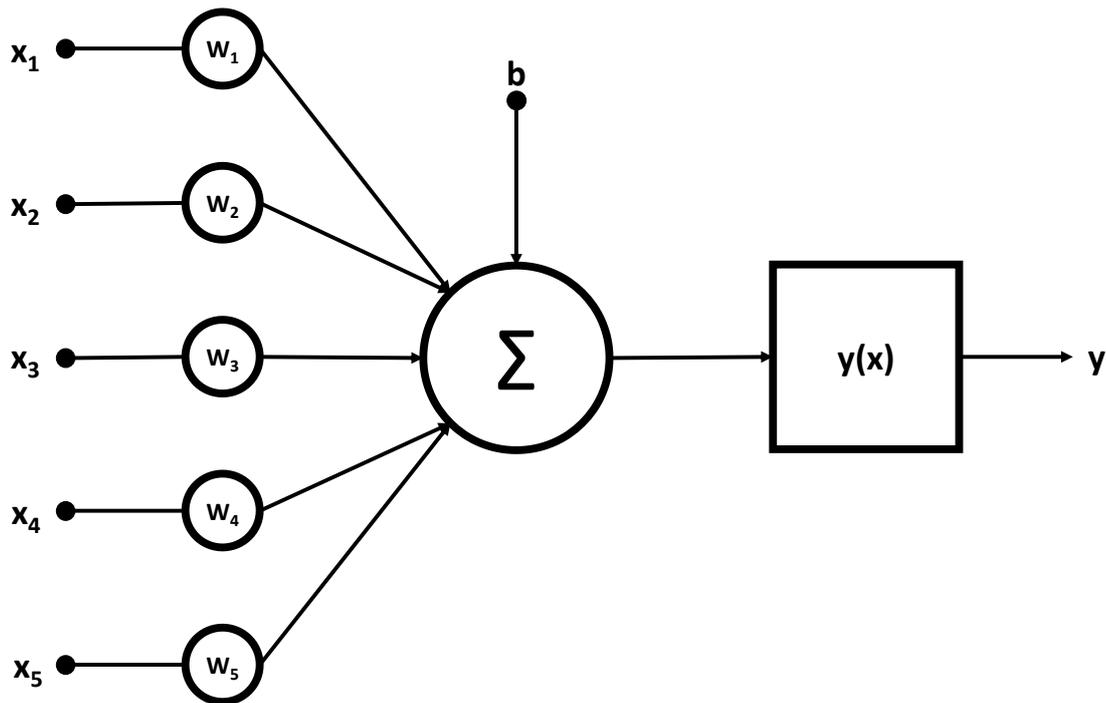
Figure 1.4: Nonlinear model of an artificial neuron.

Realising this architecture with a photonic neuron presents the potential solution to the limits of microelectronic computing, where electrical pulses are to be replace by pulses of light. In this section I will discuss machine learning and the requirements for a photonic neuron considering the design philosophy in Figure 1.4. The interconnecting of neurons, to form a network, is implied in the formation of a neural network.

1.2.4.   Deep Learning

Conventional computing systems could potentially stay stagnant at current performance levels, with research aimed at improving reliability and longevity of systems, but there are many applications for which conventional computing systems are not adequate. The most prominent of which is machine learning. With applications in identifying objects in images, transcribing speech to text, targeting adverts and products at consumers, and much more, machine learning is a major industry on the rise [40].

Many of the applications discussed above use a class of machine learning known as deep learning. Unlike machine learning, which is limited in its ability to process raw data and requires significant pre-treatment, deep learning methods transform raw data into more abstract levels

that can be processed. The key is that deep learning layers are not human designed but are instead learned from a training protocol with engineered feedback on performance.

Deep learning is the ideal tool for image and speech recognition, where it is already outperforming other machine learning techniques [41]–[44]. The pattern recognition capabilities are applicable across several industries including security, translation, and driverless vehicles.

1.2.5.   Convolutional Neural Networks

A deep learning architecture is typically a multilayer stack of simple modules, where each module in the stack transforms its input to increase the selectivity and invariance of the representation, thus forming a network. One type of deep network that has proved significantly easier to train and can be applied to more general purposes Is the convolutional neural network (CNN) [45]. A CNN is a type of deep learning model for processing data with a grid pattern, such as images, and is designed to learn spatial hierarchies of features automatically and adaptively.

A CNN consists of a convolution and pooling layers for feature extraction, and a fully connected layer for output classification. The convolution layer is the most computationally taxing process as the mathematical convolution between the input image and stored training images, known as "kernels", is performed at each image position as an object could be anywhere in an input image [46].

$$c = \mathcal{F}^{-1}\{C\}, \qquad C = A * B = \left\{\mathcal{F}\{a\} \cdot \mathcal{F}\{b\}\right\} \qquad \textit{Equation 1.1}$$

The computationally heavy kernel function for the convolution of image data can be expressed as matrix multiplication of the 2D Fourier transform ($\mathcal{F}$) of both object (a), A, and the kernel function (b) in the Fourier domain, B. The inverse Fourier transform ($\mathcal{F}^{-1}$) of the resultant matrix (C) produces the image convolution, c (Equation 1.1). The fact that convolutional neural networks can be expressed as Fourier transforms opens the door for an optical approach where, in the optical domain, Fourier transforms are easy to perform.

1.2.6.   State-of-the-Art Neuromorphic Computing

While the details of neural network architectures, algorithms, and performance metrics are beyond the scope of this thesis, to summarise this section I will give a brief overview of some of the CMOS-based neuromorphic chips that are currently at the forefront of neuromorphic computing. The performance metrics amongst the literature is often quoted in system specific units, here I have attempted to give fair quotations of the respective metrics, and where appropriate a comparative measurement of energy consumption per synaptic event, i.e., the energy required to process one piece of information.

The first for comparison is TrueNorth [33], [47], [48], a digital CMOS chip created by IBM using 1272 application-specific integrated circuits (ASICs) on printed circuit board (PCB), consisting of one million distributed digital neurons exhibiting spiking behaviour. Each die holds 4096 cores, with 256 digital neurons per core, and 256 synapses per neuron. Each die consumes 72 mW of power, but the true power of TrueNorth comes in stacking 16 die on a chip, which consumes 1 W of power and operates at 1 kHz speed. TrueNorth is reported consuming 26.6 nJ per synaptic event. Neurogrid [49] is a chip created at Stanford University, and uses analogue neural circuits to compute real-time large-scale neural simulations using digital spike communication between neurons. The chip consists of 16 cores containing 65000 neurons, totalling 1 million neurons. The 16-core chip has an energy consumption of 3 W of power, or approximately 150 mW per core. Neurogrid has a reported figure of merit of 941 pJ per event.

Catholic University Louvain demonstrated a digital neuromorphic chip known as ODIN [50], a fully depleted silicon-on-insulate (FDSOI) CMOS chip, with 256 neurons supporting spike-driven synaptic events implementing neuron dynamics in sequential fashion, and consuming 15 nJ per event. SpiNNaker [34], [51]–[53] is a massively parallel multi-core computing system that implements neural and synapse models in real-time built using 18 ARM9 processors, which are a reduced instruction set computer (RISC) CMOS chip on PCB. Currently, the SpiNNaker node contains 18 processor cores with 16,000 digital neurons and 16 million synapses, with the single node consuming 1 W of power. The real power of the SpiNNaker project is that they aim to combine 57,000 nodes on a chip to increase the multi-parallelism of the system, which is quoted as consuming 90 kW of power. This is achieved through clever use of instruction, local, and shared memory, and the transfer of information through the system in packets. The SpiNNaker project is reported having a 11.3 nJ per event power consumption, with 2.2×10^6 instructions per second per watt (IPS/W).

BrainScaleS [54]–[56] is a neuromorphic wafer-scale chip, utilising a mixed signal ASICs platform on PCB, combining analog neural circuits and digital spiking communication, the same architecture as Neurogrid. BrainScaleS is, however, unique in that it operates using a spiking network speedup factor, 10^3-10^4, which means the network operates 10,000 times faster than the equivalent biological system. BrainScaleS consists of 48 modules containing 8 cores, with each core containing 512 neurons, with approximately 14000 synapses per core. BrainScaleS has a reported power consumption of 2 kW, equivalent to 83 W for 16 cores, and operates at 40 Hz spiking frequency. The true power of BrainScaleS is its scale, the number of cores per module and synapses per core dwarfs the competition, allowing for a larger data size and simultaneous processing.

1.3.   Optical Processing

The limitations of microelectronics have led to the emergence of optical replacements, and specifically large-scale all optical systems relying on optical interconnects. Specifically, silicon photonics has emerged as the leading platform due to its combination of low fabrication costs, electronic-photonic integration, and CMOS compatibility. Silicon photonics optical interconnects also offer the key functionality of optical modulation, which can be done directly and is readily integrated with electronic CMOS driving circuits. Before I discuss the inclusion of silicon photonic interconnects, I will first introduce the concept of the coherent optical processor.

1.3.1.   Coherent Optical Processor

Optical systems are typically one of two categories: a focused optical system, or an image processing system. At the basic level, an optical system consists of three components; an input plane, an output plane, and a set of components between the two planes that transform the input image (f) into the output image (g). Th output image is related to the input image through an optical impulse response function (h), such that

$$g(x, y) = h(x, y) * f(x, y)$$

Equation 1.2

defines the behaviour of the optical system. For optical image processing systems, the impulse response function is a close replica of the kernel in the input plane. The convolution of image and kernel, i.e., the Fourier transform of Equation 1.2 into Equation 1.1, is computationally

expensive and throttled by the architecture of a conventional processor, this process is trivial in the optical domain. An optical lens forms the complex two-dimensional Fourier transform of a coherent wave in its back focal plane [57]. The phenomenon allows for a two lens 4-focal length coherent optical processor, illustrated in Figure 1.5, to implement a two-dimensional convolution between a reference kernel and an input image in the complex filter. The image in the output plane indicates the presence and location of the reference object in the input scene.
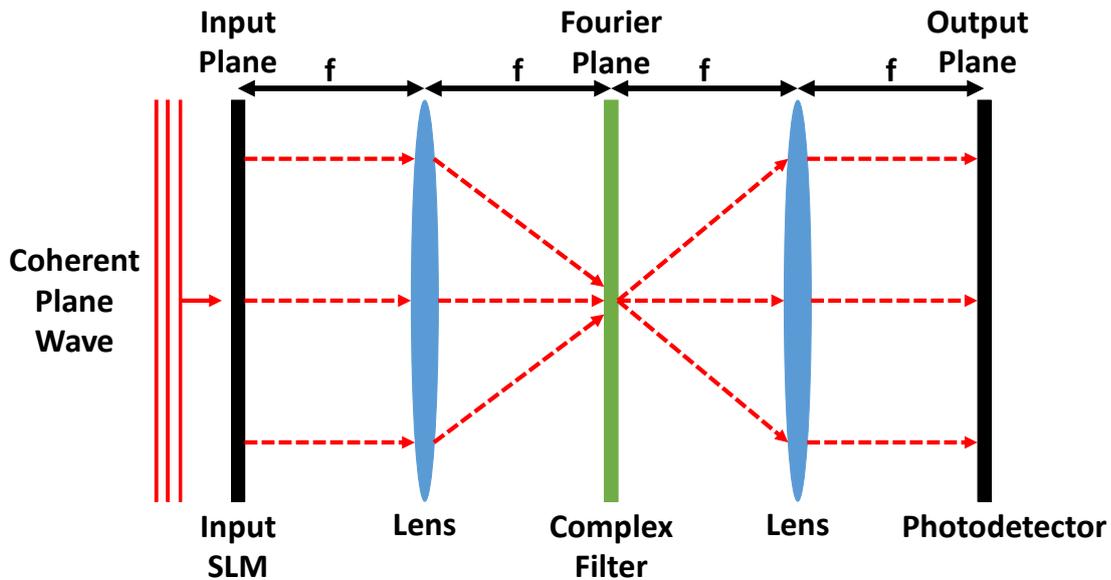
Figure 1.5: 4f coherent optical processor.

Throughout my research, I have explained the coherent optical processor using the analogy of Where's Wally, by comparing the operation of searching for a kernel in an image to looking for Wally in the popular children's puzzle book. Here, I will use the same analogy. Most deep learning systems undergo a process called supervised learning, which is to correct the machine's learning without direct input from the engineer into manipulating the raw data.

Consider a machine designed to classify images. First, a large collection of images labelled with their categories are shown to the machine. In the example shown here, this would be an image of Wally, shown in Figure 1.6(a). Naturally, in a real system, the machine would be shown multiple images of this character, and scored on how well it categorised the character as Wally. Equally, it would be shown different characters and scored well if it didn't classify them as Wally, such as those characters shown in the example scenery in Figure 1.6(b) [58].
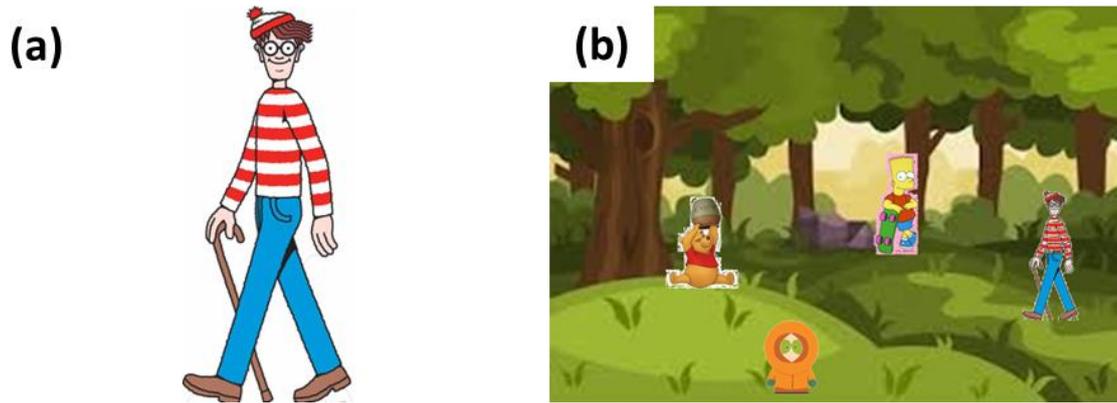
Figure 1.6: (a) An example kernel image, and (b) an example scene containing the kernel image.

Once the system has classified the target character, it stores the Fourier transform of the image, so that when it is shown a scene like Figure 1.6(b), it has the reference image (kernel) ready. To input the data, a spatial light modulator (SLM) is located at the input of the 4f optical system, illuminated by a laser source to modulate the input image. Assuming the system is then shown an image of the example scenery (Figure 1.7 (a)) and asked to find the character Wally, the scenery image is transferred to an optical signal and passed through the 4f coherent optical processor. The Fourier transform of the input image is multiplied by the Fourier transform of the kernel in Fourier space at the complex filter in Figure 1.5, where the Fourier transform of the kernel is projected, before passing through a second lens into real space.
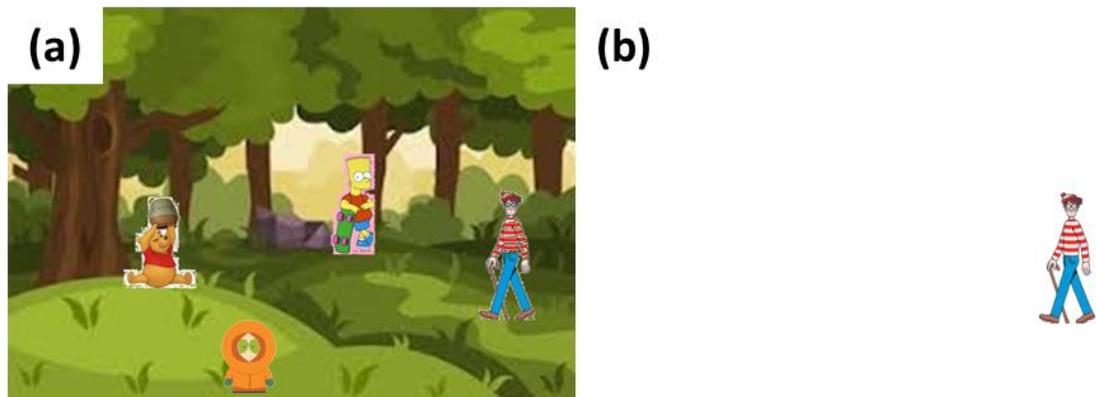
Figure 1.7: (a) An example scene containing the kernel image and (b) the returned information of the 4f coherent optical processor.

If the character Wally is present in the scene, the only returned information is Wally in the scene, as shown in Figure 1.7(b). But, if Wally is not present in the scene (Figure 1.8(a)), the returned information is nothing (Figure 1.8 (b)).
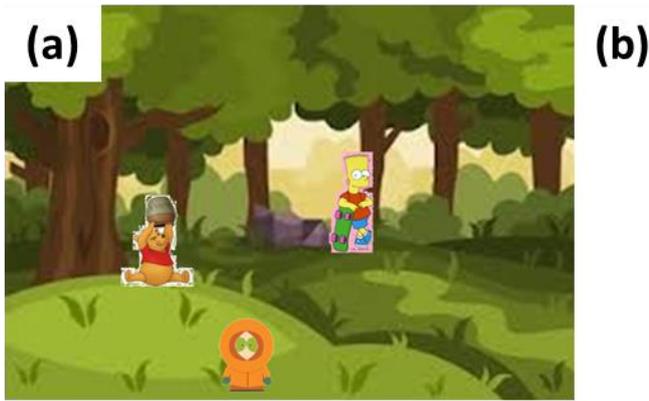
Figure 1.8: (a) An example scene that does not contain the kernel image and (b) the returned information of the 4f coherent optical processor.

The binary response from such a system, and the speed at which the convolution can be performed, highlight the power of the coherent optical processor. Naturally, in the example I have shown here, it is extremely easy to find the kernel image.
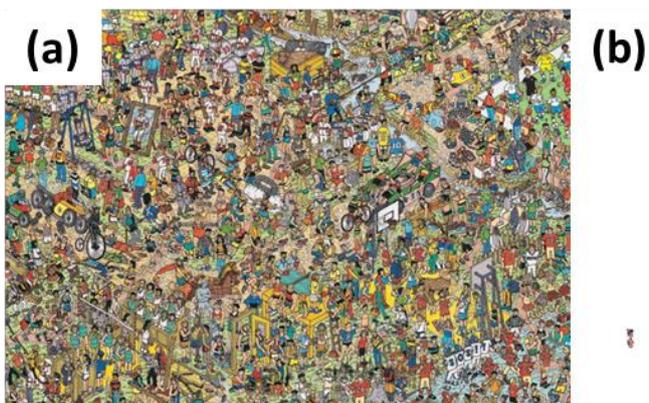
Figure 1.9: (a) An example scene containing the kernel image and (b) the returned information of the 4f coherent optical processor.

But in a scene, such as that in Figure 1.9(a), even a conventional electronic deep learning machine would take several seconds to find the character Wally. The output of the 4f coherent optical processor is shown in Figure 1.9(b) and can be calculated several orders of magnitude faster than conventional systems. The modulators discussed in this thesis serve as a spatial light modulator for the transfer of electrical data to the optical regime in Figure 1.5.

It is apparent that the 4f approach in Figure 1.5 has significant advantages compared to conventional computing methods, including the speed of light operation and passive image convolution, but it relies heavily on methods of transferring information from the electrical to the optical domain. As a single-purpose functional process, the photonic system can perform at with a significant speed increase, as well as the increase in efficiency offered by a decrease in

power consumption, the number of operations achievable per CPU clock cycle may be expressed as

$$number\ of\ operations = 6Cn^2 \log(n) + 4n^2$$

The 4f approach has the potential to be incredibly powerful. The number of operations per clock cycle is given by Equation 1.3, where n is the image size (proportional to the kernel size k), and C is a constant in the Fourier transform [59] of approximate value C = 1/log(2). Considering a pixel array of 64 x 64, driven by a top spec processor [60] at 4.8 GHz, yields a number of operations of 163,840 per processor cycle, where a conventional engine would work at 1 operation per cycle.

1.3.2.   Performance Metric of Spatial Light Modulators

In 2019 Optalysys [61] announced the world's first optical co-processor system for machine intelligence computing. They used the 4f optical processor and a deep learning framework to form a convolutional neural network architecture. Optalysys utilise spatial light modulators as the interface between electronic and optical domain, and CMOS camera sensors as illustrated in Figure 1.5. Their FT:X system operates approximately 30 times faster than state-of-the-art graphics cards, for like-for-like tasks, operating at only a quarter of the power. While commercially available SLMs can provide fine control of phase or amplitude of light, over a high resolution array of pixels (on the megapixel order) through the modulation of a coherent light source, most operate with a frame rate in the Hz – kHz range [62]. Wang et al. [63] report the use of a 4f optical processing system utilising a SLM to modulate an input signal to interfere with a radar (microwave) signal for analysis. Their system reports a modulation speed of 60 Hz. This performance limitation of digital SLMs provides an opportunity for the introduction of high-speed low-power photonic modulators, which are discussed in the next section.

1.4.   Review of Electro-Optic Modulators

An optical modulator is a device used to modulate propagating light in either free space or an optical waveguide by altering the amplitude, phase, or polarisation of the incident light. The field is extensive, so to narrow the parameter space further, we consider the detector of such a system measuring a change in incident intensity, i.e., amplitude modulation that can be achieved

through the change of a material's refractive index or the superposition of a phase difference between two paths of a coherent system, since this is practically achievable using a photodetector.

Amongst the literature, optical modulation has been achieved through electrical gating (electro-optic modulation) or thermal-gating (thermos-optic modulation), with the former preferable due to the relatively slow response rate and high-power consumption of thermo-optic modulators. The electro-optic effect is typically achieved through the Pockels and Kerr or carrier density effects [64], [65]. Electro-optic modulators (EOMs) inherently rely on the interaction strength between the signal and the modulation material, i.e., the light-matter interaction (LMI). The key to an ultrafast, compact, and power-efficient EOM is to increase the LMI to allow for a reduction in device size (L) and the driving voltage (V). The LMI can be increased, for instance, by increasing the overlap (Γ) of a waveguide mode with the material responsible for the modulation. For a modulator that induces an intensity change through a phase shift, this is represented as the product of the voltage required to obtain a π-phase shift and the modulator length, known as the half-wave voltage length ($V_{\pi}L$).

1.4.1.   State-of-the-Art Electro-Optic Modulators

The Pockels and Kerr effects are well established modulation mechanisms in the field, with the effects used primarily in lithium niobate ($LiNbO_3$)-based phase modulator designs, though this can easily be represented as an intensity modulation in an interferometric system layout. Both effects are an electric field induced refractive index change, where the Pockels effect is an anisotropic effect proportional to the electric field and he Kerr effect is a second-order electric field effect [17].

Janner and Tulli et al. [66] demonstrated a lithium niobate-based phase modulator with a Mach-Zehnder design layout operating with a half-wave voltage length of 800,000 Vμm. Wang et al. [67] report a thin film lithium niobate on insulator modulator with a figure of merit of 36,000 Vμm. Huang et al. [68] report a sub-millimetre hybrid silicon-rich nitride and thin film lithium niobate modulator with a half-wave voltage length of 67,000 Vμm. Lithium niobate modulators typically have a small footprint, but suffer from a relatively large power consumption, thus decreasing their efficiency.

The primary focus of this research, however, is carrier effect modulators. The carrier effect relies on the electrostatics of a metal-oxide-semiconductor (MOS) capacitor, whereby a bias voltage

can change the carrier concentration in an accumulation/depletion layer of a semiconductor material. The change in carrier concentration results in a refractive index change, described as the free carrier plasma dispersion effect [69].

The Intel group [70] demonstrated the first on-chip high-speed modulation on monolithically integrated silicon with an operating speed of >1 GHz, opening the door for silicon photonics in the field of electro-optic modulation. The modulator consists of a p-doped polysilicon rib waveguide separated from an n-doped crystalline silicon layer by a 12 nm oxide layer. A positive bias voltage is applied to the polysilicon waveguide, a thin layer of charge accumulates on both side of the oxide layer, modulating the refractive index of the silicon. Other examples of SiP EO modulators include an integration with Mach-Zehnder interferometer (MZI) (section 2.5.1) [71] and micro-ring structures [72]. Recent advances in the field of silicon EO modulators include Liu et al. [70] who report a silicon-based metal-oxide-semiconductor (MOS) capacitor modulator with 80000 Vμm performance.

The restricting factor of SiP phase modulators is power consumption. Generally, increasing the capacitance of the modulator reduces the power consumption as a lower drive voltage is required for the same carrier accumulation. However, this results in a greater space charge region, and overall higher dopant concentrations, which increases the overall insertion loss of the waveguide, and so there is a trade-off between phase shift efficiency and insertion losses. One might assume that simply increasing the input power would allow for improved efficiency due to a decreased effect of the insertion losses, but as will be discussed in section 3.1.1, silicon experiences nonlinearities at high throughput powers [73].

1.4.2.   Improving Light-Matter Interactions

Enhancement of the silicon platform to incorporate surface plasmon polaritons (SPPs) has attracted a lot of attention in the field in recent years [70], [74], [75] as the principle serves to overcome the intrinsic losses involved with higher phase-shift efficiency of silicon-based carrier modulation. SPPs are optical waves that propagate at the interface between a metal and a dielectric. Below the plasma frequency of the metal, the real component of its permittivity is negative and can be very large values far below the plasma frequency. As the frequency of the input signal approaches the plasma frequency of the metal, the magnitude of the metal's permittivity decreases and is opposite to the dielectric's permittivity. The dispersion of the SPP diverges with the wave number, resulting in very large wave numbers, and the group velocities

at the interface tend to zero due to the derivative relationship to the wave number. The large wave numbers and small group velocities are combined with a very short evanescent decay length of the light in the dielectric.

Consequently, SPPs are able to dramatically reduce the effective mode area, thus reducing the required volume in which the refractive index needs to be changed whilst maintaining a large mode overlap [76]. New device concepts have emerged by implementing a highly doped transparent conductive oxide in place of the metal layer, in which modification of the free-carrier concentration shifts the plasma resonance and further enhances the efficiency of the tuning mechanism. Indium tin oxide (ITO) is a highly promising material in the field [77]–[80], and it is here where this research focuses.

Real performance gains have been seen with the switch to ITO, as shown by Amin et al. [74], [77] who have demonstrated the capabilities of an ITO and aluminium oxide MOS capacitor atop a silicon waveguide, with a reported figure of merit of 52 Vμm. The dramatic increase in performance has highlighted the need for an ITO-based modulator, with further improvements lying in reducing overall insertion loss, i.e., moving beyond the silicon platform, as is proposed in this research.

1.4.3.   State-of-the-Art Optical Systems

In the wider field of all optical processing systems, Shen et al. [81] present an optical neural network (ONN) utilising 56 Mach-Zehnder interferometers (section 2.5.1) on a silicon waveguide platform. The modulation technique used is a 50thermo-optic modulation of the silicon platform. The authors report an operating speed in the GHz regime and a 10 mW per modulator operating power demand, though they also report that thermal cross-talk between modulators contributes largely to photodetection errors. Bueno et al. [82] demonstrate a 2025 node photonic neural network operating at approximately 5 Hz, and report that their biggest setback is currently the update speed of their SLMs, suggesting that high-speed EO modulators could yield significant performance gains. Cheng et al. [83] report an all optical reservoir computing neural network with a maximum 10 W power consumption, operating at the telecommunication wavelength of 1550 nm, and thus utilising a silicon platform.

1.5.   Scope of the Thesis

The core research question of this thesis is: can we develop a new, highly efficient, carrier-based electro-optic modulator platform suitable for large-scale coherent optical processors?

The concern with the silicon platform is that it is unsuitable for high-throughput optical systems, and the increased insertion losses at wavelengths lower than the typical telecommunications wavelength of 1550 nm places restrictions on the system that are not necessary. We propose a silicon nitride waveguide platform as a more suitable prospect given the low loss through the visible and infra-red spectrum [84] and the higher available throughput power making it far more suited to large-scale systems. As silicon nitride does not exhibit carrier accumulation under electrical gating, the modulation technique is through carrier accumulation in indium tin oxide, thus changing a local refractive index and the effective index of the waveguide mode. There are, naturally, several smaller research goals that are covered in this thesis, and these are outline here along with the overall scope of the thesis.

Chapter 2 provides the background theory and information for my research, specifically the optical background for guiding and manipulating light and the governing physics behind metal-oxide-semiconductor capacitors. The smaller research question discussed in this chapter is how to accurately determine the capacitance of a highly degenerate semiconductor such as indium tin oxide.

Chapter 3 provides information on how I designed and simulated my modulator structures to obtain the fabrication dimensions required to achieve my research goals. Here I discuss how valuable simulation data can be and how it can be used to reduce fabrication time. The smaller research question in this chapter is how simulation data may be applied to real fabrication techniques, and why it is important to consider the fabrication tolerances prior to simulation.

Chapter 4 details a collaborative research goal to characterise the dependency of the optical and electrical properties of indium tin oxide on the oxygen concentration of the material. In this chapter I present a novel approach to the characterisation of the material that is not discussed in the literature. The approach discussed here is necessitated by the overuse of assumed literature values.

Chapter 5 contains the relevant fabrication techniques and parameters used to create the modulators discussed in this research. Here I discuss my fabrication choices and provide a

detailed schematic of how the structure comes together. The smaller research goal outlined in this chapter is the development of a fabrication protocol.

In **Chapter 6** I use the characterisation method for indium tin oxide discussed in chapter 4 and apply it to the indium tin oxide used in my modulators to obtain the optical and electrical properties of the semiconductor layer in the capacitor. Using this information, I then characterise the capacitor and predict its performance metrics. In this section, I answer a wide literature question of the potential for metal/semiconductor penetration into the oxide layer of the capacitor and detail my findings.

In **Chapter 7** I present amplitude and phase modulation results and the details of the optical setup used to obtain them. I also discuss an unpredicted phenomenon encountered during my research.

In **Chapter 8**, I summarise my research and determine if the research question has been answered. I will discuss, critically, how I have developed and how the project could be improved, whilst giving an outlook to the future of the field.

2. Background and Theory

Chapter Synopsis

This chapter presents the textbook analysis of the background theory required and referenced throughout this thesis. It contains the relevant theory for describing light in free-space, periodic structures, resonant gratings, and slab waveguides. I also outline how the optical modulation can be achieved and measured, with specific reference to- and analysis of- the space charge region in a gated metal-oxide-semiconductor structure. Specifically, I will focus the discussion on ITO-based modulators, which have demonstrated their capabilities in the literature in recent years, with many authors utilising the material for phase modulation applications in the telecommunications regime. Before I get into designing, fabricating, and characterising my own modulators, I will detail the background of ITO modulators and the theory on which many of the models, predictions, and calculations in this thesis are built on. The chapter culminates with an analytical model for the predicted performance of the modulators discussed in this thesis, and a comparison to a literature device. The analytical model was implemented in python in original code, with reference to a textbook derivation. I would like to acknowledge Dr George Duffett for his assistance in developing this model with me.

2.1.   Describing Light

It is conventional for a discussion of the classical description of electromagnetic fields to begin with Maxwell's equations, and so that is where I shall begin. Throughout this work, I will discuss the manipulation of light in both waveguide structures and photonic crystals; therefore, in this section I will discuss how light can be guided in these structures, before moving to an application setting and how we can use Maxwell's equations.

2.1.1.   Maxwell's Equations

Light, considered as a propagating plane electromagnetic wave, consists of an oscillating electric and magnetic field, where the fields are in phase and orthogonal to each other and the propagation direction, described by the wave vector (**k**). The magnitude of the wave vector is given as a function of the wavelength (λ) in Equation 2.1. All macroscopic electromagnetism is governed by the four Maxwell equations.

$$|\boldsymbol{k}| = \frac{2\pi}{\lambda}$$

Equation 2.1

$$\nabla \cdot \boldsymbol{B} = 0, \qquad \nabla \cdot \boldsymbol{D} = \rho$$

Equation 2.2

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}$$

Equation 2.3

The four Maxwell equations (Equation 2.2 and Equation 2.3) describe the relationship between the electric field (**E**), magnetic field (**H**), displacement field (**D**), and magnetic induced field (**B**) with free charge (ρ) and current (**J**) densities. A full treatment of Maxwell's equations, especially with respect to periodic structures, is given by Joannopoulos et al. [85]. Typically, for the description of light propagating in a dielectric structure, we make certain assumptions to simplify Maxwell's equations. Firstly, we assume that there are no sources of light in the structure, and that it is free from free charges and currents, i.e., ρ = 0 and **J** = 0.

$$\boldsymbol{D} = \varepsilon_0 \varepsilon \boldsymbol{E}, \qquad \boldsymbol{B} = \mu_0 \mu \boldsymbol{H}$$

Equation 2.4

We assume that the field strengths are small enough so that we can operate in the linear regime, and the non-linear effects that arise from the displacement field relationship to the electric and magnetic fields can be ignored, and that the materials are macroscopic and isotropic. This means that the electric and displacement fields are related through the vacuum permittivity (ε_0) and a

scalar dielectric function (ε), known as the "relative permittivity", as shown in Equation 2.4. Similarly, the magnetic induction field is related to the magnetic field through the vacuum permeability (μ_0) and the relative permeability (μ). However, we assume that the materials are non-magnetic, which is typically the case at optical wavelengths (μ=1) [85].

$$\nabla \cdot \boldsymbol{E} = 0, \qquad \nabla \cdot \boldsymbol{B} = 0 \qquad \text{Equation 2.5}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \nabla \times \boldsymbol{B} = \mu_0 \varepsilon_0 \varepsilon \left(\frac{\partial \boldsymbol{E}}{\partial t}\right) \qquad \text{Equation 2.6}$$

Finally, we assume that there are no explicit material dispersion effects, i.e., the dielectric constant is not frequency dependent but instead, assumes an appropriate value for the frequency range of interest; and we assume that it is real and positive, and the material is therefore transparent. With all these assumptions, the Maxwell equations can be represented by Equation 2.5 and Equation 2.6. These equations form the basis for electromagnetic theory, of particular interest are the curl equations, which indicate that electromagnetic waves do not need a medium through which to propagate.

$$\nabla^2 \boldsymbol{E} = \mu_0 \varepsilon_0 \left(\frac{\partial^2 \boldsymbol{E}}{\partial t^2}\right) \qquad \text{Equation 2.7}$$

Maxwell's equations are coupled first-order partial differential equations, but using some well-known vector identities, they can be decoupled to yield the second-order equation (Equation 2.7) [86], also known as the "wave equation". This equation describes a wave with propagation velocity $1/\sqrt{\mu\varepsilon}$, i.e., the speed of light in a vacuum c ≈ 3x10^{8}ms^{-1}.

$$\frac{1}{\sqrt{\mu\varepsilon}} = \frac{c}{n}, \qquad where\ n = \sqrt{\frac{\mu\varepsilon}{\mu_0\varepsilon_0}} \qquad \text{Equation 2.8}$$

When light propagates through a dielectric material, the atoms are polarised and magnetised in response to the oscillating fields within the material, and in turn create their own fields. The result is a single wave with its own velocity, and the only change to the Maxwell equations is to replace μ_0 and ε_0 with μ and ε. The propagation velocity of the wave through the material is then given by Equation 2.8. The refractive index, n, acts as a velocity scaling factor and describes how the velocity is reduced compared to the vacuum propagation velocity. For the transparent non-magnetic dielectric materials discussed here, refractive index can be approximated as $n = \sqrt{\varepsilon}$.

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}(\boldsymbol{r})e^{-i\omega t}, \qquad \boldsymbol{H}(\boldsymbol{r}, t) = \boldsymbol{H}(\boldsymbol{r})e^{-i\omega t} \qquad \text{Equation 2.9}$$

Exploiting all the Maxwell equation assumptions, the temporal (t) and spatial (**r**) terms can be separated and Equation 2.8 can be expressed as "harmonic modes". Equation 2.9 describes

these harmonic modes, where the result is a spatial mode profile multiplied by a complex exponential with angular frequency (ω). This allows the spatial mode profiles to be calculated by treating Maxwell's equations as an eigenvalue problem, where the resulting mode profiles are the eigenfunctions of the system and the angular frequencies are the corresponding eigenvalues [87].

2.1.2.   Light in Waveguides

The origin of the harmonic mode solutions to the Maxwell equations becomes apparent when we look closer at how light behaves in an optical waveguide, a structure used to confine and guide light through the principle of total internal reflection (TIR). The most common optical waveguide is the optical fibre, which usually has a circular cross section and visible cladding around the waveguide to protect the delicate structure and maintain TIR. However, in the field of integrated optics, it is often planar structures such as thin film guides or waveguide strips that are used, and it is here that I will focus my discussion.
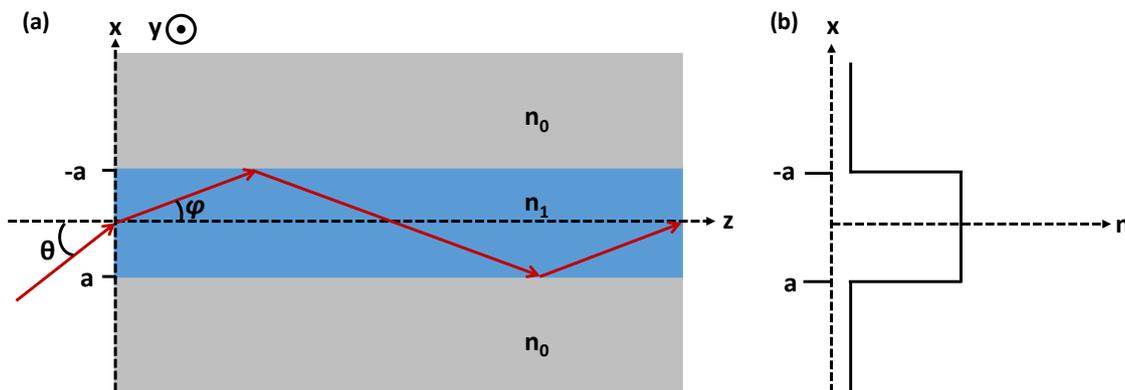
Figure 2.1: (a) Basic structure of the optical waveguide with indicated refractive indices and (b) refractive index profile of the optical waveguide. Light ray is indicated by the red arrow in (a), guided in the blue waveguide layer.

The basic optical waveguide consists of a core material, in which the light is confined, surrounded by a cladding layer, as illustrated in Figure 2.1(a). The cladding layer (n_0) may be air in some cases, but for the light to be guided in the core material, the refractive index (n_1) must be higher than that of the cladding layer, as illustrated in Figure 2.1(b).

$$n_1 \cos(\varphi) \geq n_0, \qquad \sin(\theta) = n_1 \sin(\varphi) \leq \sqrt{n_1^2 - n_0^2} \qquad \text{\textit{Equation 2.10}}$$

The condition for TIR at the core-cladding interface is given as a function of the refractive indices (n_1 and n_0) and the internal reflection angle (φ), which in turn is related to the incident angle (θ) as shown in Equation 2.10.

$$\theta \leq \sin^{-1}\left(\sqrt{n_1^2 - n_0^2}\right) \equiv \theta_{\mathrm{m}}, \qquad \theta_m \cong \sqrt{n_1^2 - n_0^2} \qquad \text{\textit{Equation 2.11}}$$

Equation 2.11 can be used to determine the limit of the incident angle (θ_m), known as the critical angle. This denotes the maximum angle at which light may be coupled into the waveguide and be fully guided through TIR and is often referred to as the "numerical aperture" (NA) of the waveguide.

$$NA = \theta_m \cong n_1\sqrt{2\Delta}, \qquad \Delta = \frac{n_1^2 - n_0^2}{2n_1^2} \qquad \text{\textit{Equation 2.12}}$$

The relative refractive index difference (Δ), commonly expressed as a percentage, is related to the numerical aperture as detailed in Equation 2.12 [88]. It should be noted, however, that the acceptance angle is discrete, and each associated mode can be determined through the following analysis.

Consider the waveguide mode illustrated in Figure 2.1, propagating in the z-direction with an inclination angle (φ) with perpendicular phase fronts. We know that the wave velocity is given by the vacuum wavelength (λ) and the core index (n_1), and we know the wave vector magnitude is given by Equation 2.1.
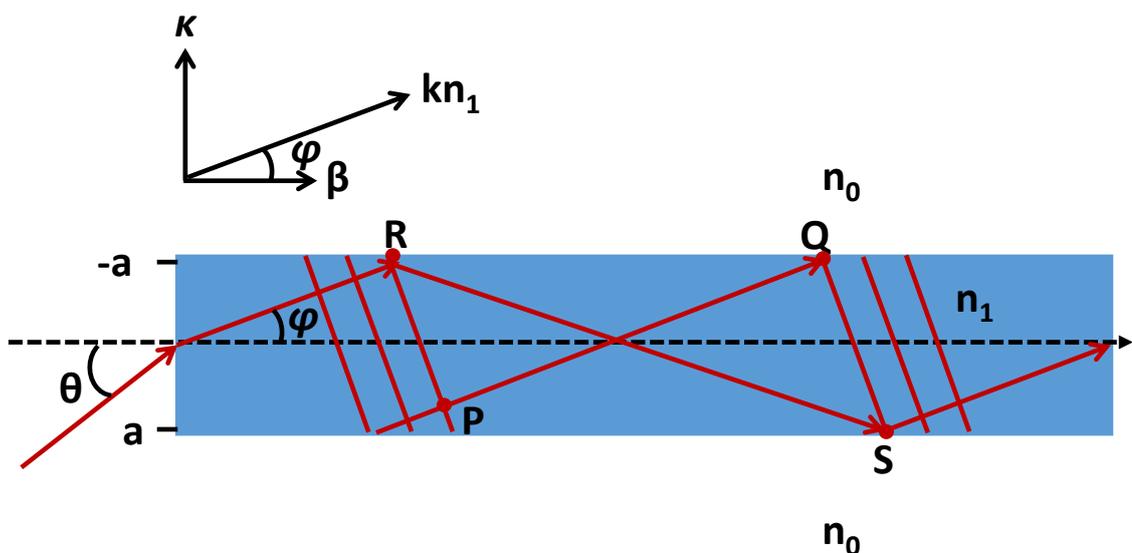
Figure 2.2: Light rays and their phase fronts in an optical waveguide.

$$\beta = kn_1 \cos(\varphi), \qquad \kappa = kn_1 \sin(\varphi) \qquad \textit{Equation 2.13}$$

Such a scenario is illustrated in Figure 2.2. As the wave is sinusoidal, we can use Euler's formula [89] to represent the propagation constants that determine the light ray vector (kn_1). The z- and x-directional components are given by β and κ, respectively, in Equation 2.13. Hence, we can adjust our expression for an electromagnetic wave (Equation 2.9), for the case where it is propagating with a constant β to Equation 2.14.

$$\boldsymbol{E}(\boldsymbol{r},t) = \boldsymbol{E}(\boldsymbol{r})e^{i(\omega t - \beta z)}, \qquad \boldsymbol{H}(\boldsymbol{r},t) = \boldsymbol{H}(\boldsymbol{r})e^{i(\omega t - \beta z)} \qquad \textit{Equation 2.14}$$

Let us consider the phase difference between the two light rays on the same plane illustrated in Figure 2.2 as |PQ| and |RS|. Despite the extra core-cladding reflections experienced by ray RS, the optical paths of the two rays should differ by a multiple of 2π as they are on the same phase front.

$$l_1 = 2a\left(\frac{1}{\sin(\varphi)} - 2\sin(\varphi)\right), \qquad l_2 = \frac{2a}{\sin(\varphi)} \qquad \textit{Equation 2.15}$$

$$(kn_1 l_2 + 2\Phi) - kn_1 l_1 = 2m\pi \qquad \textit{Equation 2.16}$$

The distance between points P and Q, and R and S, can be expressed using Equation 2.16 and the phase (ϕ) matching condition then becomes Equation 2.17, where m is an integer.

$$\tan\left(kn_1 \operatorname{asin}(\varphi) - \left(\frac{m\pi}{2}\right)\right) = \sqrt{\left(\frac{2\Delta}{\sin^2(\varphi)}\right) - 1} \qquad \textit{Equation 2.17}$$

Combining Equation 2.16, Equation 2.17, and the Goos-Hänchen [90] shift, yields the condition for propagation angle (φ) shown in Equation 2.17. This shows that the propagation angle of a light ray is discrete and is determined by the waveguide structure and the wavelength of incident light [91]. The optical field distribution that satisfies the phase-matching condition of Equation 2.17 is called the mode. The allowed value of the propagation constant β is also discrete and is denoted as an eigenvalue. The mode with the minimum internal reflection angle occurs at m=0 and is known as the fundamental mode, with higher order modes thus having larger internal reflection angles.

2.1.3.   Light in Periodic Structures

To this point I have only considered a homogenous medium, where the harmonic solutions to the wave equations are plane waves and the permittivity (ε) has been considered only as the real component (ε_r) of the refractive index (**n**).

$$\boldsymbol{\varepsilon} = \varepsilon_r + i\varepsilon_i = \boldsymbol{n}^2 = (n + ik)^2 \qquad \textit{Equation 2.18}$$

While the use of complex materials (Equation 2.18) is reserved for later sections, a special case for solving the wave equations arises when we consider a spatial distribution of the permittivity with a period (a) of comparable size to the material wavelength of light; it is this structure that I will discuss here. Such a structure is known as a photonic crystal, and may have a periodic permittivity in one-, two-, or three-dimensions. For simplicity and relevance to this work, I will only consider a photonic crystal with periodic permittivity in one dimension.

$$\boldsymbol{E_k}(\boldsymbol{r}) = u_k(\boldsymbol{r})e^{i(\boldsymbol{k}\cdot\boldsymbol{r})} \qquad \textit{Equation 2.19}$$

Bloch's theorem [92] states that when the permittivity is periodic, the spatial mode profiles (**E$_k$(r)**) can be expressed as a plane wave with an amplitude modulated by a function (u$_k$(**r**)) with the same periodicity as the permittivity, as detailed in Equation 2.19 [93].

$$\boldsymbol{G} = \frac{2\pi}{a} \qquad \textit{Equation 2.20}$$

$$\boldsymbol{E_k}(\boldsymbol{r}) = \sum_{m=-\infty}^{\infty} E_m e^{i(\boldsymbol{k}+m\boldsymbol{G})\cdot\boldsymbol{r}}, \qquad m = \cdots, -2, -1, 0, 1, 2, \ldots \qquad \textit{Equation 2.21}$$

The Bloch states given by Equation 2.19 are distinguished by their k-vector but can be expressed using a grating vector (**G**), which is a function of the grating period (a). This leads to Equation 2.21, an expression for a plane wave coupling to the Bloch modes in a periodic structure by adding integer multiples (m) of the grating vector to the incident k-vector.
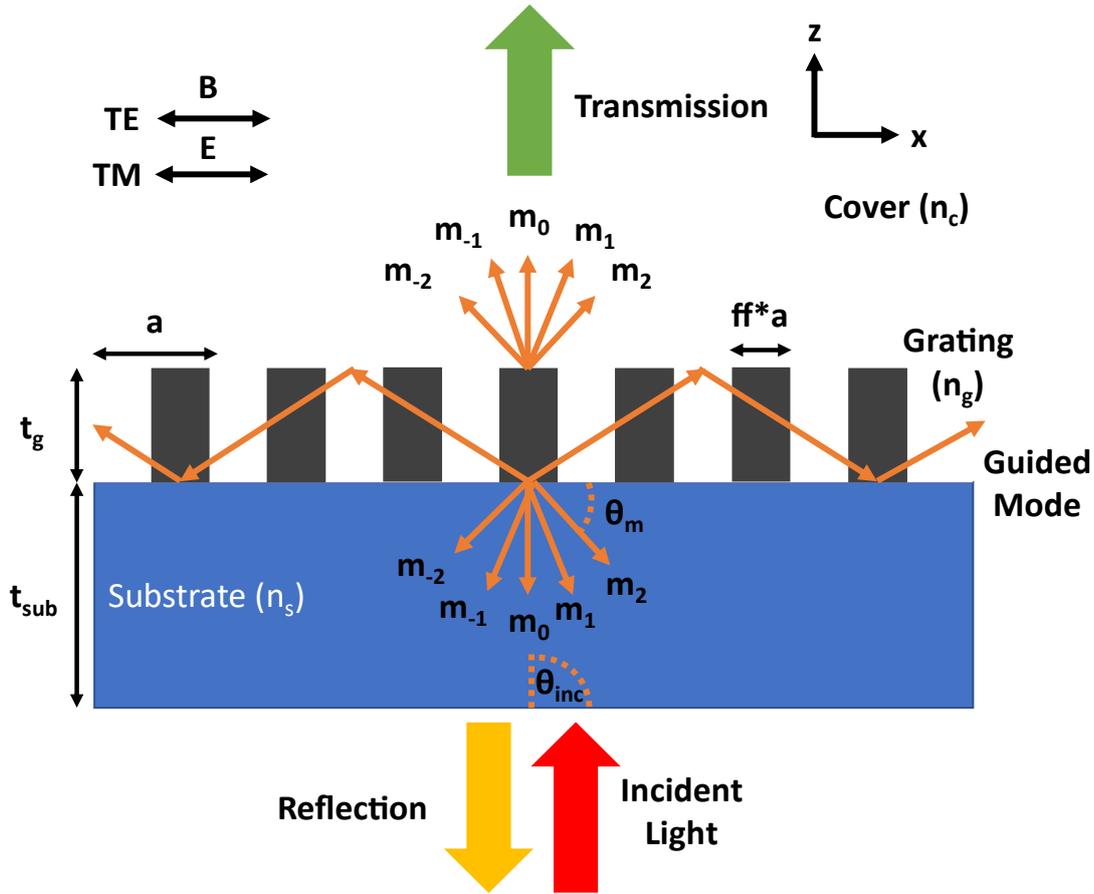
Figure 2.3: Illustration of a typical 1D periodic structure used in this work, with guided and diffracted modes indicated.

This means that the incident wave splits into diffraction orders, each with their own k-vectors as illustrated in Figure 2.3, a diagram of the typical 1D periodic structure used in this work. Light is incident on the grating from below, with a k-vector (k_{inc}), through the substrate material (n_s) at normal incidence. The grating (n_g) and cover (n_c) indices are marked, with the grating having a thickness (t_g), period (a), and ridge width (ff*a), where the fill factor (ff) is a percentage of the period. There are diffracted modes and a guided mode as illustrated in Figure 2.3, with the diffraction orders individually marked with their order up to, but not limited to, m = ±2.

$$k_{x,m} = k_{x,inc} - mG, \qquad m = \cdots, -2, -1, 0, 1, 2, \ldots \qquad \textit{Equation 2.22}$$

$$n_{eff} \sin(\theta_m) = n_s \sin(\theta_{inc}) - m\left(\frac{\lambda_0}{a}\right) \qquad \textit{Equation 2.23}$$

The wave vector describes the propagation of the wave, and we can impose field boundary conditions and the conservation of momentum to dictate the phase-matching condition in the direction of the periodicity (x) (Equation 2.22). Substituting more suitable values for the k-vectors and introducing an electric field weighted average refractive index (n_{eff}), known as the

"effective index", we can rearrange to yield the familiar grating equation (Equation 2.23). Here I have used θ_m and θ_{inc} to indicate the diffraction angle of the mth order mode and the incident light respectively, *a* represents the grating period, and λ_0 refers to the vacuum wavelength of the incident light [94]. It has been shown that the effective grating index can be approximated from the grating parameters [87], but a more in-depth analysis is beyond the scope of this thesis.

2.2. Describing Permittivity

Permittivity is simply the ratio of electric displacement to the electric field in a material, i.e., it describes the opposition a material provides against an electric field. The permittivity is a representation, then, of how electric fields propagate through a material, and hence the relationship between permittivity and refractive index becomes apparent. It follows that the permittivity of a material is dispersive with respect to the electric field strength, and it is commonplace to plot the dispersive permittivity as a function of incident wavelength or frequency.

For this work, the permittivity of the dielectric (insulating) materials has been well studied and characterised in the literature [84], [95], [96]; but the problem becomes more complex when considering metallic films, such as indium tin oxide. Here, the permittivity is strongly linked to the material's electron response to the electric field, and electron behaviour in solids is not trivial. In metallic materials, the intraband and interband transitions from the free and bound electrons, respectively, contribute to the permittivity and can be modelled by classical forces acting on a negatively charged gas of electrons. This model is known as the "Drude Model" [97].

2.2.1. The Drude Model

The Drude model [98] is an application of kinetic theory and aims to link electron motion to a conductor's macroscopic material properties. The model was extended by Lorentz, who linked the periodic motion of electrons to the current density equation derived by Drude to yield a permittivity [99]. To derive the Drude-Lorentz equation, we first must make some assumptions; firstly, that the conduction electrons are free to move against a background of positive ion cores and have a mass m*, an adjusted value based on their motion between the ion cores. We assume that the electrons collide with the ion cores and other electrons with a frequency (Γ), proportional to their velocity d**x**/dt, which damps their oscillation, known as the "collision

frequency". As the electrons move in an electric field, they experience a force (F) which is proportional to their mass and acceleration (m* d^2**x**/dt).

$$m^*\left(\frac{d^2\boldsymbol{x}(t)}{dt^2}\right) + m^*\Gamma\left(\frac{d\boldsymbol{x}(t)}{dt}\right) = -e\boldsymbol{E}(t)$$

Equation 2.24

If we assume that the electric field has a time-harmonic dependence as given in Equation 2.9, then the description of the electron's motion becomes Equation 2.24, where e represents the elementary charge [79]. The most convenient way to solve Equation 2.24 is to consider it as a function of frequency.

$$\boldsymbol{x}(\omega) = \left(\frac{e}{m^*}\right)\left(\frac{\boldsymbol{E}(\omega)}{\omega^2 + i\omega\Gamma}\right)$$

Equation 2.25

The time-harmonic electric field has a solution of the form $\boldsymbol{x} = \boldsymbol{x}_0 e^{-i\omega t}$, which yields the general solution for the equations of motion, shown in Equation 2.25. Now that we have this solution, we can find the susceptibility and conductivity relationships and derive the permittivity equation.

$$\boldsymbol{P}(\omega) = -ne\boldsymbol{x}(\omega) = \varepsilon_0\chi(\omega)\mathbf{E}(\omega)$$

Equation 2.26

If we assume that the electrons do not interact with each other, the polarisation density (**P**) is given as a function of the electron density (n) and has a linear relationship to the susceptibility (χ) and the permittivity of free space (ε_0), shown in Equation 2.26. The susceptibility is the component of the permittivity that is frequency dependent and is given by Equation 2.27.

$$\chi(\omega) = -\left(\frac{\omega_p^2}{\omega^2 + i\omega\Gamma}\right)$$

Equation 2.27

$$\omega_p^2 = \frac{ne^2}{\varepsilon_0 m^*}$$

Equation 2.28

Here we have incorporated the electron charge, charge density, and electron mass into a parameter known as the plasma frequency (ω_p^2), shown in Equation 2.28. Now we have a function of the relative permittivity of the material (ε_r), given in Equation 2.29.

$$\varepsilon_r = 1 + \chi(\omega) = 1 - \left(\frac{\omega_p^2}{\omega^2 + i\omega\Gamma}\right)$$

Equation 2.29

Finally, we consider the conductivity of the material using Ohm's law [100], specifically the current density (**j**), which is proportional to the velocity of the electrons and the material conductivity (σ) (Equation 2.30).

$$\boldsymbol{j}(\omega) = -ne\left(\frac{d\boldsymbol{x}}{dt}\right) = i\omega\left(\frac{\varepsilon_0\omega_p^2\boldsymbol{E}(\omega)}{\omega^2 + i\omega\Gamma}\right) = \sigma(\omega)\boldsymbol{E}(\omega)$$

Equation 2.30

$$\boldsymbol{\varepsilon}(\omega) = \varepsilon_\infty - \frac{\sigma(\omega)}{i\omega} = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\omega\Gamma}$$

<div align="right">*Equation 2.31*</div>

This leads to the general permittivity equation (Equation 2.31) for the Drude-Lorentz model with the additional contributions from the ion cores manifesting as an offset in permittivity known as the "high frequency permittivity" (ε_∞) [101], [102].

$$\boldsymbol{\varepsilon}(\omega) = \varepsilon_r(\omega) - i\varepsilon_i(\omega)$$

<div align="right">*Equation 2.32*</div>

$$\varepsilon_r(\omega) = \varepsilon_\infty - \left(\frac{\omega_p^2}{\omega^2 + \Gamma^2}\right), \qquad \varepsilon_i(\omega) = \frac{\omega_p^2 \Gamma}{\omega^3 + \omega\Gamma^2}$$

<div align="right">*Equation 2.33*</div>

Using the complex permittivity relationship (Equation 2.32), it becomes clear that the real component of the permittivity (ε_r) is dominated by the electrical response of the material in an electric field, and the imaginary component (ε_i), i.e., that which contribute to the absorption of the material, is dominated by the collision frequency, and therefore the electron density. I have used the Drude-Lorentz model for permittivity above the plasma frequency, where the Drude model is valid for the material, throughout this work to characterise the optical response of thin film indium tin oxide in an electric field and have gone into more detail in chapter 4.

2.3.   Resonant Gratings

While resonant gratings are only a tool used in this work to further understand the characteristics of thin film indium tin oxide (section 4.6), it would be remiss of me not to discuss the theory of guided-mode resonance in high-contrast gratings to further the understanding in this thesis. High-contrast gratings allow for a higher reflectivity and quality factor in their resonance, hence the inclusion in this work as a characterisation tool. Here I will briefly discuss the principle of guided-mode resonances, and how the grating parameters can affect the resultant optical response. The operation of these structures is derived from the theory discussed in section 2.1.3.

2.3.1.   Guided-Mode Resonances

Guided-modes resonances (GMRs) are supported by a grating when the grating layer also operates as the waveguiding layer. Here I will refer to Figure 2.3 in my description of the grating. The gratings can be designed to exhibit arbitrarily narrow resonance linewidths or complete reflectance/transmittance by tuning the refractive index contrast between the grating and cover

layers, or the fill factor [87]. The resonance wavelength is highly sensitive to the refractive index contrast between the grating index and cover index, which what makes GMRs useful in this work as we can use them to reverse engineer the refractive index value from a resonant response (section 4.6).

$$n_{eff} > n_s, \qquad \frac{\lambda_0}{n_{eff}} < a < \frac{\lambda_0}{n_s} \qquad\qquad \text{\textit{Equation 2.34}}$$

The resonance in these structures is governed by Equation 2.34, which states that the effective index of the mode must be higher than the refractive index of the surrounding layer, which I have assumed to be the substrate layer in Figure 2.3. This condition is a direct result of the Bragg condition at the boundary between two permittivities, and the solution to the Bloch mode equation detailed in Equation 2.23. The condition ensures that the grating can act like a waveguide, and support guided modes; like a waveguide, the number of modes it can support is dictated by the grating thickness (t_g). The other condition in Equation 2.34 states that the grating period (a) must be smaller than the wavelength in the incident material (again indicated as the substrate layer) but higher than the wavelength of the guided mode. This condition ensures a strong resonance and is often referred to as the "near-lambda" condition [103].

2.3.2.   Operation Modes

From Equation 2.34, we can extract the three operating modes of a GMR grating, determined by the grating thickness (t_g) and the grating period (a). To demonstrate this, I used a rigorous coupled-wave analysis (RCWA) technique using an implementation known as "S4" [104], which I will discuss in detail (section 3.2). I used an indium tin oxide grating (n_g=2.0) and a borosilicate substrate (n_s=1.46) with air as the cover material (n_c=1). The ridge width is determined by a scaling factor, known as "fill factor", which was set to 0.7, and the grating period was set to 500nm; this yields a ridge with of 350nm. The grating response is in the transverse electric (TE) polarisation. The resulting resonant response is shown in Figure 2.4(a), with the three operating modes.
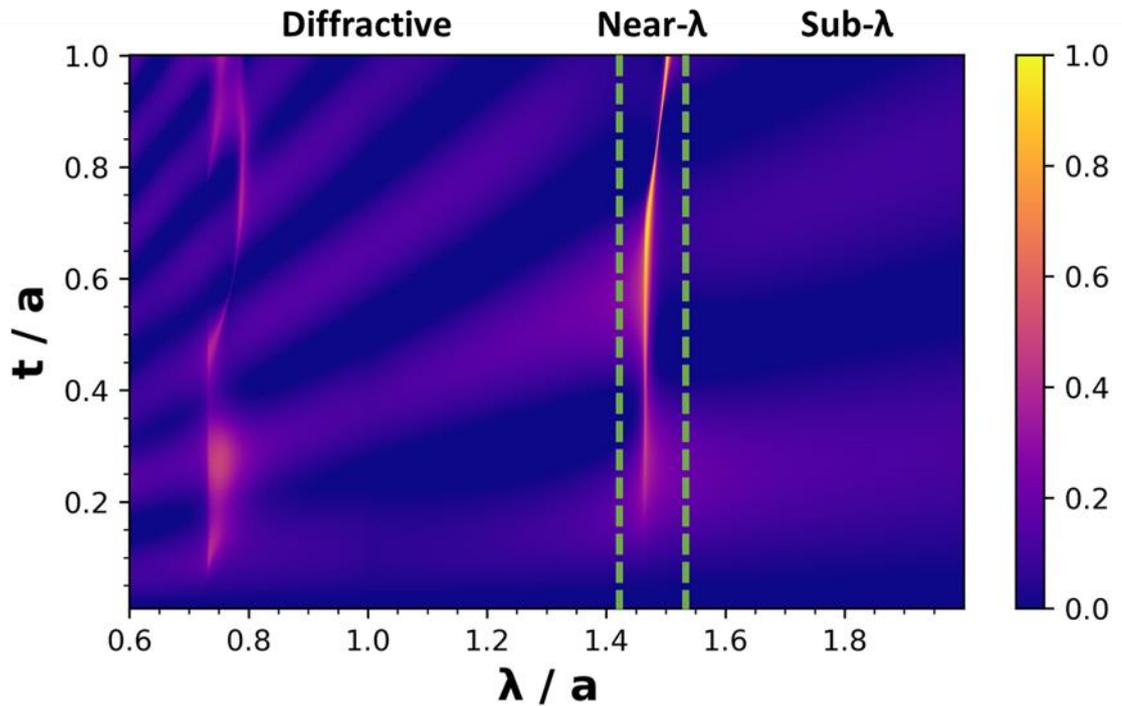
Figure 2.4: Reflectance as a function of wavelength (λ) and grating thickness (t), both normalised to grating period (a), showing the diffractive, near-λ, and sub-λ operation modes.

The three operating modes shown in Figure 2.4 are as follows:
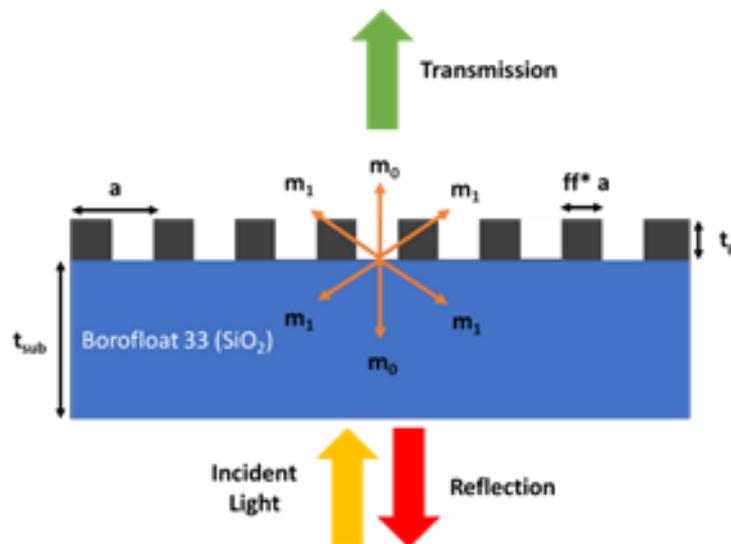
Figure 2.5: Schematic of diffractive operating mode for a resonant structure.

Diffractive: The grating is purely diffractive and generates modes at different angles, observed by the high reflectivity around 0.7-0.8 on the x-axis in Figure 2.4. This occurs at λ/a ≤ 1.4 (approximately the index of the substrate), as dictated by the grating equation (Equation 2.34), and in this regime it is the Fabry-Perot response of the thin film that can be observed in Figure

2.4 by the horizontal light and dark pattern in the reflectance. As illustrated in Figure 2.5, the non-zero order modes carry energy away from the grating at different diffraction angles, and as such there are no high reflectance areas or resonances in the diffraction region.
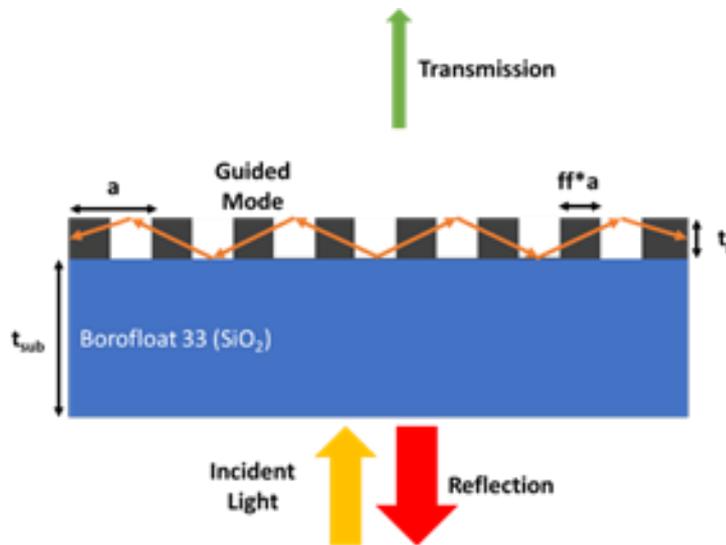
Figure 2.6: Schematic of near-λ operating mode for a resonant structure.

Near-λ: For this grating there is one mode of high intensity, indicated by the strong reflectance bar between the green dotted lines in Figure 2.4. This is a guided mode resonance, occurring in the $1.4 \leq \lambda/a \leq 1.6$ (approximately the index of the substrate and the effective index of the structure), as dictated by the grating equation (Equation 2.34), region. Here we are observing the fundamental mode ($m = 0$, $\theta_m = 0°$), with the other diffractive modes coupling into the grating layer due to a high effective index. As illustrated in Figure 2.6, the reflectance within the grating is a result of the thin-film interference with guided mode excitation. The reflectance in Figure 2.4 is a result of the waveguide thickness beneath the grating, and more guided modes would be visible for a thicker waveguide layer.
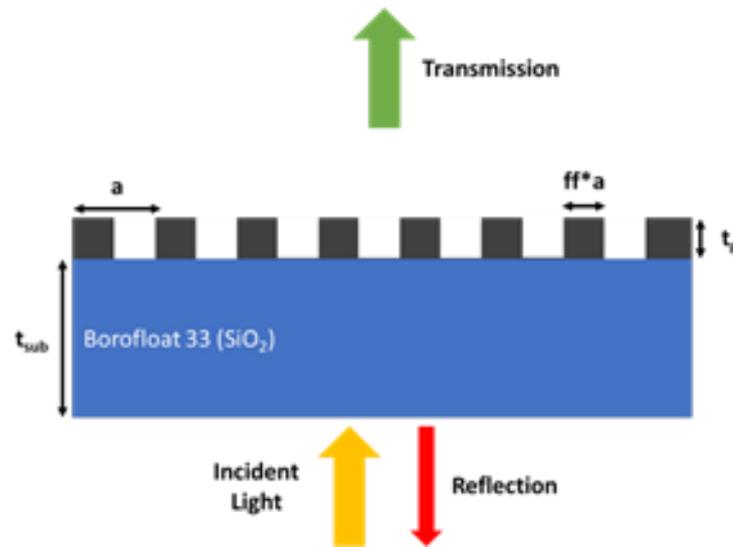
Figure 2.7: Schematic of sub-λ operating mode for a resonant structure.

Sub-λ: When the grating period is sufficiently small compared to the wavelength, the GMR operates in a sub-λ mode, where the light behaves as if it was incident on a thin film with the same effective index, the response is described by the Fresnel equations in this regime. We observe the sub-λ operating mode in Figure 2.4 at λ/a ≥ 1.6 (approximately the effective index of the structure) as dictated by the grating equation (Equation 2.34). As illustrated in Figure 2.7, in the sub-λ regime, only the zero-order mode can propagate away from the grating, and we observe the thin film effect.

It is clear from the above discussion that the ideal operating mode of the GMR grating is in the near-λ, where the 0th order mode exists in the cover and substrate layers, but the higher order modes are coupled and propagate laterally through the grating layer. The light scatters at each grating interface, since unlike a normal waveguide, the permittivity is inhomogeneous. These scattered modes are often referred to as "leaky modes" [105]–[108], and allow light to be coupled out of the grating. When these leaky modes are phase matched to the 0th order mode propagating through the structure, they produce a strong reflectance or transmittance response, which we observe in Figure 2.4 in the near-λ regime.

The line shape of the grating resonance is asymmetric due to the interference between the slowly-varying thin-film response of the grating layer and the quickly-varying Bragg resonance of the in-plane guided modes. As such, it is common to observe a Fano line shape [109], where the asymmetry is a result of resonant scattering and a background intensity. Fano resonances will be discussed in more detail in chapter 4. This has been a brief foray into GMR gratings, there

is a significant body of literature concerning these structures [87], [94], [110]–[115], and I have merely scratched the surface of what is necessary to understand this topic.

2.4.   Slab Waveguide

The bulk of this thesis is focused on electromagnetic waves propagating in optical waveguides, and the manipulation of the waveguide mode. It is with the former that I shall begin with. I have briefly discussed a 2D analysis of the effective index of a waveguide, now I will look at the derivation of an effective index method for the slab waveguide; a waveguide that is not only invariant in the direction of propagation, but also perpendicular to the direction of propagation, like that illustrated in Figure 2.8.
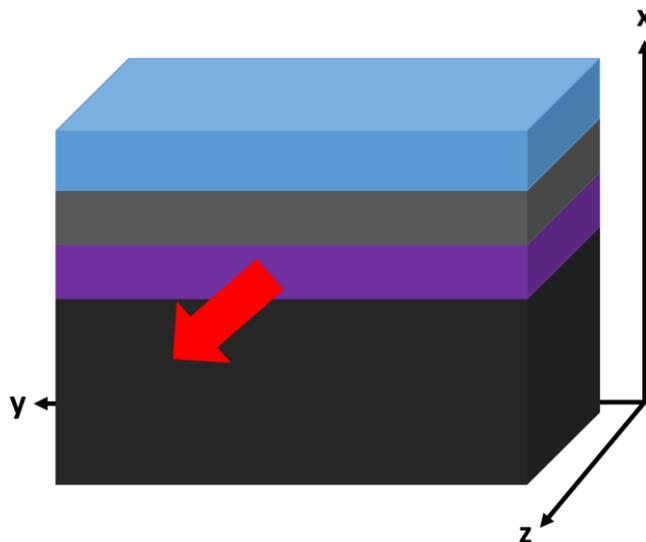
Figure 2.8: Illustration of slab waveguide, layer materials, permittivities, and dimensions are not important, note that the y scale is significantly larger than the wavelength of propagating light to ensure invariance in y. Propagation is in the z-direction, indicated by the red arrow.

In this instance, each layer in considered invariant in both the z- and y-direction compared to the supported waveguide mode. While this is an oversimplification, the analysis can be used to approximate the waveguide mode propagation through simple numerical analysis.

2.4.1.   Multi-Layer Slab Waveguide

The derivation of an analytical model for a multi-layer slab waveguide, as illustrated in Figure 2.8, begins with Maxwell's curl equations (Equation 2.6).

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}(\boldsymbol{r})e^{-i\beta z}, \qquad \boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}(\boldsymbol{r})e^{-i\beta z}$$

Equation 2.35

Substituting Equation 2.35 into the curl equations yields two sets of eigenmodes for the transverse electric (TE) and transverse magnetic (TM) orientations, where *e* and *h* are the electric and magnetic field components.

$$\frac{d^2 e_y(x)}{dx^2} + k_0^2 n^2(x) e_y(x) = \beta^2 e_y(x)$$

$$\frac{d}{dx}\left(\frac{1}{k_0^2 n^2(x)}\left(\frac{dh_y(x)}{dx}\right)\right) + h_y(x) = \frac{\beta^2}{k_0^2 n^2(x)} h_y(x)$$

Equation 2.36

By eliminating the x- and z-field components, the eigenmodes yield two second order differentials shown in Equation 2.36.

$$e_{y,i} = A_i e^{ik_{x,i}(x-a_i)} + B_i e^{-ik_{x,i}(x-a_i)}$$

$$k_{x,i} = \sqrt{k_0^2 n_i^2 - \beta^2}$$

Equation 2.37

However, because the multi-layer slab waveguide consists of several layers with piecewise refractive index (n_i), we can consider the k-vector piecewise ($k_{x,i}$) for a layer with thickness d_i. The result is the general solution shown in

Equation 2.37, where k_0 and β are the free-space k-vector and propagation constant, respectively.
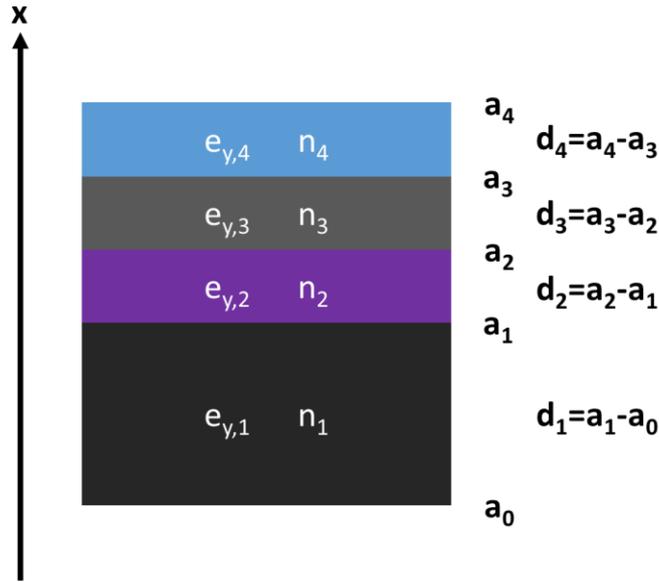
Figure 2.9: Multilayer slab waveguide with piecewise refractive index (n), electric field (e), and thickness (d).

A simplistic piecewise illustration of a multilayer slab waveguide is illustrated in Figure 2.9, where I have only detailed 4 layers, but one could easily extrapolate ad infinitum. The layer materials are, for now, not important, but are illustrated as a 2D cross-section of Equation 2.5.

$$\begin{bmatrix} A_i \\ B_i \end{bmatrix} = \frac{1}{2\alpha_i} \begin{bmatrix} (\alpha_i + \alpha_{i+1})e^{-\delta_{i+1}} & (\alpha_i - \alpha_{i+1})e^{\delta_{i+1}} \\ (\alpha_i - \alpha_{i+1})e^{-\delta_{i+1}} & (\alpha_i + \alpha_{i+1})e^{\delta_{i+1}} \end{bmatrix} \begin{bmatrix} A_{i+1} \\ B_{i+1} \end{bmatrix}$$

$$\alpha_i = ik_{x,i}, \qquad \delta_i = \alpha_i d_i$$

Equation 2.38

By applying appropriate boundary conditions,

$$\begin{cases} e_{y,i}(a_i) = e_{y,i+1}(a_i) \\ \left(\dfrac{de_{y,i}(a_i)}{dx} \right) = \left(\dfrac{de_{y,i+1}(a_i)}{dx} \right) \end{cases}$$

Equation 2.39

the piecewise eigenmode (

Equation 2.37) can be calculated using the matrix equation shown in

Equation 2.38 [116]–[118]. This method can be applied to analytical models and is often used by simulation techniques to yield the effective index (n_{eff}) of a multilayer waveguide. I will use

Equation 2.38 (section 3.4) to demonstrate the effective index of my waveguide modulators.

2.5.   Optical Modulation

As discussed in the introduction (section 1.3), there are two techniques that can be used to achieve optical modulation, i.e., phase or absorption. While the mechanism for modulation will be discussed in the next section of this chapter, it is necessary to discuss the design choices required for both these techniques. Absorption modulation can be achieved in a straight-through waveguide setup, i.e., a waveguide with an input, a modulator, and an output all in one path. Phase modulation, however, typically relies on the use of an interferometric setup, where the interference of a modulated and unmodulated signal creates a measured output. In this section I will discuss how this can be achieved, and the important considerations therein for the modulators presented in this work.

2.5.1.   Mach-Zehnder Interferometer

The Mach-Zehnder interferometer (MZI) is a particularly simple optical device used to split a light beam into two paths using a beam splitter and recombine under the same principle. While an MZI is typically considered with free-space optics, the principle can be extended to integrated optics as it relies on a very simple and well-known equation.

$$v = \frac{d}{t}, \qquad n = \frac{c}{v}, \qquad OPL = nd \qquad \text{\textit{Equation 2.40}}$$

In its simplest form, light travels through a vacuum with a velocity (c) which is given as a distance travelled (d) in a time (t). Extending our understanding of refractive index (n) from Equation 2.8, we know that this velocity scales (c/v) when light is travelling through a medium and, in optics, the distance travelled by light is referred to as the optical path length (OPL) which can be expressed as a function of the distance (d) as shown in Equation 2.40.
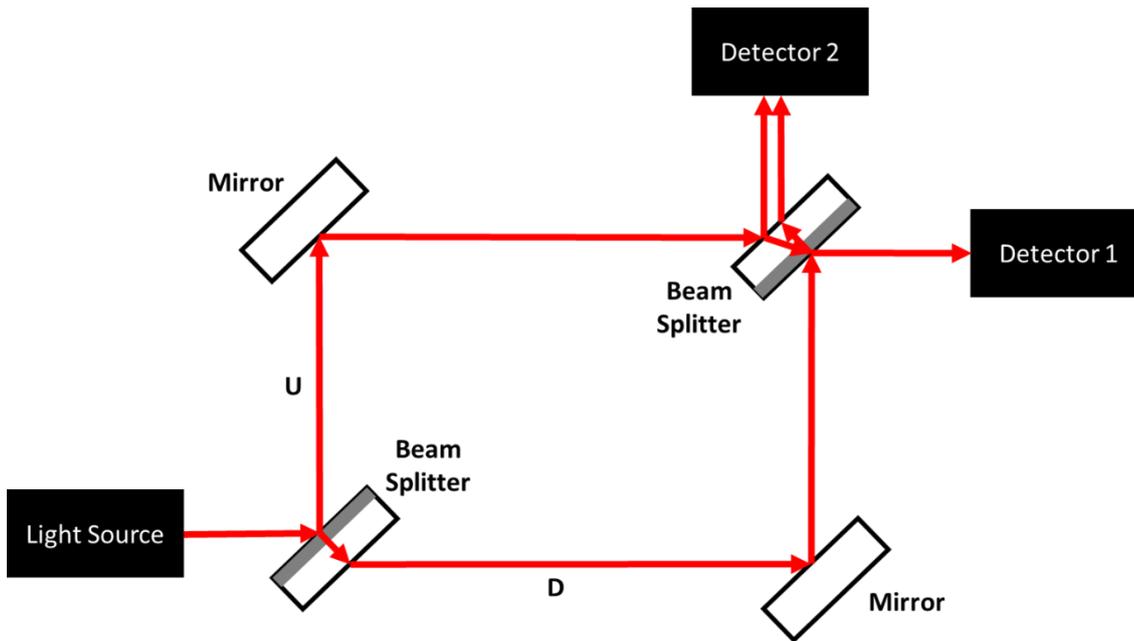
Figure 2.10: Mach-Zehnder Interferometer

As illustrated in Figure 2.10, an MZI splits a singular beam of light of wavelength (λ) into two paths (U and D) using a beam splitter. We can consider a phase (ϕ) change at each reflection equal to $\lambda/2$. At detector 1, paths U and D both experience the same number of reflections and distance travelled through the beam splitter, as such they are in phase, and we get constructive interference. At detector 2 there is an extra distance component for path D, and if we assume this to be $\lambda/2$, then there is destructive interference and no light is being detected [119].

For an ideal MZI, each reflection (and path length difference) contributes a $\lambda/2$ phase shift and results in perfect constructive and destructive interference. However, there is a great body of research [74], [75], [77], [120]–[122] describing the electrical tuning of the phase shift to produce a range of interference responses.
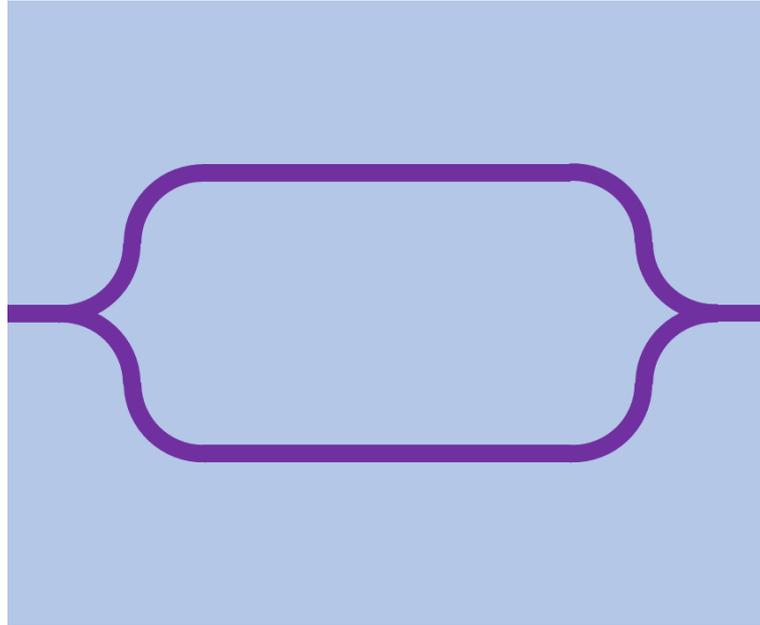
Figure 2.11: Mach-Zehnder interferometer waveguide platform.

In integrated optics, the phase shift can be achieved by changing the OPL, and the beam splitting technique can be realised via a multi-mode interferometer (MMI) or a Y-junction waveguide, as illustrated in Figure 2.11 where a single input is split into two waveguides and recombined to one output. In this instance, the optical path length is changed on the split waveguide paths. This design platform is discussed further in section 3.6.

2.5.2.   Phase Modulation

Given that a change (modulation) in phase is the desirable effect in optical modulation, we now need to consider how the optical path length can be modulated. A change in optical path length, or the difference (OPD), is the difference of the refractive index (n) and distance (d) product between two paths. The phase (ϕ) of the light is then given as:

$$OPD = n_1 d_1 - n_2 d_2, \qquad \phi = 2\pi \left(\frac{OPD}{\lambda} \right)$$

Equation 2.41

where λ is the wavelength of the input light. In integrated optics, a change in phase can be achieved as a change in the refractive index experienced by the optical mode. The resulting phase change can be calculated using Equation 2.41.
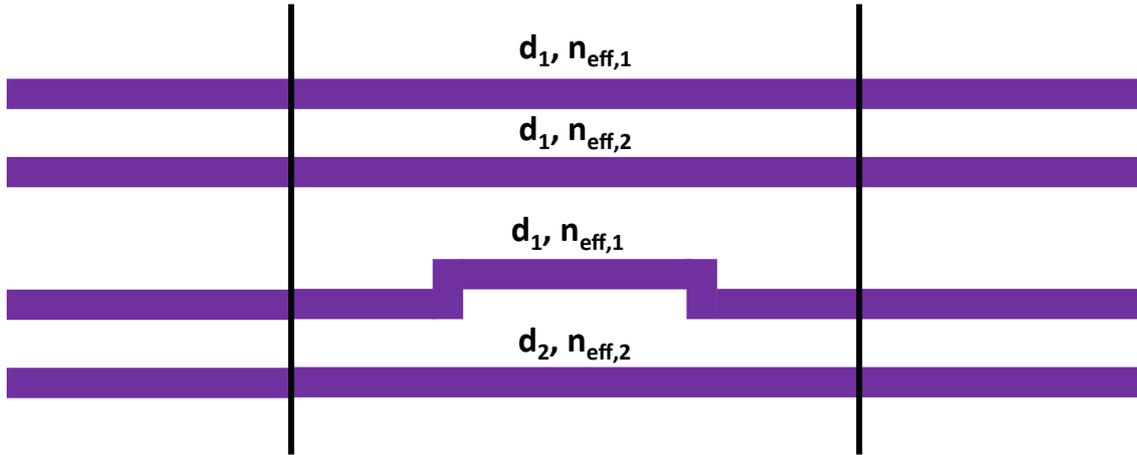
Figure 2.12: Illustration of phase shift methods in Mach-Zehnder modulator.

Physically, this can be represented as an increase in the distance travelled by the light, a change in the effective index of the waveguide mode, or a combination of both, as illustrated in Figure 2.12 where the system is identical on the left and right side of the black lines, and the optical path length only changes in the middle of the Mach-Zehnder interferometer.

$$E_{out} = E_{in}\left(a_1 e^{-i\phi_1} + a_2 e^{-i\phi_2}\right) \qquad \text{Equation 2.42}$$

$$T = \left|\frac{E_{out}}{E_{in}}\right|^2 = a_1^2 + a_2^2 + 2a_1 a_2 cos\Delta\phi \qquad \text{Equation 2.43}$$

The input and output of the interferometer is a singular waveguide, though it is not illustrated here. Using the principle of superposition, the output electric field (E_{out}) can be represented as shown in Equation 2.42, where E_{in} represents the input electric field, a_1 and a_2 are the amplitudes of the light signals in arms 1 and 2 (these are equal to 1 in a loss-less system), and ϕ_1 and ϕ_2 are the phases corresponding to the propagation through each arm, where the difference in the phases, $\Delta\phi$, can be used to calculate the intensity transmission (T), as shown in Equation 2.43. As the intensity is given as the cosine of the phase difference, as the two paths approach a π-phase difference, the intensity tends to a minimum.

2.5.3.   Absorption Modulation

So far, I have only considered a loss-less system, where the modulation is achieved through a change in the effective refractive index (n_{eff}). In absorption modulation, we consider a material with loss (α).

$$\boldsymbol{n} = n + ik, \qquad \alpha = \frac{4\pi k}{\lambda} \qquad \text{\textit{Equation 2.44}}$$

The complex refractive index (**n**) of a material is a function of the refractive index (n) and the extinction coefficient (k), with the latter responsible for the loss of a material, as shown in Equation 2.44. It follows that a waveguide mode experiences both an effective index and an effective loss (k_{eff}, α_{eff}), such that an input signal is modulated in intensity (I) as well as in phase.

$$I = I_0 e^{-\alpha_{eff} d} \qquad \text{\textit{Equation 2.45}}$$

The intensity of the output signal is given by the Beer-Lambert law [123] and is a function of the input intensity (I_0), the and the effective absorption length ($\alpha_{eff}d$) (Equation 2.45).

Figure 2.13: Illustration of absorption modulation concept.

Absorption modulation can be achieved, then, either as a stand-alone technique as illustrated in Figure 2.13, or by including absorption in the phase modulation calculation in Equation 2.41, taking the effective index as the complex effective index (n_{eff}).

2.6.   Phase Modulation

While it is possible to achieve phase modulation with a change in the path length, in this body of work I will focus on the electrical modulation of the refractive index component of the optical path length due to the integrated optics aspect of my work, in free space this could be achieved by changing the physical path length (d). As discussed in section 1.4 the electrical tuning of the refractive index can be achieved with transparent conductive oxides (TCOs), where the complex permittivity is controlled through the accumulation of carriers as described by the Drude Model. Of specific interest in this field is indium tin oxide (ITO), which has been shown to have unity order refractive index modulation [80] (in a very thin film) through the gated accumulation of

electrons using a metal-oxide-semiconductor (MOS) capacitor. Here, I will discuss the MOS capacitor and present a simplistic model of the achievable accumulated carrier concentrations.

2.6.1.   Conductivity in Semiconductors

The electrical conductivity of metals, semiconductors, and insulators can be described qualitatively in terms of their energy bands. Metals are characterised by very high conductivity, where the conduction band is either partially filled or overlaps with the valence band, i.e., there is no bandgap and electrons in the valence band can easily move to the conduction band where they are free to move around. This makes metals excellent conductors. In contrast, valence electrons in insulators form strong bonds between neighbouring atoms that are difficult to break. This yields a large bandgap and prevents conduction of currents even when the electrons are given a high thermal or electrical energy.

There are two factors that are affected by temperature in semiconductor materials, the mobility of the carriers and the number of carriers available, though the latter is more dependent on the doping level (the initial concentration of carriers). Semiconductors have a small energy bandgap, of the order of 1 eV at room temperature. At a temperature of 0K all the electrons are in the valence band and no conduction may take place, but at room temperature, the thermal energy represents a sufficient fraction of the bandgap energy such that the electrons can get excited into the conduction band. There are many empty states in the conduction band at room temperature, so a small applied potential can move some valence electrons into the conduction band, thus semiconductors may conduct a moderate current.

Temperature alone only accounts for a small increase in available carriers, doping a semiconductor is the easiest way to change the number of available carriers. A semiconductor that is doped with a group III element is labelled p-type and is labelled with n-type if the doping element is a group V element. P-type semiconductors have majority carriers as holes, and n-type have majority carriers as electrons. A degenerate semiconductor is doped to such a high level that is acts more like a conductor and a non-degenerate behaves more like an insulator. Indium tin oxide, as is discussed in the following sections, is a degenerate n-type semiconductor.

2.6.2.   The Ideal MOS Structure

The MOS structure is an important type of structure in the world of semiconductor devices, being at the core of applications such as metal-oxide-semiconductor field effect transistor (MOSFET) [5], [124]–[128], charge-coupled devices (CCDs) [129]–[132], and electro-optic modulators (EOMs) [74], [133], [134]. The latter is the focus of this work, where we use ITO as the semiconductor material. Given that the optical response of a Drude material under an applied electric field is mainly governed by the electronic response within the material, a comprehensive study of the MOS structure using ITO from an electrostatic perspective is required.

Fortunately, most of the MOS capacitor theory is identical to that of the MOS diode, which has been known for decades in the context of the MOSFET [135], [136]. The major consideration for applying the typical MOSFET analysis to a degenerate semiconductor is the space charge region width; for nondegenerate semiconductors, this width is on the order of 100 nm, but for degenerate semiconductors it shrinks to a much lower value (1-2 nm) [137].
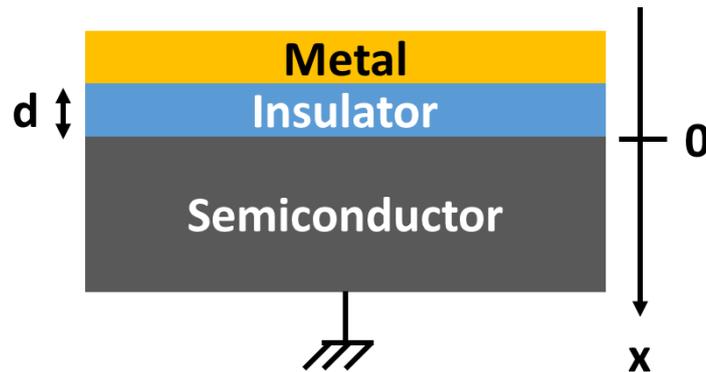
Figure 2.14: Cross-section of a metal-oxide-semiconductor structure.

An ideal MOS structure consists of three materials; a metal, an ideal insulator in which we can neglect charge accumulation and tunnelling, and a semiconductor, as illustrated in Figure 2.14. For this analysis, we assume that the insulator (oxide) has a thickness d, the semiconductor layer is connected to an Ohmic contact, and the voltage (V) is positive when the metal plate is positively biased with respect to the Ohmic contact.

[135]
$$e\phi_{ms} \equiv (e\phi_m - e\phi_s) = e\phi_m - e\chi + \frac{E_g}{2} + e\psi_B = 0$$
Equation 2.46

Furthermore, for the ideal MOS structure, the energy difference (ϕ_{ms}) is equal to the difference between the metal and the semiconductor work functions (ϕ_m and ϕ_s, respectively), and is equal

to zero at zero applied bias voltage. The work function of the semiconductor, as shown in Equation 2.46, can be calculated from the electron affinity (χ), the energy difference (ψ_B) between the Fermi level (E_f) and the intrinsic Fermi level (E_i), and the band gap energy (E_g), where the elementary charge (e) has its usual meaning. Essentially, the structure is in flat-band condition when there is no applied voltage.

2.6.3.   Space Charge Region

When the ideal MOS structure is biased, three cases may exist at the semiconductor surface: accumulation, depletion, and inversion. The boundary between accumulation and depletion is the flat-band voltage, and the boundary between depletion and inversion is the threshold voltage. The flat-band voltage is the voltage at which no band bending occurs and the energy bands are flat. Since the bands are flat, the electric field is zero everywhere and the hole concentration is the same as the acceptor concentration, and the charge density is zero.
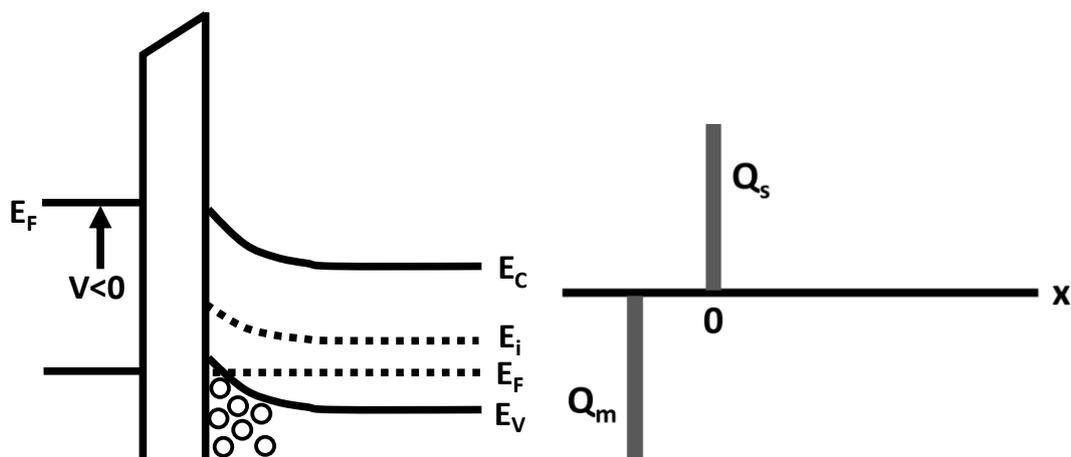
Figure 2.15: Energy band diagram and charge distribution of an ideal MOS capacitor in accumulation.

Accumulation: For the p-type semiconductor, a negative voltage at the metal causes the energy bands near the semiconductor surface to bend upward as illustrated in Figure 2.15. As no current flows in an ideal MOS structure, the Fermi level in the semiconductor remains constant. The bending of the bands causes an increase in the energy difference ($E_i - E_f$), thus yielding an enhanced concentration of majority carriers (holes for p-type).
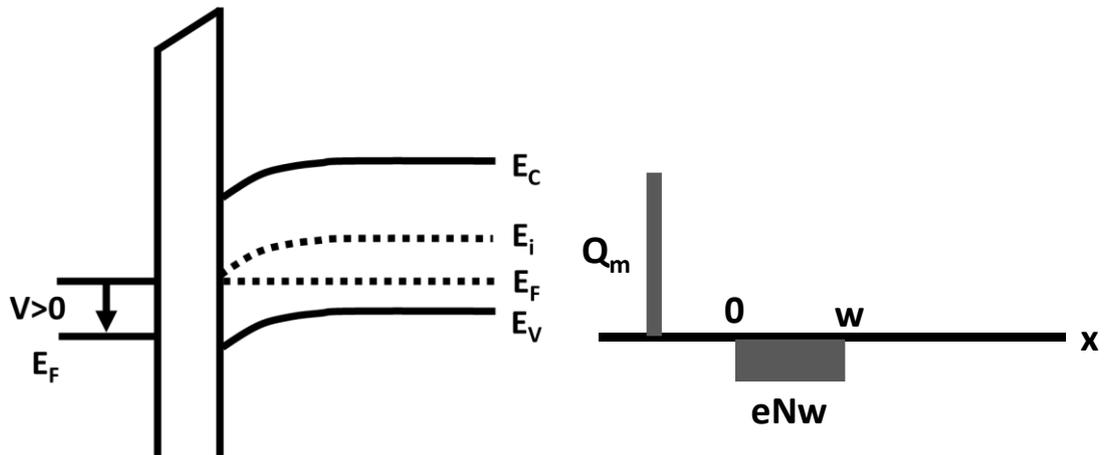
Figure 2.16: Energy band diagram and charge distribution of an ideal MOS capacitor in depletion.

Depletion: When a small positive voltage is applied at the metal, the energy bands near the p-type semiconductor surface bend downward as illustrated in Figure 2.16. This bending causes the majority carriers (holes for p-type) to be depleted.
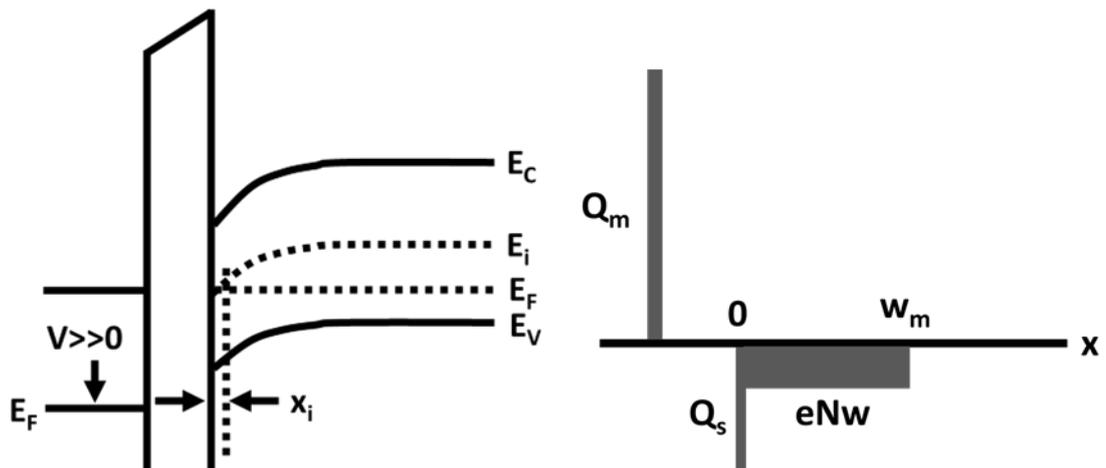
Figure 2.17: Energy band diagram and charge distribution of an ideal MOS capacitor in inversion.

Inversion: As the positive voltage applied to the metal increases, the energy bands bend further downward such that the intrinsic Fermi level crosses the Fermi level at the semiconductor surface as illustrated in Figure 2.17. The positive voltage begins to induce excess negative carriers (electrons) at the semiconductor-insulator interface. The electron concentration is then larger than the intrinsic carrier concentration (n_i) and the hole concentration is lower than n_i and the surface is inverted.

The above is a description of accumulation and depletion in a p-type semiconductor, but the principle can be applied to an n-type where the positive voltage induces an accumulation of

electrons, and a negative voltage induces the depletion of electrons. The space charge region description is applicable in both instances.

$$Q = eNw$$
<div align="right">Equation 2.47</div>

Once very strong inversion occurs, the space charge region reaches a maximum width as the bands are bent downward far enough that a small increase in band bending corresponds to a very small increase in space charge region width. As such, the charge (Q) per unit area in the semiconductor layer is proportional to the dopant density (N) and the space charge region width (w) as shown in Equation 2.47 [136].

2.6.4.   Capacitance

Capacitance is the measure of the ability of a device to store electric charge; the capacitance can be calculated either from the geometry of the conductive plates and the dielectric properties of the insulator, or from the charge stored as a function of applied electric voltage. We can quantify capacitance as the charge (q) per unit potential

$$C = \frac{q}{V}, \qquad C = \varepsilon\left(\frac{A}{d}\right)$$
<div align="right">Equation 2.48</div>

as shown in Equation 2.48 or calculate it geometrically using the material dielectric constant ($\varepsilon = \varepsilon_0\varepsilon_r$) and the area of the smallest electrical contact (A) and the distance between the two contacts (d). Given that the charge per unit area (Q) is given by

$$Q = \frac{qe}{A}$$
<div align="right">Equation 2.49</div>

we can substitute Equation 2.48 to express the charge per unit area, or 2-dimensional density (n_{2D}) as

$$Q = n_{2D} = \varepsilon_0\varepsilon_r\left(\frac{V}{de}\right)$$
<div align="right">Equation 2.50</div>

that we can use to determine the carrier concentration in ITO under an applied gate voltage [138]. The conversion to a three-dimensional carrier density is a simple division by the space charge region thickness (~a few nm). A more in-depth analysis can be found in Appendix section 9.2.

2.7. ITO for Electro-Optic Modulation

Mach-Zehnder modulators have been employed by many authors in the field of electro-optic phase modulation [70], [139]–[142]. Naturally, the figure of merit for such a device is given as a function of the achievable voltage-dependent phase shift (V_π), given as the voltage required to achieve a π-phase change, and the length of the modulator (L). This figure of merit has already been discussed in detail in section 1.3.2, but as I have discussed the mechanism for voltage induced carrier modulation and refractive index modulation in this chapter, here I will use our calculations with literature results.

2.7.1. Benchmarking Analytical Waveguide Model

I look to Sorger et al. [77] who demonstrated an ITO-based MZM on a silicon platform, operating at the telecommunication wavelength of 1550 nm. In their research, they were able to demonstrate a 0.52 Vmm figure of merit harnessing the refractive index modulation in a thin layer of ITO on top of one arm of the waveguide.
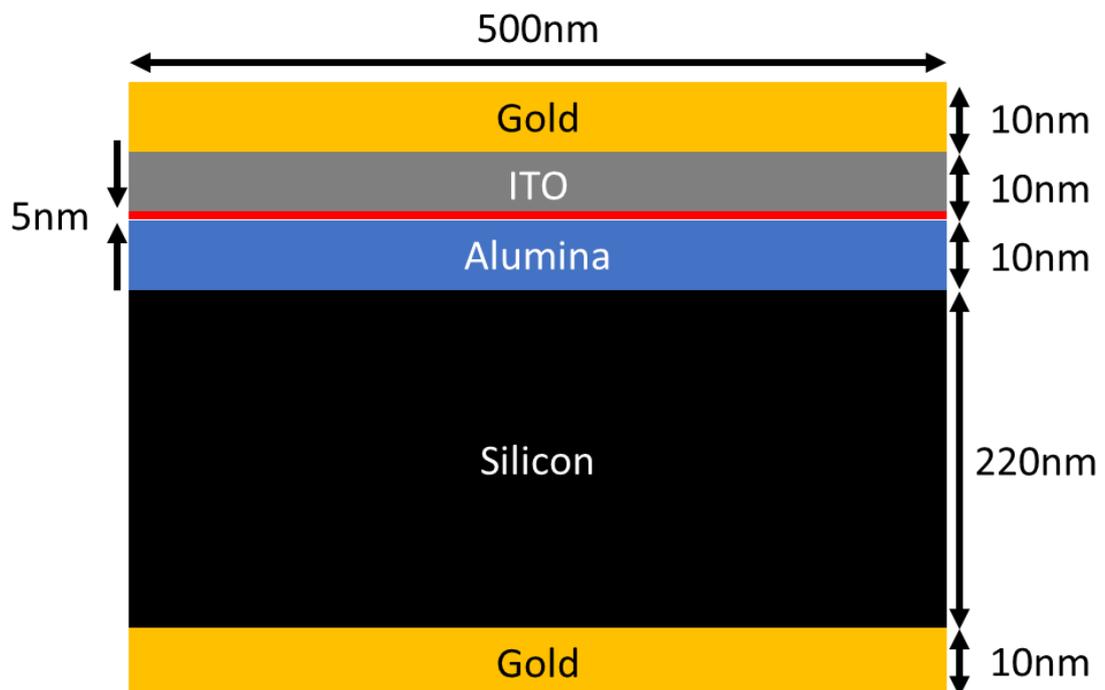
Figure 2.18: 2D schematic of the structure presented by Sorger et al.

I began by looking at the carrier (electron) concentration in the accumulation layer, and although the literature states an accumulation layer thickness of 1-2 nm, the authors here claim a 5 nm

layer. The structure presented by the authors is illustrated in Figure 2.18 and features a 220 nm thick silicon waveguide, connected to a gold electrical contact and a 10 nm insulating layer of aluminium oxide (alumina). The optical modulation is achieved through a 5 nm accumulation layer in 10 nm of ITO with a further gold electrical contact on top. The structure is 500nm wide and 32 μm long. Using Equation 2.50 and the relevant parameters, I was able to determine a voltage-dependent carrier density in the 5nm accumulation layer, shown by the red line in Figure 2.19, which naturally has a linear relationship.
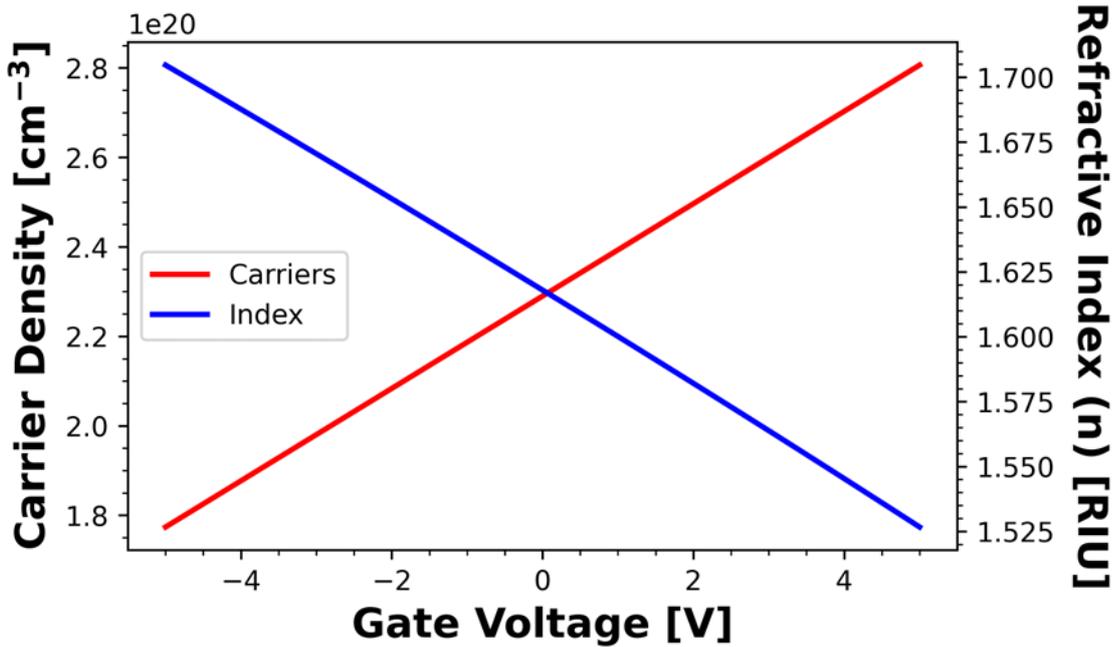
Figure 2.19: Local carrier density and refractive index within 5nm accumulation layer of ITO in [77].

The authors claim a free carrier density in their ITO of 2.29×10^{26} m^{-3} and a $\Delta n/\Delta V = 1 \times 10^{19}$ cm^{-3}V^{-1}. Our simplistic approximation agrees with their experimental results. Using this carrier density and Equation 2.31, I was then able to determine the material refractive index within the accumulation layer, leading to a local refractive index as shown in blue in Figure 2.19. Using the material indices and dimensions listed in [77], and the slab waveguide model in

Equation 2.38, I was then able to determine the effective index of the guided modes within the structure.
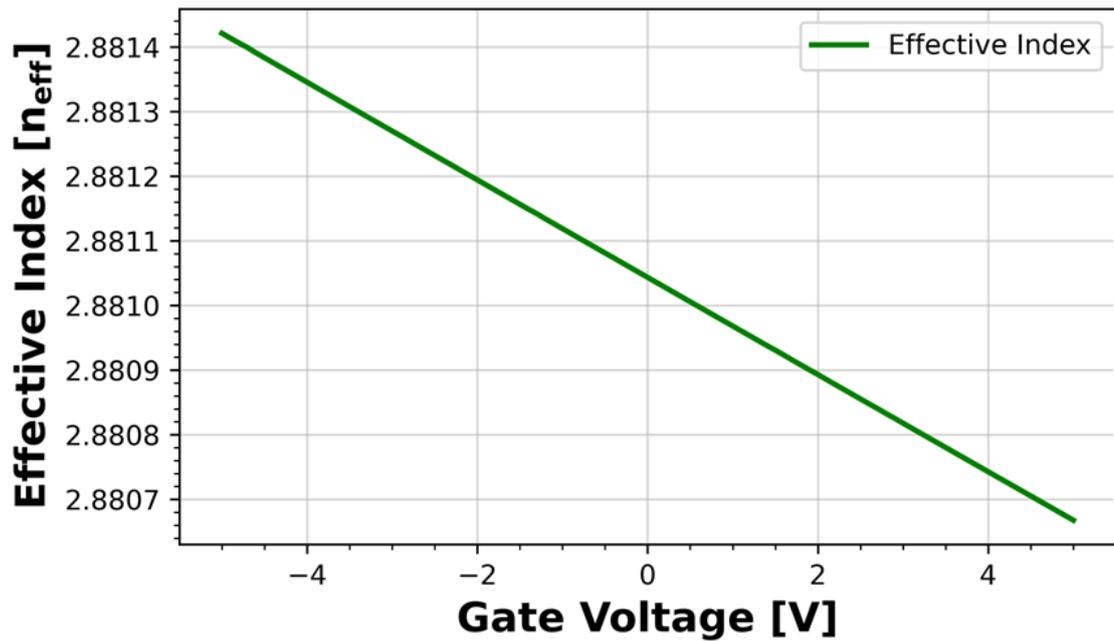
Figure 2.20: Effective index as a function of gate voltage for ITO-based modulation in [77] plotted as a function of the gate voltage.

The fundamental waveguide mode is plotted as a function of the gate voltage in Figure 2.20, and we can see that there is an approximate $\Delta n_{eff}/\Delta V$ relationship of 10^{-4} RIU/V. This value is two orders of magnitude smaller than the quoted value by the authors. Given that our model of electron accumulation predicts the carrier density accurately, a lot of these discrepancies can be explained by the simplicity of my slab waveguide model and the optimistic interpretation of the magnitude of the change in refractive index of ITO, of unity order, given by the authors.

2.7.2.   Discrepancies in the Literature

This observation leads on to an issue that will be addressed in section 4.3, which is to say that there are some discrepancies amongst the literature as to the optical properties of ITO, with many authors determining the complex refractive index using spectral ellipsometry, such as Sorger et al. [77] in the previous section, and then using these to explain the electrical properties. Some authors, like Feigenbaum [80], on the other hand use the electrical properties to determine the optical. In my brief study above, the latter technique is what I have used. I have discussed the common discrepancies in the literature surrounding ITO in more detail in section 4.3 and set this question as one of the main goals of my research.

The behaviour of the carriers in ITO is not the only question surrounding the behaviour of ITO MOS capacitors in the literature, with several authors reporting a higher modulation performance and efficiency from their modulators. It is commonplace for these authors to suggest that there are pinholes in their insulator layer, leading to "metal fingers" dangling through their insulator layer, thus decreasing the capacitor separation distance (d) and providing an increase in performance. One such example in [80] is detailed in the supporting information with little detail given. I therefore note that these papers provide unsubstantiated statements to explain that their models do not fit their experimental observations, which is highly unsatisfactory. As such, the examination of the elemental "leaking" into the insulator layer is a further question for my research.

3. Design and Simulation Methods

Chapter Synopsis

This chapter presents the material choices and structure designs as influenced by the research goals and simulations discussed herein. Benjamin Franklin once said, "if you fail to plan, you plan to fail", and with nanofabrication, this statement could not be more accurate. An overall concept is meaningless if the device will never work because the waveguide facet is too rough, the waveguide is too thin to support a guided mode, or there is an overlap between the metal contacts and thus no capacitance. Indeed, I was very fortunate to have access to the simulation and design techniques discussed in this chapter. Here I will discuss the choice of materials I will use in the modulator, the simulation techniques used in this research to sweep parameters and have an approximation of key dimensions, and the design considerations made allowing for fabrication and experimental tolerances quickly and inexpensively. Specifically, I will discuss the Fourier-Modal and Finite-Difference techniques used to design and implement the electro-optic modulators on which this thesis is built. I developed the Fourier-Modal and Finite-Difference simulations presented in this chapter over multiple years with inspiration from similar models available in and used regularly within the wider research group. With that in mind, I would like to acknowledge Dr Christopher Reardon, Dr George Duffett, Dr Alex Drayton, Dr Manuel Deckart, Dr Kezheng Li, and Samuel Blair for their assistance developing and refining these models.

3.1.   Choice of Materials

The electro-optic modulators in this body of work are planar, built on the idea that light may be guided and manipulated using an optical waveguide. As detailed in chapter 2, the effective index of the waveguide mode may be altered using a gated semiconductor with an electrically tuned permittivity, an illustration of this concept is shown in Figure 3.1.
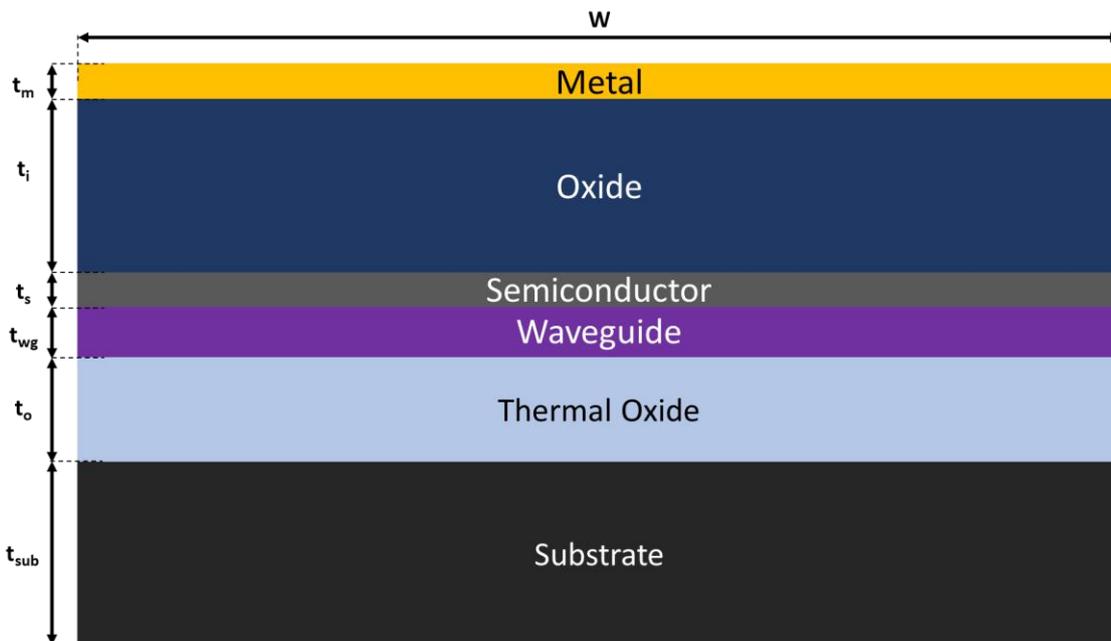
Figure 3.1: Schematic of proposed MOS EO Modulator where width (W) and thicknesses (t) are discussed in this chapter.

There are several choices for the materials that form the metal-oxide-semiconductor (MOS) sandwich illustrated in Figure 3.1. The choice of materials has been heavily influenced by the collective knowledge of the research group, fabrication tolerances and limitations, and overarching research aims. Here I will summarise and explain my choice for each layer.

3.1.1.   Waveguide

Most of the ITO-based modulators described in the literature utilise the silicon platform, due to its technological maturity and corresponding popularity for datacomms applications [74], [77], [143], [144]. We decided not to go down this route and to use a silicon nitride platform instead because a) silicon nitride is transparent into the visible regime, which is more attractive for the optical neural network application targeted here [145]; b) similarly, silicon nitride is able to carry

higher power signals than silicon [73], which is essential for coherent optical processing requiring 100s of channels; c) finally, alternatives for silicon modulators already exist based on carriers [70], [76], while there are no known high-speed modulators available on silicon nitride.

Silicon nitride (SI$_3$N$_4$) on borofloat 33 substrate is also the primary material used in the research group [111], [146], so there is a significant body of knowledge available on how best to process this material. The silicon nitride thickness used in the research group is 150 nm but, as will be discussed in this chapter, a new substrate material was required for this body of work, allowing me to choose a more suitable silicon nitride thickness. The chosen thickness was designed to increase mode overlap with the semiconductor layer (section 3.4.2).

3.1.2.   Substrate

Utilising Si$_3$N$_4$ on a borofloat 33 substrate has a major issue, i.e., that it does not readily form a cleaved facet (section 5.6.9). As the waveguide geometry we are targeting requires end-fire coupling through a facet, this is a major drawback. We initially tried to address this issue by partial dicing (section 5.5.1) the wafers through the back (non-coated side), but this proved difficult, with low yield. Instead, we changed the substrate to a bulk silicon wafer with a thinner layer of thermal oxide onto which the silicon nitride was deposited commercially, as illustrated in Figure 3.1.

Silicon is a cubic crystal structure with symmetry along the x, y, and z Cartesian coordinates, with a unit cell consisting of two face-centred-cubic (FCC) structures. The orientation of each face yields different chemical, mechanical, and optical properties, and is therefore an important parameter governing the material properties. Despite silicon having cubic symmetry, it can be cut along any direction, either parallel to one of its cubic faces, diagonally from one edge to another of the same cube, or diagonally through three corners of the cube [147], i.e., silicon cleaves readily and, as a result, I have been able to reliably create cleaved facets using a bulk silicon substrate.

Moreover, using a silicon platform increases sample strength during fabrication, where many of the steps require vacuum stages, thermal tape, or careful manipulation with tweezers, which all place mechanical pressure on the substrate material. Switching to a substrate that does not require partial dicing prior to fabrication will naturally increase the durability of the sample during fabrication.

3.1.3.   Semiconductor

This layer serves to modulate the optical mode by capacitively induced changes in carrier density. We chose indium tin oxide (ITO) as the material because its doping density can be controlled over a wide range, allowing for control over the refractive index; correspondingly, several efficient ITO-based modulators have already been described in the literature [74], [79], [80], which provides a good starting point for my studies. The ITO layer is described in much more detail in (chapter 4). Fabrication tolerances due to equipment limitations and design choices have limited the layer thickness, which is discussed in section 3.5.

3.1.4.   Oxide

The oxide, or insulator, layer in the capacitor has several requirements; a) it needs to be pinhole-free to avoid leakage currents while b) it needs to be thin such that a moderate voltage is sufficient to create a strong electric field and modulate the carrier distribution in the ITO. We chose SU-8, an epoxy-based photoresist, chosen for its chemically and thermally stable properties; its thickness can be controlled via spin speed and chemical dilution during application. Once exposed and cured, SU-8 exhibits optical and chemical behaviour similar to silicon dioxide (SiO_2) [148].

3.1.5.   Metal

For the metal, electrical contact, layer of the MOS structure, gold was the obvious choice because of its high conductivity, ease of processing, and resistance to corrosion. Since the structure is capacitive, it does not need to carry significant carrier density, so it does not need to be particularly thick; additionally, as discussed in section 5.4.2 an adhesion layer is required between the oxide and gold layers. An adhesion layer of chromium was chosen, due to its electrical compatibility with gold [149]; the gold and chromium layers are discussed in more detail in section 5.6.8.

3.2. Simulation Techniques

Creating a MOS EO modulator, as illustrated in Figure 3.1, requires development and optimisation; this would be nearly impossible to achieve through fabrication and measurement alone. Simulation techniques off a relatively fast method to sweep and optimise parameters by comparison. MOS EO modulators, as will be discussed in chapter 4, are not the only structures simulated in this body of work; I also simulated 2D periodic structures, that exhibit guided mode resonance (GMR). Computational electromagnetism (CEM) is used, which solves Maxwell's equations through discretisation in both space and time.

The time component of Maxwell's equations can be discretised by defining the electric field at discrete frequencies, using a finite-difference time-domain (FDTD) [150] or finite-difference frequency-domain (FDFD) [151] method. The space component can be discretised through a periodic permittivity distribution at discrete spatial frequencies in Fourier space, using the finite element method (FEM) [152] or the Fourier modal method (FMM) [153]. In my research I used a combination of FDTD and FMM to simulate and optimise the two structures detailed above, and I will discuss both methods in this chapter.

3.2.1. Fourier Modal Method

The Fourier modal method, also called rigorous coupled-wave analysis (RCWA) [104], [154], [155] or the scattering matrix method (SMM) [156], [157], is the ideal method for periodic structures of layers that are invariant normal to the periodicity. RCWA is the implementation of the time-harmonic spatial-frequency form of Maxwell's equations, through expansion of the electromagnetic fields within each layer into eigenmodes with an exponential dependence on the normal direction. Bloch's theorem [92] underpins RCWA, so electromagnetic waves in each layer are expanded to a truncated Fourier series and transmission, reflection, and field information are obtained by specifying the complex amplitude, wavevectors, and coupling coefficients of every Bloch eigenmode for each layer [158].

$$\nabla \times \boldsymbol{E} = i\omega\mu_0\mu\boldsymbol{H}$$
Equation 3.1

$$\nabla \times \boldsymbol{H} = -i\omega\varepsilon_0\varepsilon\boldsymbol{E}$$
Equation 3.2

For isotopic homogeneous nonmagnetic materials without source, the time-harmonic Maxwell curl equations with a time dependence of $e^{-i\omega t}$ can be represented by Equation 3.1 and Equation

3.2, where **E** and **H** are the three-dimensional electric and magnetic fields respectively, ω is the angular frequency, ε_0 and μ_0 are the vacuum permittivity and permeability respectively [159]. The relative permittivity and permeability (ε and μ) are material dependent spatially varying periodic functions and can be approximated by the truncated Fourier series. A simple distribution for a layer periodic in x and homogeneous in y and z is expressed in Equation 3.3 and Equation 3.4.

$$\varepsilon = \sum_{n=-N}^{N} \varepsilon_n e^{iGnx}$$

Equation 3.3

$$\mu = \sum_{n=-N}^{N} \mu_n e^{iGnx}$$

Equation 3.4

In Equation 3.3 and Equation 3.4, ε_n and μ_n are the Fourier expansion coefficients of material permittivity and permeability, and $G = \frac{2\pi}{a}$ represents the Bloch wavenumber, and a is the structure period. Equation 3.3 and Equation 3.4 define a unit cell with periodic boundary conditions, and N represents the finite number of Fourier harmonics (or truncation order) along the x-direction. The magnitude of N essentially determines the accuracy of the RCWA method.

Multilayer structures are defined by stacking layers in the z-direction with the same period in the x-direction, where each layer is represented by its own Fourier series (invariant in z). The individual layers are connected by defining boundary conditions to enforce continuity of the tangential electric and magnetic field components. The top and bottom layers are typically defined as quasi-infinite half-spaces to allow the reflected and transmitted waves to propagate in these layers [102].

$$\boldsymbol{E} = \sum_{n=-N}^{N} \boldsymbol{E}_n e^{i(k_x-nG)x}$$

Equation 3.5

$$\boldsymbol{H} = \sum_{n=-N}^{N} \boldsymbol{H}_n e^{i(k_x-nG)x}$$

Equation 3.6

Bloch's theorem yields the periodic electric and magnetic fields, respectively, in the x-direction, given by Equation 3.5 and Equation 3.6. **E$_n$** and **H$_n$** represent the complex amplitude vectors of the electric and magnetic fields of each Bloch eigenmode and k_x is the discrete in-plane wavevector determined by Bloch's theorem. The solving of Maxwell's curl equations using Equation 3.3 - Equation 3.6 requires algebra beyond the scope of this thesis, but can be found in peer-reviewed implementations of RCWA [104], [160]. I used an implementation of RCWA known as S^4, which is discussed in greater detail in section 3.2.1, to simulate 1D periodic

structures fabricated in the surface of sputter deposited thin films of indium tin oxide. This implementation, apart from being peer-reviewed, benefits from its open-source nature and easy implementation with coding techniques.

3.2.2.   Finite-Difference Methods

The finite-difference time-domain (FDTD) method can be classed in as a grid-based differential numerical modelling method (finite-difference method) and is often included as an extra feature in software options for finite-difference (FD) solvers. There are a number of software options to consider when using FD techniques, but of particular interest is Ansys Lumerical [161] and COMSOL Multiphysics [162], which are finely tuned FD implementations with simple user interfaces, a wide range of geometry and parameter options, FDTD solvers, and the potential for combination with other simulation techniques. FDTD is a time-domain computational electrodynamics technique used to solve Maxwell's equations in the frequency space [163].

COMSOL and Lumerical discretise the simulated geometry and permittivity in a grid, referred to as "Yee lattice" [163] or "mesh", defining finite elements in space over which Maxwell's equations can be applied. The mesh removes the need to solve simultaneous equations and therefore reduces the required computational power and processing time. The most common mesh is a conformal mesh [150], [164], which removes the need to solve layers simultaneously, reducing required power and processing time, but introduces a strict boundary on the spatial distribution to ensure numerical stability [165].

Both software options allow user-defined meshes and can generate a mesh automatically based on the user-defined layer dimensions and material parameters. The documentation states that the mesh is graded such that a material with refractive index 2 will have a mesh twice as small as a material with refractive index 1, because the wavelength is smaller in the higher index material. Using too few mesh elements reduces the precision of features that can be resolved, but forcing a very small spatial mesh increases the computation time by a factor dx^4 for 3D simulations due to the increase in frequency spacing.

There are further limitations to consider, namely that it is impossible to recreate the fine details and imperfections created in fabricated structures. As such, material roughness, layer misalignment, and material thickness distribution across the layer cannot be considered in the simulation domain. Therefore, the simulations presented in this work form part of the description of a device's performance, guiding the fabrication and providing context for the

experimental results, and were optimised for speed of output and manipulation of input parameters rather than outright precision.

3.3.   Grating Parameter Investigation

In section 2.3 I discussed the behaviour of guided mode resonance (GMR) in periodic structures that are used in this body of work to determine the optical parameters of our ITO. To further understand the resonance behaviour, I investigated how the resonant wavelength (λ_R) was affected by the grating parameters. I used RCWA to simulate an ITO grating on a borosilicate (n_s=1.46) substrate, varying each parameter (period, fill-factor, thickness, refractive index) individually and maintaining the others as constants. The base parameters were as follows: grating period (a) of 500 nm, grating thickness (t) of 150 nm, fill factor (ff) of 0.7 where the ridge width is given as ff*a, and refractive index (n_g) of 2.0. The resulting spectrum was measured in reflectance and plotted as a function of the varied parameter and incident wavelength. Only the transverse electric (TE) mode was measured. Figure 3.2 shows the results of this investigation, where the colour bar represents the normalised reflection.

Figure 3.2(a) shows the reflectance spectrum as a function of the grating period with a clear linear dependence between the resonance position and the period as indicated by the diagonal yellow line starting at point (220, 200) on the plot. This relationship is expected from the grating theory (2.3.2), though the plot does highlight the limited resonance behaviour of a low index grating. Figure 3.2(b) shows the relationship between fill factor and resonance for a fixed grating period, where the FF can be considered a representation of the effective index (n_{eff}) from (Equation 2.34). The plot shows that the resonance position remains constant across a large FF range, suggesting that once an effective index to sustain a guided mode is achieved, the mode remains invariant across a large ridge width range and is instead mainly dependent on period.
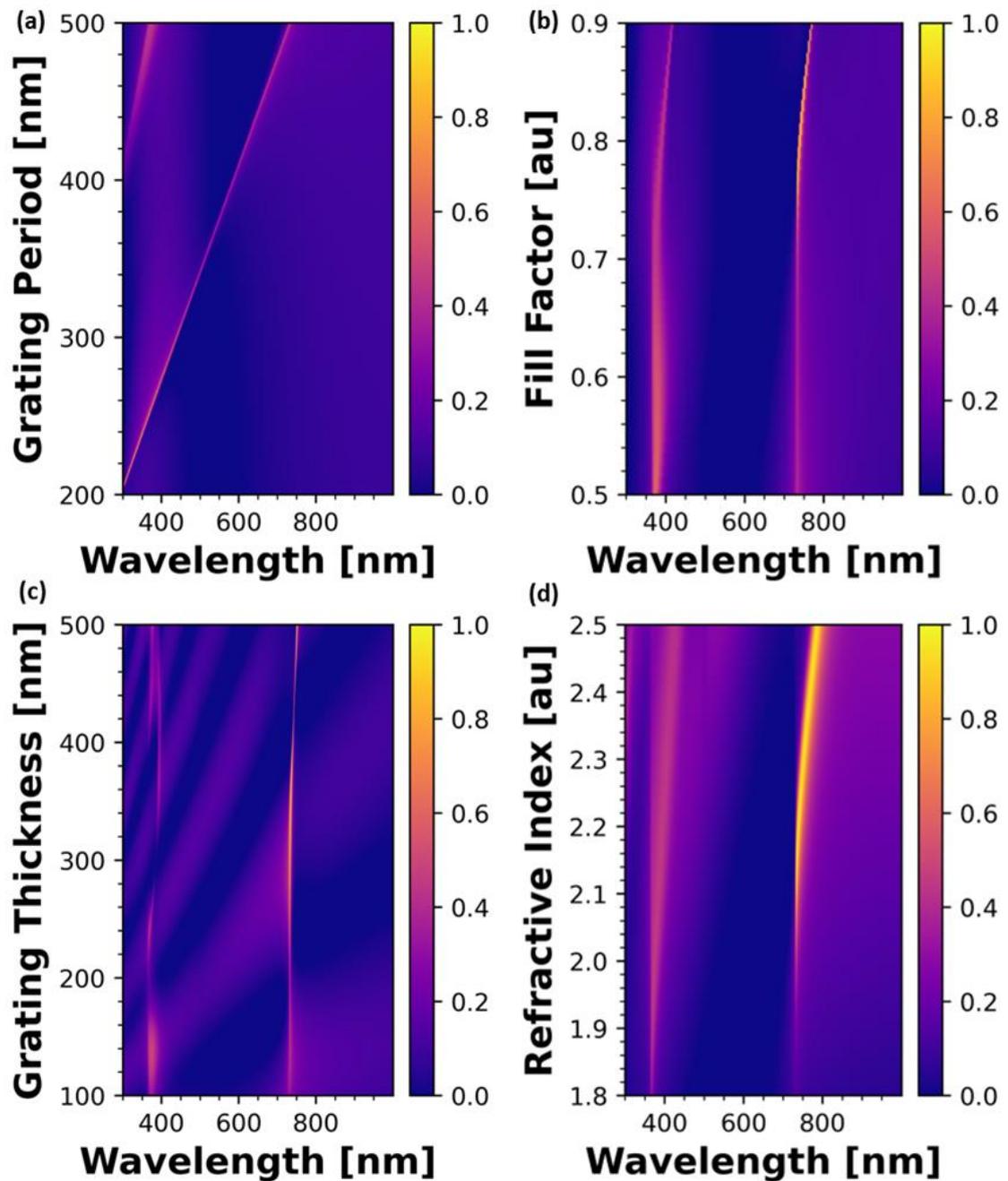
Figure 3.2: Reflectance surface plots with incident wavelength on the x-axis and (a) grating period, (b) fill factor, (c) grating thickness, and (d) grating refractive index on the y-axis. Reflectance is given by the colour bar. Simulations conducted using RCWA (S^4) software. The base parameters were a period of 500 nm, grating thickness of 150 nm, fill factor of 0.7, and a refractive index of 2.0 for the grating layer.

The effective index argument for the fill factor variation is applicable to the grating thickness and refractive index of the grating, as shown in Figure 3.2(c) and (d) respectively. Above a particular grating thickness, the grating can sustain a guided mode, above and below this thickness the grating operates in sub-λ or diffractive mode respectively. Lastly, the relationship between the grating index and the resonance position behaves similarly. One would expect that

93

the grating index is largely responsible for the effective index, and therefore increasing this index would cause a corresponding shift in the resonance wavelength. Indeed, this relationship is beginning to occur for higher grating indices, shown by the curve in the resonance response beyond n = 2.2. The relationship is clearly nonlinear and is an important one to recognise when designing these gratings.

The results of this investigation, and the relationships observed in Figure 3.2, were used to ensure that there were GMRs in the ITO gratings prior to the fabrication process discussed in chapter 4. The RCWA method to produce the reflectance surface plots can be reverse to determine the grating parameters, and this is discussed in more detail section 4.6.

3.4.   Mode Overlap Investigation

To determine the ideal layer thicknesses and waveguide width illustrated in Figure 3.1, I used a Lumerical FDTD solver to understand the layer-wise overlap of an optical mode in the silicon nitride waveguide. Here, I will briefly discuss the simulation parameters and how the data was processed to determine these layer properties.

3.4.1.   Model Parameters

The model was setup using the FDTD solver in Lumerical and consists of a silicon nitride waveguide on a glass substrate, an ITO layer, an SU-8 layer, and a gold layer, where each layer's effect on the waveguide mode can be analysed individually. As the model serves as a guide, I used the automatic mesh generation option, which created a mesh range between 10-30 nm across the structures, narrowing at thinner layers. The structure was considered to be a slab waveguide with the layer refractive indices taken as literature values for glass [166], silicon nitride [84], ITO [167], SU-8 [96], and gold [101]. The cover layer was assumed to be air, with a refractive index of 1. The incident wavelength was 509 nm.
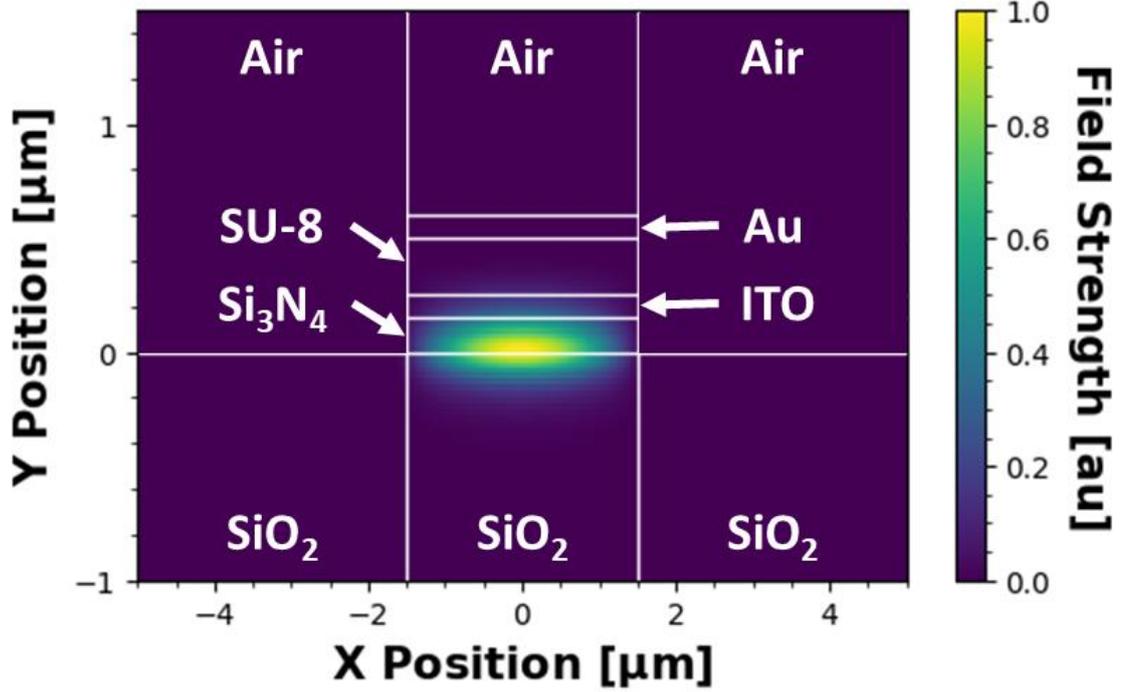
Figure 3.3: Cross-section of slab waveguide with metal-oxide-semiconductor structure surrounded by air.

The individual layers are separated by white lines and labelled in Figure 3.3 where the optical mode is predominantly in the waveguide layer and the slab waveguide is considered between -1.5 and 1.5 μm in the x-direction. The substrate layer is considered infinite from y=0 down. The FDTD frame sets the boundary conditions of 10x5 μm, centred around the waveguide. It is within this window that the electric field will be evaluated. The layer dimensions are evaluated individually, with the base thicknesses and widths as follows: W = 3 μm, t_{wg} = 150 nm, t_{ito} = 100 nm, t_{SU-8} = 500 nm, and t_g = 100 nm for width, waveguide thickness, ITO thickness, SU-8 thickness, and gold thickness respectively.

$$E_n = \sqrt{E_x^2 + E_y^2 + E_z^2}$$

Equation 3.7

The normalised electric field is shown in Figure 3.3 and is used to calculate the mode overlap throughout this section. It can be calculated as the sum of the squared electric field components, using Equation 3.7. The effect of each layer is evaluated individually by separating the structure into layers. In Figure 3.3, the top box and both sides are considered air, while the bottom box is the substrate layer. In the slab waveguide in the centre, the waveguide is the bottom with ITO, SU-8, and gold layers on top.

$$S_i = \frac{\sum_{i=1}^{N} E_{n,i}}{A_i}, \qquad O_i = \frac{S_i}{S_T}$$

Equation 3.8

The sum of the normalised electric field ($E_{n,i}$) divided by the layer area (A_i) yields a normalised sum of the electric field (S_i), that can be used to calculate the layer wise mode overlap (O_i) as a fraction of the total electric field sum (S_T) using Equation 3.8. I have only considered the mode overlap with the substrate, waveguide, and ITO layer as desirable overlap, i.e., the substrate overlap is undesirable as it can cause scattering losses, the waveguide overlap is desirable for guiding the input signal, and ITO overlap for signal modulation. Any other overlap is considered not relevant. It should be noted, only the fundamental mode was considered.

3.4.2.   Waveguide Layer

I will begin with the waveguide layer dimensions. Here the capacitor layer thicknesses remain constant while the slab width and waveguide thickness are varied separately. The mode overlaps are shown in Figure 3.4(a) and (b) for the slab width and waveguide thickness, respectively. It should be noted that the waveguiding layer is silicon nitride, as discussed in section 0.
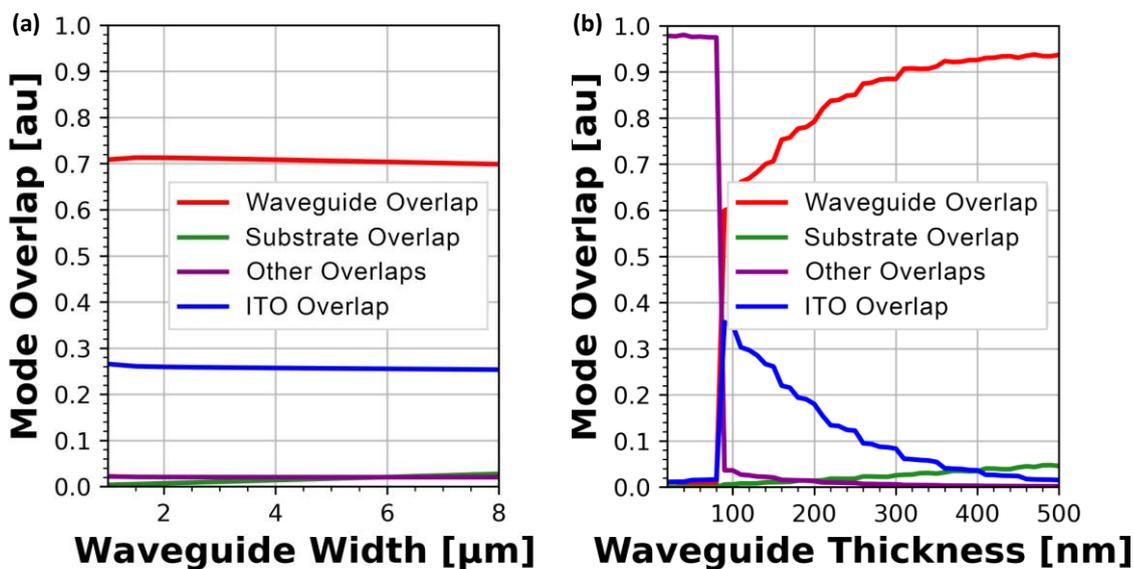
Figure 3.4: (a) Mode overlap as a function of slab waveguide width and (b) mode overlap as a function of waveguide layer thickness.

It is clear from the red and blue lines in Figure 3.4(a), that the mode overlaps in the waveguide and ITO layers, respectively, are not heavily influenced by the slab thickness. There is a minor increase around 2-3 µm, but the difference is minimal in an approximation such as this. Therefore, I can set my target waveguide thickness to 3 µm and afford some tolerance in coupling into the waveguide and in the lithography steps, discussed in section 7.1.2 and 5.6.5. As for Figure 3.4(b), it is intuitive that as the waveguide layer thickness increases, more of the

optical mode is in this layer and the inverse is true for the ITO layer. Therefore, it is a balancing act between the two to find an optical overlap for modulation. I opted for a target waveguide thickness of 150-200 nm, allowing some tolerance in the fabrication steps whilst maintaining a waveguide and ITO overlap around 70-80% and 25-20% respectively. It should be noted, below 100 nm waveguide thickness, the mode is not guided and is largely absorbed by the gold layer.

3.4.3.   Capacitor Layers

Finally, I will consider the individual layers in the capacitor structure. For each of these layers, the waveguide dimensions remained constant, and only one layer thickness was varied at a time. The mode overlaps are shown in Figure 3.5(a), (b), and (c) for the ITO, SU-8, and gold layers respectively.

Once again, we are comparing the red (waveguide) and blue (ITO) lines in Figure 3.5. For the ITO layer thickness (a), the mode overlap decreases to a minimum as the layer thickness increases. This is likely due to good confinement in the waveguide layer and a non-significant refractive index contrast between the silicon nitride and ITO preventing more of the mode being drawn into the ITO layer. Hence, we can afford to reduce the layer thickness in the ITO layer to increase the overlap. However, achieving a thin layer (<50 nm) using the deposition techniques available is non-trivial. As such, I have assumed a target thickness between 50-100 nm to achieve an ITO layer overlap of 25-20%. While overlap with the ITO layer is not the target in this work, but instead an overlap with a small portion at the ITO-SU-8 interface, measuring the overlap with a few nm of material is not feasible with the available mesh size of the simulation. The overlap with the ITO layer is sufficient when considering device dimensions.
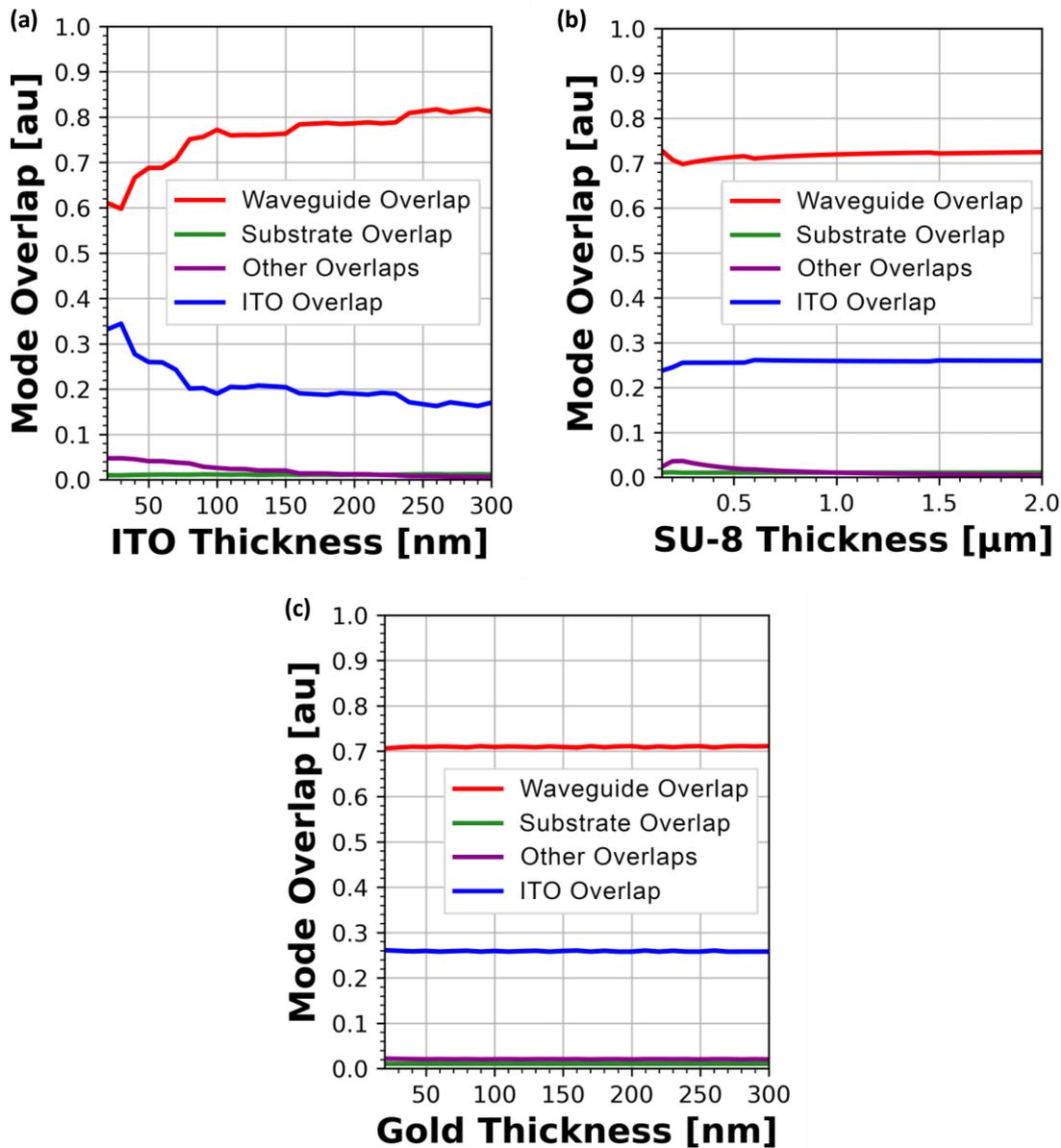
Figure 3.5: Mode overlap as a function of (a) indium tin oxide layer, (b) SU-8 layer, and (c) gold layer thickness.

The remaining capacitor layers of SU-8 and gold, (b) and (c) respectively, highlight that there is good confinement of the optical mode in both the waveguide and ITO layers, with the thickness of both these layers not having a dramatic effect on the mode overlap. This is good, as I can tune the SU-8 thickness to achieve a target capacitance (section 6.1.2), and therefore a modulation speed/depth, and I can increase the gold layer thickness to increase durability when making electrical contact. To begin with, I assumed an SU-8 thickness of around 200-500 nm and a gold thickness of 100 nm because a thinner SU-8 layer brings the gold layer closer to the waveguide mode and results in increased losses.

3.5. Modulator Design

To this point, I have considered the modulator as a slab waveguide, but this poses difficulties experimentally; namely, how would one make electrical contact to the ITO layer. To make the modulator practical and be able to place electrical contacts away from the waveguide, I needed a redesign. Here I will discuss these changes and their potential effects on the modulator performance.

3.5.1. Capacitor Design

In principle, there are two design layouts that would allow the electrical contacts to be taken away from the sensitive waveguide structure. These two layouts are illustrated in Figure 3.6, with (a) a surrounded and (b) a partial surround design.
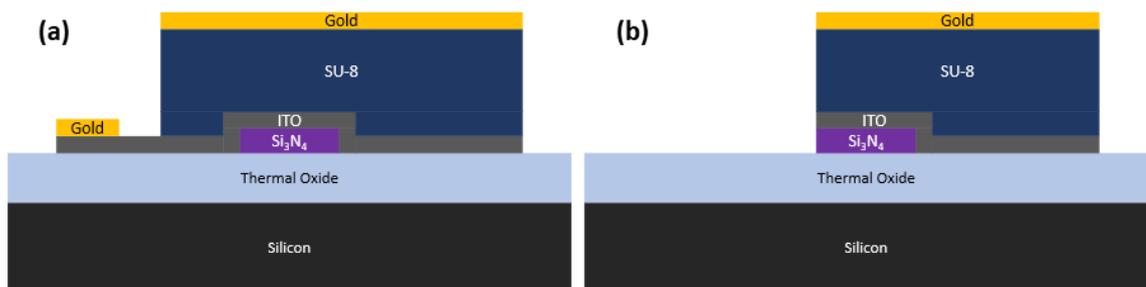
Figure 3.6: 2D schematic of a (a) fully surrounded and (b) partially surrounded modulator design.

Due to the good horizontal confinement of the waveguide mode observed in Figure 3.3, and based on the preliminary simulations shown in Appendix section 9.3, it is unlikely that either the fully surrounded or partially surrounded designs will alter the waveguide mode effective index due to the majority of the unconfined portion of the mode being vertically spread. To confirm that this is indeed the case, I used the effective index method discussed in section 2.4.1 to determine the waveguide effective index as a function of the three parameters that I showed to have the largest effect on the waveguide mode in Figure 3.5; the waveguide thickness, waveguide width, and the ITO layer thickness. The same simulation conditions were used here as were used in section 3.4, i.e., the same refractive indices, layer thicknesses, and widths where appropriate. I only varied the waveguide thickness, width, and ITO layer thickness individually, and measured the waveguide mode effective index. The results are shown in Figure 3.7.
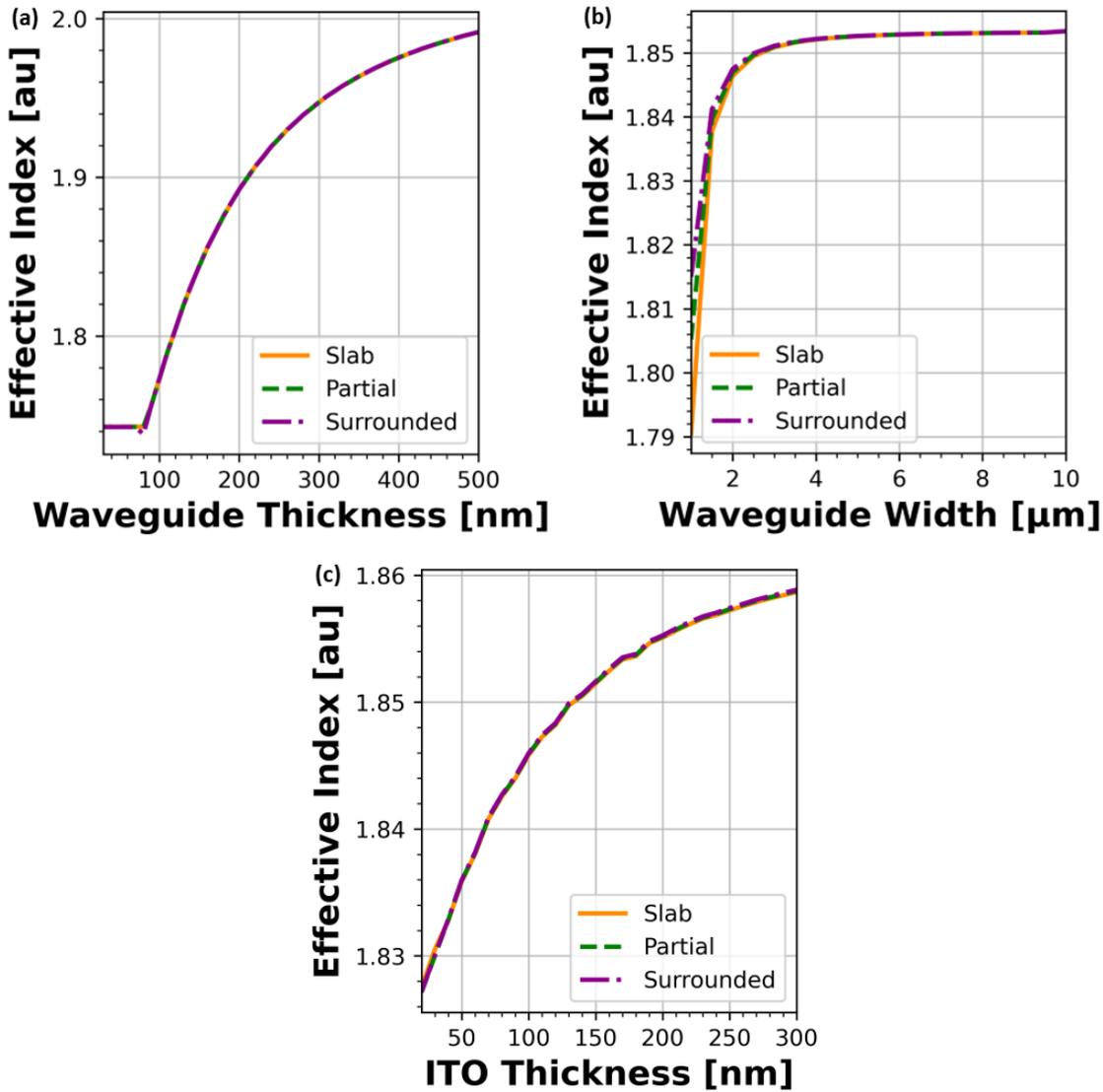
Figure 3.7: Waveguide mode effective index for the slab (orange), partial (green), and surrounded (purple) waveguide designs as a function of (a) waveguide thickness, (b) waveguide width, and (c) indium tin oxide layer thickness measured using the effective index method in finite-difference time-domain simulation in Lumerical.

The effective index method confirms the idea that the waveguide mode is well confined, as can be observed by equal effective index over the (a) waveguide thickness and (c) ITO thickness ranges in Figure 3.7. The horizontal confinement slightly deviates below 1 µm width, as is apparent from the splitting of the three lines in Figure 3.7(b). This is consistent with the mode overlap experiment and is not of concern for the waveguides discussed in this work. Hence, the design of the modulator can be, to some extent, guided by whatever makes the fabrication process easier. I chose to use the design illustrated in Figure 3.6(a).

3.5.2. Electrical Contact Design

It is common in the literature [74], [77], [78] to see phase modulator designs with the MOS capacitor as illustrated in Figure 3.6(b), with the bottom electrical contact on the unmodulated waveguide arm. For the modulators discussed in this work, the ITO layer acts as the bottom contact, and as such needs to have an open area for electrical contact. However, I chose to separate the electrical contacts to prevent the build-up of an electric field between the two contacts and concentrate the electric field between the gold and ITO layers in the capacitor.
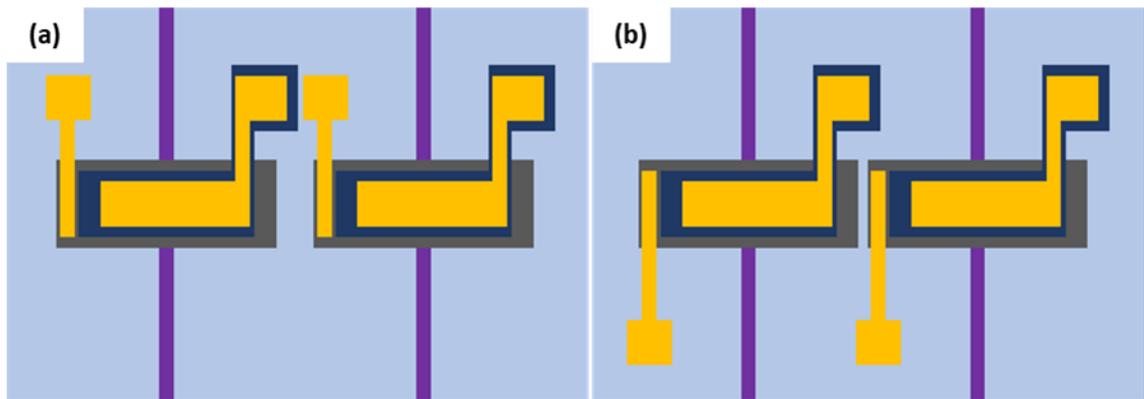
Figure 3.8: Top-down illustration of the (a) same level and (b) staggered top and bottom electrical contacts with arrayed modulators.

Initially, I used the contact layout as shown in Figure 3.8(a), where both the top (right) and bottom (left) contacts were equal but contacted each capacitor plate separately. I included an insulating layer of SU-8 between the top contact and the thermal oxide substrate to prevent any chance that a large potential could create a strong electric field between the two contacts. Fundamentally, the contact layout illustrated in Figure 3.8(a) is good enough for a single-waveguide modulator, i.e., an absorption modulator.

However, once I began measuring phase modulation, with a Mach-Zehnder Modulator (MZM) setup as discussed in section 2.5.1, I used the layout illustrated in Figure 3.8(b) to allow for a tighter packing of modulators when creating arrays of the MZMs. With the overall goal of these modulators to be as small as possible, as it is with the field effect transistor in conventional computing systems, a tighter packing of the modulators is favourable. Placing two modulators side by side, as illustrated in Figure 3.8, is significantly easier with the staggered top and bottom contact (b) than it is with contacts on the same level (a). Hence, I decided to use the staggered contacts illustrated in (b).

3.6.   MMI and MZM Designs

Finally, the waveguide design itself should be considered. The design is based on a Mach-Zehnder modulator (MZM). As discussed in section 2.5.1 there are two options when it comes to splitting or recombining one signal to two or vice versa: the multi-mode interferometer and the Y-junction waveguide. Here I will discuss the design challenges facing both options and justify my choice.

3.6.1.   Multi-Mode Interferometer

The multi-mode interferometer (MMI) is based on the principle of multimode propagation (MMP) of waveguides. The principle is derived from hollow core waveguide systems but can be applied to modes that are confined in a solid-core waveguide such as the silicon nitride waveguides in this research.
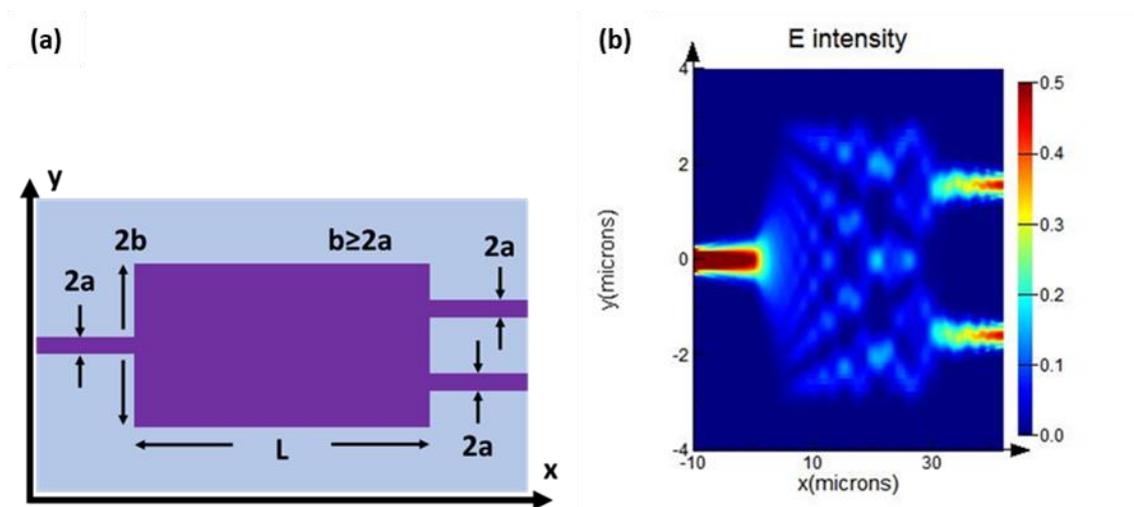
Figure 3.9: (a) Schematic of a one-to-two multimode propagation splitter in solid-core waveguide technology. (b) Electric field profile of a one-to-two multimode propagation splitter [161].

Consider a single-input system with a square waveguide of width and thickness 2a. The waveguide is fed symmetrically (around y=0) into a rectangular waveguide of width, thickness, and length 2b, 2a, and L respectively.

As the single mode input enters the larger section, of width 2b, it excites the higher order modes that have now become available due to the increased dimensions. These modes propagate and interfere to reform the original single mode input after a distance 2L, so the MMI can image the

input after a certain propagation length. For half this distance, the length L chosen here, the system produces a split image of the input, i.e., the intensity is equally distributed over two tightly confined spots. It is this property which makes the MMI useful as a 1-to-2 splitter, so the input signal is then split into two output square waveguides of width and thickness 2a. Such a system is illustrated in Figure 3.9(a), where the input signal travels from left to right in the positive x-direction. The electric field profile, shown in Figure 3.9(b), is described by:

$$\boldsymbol{E(x)} = \Sigma E_m e^{ikx}$$

<div align="right">*Equation 3.9*</div>

We can assume that input signal excites *m* modes as indicated by Equation 3.9, where x is the position at which we take the electric field, and m is the mode order. We ideally want the distance x=L at which the two output waveguides have a fundamental mode excitation.

$$L = \frac{4b^2}{s\lambda}$$

<div align="right">*Equation 3.10*</div>

In general, if we assume that b ≥ sa, where s in the number of splits in the system, and that the wavelength in the core of the rectangular waveguides is λ, then the first point at which the two output waveguides experience fundamental mode excitation is given by Equation 3.10.

$$y_s = -b + \left(\frac{b}{s}\right)[2(n-1)+1], \qquad n = 1, 2, 3, \ldots, s$$

<div align="right">*Equation 3.11*</div>

At the point L, the position of the two output waveguides (y_s) are given by Equation 3.11 as demonstrated by Jenkins et al. [168]. The problem with the MMI splitter for my research is that, as can be seen in Equation 3.10, there is a wavelength dependency on the dimensions of this structure. This wavelength dependency affects both the length and width of the rectangular guide, which in turn determines the position of the output waveguides. In my research, I wanted to demonstrate the power of a silicon nitride platform by showing modulation in multiple wavelengths. Such a task would then require individually optimised MMIs for each desired wavelength, thus removing the comparability of the results as each system may have its own defects/imperfections from the fabrication process. As such, I decided to use a different type of splitter.

3.6.2.   Y-Junction Waveguide

A far simpler and more broadband structure, but one that does not guarantee an equal division of the input signal, is the Y-junction waveguide, illustrated in Figure 3.10. The Y-junction is characterised by a one-to-two split where the split angle (θ) is the important parameter to consider. To keep the modulators in this research more broadband, I chose to use a Y-junction splitter. A Y-junction is more broadband than an MMI due to the dimension/wavelength relationship used to achieve an equal split of the input signal.
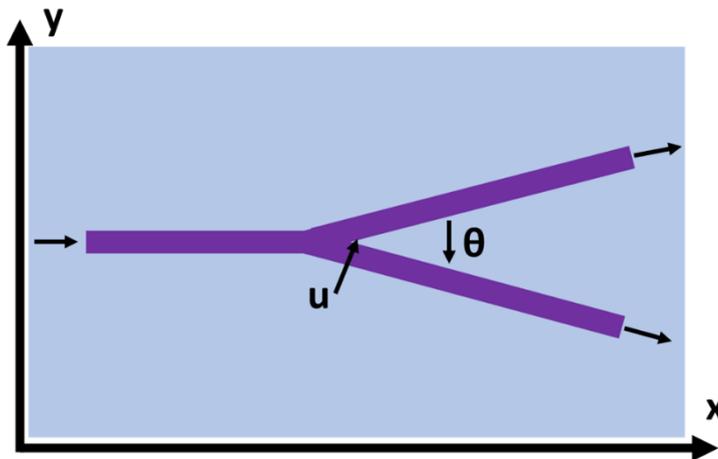
Figure 3.10: Optical power splitter with Y-junction waveguides.

In practice, the Y-junction waveguide is polarisation and wavelength independent, but the split angle should be very small to reduce scattering at the point (u). This means that Y-junction waveguides often suffer from longer device length than MMIs [169]. To achieve a smooth transition, we used a rounded point (u) in the split, specifically designed to provide a smooth transition between the single input and multiple output.

However, as can be seen by the angle of the two output waveguides in Figure 3.10, the recombination the two signals must at some point require a bend. In fact, I included a bend at the input end to ensure straight waveguides for the modulated section (for ease of fabrication), and thus created a symmetrical input-output MZM design. The bend geometry is illustrated in Figure 3.11, where the sharpness of the bend is exaggerated for illustrative purposes.
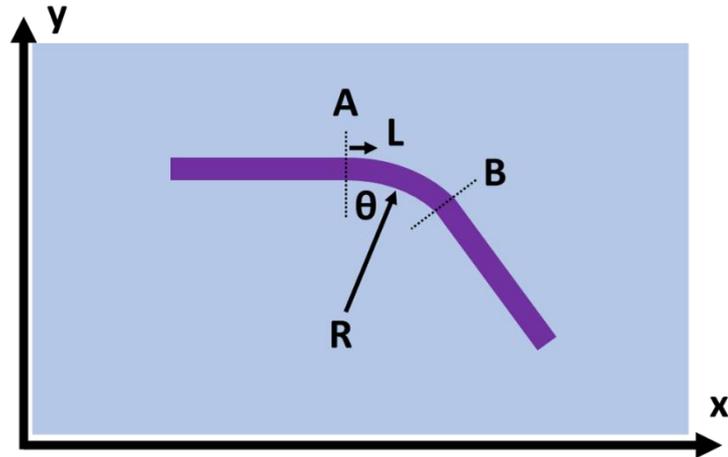
Figure 3.11: Straight waveguide – bend – straight waveguide geometry.

Given that the waveguide mode is supported through total internal reflection, we can approximate the bend angle (θ) as the square root of the refractive index contrast (~3), and thus the angle per unit length (L) is approximately 1.7°, using simple trigonometry [170]. In my research, I tested a variety of bend angles, and measured the optical power carried to ensure that each waveguide was receiving an equal split and found the waveguides with maximum bend angle of approximately 1.66° were the sharpest I could fully utilise.

3.7.   Design Summary

In this chapter I have discussed the design parameters of the capacitor materials, waveguide platform, and the interferometry architecture of the phase modulator, as well as the choice of materials for the modulator. The novelty of my research lies in combining the carrier induced permittivity change in a capacitor gated ITO layer to overlap with a waveguide mode in a silicon nitride platform. I chose to use a silicon nitride waveguide of width 3 µm and thickness 150-200 nm. As discussed in section 2.5 the modulator is suitable for phase and amplitude modulation in a Mach-Zehnder architecture or as a straight through waveguide. In this chapter, I have chosen to focus the discussion on designing a Mach-Zehnder Modulator (MZM) phase modulator. The splitting of one input into two is achieved using a broadband Y-junction splitter into two bends to separate the modulators, with a maximum bend angle of 1.66°.

The metal-oxide-semiconductor (MOS) capacitor layers consist of gold, SU-8, and ITO respectively. The capacitor is used to modulate the local permittivity in the ITO-SU-8 interface and thus the effective index of the waveguide mode. The capacitor layer thicknesses were

chosen to provide the maximum mode overlap with the ITO layer for optimum modulation depth, while minimising the mode overlap with the gold layer to reduce system loss. I used a gold layer of 100 nm thickness, an SU-8 layer thickness between 200-500 nm, and an ITO layer thickness between 50-100 nm, where the range is dependent on fabrication conditions.
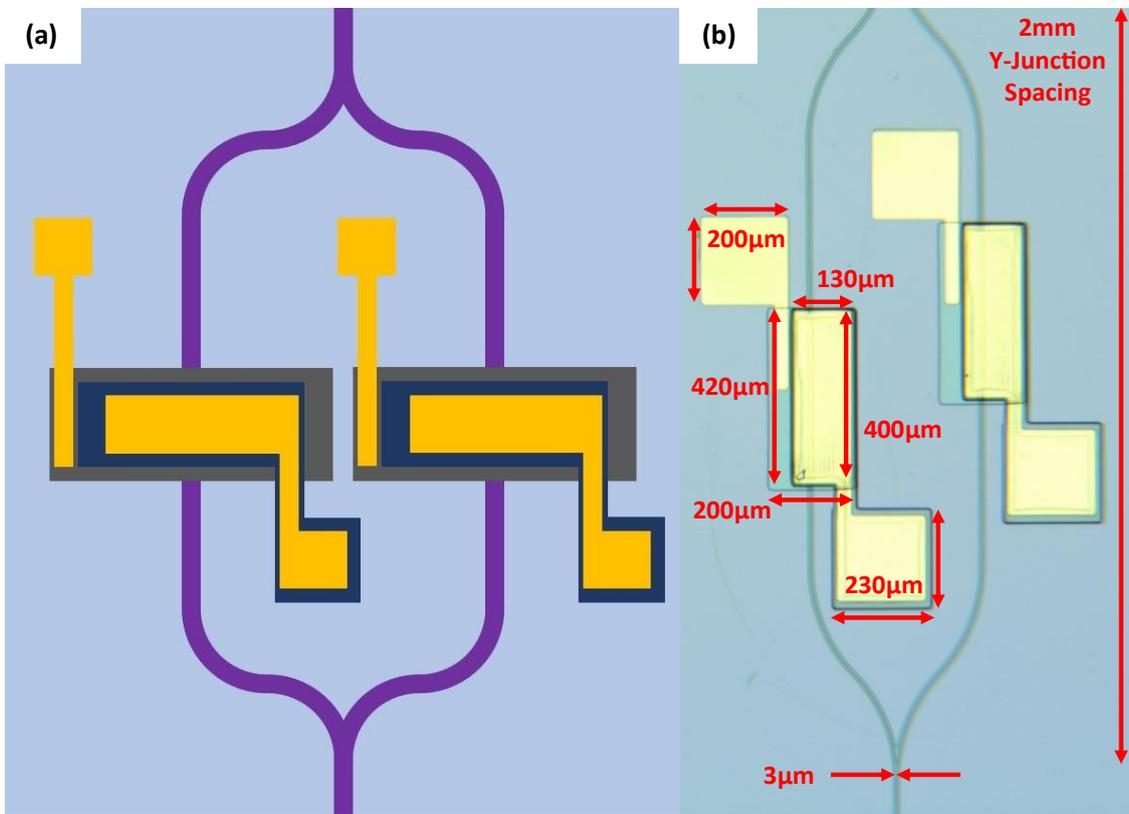
Figure 3.12: (a) Illustrate diagram and (b) micrograph of the phase modulator with a capacitor on each arm of the Mach-Zehnder architecture.

I have also discussed the design considerations of the electrical contacts to the capacitor in this chapter, opting for an asymmetric design with an insulated top contact to concentrate the electric field between capacitor plates. The system design is illustrated in Figure 3.12(a) and pictured in (b, showing the double capacitor layout to remove any 0V bias. The waveguide bend in Figure 3.12(a) is exaggerated for illustrative purposes and the SU-8 layer in Figure 3.12(b) is observed under the gold (yellow) layer on the right hand side of each capacitor, appearing as a "bold line" effect.

4. Characterising Indium Tin Oxide

Chapter Synopsis

This chapter presents an original contribution to the literature and furthers the understanding of indium tin oxide thin film deposition. In this chapter I describe the material properties of ITO and how it can be used in electro-optic modulators to introduce a signal change. I will also discuss some of the literature inconsistencies surrounding ITO that introduced many misconceptions about the material properties. The work presented here was conducted collaboratively with Professor Thomas Krauss, Dr Stuart Cavill, Dr Christopher Reardon, and Samuel Blair. I present our method of determining the optical and electrical properties of ITO that we believe is a more robust technique. The chapter discusses the electrical and optical characterisation of indium tin oxide thin films including surface profilometry, four-probe, Hall probe, X-ray diffraction, scanning electron microscopy, and resonant grating techniques. I have already discussed the interest in transparent conductive oxides (TCOs), and specifically indium tin oxide (ITO), in the field of optoelectronic technologies. This interest is based on their optically transparent yet electrically conductive properties. ITO is not a new material, having been used in heated windscreens [171], infra-red filters [172], transparent electrodes [173], and flat-panel displays [174], though it has been thrust to the forefront of optoelectronic devices since the Atwater group published the ability to achieve up to unity order refractive index change [80]. I have chosen to focus this chapter on my contributions to this body of work, namely the deposition protocol and measurement analysis. I focused my efforts on analysing the surface profilometry, RCWA simulation, scanning electron microscopy, and resonant grating analysis. Where appropriate, I have indicated this with a "we" and "I" distinction.

4.1.   Indium Tin Oxide

Indium tin oxide (ITO), a heavily doped n-type semiconductor, is a ternary mix of indium (In), tin (Sn), and oxygen (O_2). The optical and electrical properties are strongly influenced by the crystalline structure and thickness of the ITO, making the deposition of ITO significantly more challenging than metals and most dielectrics.

4.1.1.   Indium Tin Oxide Structure

The atomic structure of a unit-cell of ITO resembles that of indium oxide (In_2O_3), an ionically bound semiconducting oxide that crystallises in a cubic bixbyite-type structure with a space group Ia3 and a lattice constant of 10.118 nm [175]. The band structure of ITO is a function of the complex unit cell lattice and electronic interactions, and therefore the material properties of ITO are usually discussed in terms of an assumed band diagram consisting of an isotropic parabolic conduction band [137], [176]. Fan and Goodenough's [177] band diagram for In_3O_3 exhibits a band gap of 3.5eV, prohibiting interband transitions in the visible range and yielding over 90% transmission in this region. The conduction and valence bands arise from indium 5s orbitals and oxygen 2p electrons, respectively, and the Fermi level (E_f) is found to be above the conduction band, due to n-type doping of the tin impurities [137].

The bands differ depending on doping density; for low doping density, donor states are formed just below the conduction band, but for very high doping densities, the donor states increase and merge with the conduction band at a critical density. This critical density, known as the Mott density, is governed by the Bohr radius and for ITO is approximately 1.3 nm, corresponding to a doping density of $3.4x10^{19}$ cm^{-3} [137]. Above the Mott density, the electrons occupy the bottom of the conduction band and form an electron gas, reducing the electron mobility [178], [179].

4.1.2.   Electrical and Optical Properties of ITO

To use ITO for nanophotonic devices, a detailed knowledge of the deposition parameter relationship to the optical and electrical properties must be known. In general, the Drude-Lorentz model can be used to model the optical properties of ITO, which is discussed later in this chapter, due to the high carrier concentration. The permittivity of ITO is heavily influenced by the electrical properties, namely carrier concentration, of the material, resulting from the

doping tin donors and oxygen vacancies. The doping levels is a balancing act, as substituting an indium atom, thus freeing an electron and increasing carrier density, allows a doubly-charged oxygen vacancy and singly-charged tin donor, reducing the carrier mobility due to ionised impurity scattering [137]. It follows that tuning ITO is possible by controlling the oxygen and tin content of the material during deposition. However, it is not that simple as the properties also depend on the material thickness and film crystallinity [180].

Studies into the film colour and carrier mobility because of the oxygen/tin doping concentration and sputtering power density, respectively, during deposition further compound the notion that controlling the optical and electrical properties of ITO is non-trivial. The resistivity of ITO is shown to be high for low doping concentrations, while the resistivity is also high for increased carrier concentrations, thus there is a local minimum of parameters if the desired effect is low resistivity/high conductivity, and therefore high speed, operation. The high carrier density associated with increased oxygen content is a result of enhanced crystallisation which can also be achieved through high temperature annealing, thus adding another complication to the tunability of ITO.

4.2.   Electro-Optic Modulation in ITO

While ITO has been most commonly used as a transparent electrode in photovoltaic and touchscreen applications [181], [182], in photonic applications a lower conductivity ITO is more desirable. The lower optical losses associated with a lower conductivity is of interest for use in light modulation [74], [77], [78], tuneable metasurfaces [183]–[185], and efficient light-matter interaction in the epsilon-near-zero (ENZ) regime [186]. We are interested in using ITO for electro-optic (EO) modulation, namely the modulation of an effective index of an optical waveguide mode. This is achieved in ITO through a change in carrier concentration as described by the Drude model.

4.2.1.   The Drude Model

We start by considering the complex dielectric permittivity of ITO, which can be described by the simplified Drude model. The complex permittivity (ε)

$$\varepsilon(\omega) = \varepsilon_\infty - \left(\frac{\omega_p^2}{\omega^2 + i\omega\Gamma}\right)$$

Equation 4.1

can be expressed as a function of the angular frequency of the incident light (ω) (Equation 4.1), with the material dependent variables: high frequency permittivity (ε_∞), plasma frequency (ω_p^2), and the collision frequency (Γ). The plasma frequency, or ENZ point, represents a plasma oscillation resonance where the dielectric function changes sign and the real part of the permittivity drops to zero. Light with angular frequency below the plasma frequency is totally reflected, whereas light with angular frequency above the plasma frequency can penetrate the material. The plasma frequency

$$\omega_p^2 = \frac{ne^2}{\varepsilon_0 m_e^*}$$

Equation 4.2

is a material constant, related to the free carrier density (n) and the effective electron mass (m_e^*) (Equation 4.2). Where the permittivity of free space (ε_0) and the elementary charge (e) have their usual meaning [79]. It is clear, from Equation 4.2, that the Drude permittivity is entirely governed by the free carrier density of the material

$$n = \frac{\sigma}{\mu e}$$

Equation 4.3

which is related to the electrical properties of conductivity (σ) and electron mobility (μ) as shown in Equation 4.3.
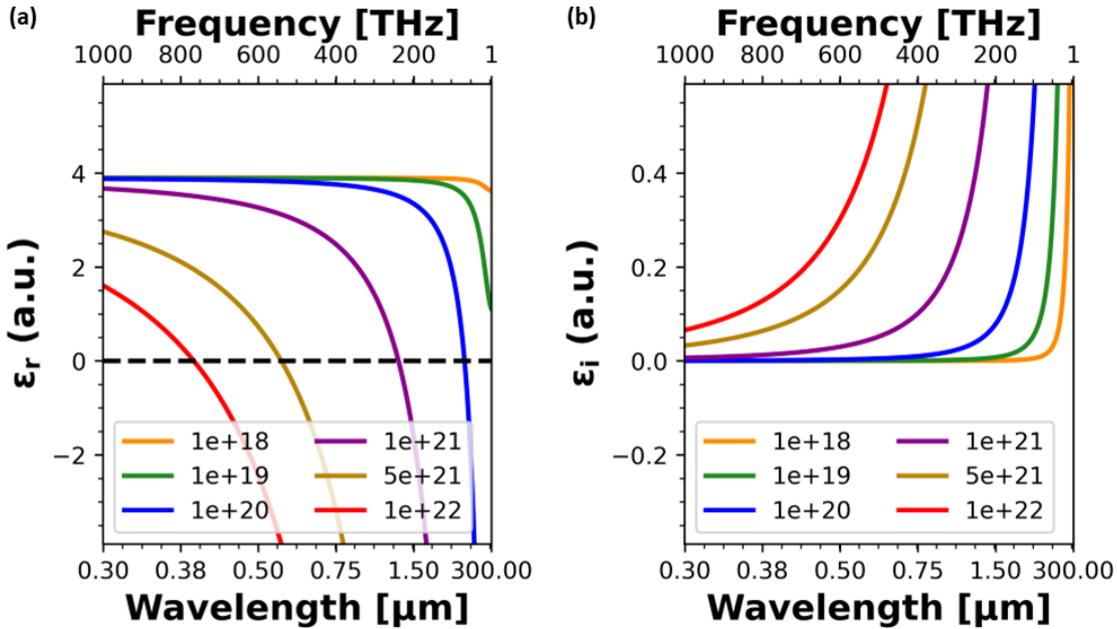
Figure 4.1: (a) Real and (b) imaginary components of the complex Drude permittivity for literature values of free carrier density. Data taken from Sorger et al. [187] and replotted. Note that carrier density is expressed in cm^{-3} as is customary in the field.

The dependence of the complex Drude permittivity on the electrical properties of ITO, given in Equation 4.3, is shown in Figure 4.1, where change the (a) real and (b) imaginary permittivities as a function of the free carrier density are highlighted by the dramatic shift in the permittivity curves. The ENZ region, shown by the crossing of the y axis at 0 (Figure 4.1(a)), is popular because it affords more efficient light-matter interaction, due to the $\frac{\Delta \varepsilon_r}{\varepsilon_0}$ dependence of many linear and nonlinear effects.

The high frequency permittivity, collision frequency, and effective electron mass used in Equation 4.1 to describe the permittivities shown in Figure 4.1 are taken from the literature [188], and are assumed to be 3.9, 180 THz, and 0.35, respectively. The elementary charge and electron mass are taken as the commonly accepted values. The refractive index (n) and extinction coefficient (k) components of the complex refractive index

Equation 4.4

$$n = \sqrt{\frac{\sqrt{\varepsilon_r^2 + \varepsilon_i^2} + \varepsilon_r}{2}}, \qquad k = \sqrt{\frac{\sqrt{\varepsilon_r^2 + \varepsilon_i^2} - \varepsilon_i}{2}}$$

can be calculated from the complex permittivity components using Equation 4.4, where $\sqrt{\varepsilon_r^2 + \varepsilon_i^2}$ represents the complex modulus [189].
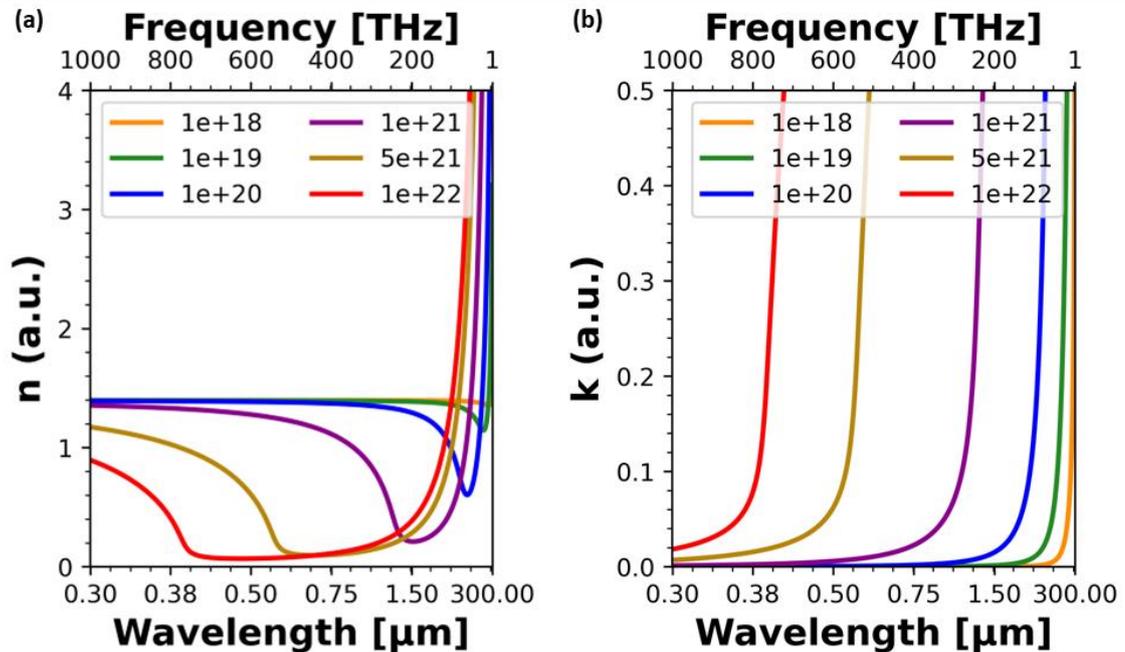
Figure 4.2: (a) Refractive index and (b) extinction coefficient components of the complex refractive index calculated from the Drude permittivity for literature values of free carrier density. Data taken from Sorger et al. [187] and replotted. Carrier densities are expressed in cm^{-3} as is customary in the field.

Using the complex Drude permittivities in Figure 4.1, I was able to calculate the corresponding refractive index components shown in Figure 4.2, where the change in (a) refractive index and (b) extinction coefficient (loss) as a function of the free carrier density are highlighted. What is remarkable about ITO is that if the carrier concentration can be modulated by 2 orders of magnitude, a unity order change in the refractive index and a 2-orders of magnitude change in the extinction coefficient can be achieved.

4.2.2.   Limitations of Performance

Despite the promising literature calculation shown in Figure 4.2, which indicates that a unity-order refractive index change is achievable in ITO, there are several factors that limit the performance of an ITO-based EO modulator. Placing ITO in a metal-oxide-semiconductor (MOS) capacitor, thus achieving electrically tuneable refractive index modulation in the ITO, places electrostatic limitations on the achievable change of the carrier concentration. Free carrier accumulation inherently leads to the formation of an accumulation layer, which is the portion of the ITO in which the refractive index modulation will be achieved. As is well known in electrostatics, the thickness of such an accumulation layer scales inversely with the carrier density, such that, only a small modal index change can be achieved in practise. This small modal index change clearly limits the achievable modulation strength.

Due to these limitations, it is imperative to process ITO in such a way that the deposited film yields a carrier concentration in a desirable region. There are two potential routes: low carrier concentration with low optical loss, and high carrier concentration with high optical loss. Typically, the photovoltaic industry requires low resistivity ITO for use in transparent electrodes, whereas EO modulation is more of a trade-off between speed and performance. A higher carrier concentration yields a greater shift in the refractive index, and thus a greater modulation depth (Figure 4.1) [79], but at the cost of increased losses.

It is well understood that the optical and electrical properties of ITO are sensitive to small variations in the deposition conditions, such as oxygen flow, operating pressures, and thin film deposition technique. Typically, ITO is deposited from a source material of diindium trioxide (In_2O_3) and tin dioxide (SnO_2) and, as such, high temperature deposition techniques tend to dissociate the oxygen, which is largely responsible for the free electron density, and therefore the source of index modulation in the material. Therefore, high temperature deposition techniques produce a low resistivity ITO desirable for high-speed applications. I have discussed

how the ITO in this body of work was deposited, and the justification for this choice, in more detail in section 5.2.9. It is unclear in the surrounding literature how the deposition conditions affect the optical and electrical properties of ITO. Here, I will show how the deposition conditions affect the electro-optic performance of the resulting film.

4.3.   Literature Inconsistencies

Many authors assume previously published literature values for important optical and electrical properties for the Drude model of ITO including carrier mobility (μ), high frequency permittivity (ε_∞), electron effective mass (m_e^*), and the collision frequency (Γ) without accounting for the deposition conditions for which these values were obtained. We propose an alternative approach to spectroscopic ellipsometry measurements to determine the dispersion curves for a complex semiconductor such as ITO, noting that many of these parameters vary as a function of the deposition conditions. To illustrate the effect on the complex Drude permittivity of these parameters, I will briefly discuss the values assumed in the literature and show the resulting Drude permittivity and complex refractive index components. I have already shown (Figure 4.1 and Figure 4.2) the effect the carrier density, and by extension the carrier mobility, have on the Drude permittivity and refractive index, respectively, though I used this to demonstrate a performance metric. While neither can be assumed independently, without ignoring material conductivity, the importance of accurately measuring the electrical properties is highlighted.

4.3.1.   High Frequency Permittivity

The high frequency permittivity is the real permittivity component at short (ideally, zero) wavelength and is determined by the bound carrier concentration. Most authors assume the high frequency permittivity to be constant, but this is an oversimplification as it reflects the degree of crystallinity of a material. The high frequency permittivity is essentially the y-intercept of the Drude curve, and is often assumed in the literature to be either 1 or 3.9 in published literature [80], [133], [188], [190]. It is common to assume a permittivity of 1 when considering the material properties below the plasma frequency, where the real component is not relevant. For the modulators considered in this research, it was important to clarify the optical and electrical behaviour of ITO in the visible spectrum. To demonstrate the effect the high frequency

permittivity value has on the optical properties of ITO, I assumed a carrier density of 1×10^{20} cm^{-3}, a collision frequency of 180 THz, and an effective mass of 0.35 [79].
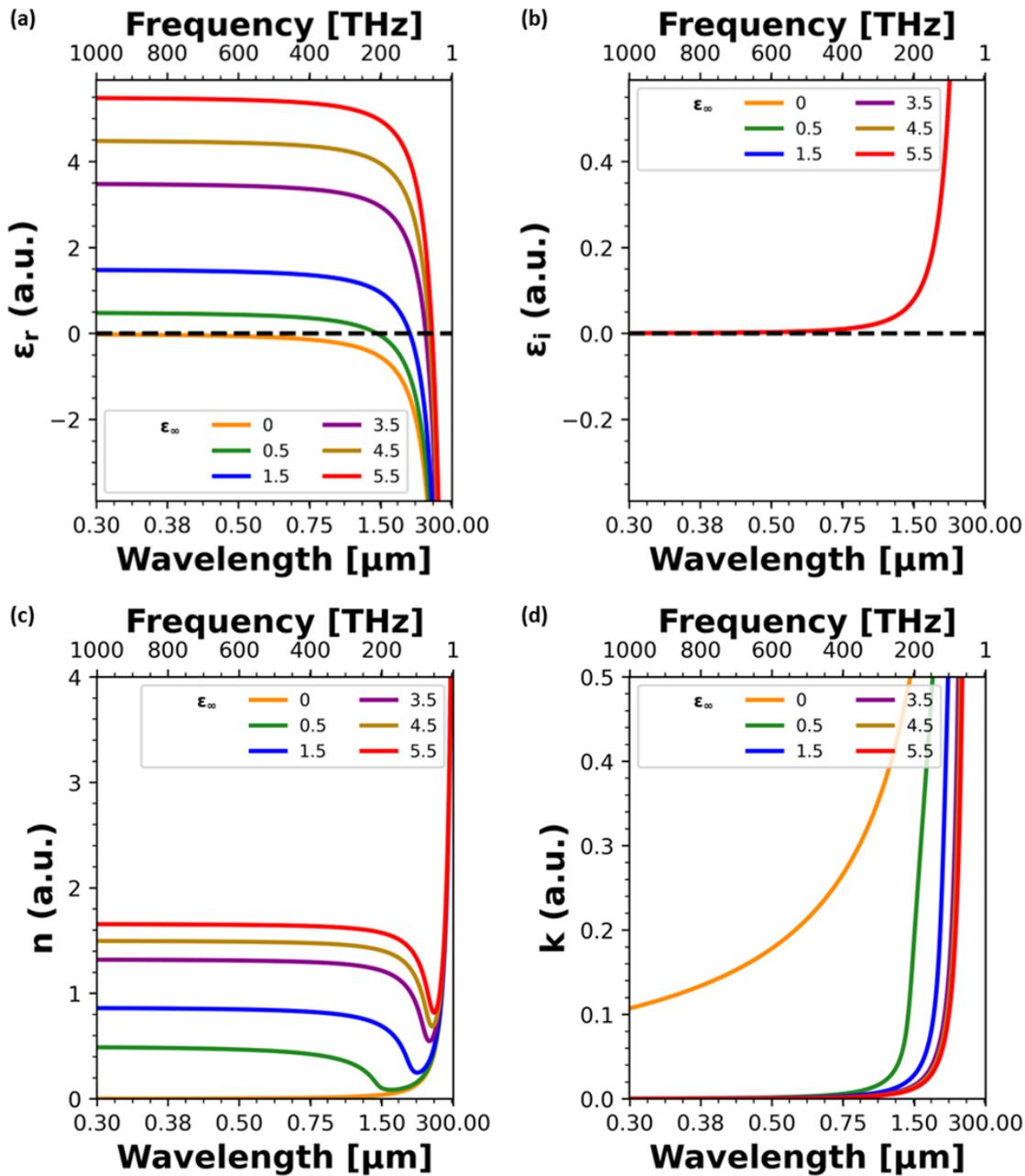
Figure 4.3: (a) Real and (b) imaginary components of the complex Drude permittivity and the corresponding (c) refractive index and (d) extinction coefficient of the complex refractive index as a function of the high frequency permittivity.

The effect of the high frequency permittivity value on the complex Drude permittivities is shown in Figure 4.3(a)(b), where the imaginary components are all approximately equal, as the imaginary component is, inherently, dependent on the parameters in the loss region of the spectrum. In the negative epsilon regime, the material becomes highly reflective, which is

desirable for tuneable reflectors. It is common in the literature to operate close to the epsilon near zero condition for the maximum shift in permittivity. As can be seen in Figure 4.3(c), the high frequency permittivity can alter the refractive index by unity order, the same level as the carrier density. Hence, accurately determining the high frequency permittivity is important for characterising the optical properties of ITO.

4.3.2.   Collision Frequency

The simplest form of the Drude model assumes that electrons moving in a constant uniform electric field have a velocity that is sufficiently high to accumulate a moment between collisions, where the time between collisions is given as the reciprocal of the collision frequency. It follows that a material with high carrier density will have a high collision frequency, and thus a highly absorptive material will have a high collision frequency; in a way, the collision frequency can be understood as a loss term with respect to electron mobility. The collision frequency is often assumed to be either 0 or 1.8 THz [80], [188]. To demonstrate the effect of the collision frequency value on the optical properties of ITO, I assumed a carrier density of 1×10^{20} cm^{-3}, a high frequency permittivity of 3.9, and an effective mass of 0.35.
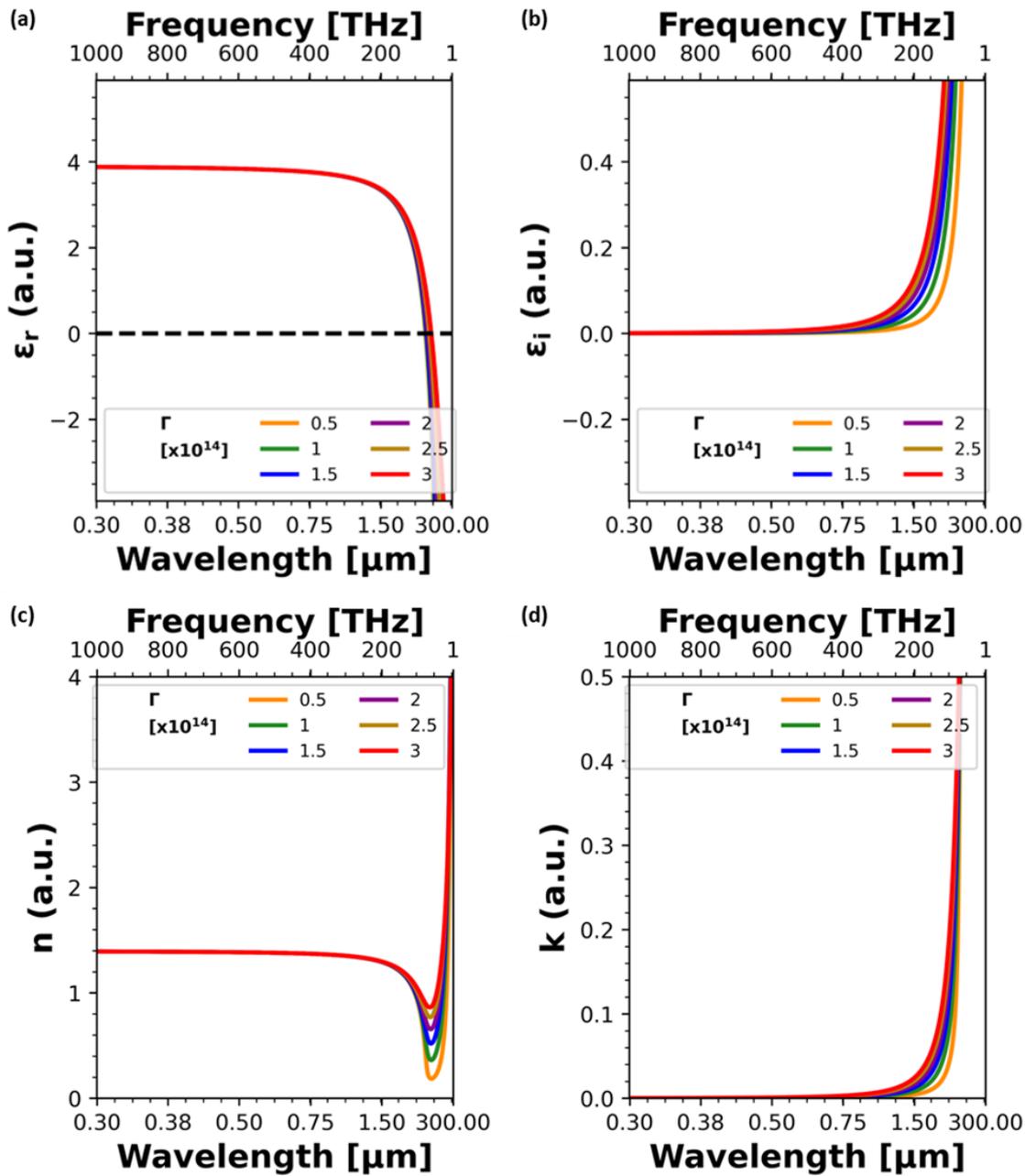
Figure 4.4: (a) Real and (b) imaginary components of the complex Drude permittivity and the corresponding (c) refractive index and (d) extinction coefficient of the complex refractive index as a function of the collision frequency.

The effect on the complex Drude permittivities of the collision frequency is shown in Figure 4.4(a)(b), where the data labels are given in THz. The effect is not as dramatic as the high frequency permittivity, especially on the refractive index components (Figure 4.4(c)(d)). This is because the collision frequency is only having a significant effect on imaginary component of the complex Drude permittivity, i.e., the region of the permittivity beyond the plasma frequency. This is demonstrated in Figure 4.4 (a) and (c).

4.3.3.   Effective Electron Mass

An electron in a crystal may behave as though it has a mass different to the free electron mass (m_0). The effective electron mass describes the mass an electron appears to have when responding to an applied electric or magnetic field, or when interactive with another electron. The description of the effective electron mass in response to an applied field is described by Newton's second law of motion and can be used to describe the transport of electrons. It follows that the effective electron mass is therefore a description of the carrier density of a material, temperature of a material, and external field strength. Hence, it would be wrong to assume that the commonly used value of 0.35 [77], [79], [133] is fixed and independent of the deposition conditions. To illustrate the effect of the effective electron mass on the optical properties of ITO, I assumed a carrier density of 1×10^{20} cm^{-3}, a high frequency permittivity of 3.9, and a collision frequency of 180 THz.
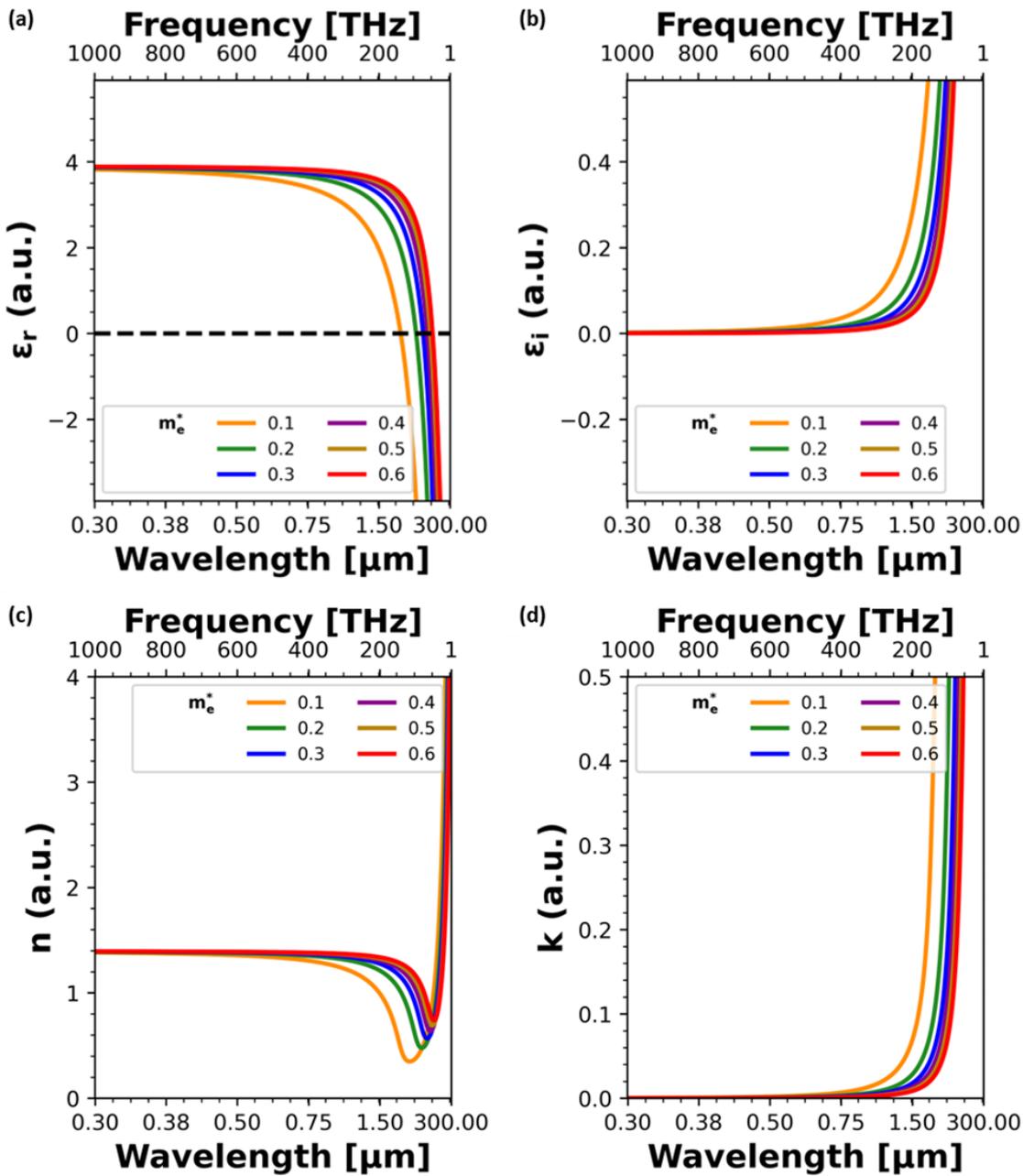
Figure 4.5: (a) Real and (b) imaginary components of the complex Drude permittivity and the corresponding (c) refractive index and (d) extinction coefficient of the complex refractive index as a function of the effective electron mass.

The effect of the effective electron mass on the complex Drude permittivities is shown in Figure 4.5(a)(b), with the corresponding complex refractive index components are shown in (c) and (d). The effect of decreasing the effective electron mass is a shift to a lower plasma frequency and lower refractive index, with an increase in the loss component at shorter wavelengths.

# 4.4.	Investigation into Photonic Properties

It is clear, from Figure 4.2 - Figure 4.5, that irrespective of the preferred operating wavelength range, the Drude parameters all impact on the dispersion curves, and thus on an accurate understanding of the optical and electrical properties of ITO. A misinterpretation of the Drude parameters, when considering a MOS EO modulator, could have significant impact on the design and thus lead to wasted development and poor performance.

To better understand how the incorporation of oxygen can affect the electro-optic properties of ITO, we conducted a comprehensive analysis of a variety of ITO thin films deposited using a range of conditions. We made no intrinsic assumptions about any of the material parameters using 1D periodic structures fabricated in thin films of ITO, which allows us to characterise the material as a device, rather than measuring the properties of an unaltered film. Measuring the material parameters in a device structure means any post-processing of the film such as lithography, etching, material deposition, or cleaning protocols that may affect the thin film properties have been accounted for in our measurements. These processes may cause displacement of oxygen or tin from the surface of the material which would alter the material permittivity. For the modulators in this research, a periodic structure is not necessarily of interest, but the protocols used to fabricate such structures are synonymous with those discussed in chapter 5. This detailed investigation will be discussed in this chapter, and we hope it will lead to a better understanding of how to control the deposition conditions to yield the desired properties of ITO.

## 4.4.1.	Investigation Constants

Clearly, if we are going to propose a new method for accurately determining the optical and electrical properties of Drude materials, specifically ITO, then we need to maintain consistent fabrication, measurement, and data processing techniques. Throughout this investigation, the thin films of ITO and fabricated gratings were kept consistent by maintaining one fabricator and a strict protocol. The data analysis and measurement techniques were conducted under similar restrictions to ensure an unbiased treatment of the characterisation. I will now discuss the fabrication and data processing constants used throughout this investigation.

4.4.2.   Deposition of ITO Thin Films

As has been discussed in section 4.2.1, it is known that the optical and electrical properties of ITO are highly dependent on the deposition and post-deposition protocols, namely that there is an increase in crystallinity with deposition and annealing temperatures [191], [192]. ITO, for thin film deposition, is typically supplied either as a sputter target or in pellets for thermal evaporation as a 90-10 wt% of diindium trioxide (In_2O_3) – tin dioxide (SnO_2), and the disassociation of oxygen from the constituents is proportional to temperature. As is discussed in section 5.2.9, we chose to use a pulsed DC reactive magnetron sputtering technique to avoid high deposition temperatures and allow for finer control of oxygen incorporation during the deposition. The deposition temperature of the sputtering technique was not controlled outside of using the same pressure and DC voltage to deposit the film. The temperature during the process is, as discussed in section 5.2.9, is proportional to the DC volage and plasma pressure, and therefore maintaining consistent conditions during the deposition reduces variation in deposition temperature. As for post-deposition protocols, specifically we are discussing the thermal annealing of the deposited films (section 5.4.1), where we can introduce further oxygen in to the as-deposited films.

Table 4.1: Oxygen flow in the sputtering process represented as percentage concentration in the chamber.

O_2 Flow Rate [SCCM]	O_2 Concentration [%]
0	0
0.5	2.5
1	5
2.5	10
3.5	15
5	20
7.5	27

ITO thin films were deposited on to a borofloat 33 substrate using the sputtering technique mentioned above, where the reactive gas (oxygen) flow was varied to adjust the as-deposited film conductivity using the flow rates: 0, 0.5, 1, 2.5, 3.5, 5, 7.5 standard cubic centimetre per minute (SCCM). Not all the oxygen concentrations are used for all the measurement techniques discussed in this chapter; as such, to keep the nomenclature simple, I will refer to the oxygen flows as percentage concentrations, as indicated in Table 4.1. The values are rounded for simplicity.

While the oxygen flow was varied throughout this experimental campaign, the base parameters for the sputtering protocol were maintained as follows: 20 SCCM Argon as the sputtering gas, 100 W supplied DC power, and a plasma pressure of 0.75 Pa. To avoid reactions between the as-deposited film and any unwanted gas, the samples are allowed to cool overnight before bringing the sputtering chamber back to atmosphere with nitrogen gas.

4.4.3.   Annealing of ITO Thin Films

The ITO films are highly amorphous post-deposition and, as such, require a thermal anneal to reintroduce oxygen in a controlled fashion and introduce some crystallinity back into the lattice. This is done using a tube furnace (section 5.4.1). Oxygen flows into the tube furnace at 500 SCCM throughout the anneal with a 5-minute precondition prior to heating. The samples begin at room temperature (~20°C), and the furnace ramps at 5°C/min to a maximum of 500°C. The samples sit at 500°C for 1 hour in the oxygen environment before cooling to room temperature overnight. Cooling is unassisted and is therefore achieved purely through thermal radiation into the room. We also investigated the ITO sheet resistance dependency on the temperature and oxygen flow during annealing. In this instance, the temperatures and oxygen flow rates were varied, while the heating and cooling ramp rates were maintained, and samples were deposited using 0 SCCM oxygen flow.

Table 4.2: Effect of oxygen flow, annealing temperature, and annealing time on the sheet resistance for sputter deposited ITO with 0 SCCM oxygen.

Gas Flow	Temperature	Anneal Time	Sheet Resistance
[SCCM]	[°C]	[hrs]	(R_S) [Ω/sq]
Air	500	1	48.3 ± 0.5
500 O_2	500	1	41.7 ± 0.4
1000 O_2	500	1	38.9 ± 0.1
500 O_2	500	2	46.6 ± 0.4
500 O_2	300	1	88.9 ± 0.3

The results of changing the anneal gas flow, temperature, and duration are shown in Table 4.2, indicating that while the oxygen flow during anneal and the duration of the anneal does not affect the electrical composition of ITO, the temperature at which the anneal is performed does. The measured sheet resistances are all approximately the same for a constant anneal time and oxygen flow, but for a decreased anneal time, the sheet resistance (4.5.2) is almost doubled. For

consistency, and a lower resistivity, we maintained a temperature of 500°C and only varied the oxygen flow during deposition during this investigation.

4.4.4.   Fabrication of ITO Gratings

To limit the number of small references to the fabrication chapter, all the fabrication methods are discussed in more detail in chapter 5, but I will give a summary of how the gratings were patterned into the ITO films and then discuss how the grating parameters were measured and used to determine the optical properties.

Due to the nature of the optical measurements (section 4.6), the ITO gratings needed to be fabricated on-to a transparent substrate. We used borofloat 33 substrates, diced into 15x15 mm sample sizes. The mechanical dicing process leaves a residue on the samples, and so the samples are cleaned in Piranha solution for 5 minutes and then rinsed in acetone and isopropanol (IPA). The ITO films are then deposited and annealed as discussed above.

The gratings are patterned using electron beam lithography (EBL), which requires a lithography resist layer. We used AR-P 13 [30], spun at 5000 rpm for 60 seconds and soft-baked at 180°C for 10 minutes. The EBL parameters used to fabricate the gratings are: 40 μm aperture, 0.6 nA beam current, 10 nm step size, 130 μC/cm^2 base dose, 1.71 base dose multiplier, and 9.5×10^{-2} ms exposure time. The samples are then developed in xylene for 2 minutes and rinsed with IPA.

Post development, samples are dry etched using reactive ion etching (RIE). The gas mixture is 21 SCCM argon and 5 SCCM hydrogen with a pressure of 3×10^{-2} mbar. The DC voltage is fixed at 450V. These conditions yield an etch rate of 5nm/min and are a combination of the physical and chemical etching component of the argon and hydrogen respectively. Regarding the chemical component, the hydrogen reduces the metal oxide to form water, which is volatile at the process pressure, whilst the argon works to physically bombard the indium and tin. Samples were etched for 30 minutes, creating an etch depth of approximately 150 nm, though of course this process is not perfectly linear. Following etching, the remaining photoresist was removed by rinsing in 1165 solvent remover at 65°C, and cleaning with acetone and IPA.

4.4.5.   Reducing Errors

A common technique to reduce variations in measurements, and to help improve precision, is to take repeated measurements and calculate average results. Where possible, throughout this investigation, we took multiple measurements of each parameter for every sample and used these measurements to calculate an average result. The formulation and technique are discussed in Appendix section 0. During the data analysis portion of the campaign, I used a combination of linear and nonlinear regression tools to fit equations to measured data points. In this instance, the average and SEOM method is unsuitable. Regression is a statistical technique, and the error analysis is typically conducted using a covariance matrix method [194]. The covariance is a vector measurement of joint variability of random variables [195], where the sign of the covariance refers to the linearity of the variables. The magnitude of the covariance is the geometric mean of the variances of the variables, i.e., the correlation between the variables can be calculated by dividing the magnitude of the covariance by the total variances of the random variables. The errors are calculated from the covariance matrix by taking the square root of the diagonal of the covariance matrix.

4.5.   Determining Electrical Properties

We begin the investigation by considering Equation 4.3 and the electrical properties of ITO, specifically the material conductivity (σ) and the carrier mobility (μ). The conductivity can be measured using the Kelvin technique (four-probe method) but requires the additional parameter of material thickness, and the mobility can be measured using a Hall probe technique. I will briefly discuss these methods and how they were applied in this investigation, but first we also need to discuss how we accurately measured the ITO film thickness.

4.5.1.   Surface Profilometry

As discussed in section 5.2, the sputter deposition is monitored using a quartz crystal microbalance (QCM), which is a common method for monitoring the rate of deposition in real time in most thin film deposition techniques. The QCM measures the mass loading effect on a piezoelectric crystal, where the resonant frequency of the crystal is a function of the crystal thickness and, as such, any change in mass can be measured as a change in the frequency

response under an applied voltage. However, there are many factors such as temperature, calibration, film stress, and deposition energy that can affect the accuracy of the QCM [147].

As we need to know the thickness of the ITO films to accurately determine the electrical properties discussed in the next section, we need to employ another technique to measure the deposition thickness. A common, and reliable, measurement tool is the stylus profiler, which measures a step height variation by dragging a fine tip across the sample and measuring tip deflection. To measure the film thickness, the stylus profiler requires a step change between the substrate and deposited ITO film. This step can be introduced by covering a small area of the substrate with tape, marker ink, or photoresist, removed by mechanical pulling, dissolving in acetone, or photoresist developer post-deposition, respectively.
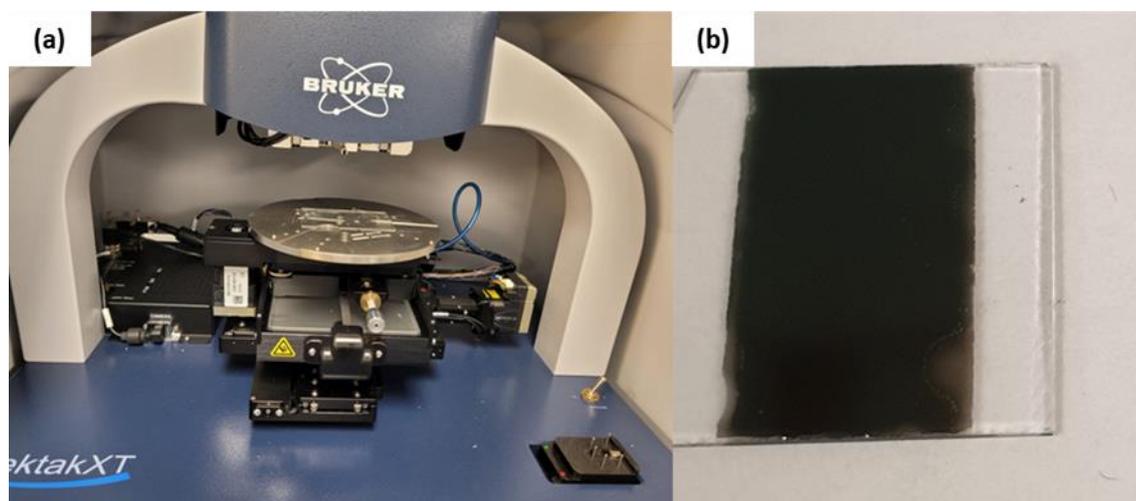
Figure 4.6: (a) Bruker DektakXT stylus profiler at York, (b) resulting step height from taped substrate during thin film deposition.

To measure the thickness of our sputter deposited ITO films, we used a Bruker DektakXT stylus profiler ("Dektak") [196]. The Dektak, shown in Figure 4.6(a), uses the step introduced by taping the substrate during deposition, shown in Figure 4.6(b), to Angstrom precision. There is, however, a problem with stylus profilometers when measuring across relatively large (few mm) lengths, namely that there is a distortion effect caused by alignment deviation [197]. This distortion is primarily caused by small damages to the stylus tip over time and manifests itself as an angle offset in the measured data, shown in Figure 4.7.
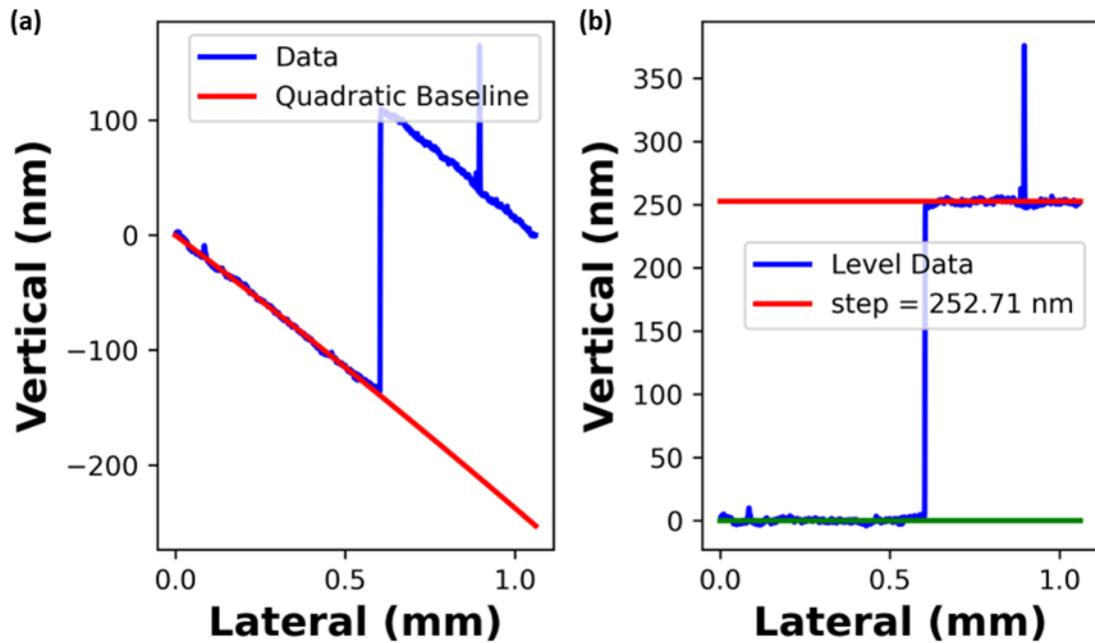
Figure 4.7: A (a) distorted and (b) levelled surface profile data for the same sputter deposited ITO film. Data line marked in blue, quadratic line marked in red, zero marked in green.

Figure 4.7 is the perfect example of the tip-affected distortion, observe how the blue data line in (a) is angled with a minor curve downward. To overcome this distortion, I use a quadratic equation ($ax^2 + bx + c$) fitted to the data and least squares regression [198] to fit a baseline to the data. A quadratic was chosen as it suits the two most common scenarios using a DektakXT system; a) the distortion creates a parabolic effect in the measurement and b) there is a linear offset, in which case a=0. The baseline can then be subtracted from the original data to produce the levelled data shown in Figure 4.7 by the blue line in (b). This quadratic method is distortion independent and can always be applied to the measured profile.

The step height indicated in Figure 4.7(b) is calculated as the difference between the averages of the upper and lower region. It is almost impossible to achieve a perfectly flat measurement due to the amorphous nature of sputtered thin films producing some level of surface roughness and the chance of a speck of dust producing spikes in the measured profile. To accurately obtain a step height, average values are calculated over a large region in both the substrate and film regions (lower and upper regions in Figure 4.7, respectively). Multiple surface profile measurements were taken for each film to calculate an average value and associated error as discussed in 4.4.5 [199].

4.5.2.   Four-Probe Method

Now that we have a method for determining the ITO film thickness, we can determine the film conductivity using the Four-probe method. Ohm's law states that when a voltage (V) is applied between two points in a circuit, an electrical current (I) will flow between them

$$I = \frac{V}{R}$$

<div align="right">Equation 4.5</div>

where the amount of current which flows is encourage by the potential difference and restricted by the amount of resistance (R) present (Equation 4.5) [100]. The electrical resistance

$$R = \rho \left(\frac{L}{A}\right)$$

<div align="right">Equation 4.6</div>

through a conductor can depend on the conductor length (L), the cross-sectional area of the conductor (A), and the material through which the current flows, which is assigned a constant known as "resistivity" (ρ). The resistance can then be expressed by Pouillet's law (Equation 4.6) [200]. The electrical resistivity

$$\sigma = \frac{1}{\rho}$$

<div align="right">Equation 4.7</div>

is a measure of how strongly the material opposes the flow of the current. It follows that a measure of how easily a current flows through the material could be expressed as the reciprocal of the resistivity, this is known as "conductivity" (σ) (Equation 4.7).
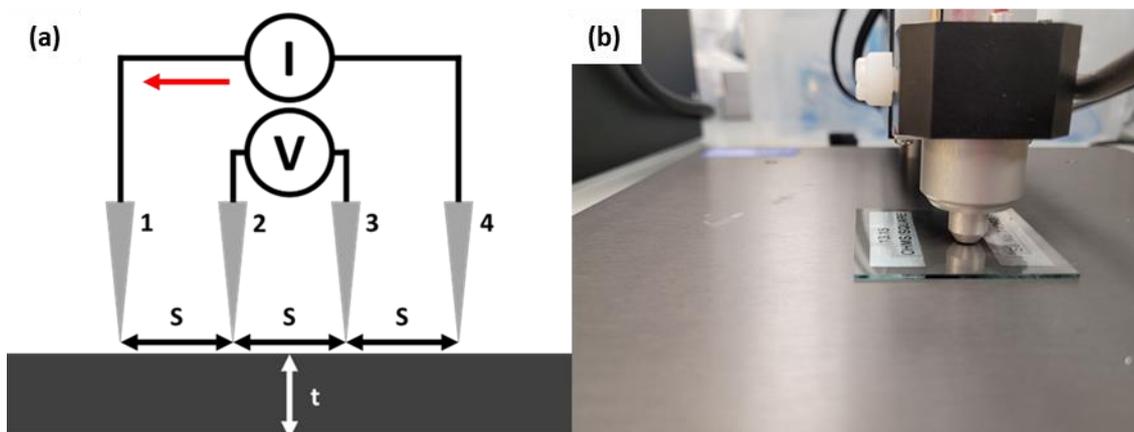
Figure 4.8: (a) Schematic diagram and (b) picture of the four-point probe system at York. Here the probe is pictured measuring a calibration sample.

A four-probe system (four-point probe), illustrated in Figure 4.8(a), consists of four equally spaced, co-linear electrical probes. The method relies on the principle that a sharp probe

injecting a DC current into a thin film of a semi-infinite hemispherical volume. The current density (**J**)

$$J = \frac{I}{2\pi r^2}$$

through the conductive material can be expresses as Equation 4.8. Where r is the radial distance from the probe tip. Applying Ohm's law, integrating over the volume, and applying a boundary condition that the voltage approaches zero as the radius approaches infinity, the voltage drop

$$V = \frac{\rho I}{2\pi r}$$

across the hemisphere can be expressed as Equation 4.9. As illustrated in Figure 4.8(a), the current is injected at probe 1 and measured at probe 4 which are in a line with equal spacing (s). The voltage at any point between the probes is equal to the sum of the voltage due to each probe separately, and hence the voltage change (ΔV)

$$\Delta V = \frac{\rho I}{2\pi s}$$

between probes 2 and 3 is given by Equation 4.10. However, four-probe systems assume a penetration thickness (t), up to a maximum of 650µm for the system used in this research and introduce a cylindrical assumption of the current density.

$$R_s = \frac{\rho}{t} = \frac{\pi}{\ln(2)}\left(\frac{\Delta V}{I}\right)$$

This assumption carried through Equation 4.8 - Equation 4.10, gives an expression (Equation 4.11) for the film resistivity. Where R_s is a parameter known as "sheet resistance". Sheet resistance is defined as the lateral resistance through a thin square of the material, introduced by the four-probe method as a size and instrument independent variable, allowing for comparison between different samples. The sheet resistance is generally measured in ohms per square (Ω/sq.), to differentiate it from bulk resistance [201]–[203].

One of the primary advantages of using a four-point probe to perform electrical characterisation is the elimination of contact and wire resistances from the measurement. We used a Jandel RM3000+ four-probe system with a Jandel cylindrical four-point probe (Figure 4.8(b)) to measure the sheet resistance of the ITO thin films and used Equation 4.11 and Equation 4.7 to calculate the corresponding conductivity, using the thin film thickness measured as detailed in section 4.5.1. As with the surface profilometry, multiple measurements are taken across each deposited film to calculate an average sheet resistance and associated standard error as discussed in section 4.4.5.

4.5.3.   Hall Probe Method

The Hall effect is the generation of the Hall voltage by the deflection of charge carriers in a magnetic field. The Hall effect is underpinned by the Lorentz force acting on an electron moving perpendicular to an applied magnetic field. The direction of deflection depends on the charge of the carrier, so holes and electrons are deflected in opposite directions. This leads to a distribution of charge and gives rise to the Hall voltage. The effect is illustrated in Figure 4.9(a), where a current is driven by a voltage source into a film under an applied magnetic field, leading to a charge build up in the film that can be measured as a Hall voltage. The Hall voltage (V_H) is given as a function of the current (I), magnetic field (B), and the elementary charge (e), and can be used to calculate the sheet density (n_s).

$$n_s = \frac{IB}{eV_H}$$ Equation 4.12

As we are concerned with measuring the Hall voltage, the sheet carrier density can be calculated using Equation 4.12 [204].
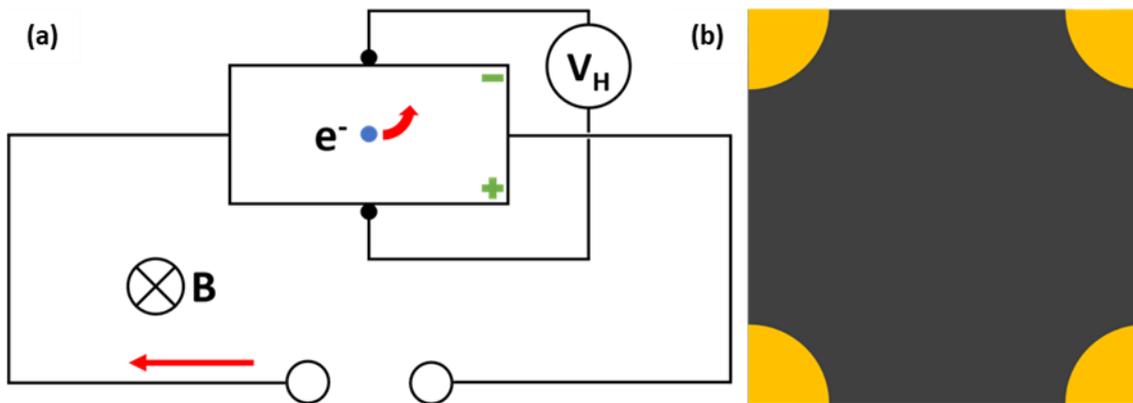
Figure 4.9: Schematic of (a) a Hall effect measurement setup and (b) Van de Pauw electrical contact configuration.

The Hall voltage is typically measured using a Van de Pauw contact configuration, illustrated in Figure 4.9(b), where arbitrarily shaped ohmic contacts are placed in the four corners of the sample. The object of the Van de Pauw configuration is to measure a sheet resistance, though as can be seen from Equation 4.13, for a known sheet resistance it can be used to determine the electron mobility [110], [205].

$$\mu = \frac{1}{en_s R_s}$$ Equation 4.13

The bulk carrier density (n) can be calculated by multiplying the measured sheet carrier density by the known film thickness (t).

$$n = n_s t$$

We used an Ecopia HMS-3000 Hall measurement system and Van de Pauw contact method with a magnetic field of 0.56T to determine the electron mobility and bulk carrier density using Equation 4.14. As with all previous measurements, multiple measurements were taken to calculate average values and use a standard error of the mean error analysis as discussed in 4.4.5.

4.5.4.   X-Ray Diffraction

X-ray diffraction (XRD) is a non-destructive analytical technique used to analyse the phase composition, crystal structure, and crystal orientation of materials, and can be used to look at the crystal structure of thin film materials. XRD is the result of constructive interference between x-rays and the crystalline structure of the target material. Given that x-rays of similar order of magnitude to the distance between the atoms in a crystal lattice produce a diffraction pattern, we can use Bragg's law to analyse the crystallinity of a material [206].

Since electron mobility can be considered the measure of how quickly an electron can move through the semiconductor in the presence of an electric field, we thought it worth examining the crystallinity of the material. One would expect a highly crystalline material to exhibit high electron mobility, while a polycrystalline material would exhibit a lower mobility, due to the presence of transport impeding grain boundaries[207], [208]. As such, we expect the ITO films that exhibit low mobility to also show a decrease in film crystallinity. We collaborated with Dr Stuart Cavill, at York, who specialises in XRD to measure the crystallinity of out 0%, 5%, and 20% oxygen flow sputter deposited ITO films.

4.5.5.   Electrical Properties – Experimental Results

It is well understood that the free carrier density in ITO arises from oxygen vacancies and tin donors [209], with the former found to make the dominant contribution to the conductivity due to the two-electron contribution compared to the single-electron contribution of tin donors [210]. We expect the conductivity to decrease with a higher oxygen content in the material due to the filled oxygen vacancies.

The sputtered ITO films were investigated electrically as a function of oxygen flow. The sheet resistances and free carrier mobilities measured using the Four-probe and Hall probe methods discussed in this section are shown in Figure 4.10(a). Using Equation 4.7, Equation 4.3, and the measured film thicknesses, I was able to calculate the corresponding film conductivity and free carrier density, shown in Figure 4.10(b).
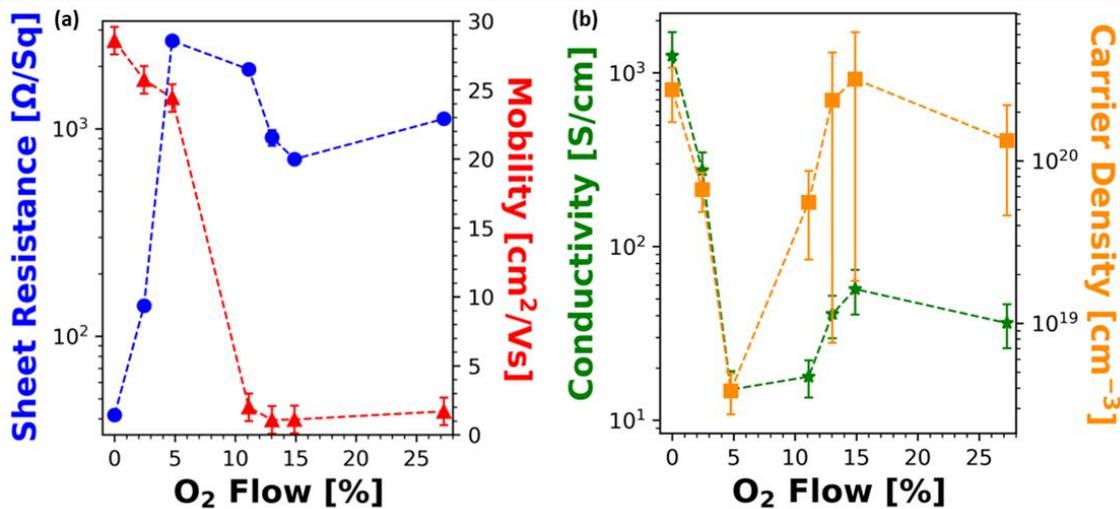
Figure 4.10: (a) Measured sheet resistance and carrier mobility for sputter deposited ITO film and the (b) calculated conductivity and carrier density as a function of oxygen concentration.

Figure 4.10 confirms the expected result, namely that increasing oxygen flow and oxygen concentration during deposition decreases the conductivity of the ITO and reduces the overall carrier mobility. This is expected due to the increase in carrier density caused by the increased oxygen incorporation, i.e., an increase in carrier density reduces mobility and increases resistance. This behaviour is observed in the literature. The conductivity increases again above 10% oxygen flow, which we associate with a higher number of substituted tin ions with increasing oxygen flow, providing additional carrier contributions, indicated by the sudden increase in the green and orange lines in Figure 4.10(b). The conductivity saturates beyond this limit due to an equilibrium state of the oxygen vacancy contributions and the tin ion contributions.

The results of the XRD data are shown in Figure 4.11(a), with the Miller indices indicating the observed diffraction planes of the Ia3 cubic space group. The visible peaks for the (222), (004), and (440) planes match the literature Bragg angles for sputtered ITO [211], which serves as a nice confirmation that there is oxygen present in the deposited material.

The degree of crystallinity is determined from the width of each reflection peak (Figure 4.11(b)-(d)) and the narrower the diffraction peak, the more crystalline the ITO and the larger the average grain size, as shown by Equation 4.15. The average grain size (δ)

$$\delta = \frac{0.9\lambda}{FWH(\cos(\theta))}$$

Equation 4.15

can be expressed as a function of the incident wavelength (λ), and the full width half maximum (FWHM) of the respective reflection and Bragg angle (θ).
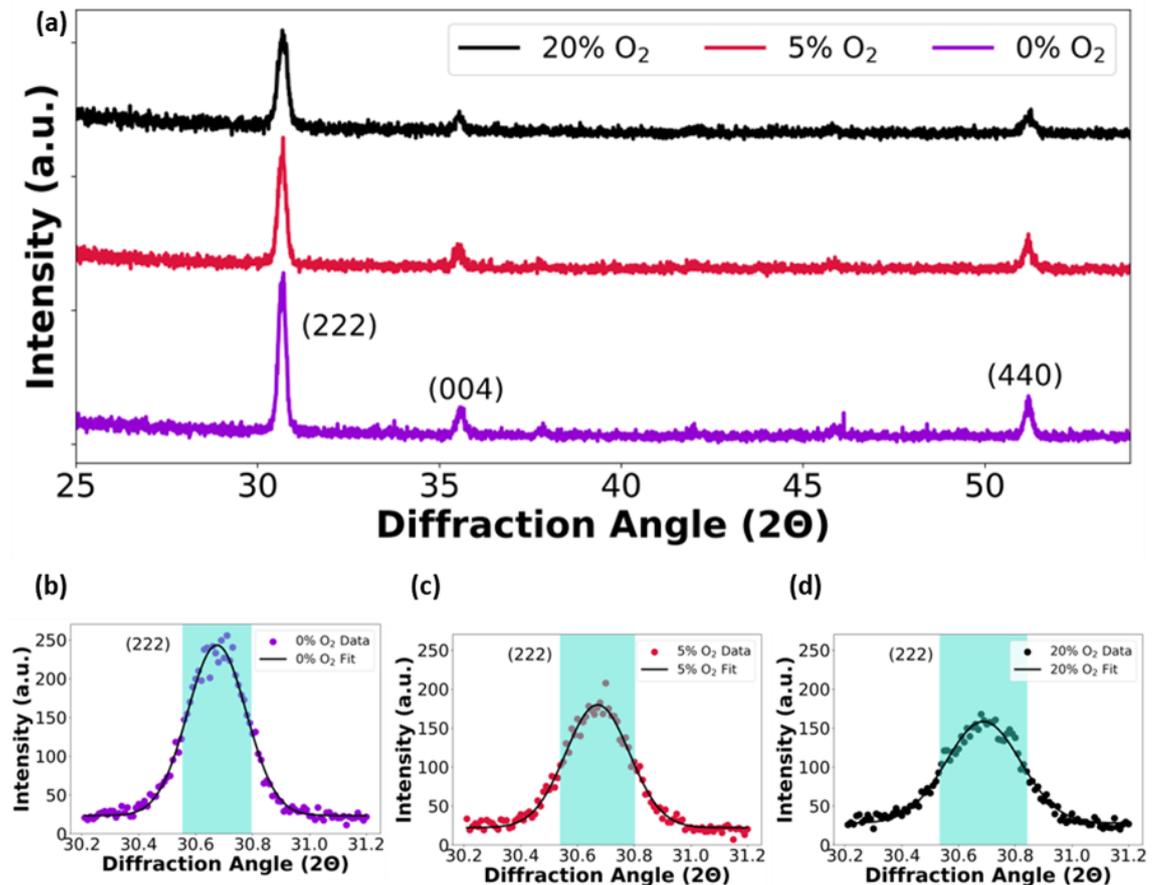
Figure 4.11: (a) XRD data for annealed ITO films of 0%, 5%, and 20% oxygen flow during deposition. (b), (c), and (d) show exemplar FWHM extractions for the respective (222) orientation for 0%, 5%, and 20% oxygen flow. The FWHM increases from 0.246° for 0% oxygen to 0.361° for 20% oxygen concentration.

An increase in carrier mobility translates to an increase in the crystallinity of the ITO because mobility is higher in a pure monocrystal compared to a polycrystalline film. Looking at the decrease in mobility with oxygen concentration (Figure 4.10(a)) and the broadening of the diffraction peak with oxygen concentration (Figure 4.11(b)-(d)), this effect is exactly what we observe. The more oxygen is incorporated into the ITO film, the poorer the crystal becomes and the lower the mobility. The average grain size of the samples shown in Figure 4.11 is represented

in Table 4.3. The grain size is seen to decrease with oxygen incorporation, further aligning with the mobility trend, i.e., the higher the oxygen incorporation, the smaller the grain size and the carrier mobility as there are more grain boundaries to impede the movement of carriers through the material.

Table 4.3: Average grain sizes for different crystal orientations for 0%, 5%, and 20% oxygen flow ITO films.

O$_2$ Flow	Conductivity	Orientation	Full Width Half Max	Average Grain Size
[%]	(σ) [S/cm]	[au]	(FWHM) [°]	(δ) [nm]
0	1300 ± 200	(222)	0.246 ± 0.001	54.3 ± 0.9
		(440)	0.261 ± 0.006	
5	21 ± 1	(222)	0.29 ± 0.02	45.7 ± 0.5
		(440)	0.313 ± 0.002	
20	57 ± 5	(222)	0.361 ± 0.007	34.5 ± 0.9
		(440)	0.389 ± 0.009	

To compound these findings with the annealing time and temperature data in Table 4.2, we see that annealing ITO is critical for achieving highly conductive films. The polycrystalline structure of the annealed ITO films is caused by the relaxation of distorted bonds in the amorphous film and a reduction of impurities [210], [212]. In the literature, it is reported that ITO has a crystallisation temperature of 400°C [213], and we found that with our equipment, a temperature of 500°C for 1 hour in an oxygen environment was sufficient to increase the conductivity to a local maximum.

4.6.   Determining Optical Properties

I have discussed how we measured the electrical properties of ITO, but to accurately determine the optical properties, we decided to use a 1D resonant grating, allowing for a guided mode resonance (GMR) (section 2.3). To relate the optical response of the grating to the complex refractive index, we need to know the grating dimensions: grating thickness, grating period, grating fill factor, and film thickness. The optical resonance, combined with these grating parameters, allows us to use rigorous coupled-wave analysis (RCWA) (section 3.2.1) to determine the complex refractive index.

Figure 4.12: Schematic of 1D periodic structure in ITO showing some of the parameters that need to be known: substrate thickness (t_{sub}), film thickness (t_{film}), grating thickness (t_g), waveguide thickness (t_{wg}), grating period (a), and grating fill factor (ff).

For the RCWA simulations of a 1D periodic structure, we need to know the parameters illustrated in Figure 4.12. We can assume that the substrate thickness (t_{sub}) is infinite due to the confinement of the light within the ITO layer. I have already discussed how we can measure the ITO film thickness (t_{film}) in section 4.5.1, and hence I will discuss in this section how we can measure the period (a), fill factor (ff), the grating thickness (t_g), and the pedestal thickness (t_p).

4.6.1.   Grating Profilometry

While we already know what the ITO film thickness is, as discussed in section 4.4.4 the grating etch is not perfectly linear and the grating may not be fully etched through to the substrate. This leaves a pedestal layer of ITO beneath the grating that shifts the optical resonance. Unfortunately, the lateral resolution on the Dektak instrument is not precise enough to resolve nanometric structures, so we need to use a different technique. We chose to use atomic force microscopy (AFM) to determine the grating thickness and to yield an estimate of the grating period and fill factor.

AFM systems achieve surface topography, and other sensing techniques, using an extremely fine tip on a micromachined probe, which is used to scan across the surface line by line in a process known as "raster scanning". There are two distinct operating modes associated with the technique, "contact" and "tapping" mode, where the distinction comes in the form of tip-sample interaction. The AFM tip is attached to a small cantilever to form a spring and, as the tip scans across the sample surface, the cantilever bends indicating a tip-sample interaction force. The

bending is detected using a laser diode and the deflection corresponds to a measurement of the surface topography. In contact mode, the tip is pressed into the surface and the tip-sample interaction force is electronically monitored to keep the deflection constant throughout the scan. In tapping mode, the contact between sample surface and the tip is limited to protect both from damage. The cantilever is, instead, vibrated near a resonant frequency and exhibits a sinusoidal motion. The motion of the tip is then altered by electrostatic forces in the tip-sample interaction to yield a topographic measurement [214].
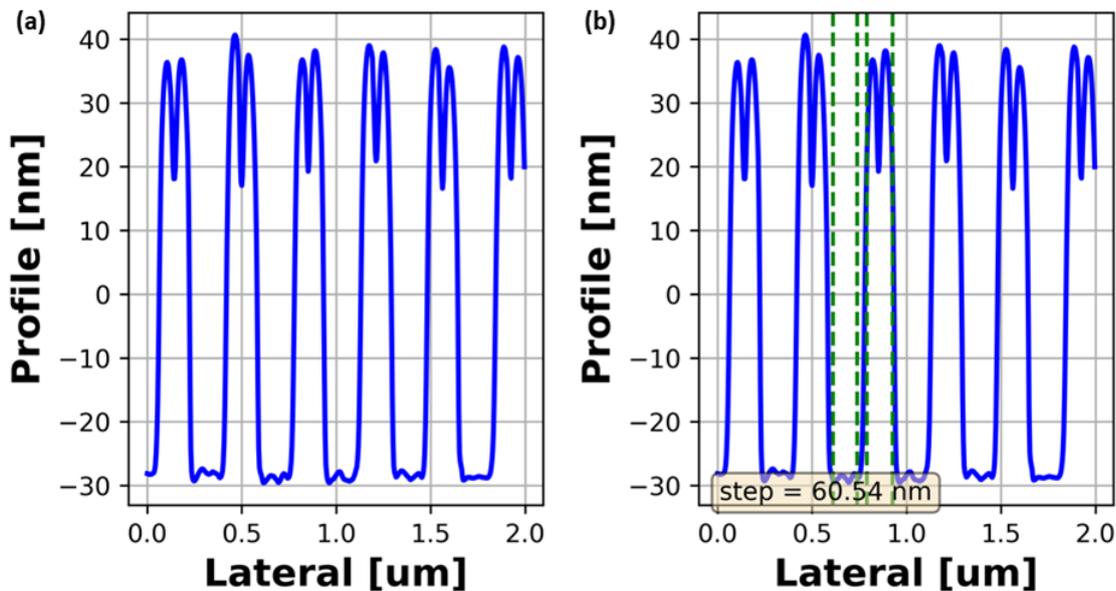
Figure 4.13: Atomic force micrographs of a 1D periodic structure in ITO showing the (a) raw and (b) step height calculated data of 60.54 nm.

To measure the grating etch depth, we used a Bruker BioScope Resolve AFM with a 2x2µm scan size with 512 lines per scan at 0.999Hz scan frequency and a Bruker RTESPA300 tip operating in tapping mode. The scans produce a topography as shown in Figure 4.13(a). To calculate the grating thickness, two regions of interest are selected (a lower and an upper region) and average values across the lateral position are taken. This method, in conjunction with the standard error on the mean method (section 4.4.5) produces two values; a lower value for the bottom of the grating and an upper value for the top of the grating, both with associated errors. The grating thickness is then calculated as the difference between the two regions and an error calculated using standard quadrature methods (Appendix section 9.1.2) [215]. The result is a grating thickness as shown in Figure 4.13(b), shown in the box at the bottom of the figure and the figure caption, and a calculated pedestal layer thickness, i.e., the remaining ITO film beneath the grating is calculated from the difference between deposited film thickness and grating thickness measurements [199].

An estimate for the grating period and fill factor is determined from the AFM measurements by taking the distance between ridges and the centre of the sidewalls, respectively. As is typical of AFM measurements, there is a component in the lateral position resolution that comes from the tip, and thus requires a complex deconvolution of the tip function and the measured topography. To complement the AFM measurements, and to get a better understanding of the grating period, we use a different method employing Fourier Transforms.

4.6.2.   Fourier Period Method

To increase the resolution of the AFM, and to determine the grating period more accurately, we decided to use a scanning electron microscope (SEM) in conjunction with a Fourier transform image analysis technique. An SEM uses electrons instead of light to form an image and yields a significantly higher resolution to image closely spaced structures. A beam of electrons is produced by an electron gun at the top of the microscope, which is then focused through a vacuum by electromagnetic fields and lenses just as in electron beam lithography (section 5.1.3). The focused electrons hit the sample and produce primary/secondary backscattered electrons and x-rays that are detected to produce an image [216].
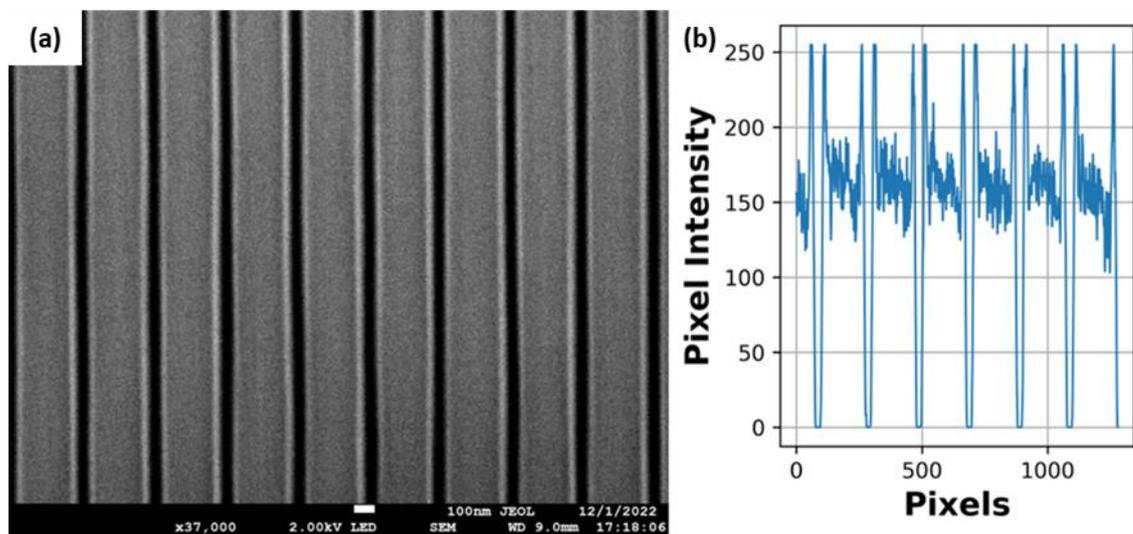
Figure 4.14: (a) Scanning electron micrograph of a 1D periodic structure in ITO, taken at 37,000x magnification, 9mm working distance, and 2kV beam power. (b) Raw pixel intensity across 1 row of an SEM image of a 1D periodic structure in ITO.

We used a JEOL JSM-7800F Prime scanning electron microscope to capture high resolution images of 1D periodic structures fabricated in ITO, with an example of these images shown in Figure 4.14(a). The SEM produces both an image and log file containing the image size (without

the border) and the pixel-distance ratio, indicated visually by the small white bar at the bottom of Figure 4.14(a). The pixel-distance ratio is important because when we take the Fourier transform of the image we obtain a frequency as a function of the pixel-space, and we want to obtain a physical distance for the grating period.

The pixel intensity is taken row-wise, Figure 4.14(b), with 960 rows per image allowing for average frequencies and standard error on the mean treatment of the frequency errors (section 4.4.5). As can be seen in both Figure 4.14(a) and (b), the edges of the grating appear brighter (white in greyscale), which indicates surface curvature altering the scattering profile of the incident electrons. The resulting edge effect on the pixel intensity introduces extra frequencies in Fourier space, thus increasing the noise of the measurement.
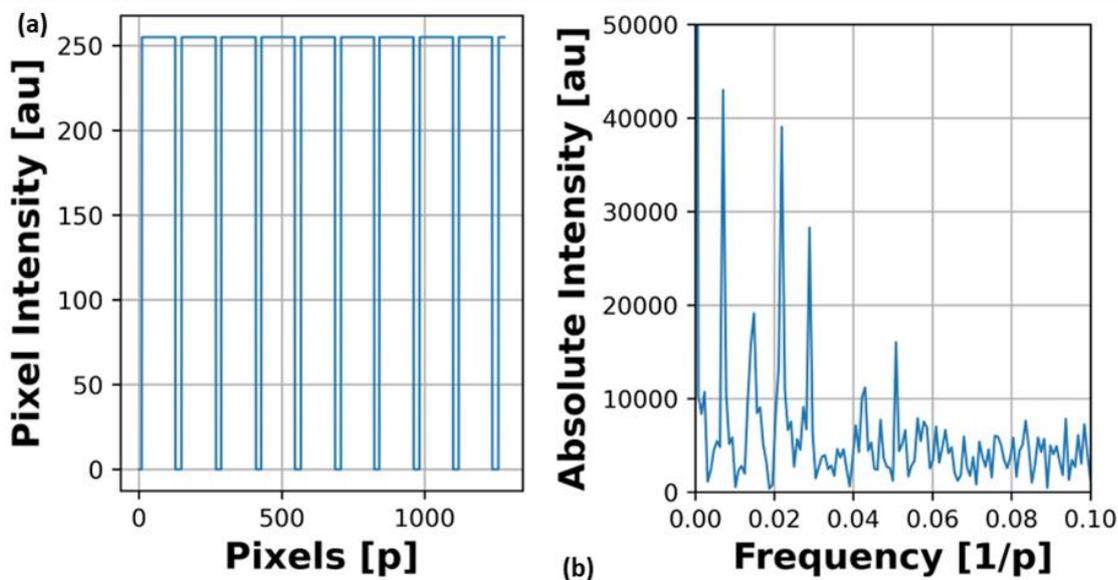
Figure 4.15: (a) Threshold treated data of 1 row of an SEM image of a 1D periodic structure in ITO and (b) the corresponding Fourier space frequencies.

Fortunately, the difference in pixel intensity between the edges of the grating and the top of the grating is minimal compared to the contrast of the bottom of the grating (Figure 4.14(a)). As such, we can implement a threshold algorithm that looks for a change in pixel intensity greater than the standard deviation of the input signal. This removes extra frequencies introduced by the edge effect, without altering the grating parameters Figure 4.15(a). The grating period is then found by taking the inverse Fourier transform of the dominant frequency in Fourier space (Figure 4.15(b)) as a function of the pixel-distance and image width (sample size), i.e., the peak corresponding to the dominant frequency in real space, which is the grating period and the usual inverse Fourier transform treatment [217].

4.6.3. Grating Resonance Method

The gratings fabricated in ITO exhibit a guided mode resonance (GMR), which I have discussed in more detail in section 2.3.1. When illuminated with a white light source, the gratings produce a Fano resonance, i.e., an asymmetric line-shape caused by the interference between the varying thin-film response and the Bragg resonance. The resonance position is determined by the grating dimensions and the refractive index of the grating material. Around the resonant wavelength, the background scattering amplitude varies slowly and the resonant scattering amplitude changes quickly, producing an asymmetric line profile.

$$f(E) = D^2 \left(\frac{(q + \Omega)^2 + \eta}{1 + \Omega^2} \right)$$

Equation 4.16

In general, the interference of states that causes a Fano resonance occurs in the absorption spectrum f(E) as a function of incident energy (E). The Fano equation (Equation 4.16) can be described by the Fano parameter (q), which is an expression of the phase shift (δ) of the continuum ($q = \cot(\delta)$), $\Omega = 2(E - E_0)/\Gamma$ where E_0 and Γ are the resonance energy and width respectively, η is a interaction constant, and $D^2 = 4\sin^2 \delta$ [218].
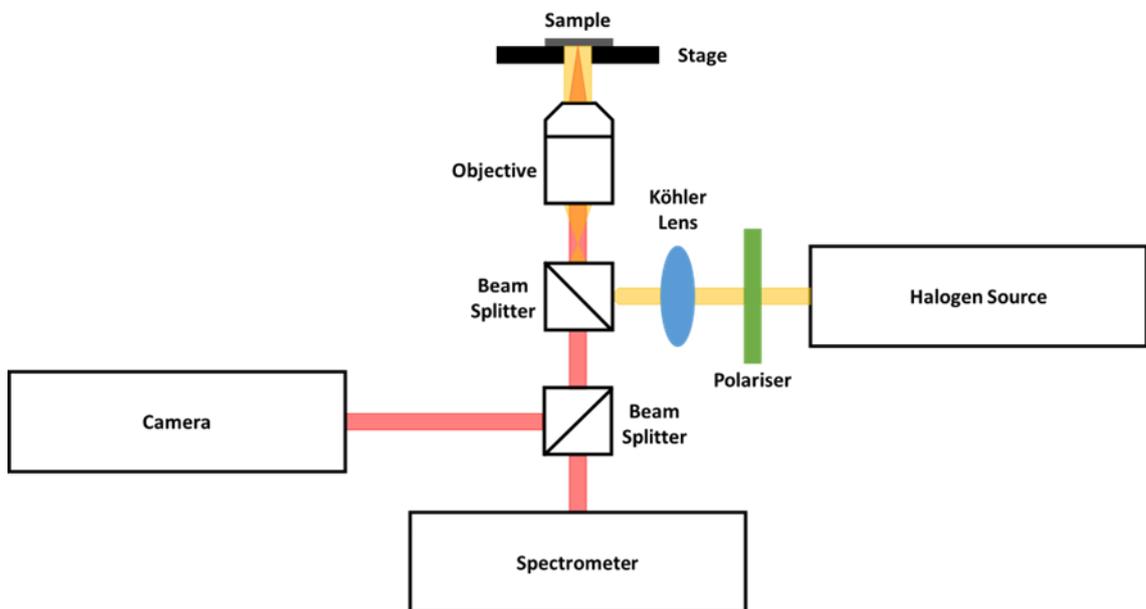
Figure 4.16: Optical setup for measuring the reflectance of guided mode resonance spectra for 1D periodic structures in ITO.

The Fano resonance response of the ITO gratings is measured using the setup illustrated in Figure 4.16. The samples are illuminated with a collimated and polarised halogen light source

being focused onto the back-focal-plane of a 4x objective. The reflected signal is collected and focused into a spectrometer. Samples are aligned in the field of view using a camera. The measured spectra are normalised to a background intensity, achieved by placing a mirror on the sample stage.
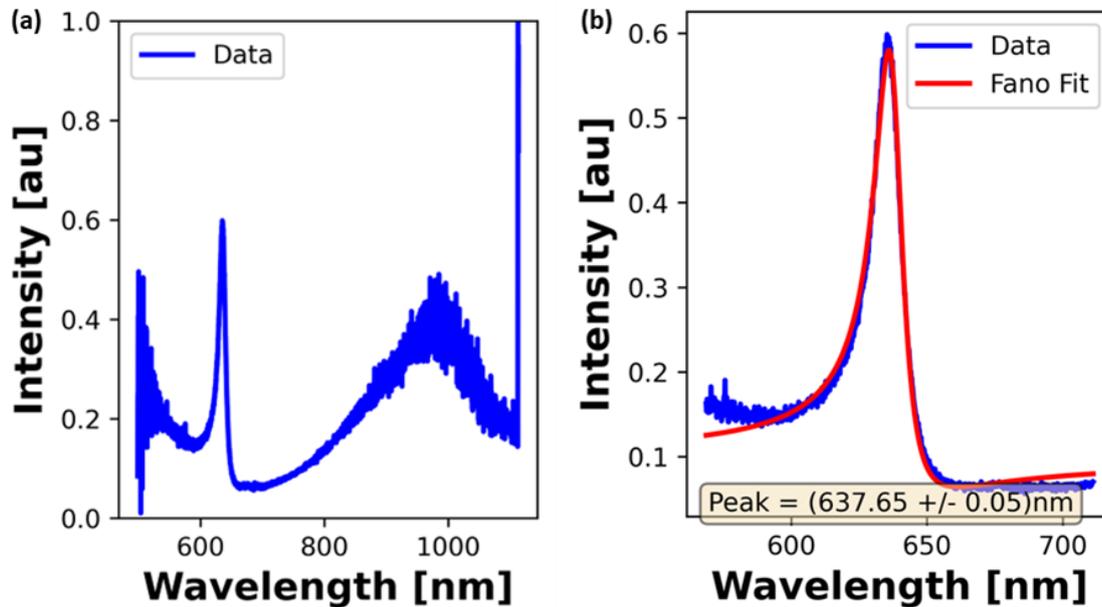
Figure 4.17: (a) Raw spectrum from 1D periodic structure in ITO and (b) the fitted Fano resonance function with peak wavelength.

An example of the normalised spectra produced by the ITO gratings is shown in Figure 4.17(a). The spectra are normalised to the background measurement with respect to individual integration times, i.e., the integration time is maximised to produce a strong peak signal and accounted for in the normalisation process. The region of interest is then trimmed down to the resonance peak, as shown in Figure 4.17(b), to allow for fast processing. The Fano resonance, Equation 4.16, is fitted to the resonance peak using nonlinear regression of the peak width, resonance position, peak amplitude, and the Fano parameter. The result is a peak wavelength and associated error calculated using the covariance matrix method (4.4.5). Data is captured for both transverse electric (TE) and transverse magnetic (TM) excitation, controlled with the polariser in Figure 4.16 [219].

4.6.4.   Rigorous Coupled-Wave Analysis Method

Here, I have employed S4 to determine the complex refractive index components (n and k) of ITO from the measured grating parameters and experimentally measured Fano resonance response. As I have discussed both GMR and RCWA elsewhere, I will instead focus on how the simulation is constructed and how I use the tool to determine n and k.
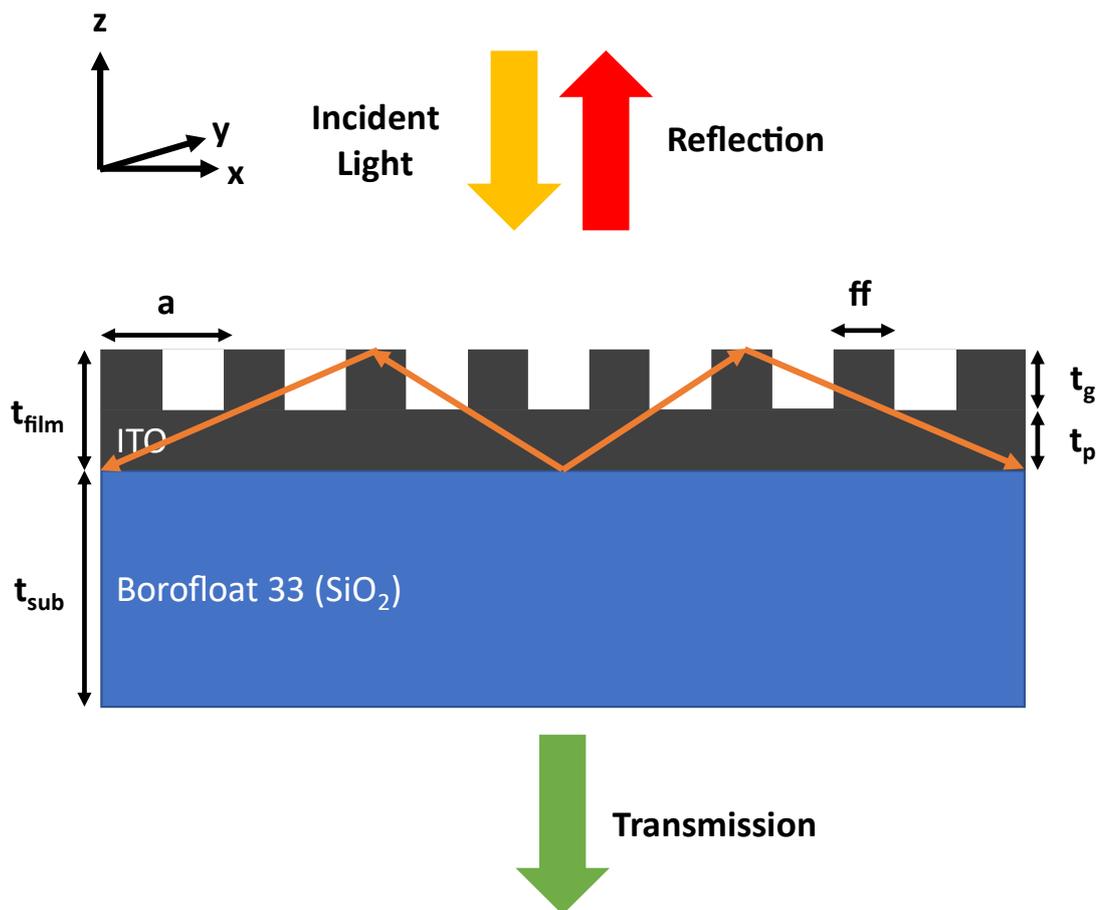
Figure 4.18: Guided mode resonance exhibited in an ITO grating, required simulation parameters labelled.

An illustration of the simulation is shown in Figure 4.18, where the grating parameters are as detailed earlier in this section. The Cartesian coordinate system is set up such that the figure is drawn in the x-z plane with y into the figure, the grating is illuminated from the top down, and hence reflection and transmission are measured from the top and bottom respectively. Each of the grating parameters, including n and k, have the potential to drastically change the position of the resonance in the spectrum. Given the volume of potential parameters, I used the S4 simulation to produce estimates of the grating parameters using a linear regression algorithm to optimise the structure and optical properties to values as close to true as is possible, and we then adjusted manually where required.

Equation 4.17

$$f(x) = \sum_{i=1}^{N-1} [100(x_i + 1 - x_i^2)2 + (1 - x_i)^2], where\ x = (x_1, \dots, x_N) \in \mathbb{R}$$

The linear regression of a multivariable system (x) can be achieved using the Rosenbrock function (Equation 4.17), utilised by many software tools to find the global minimum of the parabolic valley of the function [220]. This can be solved using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [221], which is a quasi-Newton iterative method. Essentially the algorithm comes down to finding a minimum (ideally, zero) of the differential function, the Hessian function [222], of the Rosenbrock function.

The process required several simulation variables, constants, and a figure of merit to be reduced in the regression algorithm. To begin with the variables, these were: the refractive index (n), extinction coefficient (k), grating fill factor (ff), grating period (a), grating thickness (t_g), and film thickness (t_{film}), where the waveguide thickness (t_{wg}) is then calculated as the difference between film and grating thickness values. All the structural grating variables were given starting values informed by the measured values discussed in this chapter, with upper and lower limits determined by their associated error (σ) on the order of a 10σ range. Starting estimates for n and k were taken from literature values for ITO [223] where the boundary conditions are placed at the extremes (1.4-2.5 for n, and 0.0001-2.5 for k).

The substrate material is assumed to be glass, with no loss and a refractive index of 1.45 [95], and is assumed to be infinite in all directions. This is an assumption based on the relative dimensions between the incident light (400-1000 nm) and the physical substrate thickness (~5 mm). The grating cover, providing the index contrast in the grating for resonance to occur, is assumed to be air, with no loss and an index of 1 [224]. The cover layer is also assumed to be infinite in all dimensions and is set to fill the grating gaps. S4 operates based on Fourier factorisation and requires a set number of spatial harmonics which are not indicated in the schematic. For 1D structures, 21 spatial harmonics is sufficient to achieve sub-nanometre resolution for the TE mode and 35 spatial harmonics for the TM mode.

S4 uses the complex material permittivity (ε(ω))

$$\varepsilon(\omega) = \varepsilon_r(\omega) - i\varepsilon_i(\omega) \qquad\qquad \textit{Equation 4.18}$$

which is a function of the real ($\varepsilon_r(\omega)$) and imaginary components ($\varepsilon_i(\omega)$) as shown in Equation 4.18 where the components can be expressed as a function of the refractive index (n) and extinction coefficient (k) as detailed in Equation 4.19.

$$\varepsilon_r(\omega) = n^2(\omega) - k^2(\omega), \qquad \varepsilon_i(\omega) = 2n(\omega)k(\omega) \qquad \text{Equation 4.19}$$

Hence, while I discuss the optical constants and variables for the cover, grating, and substrate layers, the simulation uses the complex permittivities to calculate the resultant resonance.

4.6.5.　　　　Optical Properties - Experimental Results
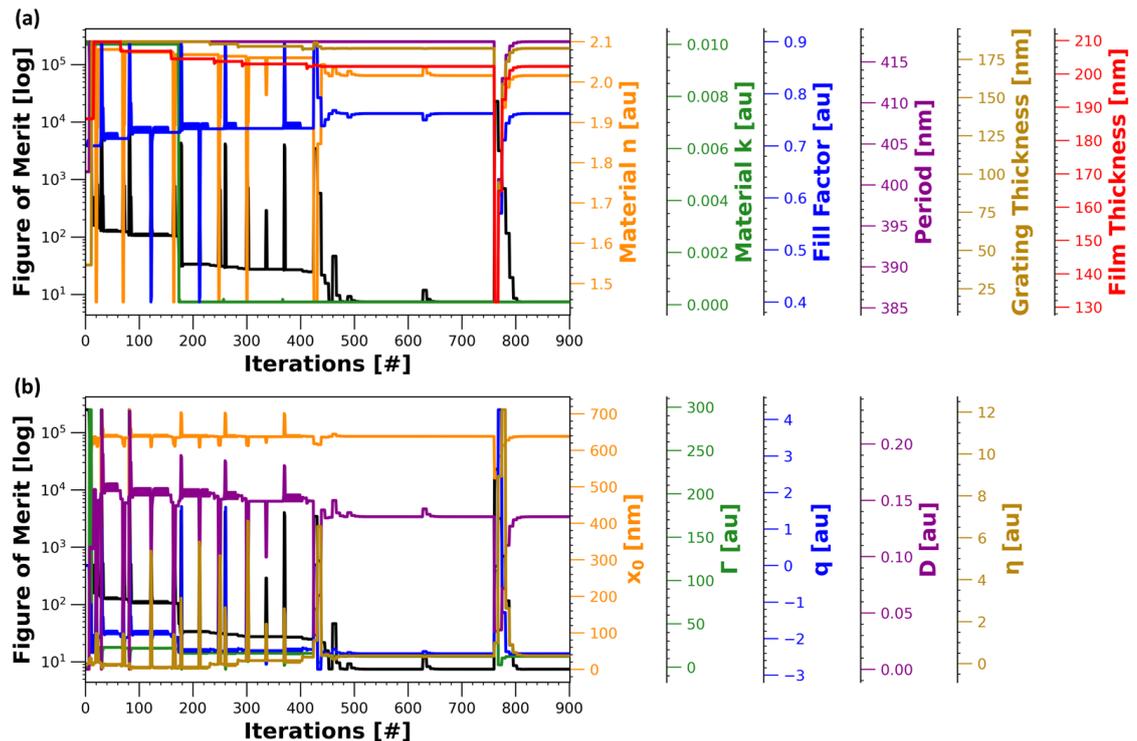
Figure 4.19: (a) Individual grating variables changing through simulation iterations and the corresponding figure of merit and (b) the simulation resonance peak Fano resonance equation parameters changing as a function of simulation iteration and the corresponding figure of merit.

The Rosenbrock regression function figure of merit, shown in Figure 4.19(a)(b), is calculated from the sum of the Fano resonance equation for both simulation and experimental resonance peak, and the overlap integral of the two curves. The former is calculated as the square of the difference between the Fano equation parameters (Figure 4.19(b)) for both the experimental and simulation Fano responses; the latter is calculated as the square of the difference between the two intensity curves. The combination of two fitting methods in the Rosenbrock function has the effect of dramatically increasing the figure of merit when there is little-to-no correlation between the two spectra, forcing the software to drastically vary the input variables shown in Figure 4.19(a), which results in the spikes observed in the black line in Figure 4.19(a)(b).
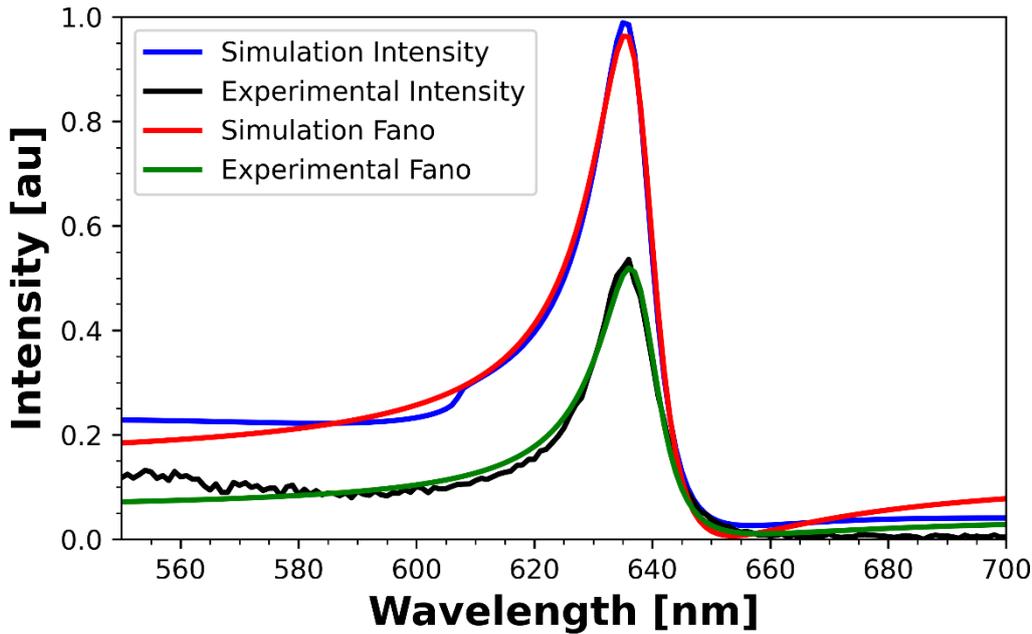
Figure 4.20: Comparison between simulated and experimental resonances for an ITO grating with fitted Fano resonances, found using minimising algorithm.

As the two errors are added up to produce one figure of merit for the algorithm, the chance that a good Fano resonance fit would dominate and produce a good peak elsewhere in the spectrum is limited, and an equally a good overlap and bad resonance cannot dominate. However, as shown in Figure 4.20, a situation may occur whereby the fits and peak positions are good, but the overall correlation is not. In this instance, we require manual adjustment of some of the input variables to obtain accurate values. Overall, this method is far quicker than adjusting the input grating parameters manually from the outset [225].

In Figure 4.20 specifically, it is the extinction coefficient (k), largely responsible for the absorption loss, which is the parameter that the algorithm has failed to optimise, thus resulting in a larger peak amplitude than expected. We found this error to be a common trend amongst many of the gratings in ITO and realised that there was more than just absorption loss in the experimental measurements. The simulated gratings are considered perfect, but the experimental measurements will have an additional scattering loss component due to surface roughness in the ITO and any minor imperfections in the gratings. To account for the scattering loss, we used the experimental peaks for the low-loss ITO material and took the amplitude difference, which is more sensitive to scattering than the sharpness (Q-factor) of the peak, to perfect reflection of the peaks and assumed that the loss exhibited by these gratings was due to scattering losses.

4.7.   Determining Drude Properties

The final step in determining the Drude equation parameters for ITO is to combine the electrical, optical, and simulation measurements in Equation 4.1 and to determine the collision frequency, effective electron mass, and high frequency permittivity. To achieve this, we calculated the expected Drude permittivity values for the measured films, at specific frequencies, and then use nonlinear regression with the measured electrical properties to determine the remaining unknown values.

4.7.1.   Nonlinear Regression Method

The key to this technique is that we only know the Drude parameters for the ITO films at the resonance frequencies of the ITO gratings, yet we want to determine the properties across the full spectrum. I use Equation 4.19 and the measured values n and k from the S4 simulation to calculate the real and imaginary components of the Drude permittivity at the resonant frequency measured using the Fano resonance in 4.6.3. The errors in the optical measurements then give an upper and lower limit in the permittivity for the nonlinear regression algorithm.

Table 4.4: Electrical parameters determined by the nonlinear Drude regression technique. The parameters correspond to the measured optical responses from ITO gratings.

O_2 Gas Flow	Sheet Resistance	Conductivity	Electron Mobility	Carrier Density
[%]	(R_S) [Ω/sq]	(σ) [S/cm]	(μ) [cm^2/Vs]	(n) [cm^{-3}] (x10^{18})
0	38.4 ± 0.5	1300 ± 200	28 ± 2	280 ± 70
5	1940 ± 30	21 ± 1	24 ± 3	4.5 ± 0.2
20	710 ± 30	57 ± 5	1.1 ± 0.7	13 ± 6
27	1120 ± 20	36 ± 3	1.4 ± 0.9	8.0 ± 0.9

The regression algorithm uses the Drude function (Equation 4.1) with the measured electrical properties of the ITO films shown in Table 4.4 . The errors in the electrical measurements provide the bounds of the nonlinear regression, allowing only minor variations in the measured parameters. Regarding the unknown parameters of the collision frequency, the high frequency permittivity, and the effective electron mass, these are given starting values from the literature [188] of 180THz, 0.35, and 3.9 respectively. The bounds placed on the unknown parameters are

set to sensible error ranges: a few orders of magnitude, 0.2, and 2, respectively; these bounds were set through experimentation and could vary depending on material [226].

4.7.2.   Drude Properties - Regression Results

Only oxygen concentrations of 0%, 5%, 20%, and 27% flow had gratings fabricated in the ITO films and are, therefore, the only oxygen concentrations that could be included in the Drude permittivity plots. The results of the nonlinear regression to determine the Drude parameters are shown in Figure 4.21.
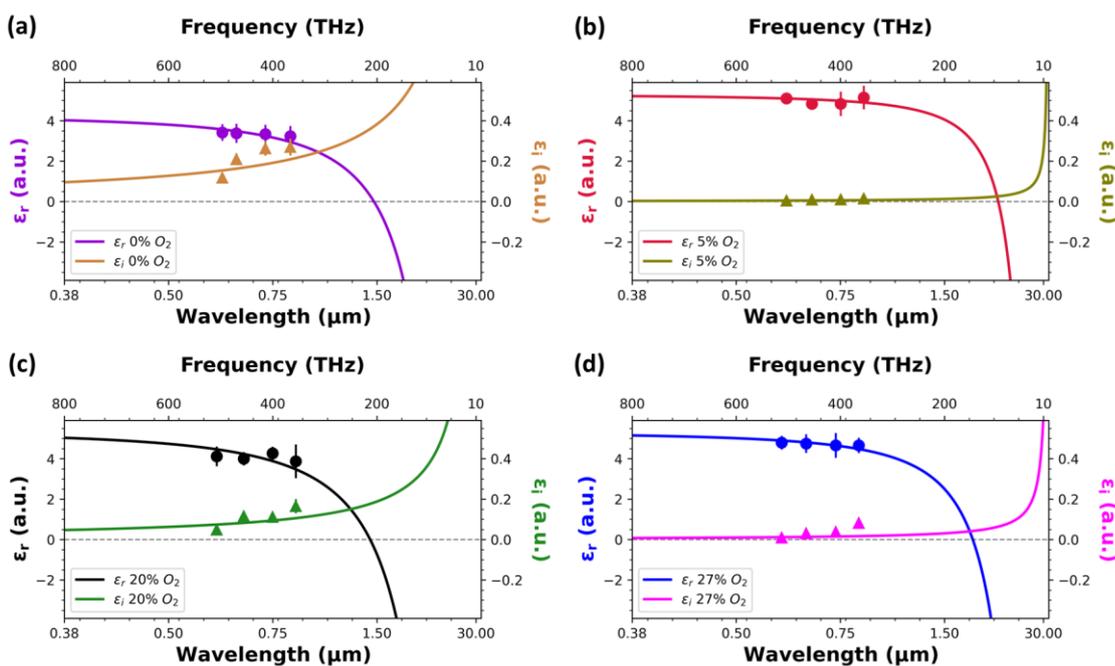
Figure 4.21: Real and imaginary dielectric permittivity as a function of wavelength (frequency) for ITO films deposited with (a) 0%, (b) 5%, (c) 20%, and (d) 27% oxygen flow.

Figure 4.21 shows how different the dispersion curves turn out as a function of oxygen incorporation during deposition. It is not as simple as x% yields a plasma frequency of y, or a high frequency permittivity of z. Each concentration of oxygen yields a different set of Drude parameters. To better understand the effect of the oxygen flow during deposition on the Drude parameters, I have summarised the results in Table 4.5, where the error values are calculated using the covariance method discussed in section 4.4.5.

Table 4.5: Summary of Drude parameters for 0%, 5%, 20%, and 27% oxygen flow concentration ITO films.

O$_2$ Gas Flow	Effective Electron Mass	High Frequency Permittivity	Collision Frequency
[%]	(m_e^*) [au]	($\varepsilon\infty$) [au]	(Γ) [THz]
0	0.5 ± 0.3	4.3 ± 0.4	147 ± 4
5	0.4 ± 0.3	5.3 ± 0.5	993 ± 2
20	0.5 ± 0.3	5.4 ± 0.6	390 ± 1
27	0.5 ± 0.3	5.3 ± 0.4	1478 ± 1

4.8.   Summary of the Investigation

The effective electron mass, defined as the apparent mass of an electron in response to an electric or magnetic field, is related to the real component of the material permittivity. One expects the effective electron mass to be higher in a material with a higher carrier density, as an increase in carriers produces a large potential in which the electron moves. A variation in effective electron mass has a similar, though inverse, effect to that of the electron mobility, shifting the epsilon near zero point to shorter wavelengths as the mass decreases. We observe this trend, in that the effective electron mass, and carrier density, is at a minimum for the ITO film with the lowest conductivity. The literature-assumed value of 0.35 is within the error range of our result, though as the effective electron mass can move the ENZ points to shorter wavelengths, we note that it is important not to assume its value.

The high frequency permittivity, the real component of the permittivity at short (ideally, zero) wavelengths is determined by the bound carriers. The assumption made by most authors is that the high frequency permittivity is constant for all ITO films with varying carrier concentrations. This assumption is clearly an oversimplification. We show that the electrical behaviour of ITO, determined by the oxygen and tin content of the material, affects the value of the high frequency permittivity drastically. The value, while not affecting the ENZ point, where many authors like to operate, affects the real and imaginary components of the complex Drude permittivity, and therefore is an important parameter to determine when operating in the visible spectrum.

Finally, we have shown that the collision frequency, assumed in the literature to be 180 or 0 THz in the literature, increases with oxygen content. We note that the high frequency permittivity has an inverse relationship with carrier mobility, which is to be expected as an increase in carrier

density is likely to lead to more collisions between carriers and thus disrupt the motion of carriers through the material.

We note that all the Drude parameters significantly affect the Drude permittivity dispersion, and simply assuming the values of any of the parameters could lead to significant errors in expected performance and device design. Using the quadrature formulation with the regression results, I estimate a 2.2% error in the fits shown in Figure 4.21. We can use the method presented in this chapter to determine the optical and electrical properties of ITO going forward, and I will use this technique to characterise the ITO used in my MOS EO modulator (section 5.7.2).

While the experimentally determined real and imaginary dispersion trends shown in Figure 4.21 are representative of those obtained by ITO ellipsometry reported recently in the literature, it would be remiss not to comment on the validity and adaptability of the methodology discussed in this chapter [80], [188]. Variable angle spectroscopic ellipsometry (VASE) is an alternative and robust measurement technique that has frequently been used to determine the optical properties of ITO [227]–[229]. VASE can provide insights into the graded nature of ITO, which our method cannot, although as with our model, VASE requires material parameter assumptions and a similarly detailed, albeit in a different field, level of understanding and knowledge to obtain the results. For a device application, such as the 1D periodic structures examined here or the modulators discussed later in the thesis, it is more practical to consider the effective index of the material, which our method readily provides. Similarly, we believe that this technique is of interest to other members of the transparent conductive oxide family, such as the crystallinity of zinc oxide films as a function of oxygen concentration during deposition [230] or the high frequency permittivity of gallium zinc oxide as a function of the deposition conditions [231].

5. Fabrication Methods

Chapter Synopsis

This chapter presents my development of the fabrication methods required to create the electro-optic modulators discussed in this thesis. I have already discussed the design and simulated the dimensions required to achieve a metal-oxide-semiconductor (MOS) electro-optic (EO) modulator based on the epsilon-near-zero (ENZ) behaviour of indium tin oxide (ITO). Simulated structures, however, very rarely translate their behaviour to real structures without a deep understanding of the available fabrication techniques and the respective limitations. The creation of the MOS EO modulator in this research required a knowledge of the material and optical theory and a creative problem-solving attitude to engineering on, at times, the nanometre scale. Without the multitude of techniques discussed in this chapter, the modulator could not have been realised. In this chapter I will discuss the fabrication and characterisation techniques used to create the MOS EO modulator. Many of the individual protocols are either well documented or known amongst the wider research group but required adapting for the materials I used or optimising to achieve the shapes, consistency, or dimensions required. Therefore, I would like to acknowledge the Photonics Research Group at the University of York for the foundations on which I could develop my protocol. I would also like to acknowledge Christina Ingleby for her assistance with the enthalpy calculations later in the chapter. Throughout this chapter I will use a specific terminology to distinguish between different aspects of the fabrication process, which I have summarised below.

Terminology

Wafer refers to the industrially manufactured substrate materials, irrespective of which materials this may refer to, all the wafers discussed in this chapter are commercially manufactured. Sample refers to a 15x15 mm square of the wafer whilst still being processed. Chip refers to the final 5x15mm cleaved sample on which the modulator is fabricated.
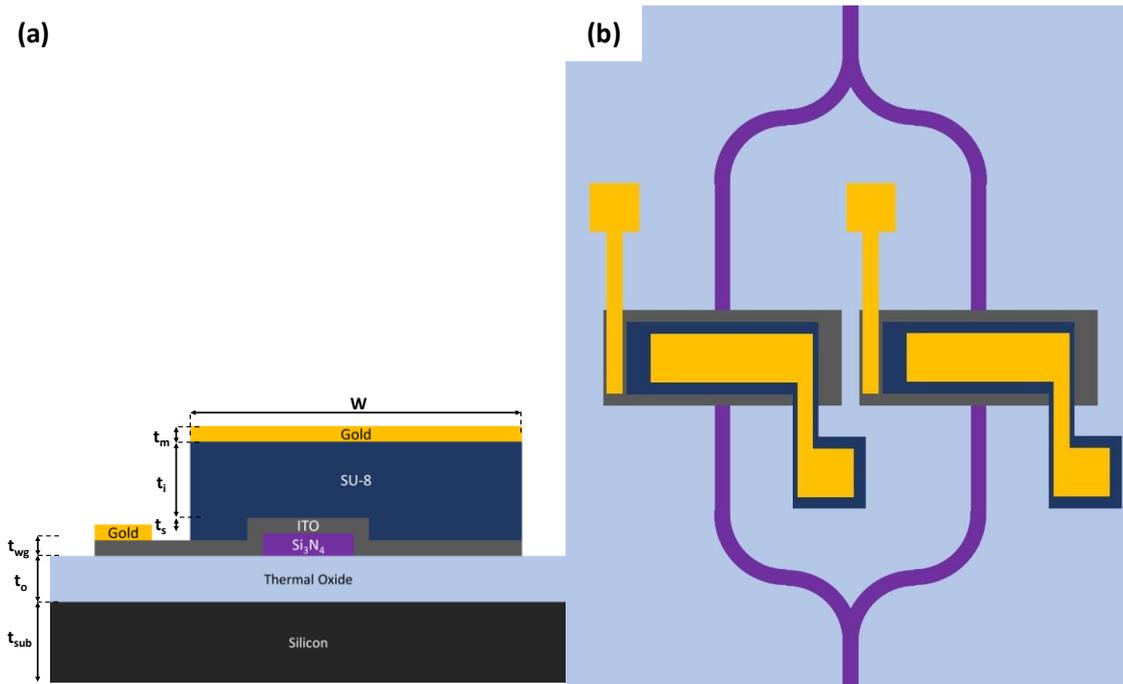
Figure 5.1: (a) Layer structure of the MOS EO modulator with dimensions: W ~ 3μm, t_{sub} ~ 525μm, t_{wg} ~ 100nm, t_i ~ 200nm, and t_m ~ 100nm. (b) Device structure of the phase modulator.

The primary design of the device, illustrated in Figure 5.1(a), includes the following, approximate, thicknesses and dimensions, discussed in section 3.5, on the order of width 3 μm and thicknesses for the waveguide 150 nm, semiconductor 100 nm, insulator 200 nm, metal 100 nm. The choice of materials from a design perspective is discussed in section 0, while here, I discuss the choice of materials from a fabrication perspective.

5.1.   Lithography Techniques

To achieve the structure shown in Figure 5.1(b), several lithographic patterning steps are required. Lithography, which will be discussed more in this section, is a process used to create structures in the horizontal plane, often referred to as patterning. In my research I have used photolithography, which is a subsection of lithography in which photons define the lateral structures, and electron beam lithography (EBL), whereby a beam of electrons defines the lateral structures. There are several steps involved in the lithography process that will be discussed in this section from both an EBL and a photolithography perspective.

5.1.1.   Application of Photoresists

Photoresists are typically long-chain polymers containing a photon or electron sensitive chemical which, when exposed, reacts to either weaken or strengthen the polymer. The reaction to exposure is dependent on positive or negative behaviour respectively, where a positive photoresists is weakened in the exposed area and a negative photoresist is strengthened in the exposed area and in both instances the weaker bonded polymer is removed during development [232].
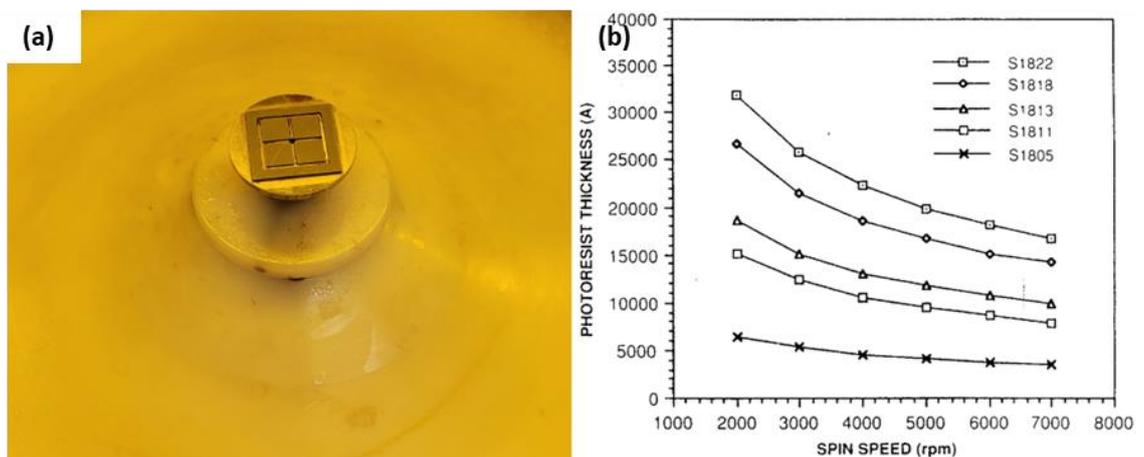
Figure 5.2: (a) Spin coater available at York and (b) photoresist thickness as a function of spin speed for S1800 photoresist series [233].

Lithography resists can be applied to a sample either through spin coating or spray coating, where the former is preferential due to the uniformity offered over larger areas [234], [235]. In spin coating, the sample is held in place using a vacuum stage (Figure 5.2(a), and the solvent is dispensed on to the sample and spun at high speeds to evenly distribute the resist across the

surface. I apply the resists using a pipette, to create a droplet of resist across the sample surface. This droplet should cover the entire surface of the sample, as it is always better to have too much photoresist than too little.

The sample is then spun at a speed that will yield a desired film thickness, as dictated by wetting properties and the viscosity of the resist; a typical set of spin speed vs thickness curves is shown in Figure 5.2(b), taken from the S1800 series datasheet [233]. Once the resists have been spun, the samples are placed onto a hotplate to allow the solvent to evaporate. This process is known as a soft bake as it removes as much solvent as necessary while maintaining a viscous, exposable layer.

Once the photoresists have been applied, the samples are covered with foil to block external UV light pre-development; while this is strictly only necessary for optical resists (not EBL resists), the practice was maintained throughout this work, as it also minimises contamination from dust and moisture.

5.1.2.   Common Spin Coating Problems

A common problem with spin coating resists is that, as the resist approaches the edge of the sample, surface tension prevents fluid flow and results in a thicker resist layer at the edge of the sample, known as "edge beads". While this effect is often minimal, it can cause under-exposed areas on the sample edge, as thicker resist requires higher exposure. To minimise the effect of this edge bead effect, it is good practice to ensure that any structures that require a high level of precision during exposure are placed near the centre of the sample, thus promoting an evenly distributed resist layer in the target area.

Another problem is known as the comet effect, where debris on the surface of the sample can cause the resist layer to streak across the sample surface as the debris moves under the tangential acceleration of the spin coater. To avoid this effect samples should be cleaned thoroughly prior to spin coating and kept in a laminar flow cabinet. To reduce the comet effect further, samples can be sprayed with high-pressure nitrogen prior to resist application.

5.1.3.   Electron Beam Lithography

EBL is a top-down nanometre-scale pattern-transfer fabrication technique, with significantly higher resolution than UVL, which typically operates on the micrometre-scale [236], [237]. EBL systems use magnetic lenses to produce a fine beam of electrons, and then steer the beam across the sample electrostatically within a fixed exposure area, known as a writefield. It is common for EBL systems to have an interferometrically controlled stage that allows the system to stitch multiple writefields together to increase the maximum exposure area. The Raith Voyager system we use has a writefield size of 500 μm and a stitching accuracy of 30-50 nm.

EBL requires the use of specifically designed e-beam resist and, in the case of patterning a dielectric, a complementary charge dissipation layer to prevent the build-up of charge on the surface negatively impacting the patterning. EBL is capable of its nanometre-scale feature sizes due to the nature of accelerated electrons, namely the wave-particle duality, where the wavelength

$$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{h}{\sqrt{2mE_{kinetic}}} \qquad \text{Equation 5.1}$$

is governed by the de Broglie relationship (Equation 5.1), where h is Planck's constant, p the linear momentum, m the electron mass, $E_{kinetic}$ the kinetic energy of the electron, and λ the de Broglie wavelength. Using the appropriate literature values and the kinetic energy of the Raith system (50 keV), the de Broglie wavelength is approximately 5 pm.

$$r_0 = \frac{1.22\lambda}{NA} \approx \frac{0.61\lambda f}{D} \qquad \text{Equation 5.2}$$

The electrons pass through a uniform circular aperture, and therefore experience Airy diffraction where the first minimum is given by Equation 5.2, where NA is the numerical aperture given as a function of the aperture (D) and focal length (f). For the Raith system we use, with apertures on the order of 10s of micrometres and a focal length of approximately 20 mm, the NA is of the order 2×10^{-3}, resulting in a spot size of order 5 nm [238], [239].
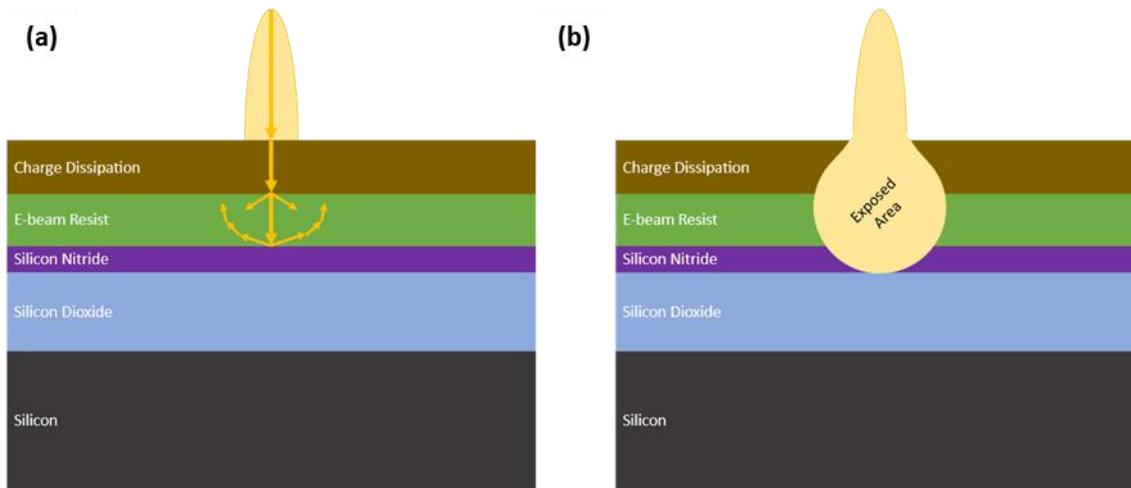
Figure 5.3: (a) Interaction of electrons with resist illustrating proximity effect and (b) the exposed area remaining.

The Raith system would form a spot this size on a bare substrate, but when incident on a resist layer, the electrons scatter both within the resist layer and in the substrate leading to a reduction in resolution, known as the "proximity effect" [240], [241]. The proximity effect is illustrated in Figure 5.3 where the scattering is shown in (a), and the resulting exposure is shown in (b). The electrons are accelerated at a high velocity, allowing them to travel fully through the charge dissipation layer and resist, where backscattered electrons and secondary electrons caused by the interaction with each layer interact as illustrated in Figure 5.3. This interaction dictates the minimum feature size, which for the Raith system is on the order of 10 nm. The exposure dose is determined by several factors; acceleration voltage, spot size, and beam dwell time which can all be controlled to optimise the dose for the target resist and resist thickness.

5.1.4.   Ultraviolet Lithography

UV-lithography (UVL) is a top-down micrometre-scale pattern-transfer fabrication technique utilising narrow-bandwidth UV light to impart energy into a photoresist to change its chemical properties. Unlike the EBL technique, the UVL available at York uses a fixed-beam moving-stage system to allow large areas to be exposed in a relatively short time. The Kloe Dilase 650 system we use has a stage size of 6 inches and can expose accurately between 1 and 100 mms^{-1} [237].
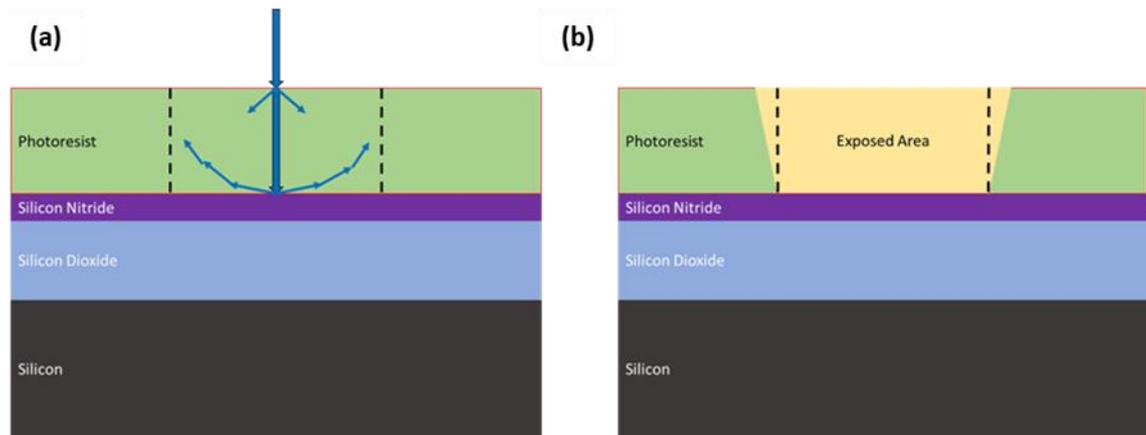
Figure 5.4: (a) Interaction of UV with resist illustrating proximity effect and (b) the exposed area remaining.

UVL is common practice in photolithography, with many specific techniques including deep- and extreme- UV lithography at very short (~100s of nanometres and ~10s of nanometres respectively), but most UV emission lines of interest are in the 300-450 nm range. The shorter wavelengths, as with EBL, are more desirable due to smaller diffraction effects leading to greater spatial resolution. UVL photoresists are typically designed for use in a set wavelength range, and it is therefore important to match the resist to the exposure wavelength. Similarly, to EBL, the focussed UV scatters on interaction with the photoresist layers and substrate to produce an angled exposed area via the proximity effect as illustrated in Figure 5.4(a) and (b).

The Kloe system we use has a wavelength of 375 nm, focussed using an aperture and 10x objective lens. The optics and wavelength can produce a spot size between 5-10 μm while the stage is stationary, where the spot size can vary with stage velocity. Exposure dose is determined by several factors; the UV source modulation power (0-100%), the stage velocity (0-100 mms^{-1}), and we can add a neutral density (ND) filter to the aperture to further reduce dose. In this work, I use UVL in conjunction with a bilayer lift off process which is detailed in section 5.3.3.

5.1.5.   Development

Exposed positive, and unexposed negative, photoresist is highly soluble in a chemical developer, which is often specifically designed by the manufacturer to dissolve the resist. The difference in solubility between the exposed and unexposed areas can be several orders of magnitude depending on the resist, leading to sharp edges created post-development [242]. Developers are specifically chosen to complement the photoresist, though typically contain a weak solution of sodium hydroxide, potassium hydroxide, or tetramethylammonium hydroxide, where the key

difference is the content of metal ions in the developer. Metal ion-based developers are typically not compatible with MOS devices due to metal contaminants and the potential for salt deposits on the sample surface. Hence, it is important to choose a suitable developer.

5.1.6.   Photoresist Outgassing

Photoresists can produce outgassing when subjected to heat or plasma during a thin film deposition under vacuum conditions. This can occur when the photoresist is not sufficiently hard baked (section 5.1.7) as there remains some solvent within the resist even after exposure and development. Outgassing will increase the scatter rate of deposited species and, potentially, cause uneven build up on the sidewall of the exposed photoresist [147]. A similar concern arises from the content of the outgassing during reactive deposition techniques, which are discussed in section 5.2. Outgassing introduces a reactive species into the vacuum chamber during deposition and may present a possible reaction between the outgassing solvent and deposited material.

5.1.7.   Hard Baking

Hard baking, as with soft baking post-spin coating, is a process designed to remove any remaining solvent in the photoresist. Typically, this process results in an increased chemical resistance and mechanical strength of the photoresist and is performed after the patterning process is complete. Baking above ~120°C results in a reshaping of the photoresist and aids in smoothing sidewalls and rounding corners of the remaining resist. SU-8 is of particular interest for this research as the result is a glass-like material that is chemically inert and structurally strong enough to withstand further processing. SU-8 is discussed in more detail in section 5.6.1, but becomes very difficult to remove once hard baked, to the point where aggressive organic removal techniques are required, such as Piranha cleaning (section 5.5.2) [243].

5.2.   Thin Film Deposition

A MOS-EO modulator, such as Figure 5.1(a), requires at least 2 layers that cannot be spun, exposed, and hard baked into place; namely the semiconductor and electrical contact layers. These layers need to be deposited using a physical vapour deposition technique such as sputtering or evaporation.

5.2.1.   Physical Vapour Deposition

Physical vapour deposition (PVD) techniques, involves the vapour phase deposition of a solid material at a controlled rate. PVD techniques include thermal evaporation and sputter deposition, where the distinction is the energy used to vaporise the solid material. As the name suggests, thermal evaporation techniques involve the transfer of thermal energy to the solid material and, as such, should not be used with materials where the deposition temperature can affect the resultant thin film. Sputter deposition is a cold deposition technique as the exchange of energy is due to a bombardment of high energy particles [147].

I use a sputter deposition technique to deposit thin films of indium tin oxide (ITO), due to the thermal sensitivity of the material reported in the literature [244]. As for the electrical contacts to the MOS capacitor, which are predominantly gold, a thermal evaporation is more appropriate. The justification for this decision is detailed in section 5.2.9, but here I will detail both techniques.

5.2.2.   Sputter Deposition

Sputtering is a thin film deposition technique common in the fabrication of semiconductor and optical devices and can be considered a sublimation process, as the solid material is turned directly into a vapour without melting. I will give a brief overview of the sputtering process performed on the bespoke pulsed direct current (DC) reactive magnetron sputterer available to my research group at York, shown Figure 5.5(a). A schematic of the bespoke system is shown in Figure 5.5(b) and is explained below.
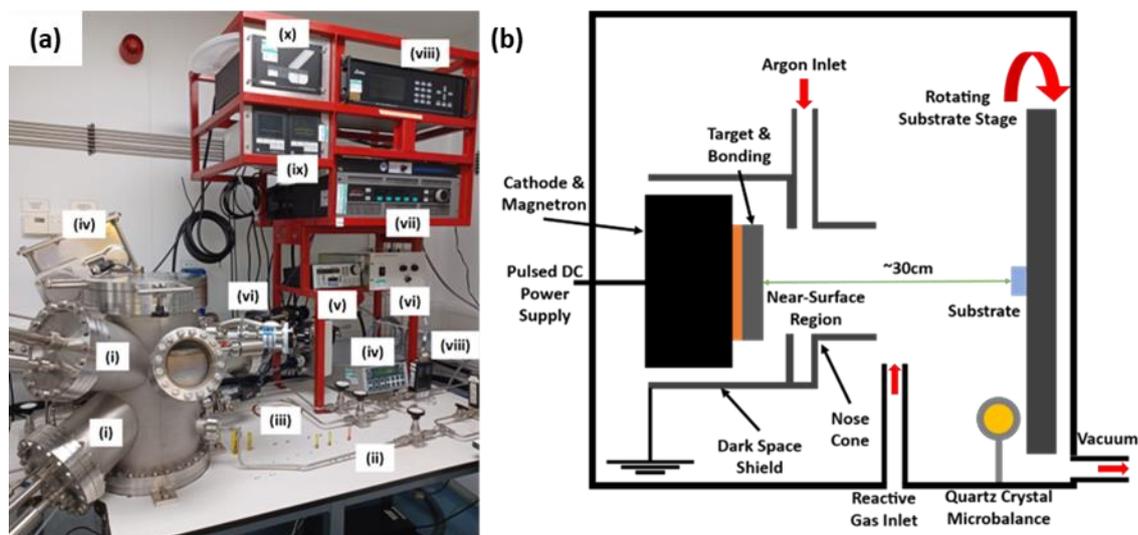
Figure 5.5: (a) Photograph and (b) schematic of the bespoke pulsed DC reactive magnetron sputterer at York, schematic not to scale.

Sputtering is the process of removing surface atoms from a target material through the bombardment of high energy inert gas ions, due to an exchange in kinetic energy between ion and target atom. If the kinetic energy of the ion is larger than the binding energy of the atom, the target atom is ejected upon collision. Momentum and energy are always conserved during the process. A negatively charged sputtering target allows free electrons, within the target, to flow away from the cathode and in to the near-surface region in Figure 5.5(b). An injected inert gas, typically argon, is present in the same region. Collisions between the target electrons and outer shell argon electrons, and the conservation of momentum, results in positively charged gas ions (a plasma).

The resulting plasma is accelerated toward the target due to the potential difference between the dark space shield and cathode. The positively charged ions impact upon the negatively charged target and eject surface atoms at high velocities, through the exchange of momentum, towards the anode. The ejected target atoms are projected across a vacuum chamber towards a substrate, where they condense and form the desired thin film [245].

Sputtering is an inherently omnidirectional process and can have a low target-to-film yield [102]. To improve efficiency, our bespoke system in Figure 5.5(b) includes a nose cone aimed at the rotating substrate stage. Conduction between cathode and target is aided by the inclusion of a copper backing plate, highlighted in orange in the schematic. The sputter target is held in place by a locking ring, highlighted in Figure 5.6(b). Deposition rate and thickness through the use of a quartz crystal microbalance (QCM) [246].

The bespoke sputtering system in Figure 5.5(a) is labelled to include; (i) sputtering guns, (ii) reactive gas inlet, (iii) argon gas inlet, (iv) gate and gate valve controller, (v) QCM, (vi) substrate stage and rotation controller, (vii) pulsed DC power supply, (viii) gas flow controller, (ix) vacuum pressure monitor, and (x) vacuum pump controller.

5.2.3.   Magnetron Sputtering

In conventional DC sputtering, electrons are released from the cathode through secondary electron emission and accelerated toward the anode. These electrons create more electrons and ions through impact ionisation, and the ions return to the cathode to release secondary electrons to sustain the plasma. This requires many impact ionisation events to sustain the plasma, therefore reducing the energy of the ions. The result is a low sputter yield and heating of the plasma and substrate, which is not ideal for thermally sensitive materials.
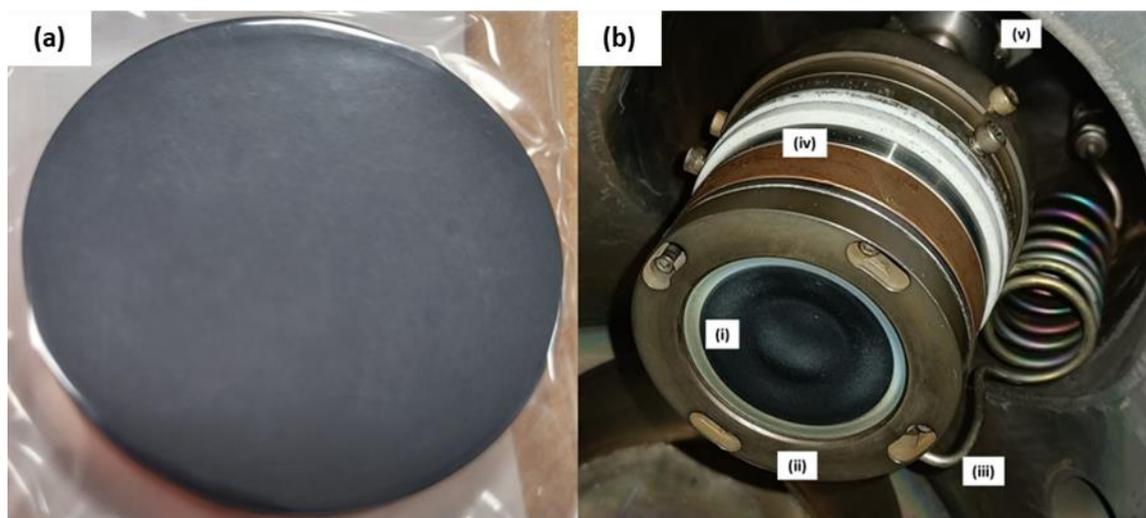
Figure 5.6: Indium tin oxide sputter target (a) new and (b) placed in the sputter gun, clamped into place, with dark space shield removed.

The internals of the bespoke sputtering system in Figure 5.6(b) are labelled to include (i) indium tin oxide sputtering target, (ii) target retaining ring, (iii) argon gas inlet (unattached), (iv) magnetron and cathode, and (v) pulsed DC power supply and water cooling into the chamber.

The sputtering yield and the energy exchange between gas ions and target atoms have been extensively studied in the literature [247]–[250] but are critical for the thin film deposition. For thermally sensitive materials, such as ITO, it is preferential to operate the plasma at low pressure and temperature, which requires the use of a magnetron which is placed behind the target in Figure 5.6(b) [251].

The magnetic field exerts a force on the electrons proportional to their velocity but with perpendicular direction, thus confining them to circular motion above the surface of the target. At low pressure, the electrons move significantly faster than the argon atoms, as such the electrons move circularly with their motion dominated by the magnetron, whereas the argon atoms (ionised by the electrons) have their motion dominated by the electric field provided by the sputter gun. Thus, ionised argon atoms are accelerated toward the target in a circular path, creating the commonly name "racetrack" effect observed in Figure 5.6(b) [250].

5.2.4.   Sputtering of Compounds

When compounds are sputtered, ion bombardment will cause the compounds to be ejected both as atoms and as entire molecules. Different atoms also have different sputter yields due to their different masses and diameters. If one type of atom has a higher yield than the other, it will be preferentially ejected over the other types. This may seem like a serious problem when sputtering compound materials, but it has been found that a certain steady-state composition and depth profile are quickly reached where the removal of both elements occurs at the original stochiometric composition [252]. This should not be surprising because the only way to consume a target to its entirety is by depleting its constituent elements in their original proportions.

5.2.5.   Reactive Sputtering

Sputtered ITO is the critical layer that dictates the performance of the EO modulator, as such the deposition parameters need to be finely controlled. We use an indium tin oxide target (In_2O_3-SnO_2 90-10 wt%) as the base material. Since the oxygen is volatile, some of it is lost during the process, which is compensated for by introducing oxygen during deposition. We have also found that oxygen backpressure impacts on the electrical and optical properties of the deposited film [253]. Oxygen can be introduced to the sputter process as a reactive gas either in the plasma (argon) gas inlet or the reactive gas inlet as an oxygen backpressure (Figure 5.5(b)).

Firstly, let's consider oxygen in the plasma gas. Oxygen tends to form negative ions, which is a concern during sputtering, as the ions are accelerated away from the target by the same electric field accelerating positive argon ions to the target. High energy negative ions impacting on the growing film can damage or alter the properties of the ITO film [254]–[256].

Introducing oxygen as a backpressure in the chamber, via the reactive gas inlet, doesn't come without its own complications. Oxygen in the sputter chamber can lead to the formation of a compound layer (e.g., an oxide layer) on the surface of the target, also known as "target poisoning". Target poisoning changes the sputter yield as the surface of the sputter target is now of different composition, and affects all discharge parameters within the chamber [250]. A compound layer forms for many reasons, primarily due to low energy atoms from the reactive gas bonding with the surface atoms of the target, but also due to recoil-implantation of reactive gas atoms due to interactions with the high energy ions above the target surface [257], [258].

Reactive gas atoms are not the only source of compound layers, it has also been shown that directly implanting reactive elements into the chamber can cause compound layers to form [250], [259]. An example of directly implanting reactive elements into the chamber would be the inclusion of an out-gassing photoresist (section 5.1.6). Another source of target poisoning is the oxygen that is introduced as a background pressure in the sputtering chamber through the reactive gas inlet in Figure 5.5(b). To reduce the effect of oxygen poisoning the target, the target is cleaned both mechanically, prior to sputtering, and with an argon plasma in the chamber prior to sputtering. Similarly, a minimal layer of ITO is deposited to reduce the contamination.

The main concern about oxygen incorporation into the sputtered ITO film, however, is whether the oxygen and the sputtered ITO particles leaving the target have a chance to interact before the ITO particles reach the sample. This chance of interaction is governed by the mean free path.

$$\lambda = \frac{k_B T}{p \pi d_m^2}$$

Equation 5.3

For a typical sputtering pressure (p) of 0.75 Pa and temperature 300K, with the Boltzmann constant (k_B), the atomic diameter of a tin atom of 300 pm and an indium atom of 400 pm, the mean free path can be estimated at 1-2 cm. This mean free path is significantly smaller than the 30cm distance between target and substrate, so there is plenty of opportunity for interaction and reactive gas incorporation. In fact, the mean free path is likely to be smaller than the approximate 2 cm due to target atom cluster ejections rather than the assumed single atom ejection in the mean free path equation (Equation 5.3) [260].

The introduction of a reactive gas can change the plasma pressure which, for a constant DC power, changes the ion density in the plasma proportionally. For high pressures, the sputter removal rate on the target will decrease due to a drop in ion voltage and an increased scattering distribution of the target atoms. To maintain a constant pressure in the sputtering chamber, a gate valve on the vacuum outlet in Figure 5.5(b) is therefore used.

5.2.6.   Pulsed DC Sputtering

Sputtering is a technique suitable for metal and dielectric material depositions, with the latter resulting in a non-conductive coating on every internal surface of the sputtering chamber due to the inherent omnidirectionality. Free charges from the plasma can build up on these surfaces, creating regions of high charge and potentially resulting in charge arcing. Charge arcing can strike the sputter target and cause nonuniform removal of target material and thus nonuniformities in the deposited films [261].

To mitigate arcing and prevent potential damage to the target, there are two competing methods, pulsed DC sputtering and radiofrequency (RF) sputtering. Both use a voltage-off (or voltage-negative for RF) regime to allow dielectric regions to discharge to prevent arcing. Pulsed DC sputtering, however, offers a greater control within the manufacturing process as it has a greater versatility on sputtering pressure, and therefore sputtering rate. RF sputtering typically, as well as having a slower deposition rate, requires significantly more power and a more complex integration with the sputtering system. To mitigate the arcing and prevent potential damage to the target, we use a pulsed DC signal provided by the power supply that oscillates at 120 kHz [262], [263].

5.2.7.   Thermal Evaporation

Thermal evaporation exploits the principle of heating a material until it evaporates at a sufficiently high rate to produce deposition on an adjacent surface through condensation. The procedure is quite simple as the temperature required for the process is only dependent on the deposition material and the pressure. The evaporation is therefore performed under vacuum condition, also to prevent any unwanted background gases interfering with the evaporation or reacting with the material.

Thermal evaporation aims to achieve a sufficiently high vapour pressure to evaporate the material and condense on a substrate and not to necessarily melt the material, though this may still happen depending on the material. Naturally, as melting is always a possibility, the evaporation direction must be vertically upwards to avoid unnecessary material loss. Most thermal evaporation techniques involve heating a container of material through resistive heating methods, but less energy consumption and higher yield can be achieved if we heat from the top surface of the material with an electron beam or via laser ablation [147].

5.2.8.   Electron Beam Evaporation

To increase efficiency, we can heat the material, which is typically stored in a crucible as a pellet, using a beam of electrons from the top down. This allows the top surface to be both the evaporative surface and the heated surface. We can further protect the bulk material by water cooling the underside of the crucible.
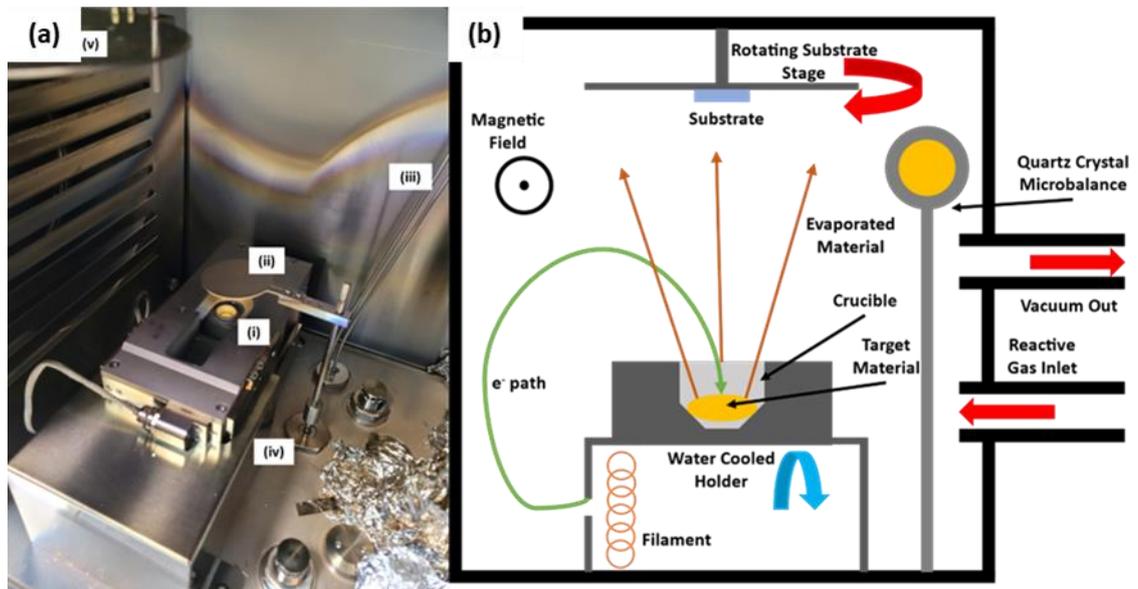
Figure 5.7: (a) Photograph of material crucible, electron beam filament, and water-cooling stage at York, (b) schematic of electron beam evaporator at York.

A schematic of the MBRAUN EVAP [264] electron beam (e-beam) evaporator available at York is illustrated in Figure 5.7(b), highlighting some key features of e-beam evaporation. A tungsten filament, held at negative potential, generates free electrons, and accelerates them away from the filament where they are guided toward the centre of the crucible by a series of beam forming plates, static magnets, and electromagnets. The latter is typically used to make small adjustments to the beam spot to evenly heat the surface of the material in the crucible. The background pressure in the chamber needs to be significantly low to allow a long mean free path for the electrons to form a focussed beam spot, typically on the scale of 1 mm^2. A spot size on this scale, with an electron penetration depth on the micrometre scale, allows for a very high power-density and a high deceleration of electrons on the material surface, resulting in very low ohmic heating of the bulk material.

The high-energy electrons displace surface atoms through the transfer of kinetic energy, where the evaporation rate is proportional to the electron energy. Naturally, the ejection of surface atoms in a vacuum chamber is omnidirectional, so the sample onto which we want to deposit

must be directly above the crucible, and the deposition rate and thickness is monitored with a QCM. The sample is affixed to a rotational stage to encourage a more even surface deposition.

Temperatures within the crucible can reach values on the order of 500°C, and so to protect the crucible and bulk material we need to introduce a water-cooling system. Crucibles are usually made of copper for the thermal conductivity to increase the cooling efficiency, and the evaporation material is placed in the crucible as pellets. The problem with using a thermally conductive crucible is that a higher beam power is required to maintain surface temperature, so often a graphite crucible liner is used to reduce the thermal conductivity and transfer of heat from beam to copper crucible [147], [265], [266].

5.2.9.   Deposition Technique Justifications

Throughout section 5.2, I have commented that I used a sputtering technique and e-beam evaporation technique for the deposition of ITO and gold respectively, though the two techniques I detailed are both physically suitable for the deposition of both materials. In principle and discounting the efficient material use in e-beam evaporation and the economic benefits therein, I could have used the commercially designed and programmed e-beam evaporator to deposit both materials. While ITO is discussed further in chapter 4, it is known that the optical and electrical properties are highly dependent on the deposition and post-deposition protocols, namely that there is an increase in crystallinity with deposition and annealing temperatures [191], [192]. It is therefore important that we consider the deposition parameters when choosing the technique.

I will first consider the rate equations. As has been discussed, e-beam evaporation is a thermal process, where a constant rate of emission requires the thermal energy transfer to overcome the material's binding energy.

$$Z_A = 3.5x10^{22} \left(\frac{P}{\sqrt{MT}} \right)$$ *Equation 5.4*

The atoms leaving the materials surface is given by Equation 5.4, where the rate of atoms per square centimetre per second (Z_A) leaving the surface is a function of the vapour pressure (P), the molar mass (M) and the temperature (T). It follows that, when multiplying the removal rate and the evaporation surface area (A), which we can assume is approximately 10x that of the spot size, we obtain an atom removal rate (R_T)

$$R_T = Z_A A \qquad\qquad \textit{Equation 5.5}$$

described by Equation 5.5. Naturally, in an omnidirectional procedure such as e-beam evaporation and sputtering, the evaporation rate from the source material will not be the same as the deposition rate. The angular distribution of the evaporation flux, the source-to-substrate distance, and collisions in the vapour will naturally interfere with the angle of deposition. However, we can assume that the source-to-substrate distance is significantly greater than the evaporation spot size, and thus an approximate point source treatment yields a flux rate (ϕ)

$$\Phi = \frac{R_T}{\pi D^2} \qquad\qquad \textit{Equation 5.6}$$

which is inversely proportional to the square of source-to-substrate distance (D) (Equation 5.6). Finally, the film growth is given by the incident flux as a function of the material molar mass (M) and the material density (ρ) (Equation 5.7).

$$r = \frac{\Phi}{\left(\frac{\rho}{M}\right) N_A} \qquad\qquad \textit{Equation 5.7}$$

Where N_A is the usual notation for Avogadro's constant. It is clear then, for e-beam evaporation, that the deposition rate is highly dependent on both the material and the evaporation temperature. Sputtering, however, is considered a low temperature process and is instead dependent on the pressure-voltage relationship of the plasma. The rate of removal (R_T)

$$R_T = \frac{I}{q} S \qquad\qquad \textit{Equation 5.8}$$

in a sputtering system is given by the incident ion energy (I) and the sputter yield (S) (Equation 5.8). Where q is the usual notation for the electron charge. The same flux treatment as with e-beam evaporation can be used as an approximate here, giving a flux rate as in Equation 5.6, and a deposition rate as in Equation 5.7 [147]. To determine a relationship between evaporation temperature and evaporation rate, for the available materials in our e-beam evaporator, I first need to consider the vapour pressure.

$$\ln\left(\frac{P_s^{sub}}{P_l^{sub}}\right) = -\left(\frac{\Delta_{fus}H}{R}\right)\left(\frac{1}{T_{sub}} - \frac{1}{T_{fus}}\right) \qquad\qquad \textit{Equation 5.9}$$

During a deposition, the vapor pressure will be in equilibrium, meaning we can consider the vapor pressure to be given as a function of the material enthalpy and temperatures of phase transitions (Equation 5.9) [267]. Where P_s^{sub} and P_l^{sub} are the sublimation pressure of the solid component and the vapour pressure of the liquid component respectively, $\Delta_{fus}H$ is the enthalpy of fusion, R is the gas constant, and T_{sub} and T_{fus} are the sublimation and melting temperature respectively. Rearranging Equation 5.9 for P_s^{sub}, and using literature values for the temperatures

and enthalpies [268]–[275], yielded the material dependent vapour pressure, which can be substituted into Equation 5.4.

The results of the approximate temperature-rate relationship for the available materials in our e-beam system are shown in Figure 5.8(a). The removal of an atom from a sputter target requires the incident ion energy to be larger than the binding energy of the target atoms, and the energy delivered upon collision is a function of the mass ratio between ion and atom. Argon is the most common gas used in sputtering systems and is the one I have considered here. The sputtering yield, S,

$$S = A(m)\left(\sqrt{I} - \sqrt{I_{th}}\right)$$ *Equation 5.10*

is the number of ejected target atoms for each incident ion, and it can be calculated using the ion energy (I), a threshold energy (I_{th}) and a mass proportional constant (A) (Equation 5.10) [276], [277]. Literature values for the sputtering yield of different materials were used to approximate the ion energy dependency on the yield [278], but the ion energy is itself dependent on the potential gradient between the anode and cathode. Therefore, I will assume a constant DC power of 100W, and use the Paschen relationship [279]

$$V_B = \frac{Bpd}{\ln(Apd) - \ln\left[\ln\left(1 + \left(\frac{1}{\gamma}\right)\right)\right]}$$ *Equation 5.11*

to calculate the ion potential. Where A and B are gas constants, and γ is the secondary scattering coefficient. The constants are well documented, and the cathode-anode separation (d) is fixed, so the plasma pressure becomes the driving variable in ion potential, and eventually deposition rate.

The constants A and B in Equation 5.11 are temperature dependent, though standard literature values for room temperature plasma (T = 293K) were used in this approximation. Using a simple P = IV relationship between power (P), current (I), and potential (V), I am able to approximate the ion energy and use the respective sputter yield to calculate rate using Equation 5.8, Equation 5.6, and Equation 5.7. The resultant approximation is shown in Figure 5.8(b).
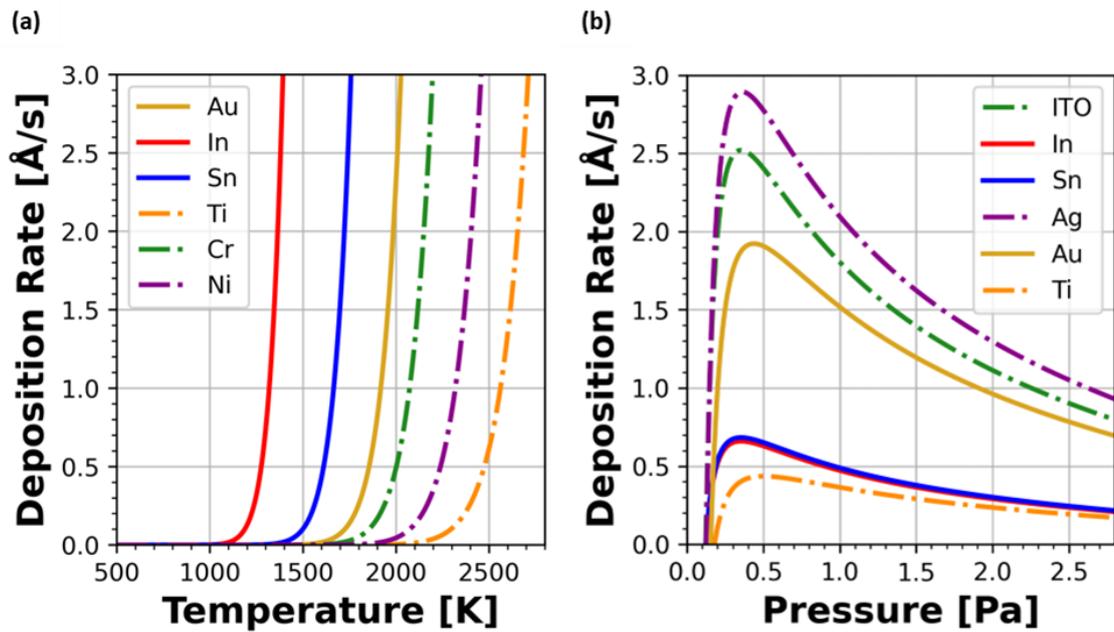
Figure 5.8: Approximation of deposition rate as a function of (a) evaporation temperature for electron beam evaporation of common materials, and (b) plasma pressure for pulsed DC reactive magnetron sputtering of common materials.

The approximate effect of temperature and plasma pressure on the deposition rate in e-beam evaporation and sputtering is shown in Figure 5.8, (a) and (b) respectively. The materials represented in Figure 5.8(a) are all common materials we can evaporate using our e-beam evaporator, specifically those shown with solid lines are the desired materials for this work, while the dashed lines are available adhesion layers. The materials represented in Figure 5.8(b) are all materials readily available for sputtering using our bespoke system, though silver is only represented to highlight the significant rate when considering ITO and not its constituents. While these values may only be approximate, as the true deposition rate is also dependent on the mean free path of the atoms, the random nature of collisions in a vapour, and angle of the removal flux, the values do agree with measure deposition rates. For instance, I sputter ITO at 0.75 Pa and achieve a rate of 0.3-0.5 Å/s which corresponds to the rate of indium and tin in Figure 5.8(b) of 0.6 Å/s. The difference in rate, between expected and achieved, is negligible when considering the assumptions made, thus the results highlight the low temperature nature of sputtering. The rate of ITO in Figure 5.8 is of similar magnitude to that of silver, which highlights the earlier discussion of compound sputtering (section 5.2.4). As for e-beam evaporation, which is a commercially controlled evaporation, a rate of 0.3 and 0.5 Å/s is set for gold and ITO respectively. Using the data in Figure 5.8(a), I can say that the gold is evaporated

at approximately 1800K and ITO is evaporated at approximately 1200-1500K, assuming that the ITO equilibrium is between indium and tin evaporations.

The deposition temperature of ITO is so important as it is largely proportional to the rate of disassociation of oxygen from the indium and tin oxides within the target material. As the temperature is significantly higher in e-beam evaporation, it follows that more oxygen is lost from the source material through this deposition technique. As oxygen in ITO is largely responsible for free electron density, and thus the electrical and optical properties desirable for an EO modulator (see section), depositing using e-beam evaporation increases the difficulty of optimising oxygen content. Of equal consideration is the mean free path in both techniques. I have shown that the mean free path of an indium or tin atom is of the order 1-2cm in the sputter chamber. Due to the decreased pressure during evaporation, caused by the increase in deposition temperature, the mean free path of an indium or tin atom in the evaporation chamber would be 2-3 orders of magnitude larger. This increase in mean free path significantly hampers oxygen reincorporation during deposition.

5.3.   Lithographic Patterning

Lithographic patterning is more than simply spin coating photoresists and depositing metals, however. Typically, the photoresist serves as a sacrificial layer and is removed through development (section 0), so to transfer the pattern to the substrate we need to employ two other techniques. These techniques involve either lifting off the deposited materials from areas where it is not needed or using the photoresist as a top-down etch mask. These two methods are commonly known as lift-off and etch-down, respectively, although specifically I use a bilayer lift-off and an etching technique known as reactive ion etching, that I will detail in this section.

5.3.1.   Reactive Ion Etching

Reactive ion etching is an etch-down method, where the pattern exposed lithographically is etched into the film beneath the photoresist. The method of deposition is not important in etch-down, but in the interest of this research, I will discuss the etch-down method of transferring a pattern into the silicon nitride layer in Figure 5.1(a). The photoresist is applied, exposed, and developed as detailed in section 0, and the remaining photoresist acts as a chemical barrier allowing us to etch into the underlying material. Plasma etching, or dry etching, is the most

suitable method of etching into materials deposited using plasma enhanced chemical vapor deposition (PECVD), due to the significant reduction in etching temperature required to remove a material, using a plasma, that was deposited using a plasma. In any chemically reactive process, deposition and removal take place simultaneously, therefore plasma etching is an ideal technique for the targeted removal of the silicon nitride layer used in this research, as it was deposited using PECVD.
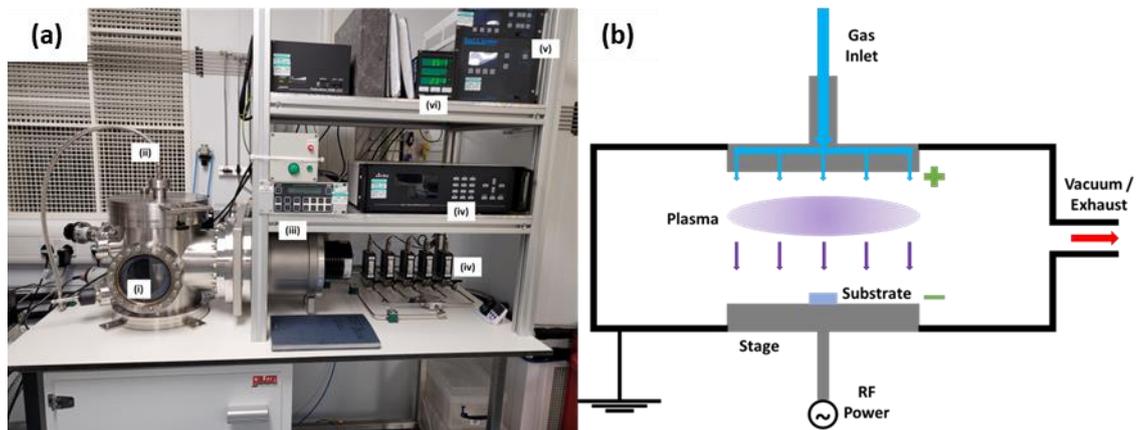
Figure 5.9: (a) Photograph and (b) schematic of the bespoke reactive ion etcher at York.

The basic plasma etching system consists of a vacuum chamber, controlled gas inlet, a gate-controlled exhaust to maintain chamber pressure, cathode, anode, and an RF power supply, as shown in Figure 5.9(b). The cathode is driven at 13.56 MHz at 10-100 W, which is a sufficient power to ionise the process gas and to produce free radical species that are responsible for the low substrate temperature. In the plasma there are ions, free radicals, and unexcited feed gas molecules, where any positive ions generated bombard the cathode and produce secondary electrons to sustain the plasma and the free radicals are primarily responsible for the substrate etching.

Dry etching can be considered as a combination of mechanical and chemical etching, as the free radicals are not driven by any force within the chamber, but instead diffuse toward the substrate. The result is an isotropic etch, just like wet chemical etching techniques. However, there is no way to produce these neutral free radicals without creating charged species, which are required to sustain the plasma. The charged species bombard the cathode to produce secondary electrons, but if they bombard on the substrate a sputtering effect is achieved. In this instance, dry etching behaves as a mechanical etching technique, hence the duality. In a typical plasma, the free radical density far outweighs the ion density, so we can consider the etch to be driven by free radicals, hence the name reactive ion etching (RIE).

The etch rate, a key parameter in RIE, is not as simple as the sum of a sputter rate and chemical rate, however. In addition to sputtering, the ion bombardment can break or loosen substrate bonds, increasing their reactivity with the free radicals, and causing the etch rate to be significantly larger than the sum of the components. The reduction in activation energy, and increase in etch rate, caused by ion bombardment only occurs on the horizontal surface of the substrate, resulting in an anisotropic edge effect. Adding to the multitude of effects is a deposition process caused by the ion bombardment on the substrate, which is dependent on the gas and process conditions (section 5.2.2) and forms a protective layer on the film and thereby reducing the etch rate. The result of all these processes is a significant reduction in etch rate on vertical surfaces, which leads to the formation of vertical sidewalls of the etched pattern if the process conditions are suitably adjusted.

The arrival energy of the ions is determined by the DC bias voltage and the gas pressure, where the former is proportional to bombardment energy and the latter inversely proportional to bombardment energy. Increasing pressure while holding a constant discharge power will lower the DC bias due to the higher ion density and higher ion current toward the cathode, and increasing RF power while holding pressure constant will result in an increase in DC bias. The parameters for RIE are all interrelated and it is not possible to control them independently. It is, however, possible to find a protocol that results in a repeatable etch rate and profile (section 5.6.5) [280].

5.3.2.   Silicon Etching

When etching silicon, the key requirement is to create silicon compounds that are volatile at the working substrate temperature and pressure, so the end products either have a high vapour pressure and evaporate, or they have a high sputter yield to be removed by ion sputtering. The most common of these products is silicon tetra fluoride (SiF_4), created by etching silicon with fluorine. Etching silicon with fluorine has several problems, primarily is the chance of producing hydrogen fluoride in the presence of moisture, and the chance that fluorine will uncontrollably react with silicon at room temperature.

Therefore, while pure fluorine is not an option as an etchant, we can use another gas containing fluorine, such as trifluoromethane (CHF_3) or sulphur hexafluoride (SF_6), both of which are available on the bespoke RIE shown in Figure 5.9(b) and are fairly inert until excited when they produce atomic fluorine free radicals [281]. Due to the addition of other elements into the

enchant, there is the chance that a non-volatile and inert material may form on the surface of the substrate. Fortunately, the ion bombardment sputtering aids to remove any developing film from the substrate surface, with only a slower etch rate as the by-product. In some cases, the growth of inert films on the substrate surface can increase the smoothness of the surface and sidewalls of the etched material.

The complications of silicon etching are minimised based on the body of experience available within the research group [94] who have previously developed a protocol for etching silicon nitride. This recipe involves a gas mixture of CHF_3 and oxygen (O_2) at 58 SCCM and 2 SCCM respectively. The fluorine radicals are provided by the CHF_3 [281], and the O_2 reacts with hydrocarbons and hydrofluorocarbons, created in the reaction, to prevent recombination and keep the reactant concentration high [282]. Referring to our bespoke system shown in Figure 5.9(a), the gases are introduced through the anode (ii), with their flow controlled using a mass-flow controller (iv). Power is supplied through an RF power supply (v) to the cathode stage (i). The chamber pressure is maintained and monitored (vi), and the chamber pressure is controlled using a gate valve on the exhaust (iii). All these elements are shown in the schematic in Figure 5.9(b).
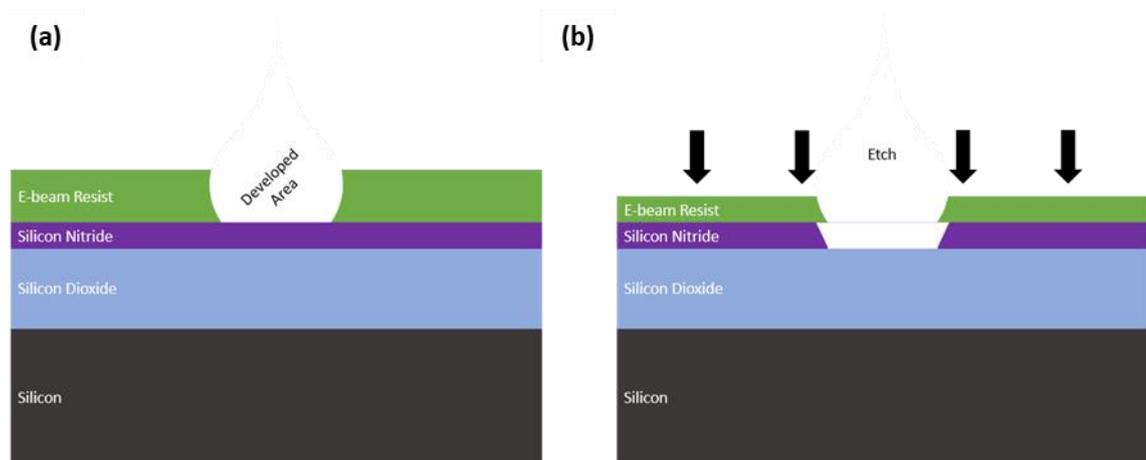
Figure 5.10: (a) Post exposure development and (b) dry etch patterning.

Following from exposure and development, discussed in sections 5.1.3 and 5.1.5, plasma etching is illustrated in Figure 5.10. The differential etch rate between surface and sidewall is highlighted in Figure 5.10(b), where the sidewalls of the silicon nitride layer appear chamfered and are also affected by ion bombardment sputtering.

5.3.3.   Bilayer Lift-Off

In the lift-off technique, the patterning is completed by removing unwanted areas of a deposited material by dissolving the underlying, unexposed, photoresist. A photoresist film is spun, exposed, and developed, and a thin film is deposited on top of the entire sample using any thin film deposition technique, including but not limited to those discussed in section 5.2. The remaining photoresist is then dissolved in a stripper or solvent, releasing the deposited film in those regions from the surface and leaving behind only patterned areas of the deposited thin film. While a lift-off process has the distinct advantage that the etch chemistry of the material is now irrelevant to the process, and all the chemistry lies in the photoresist-stripper relationship, the key to this process is that the stripper has access to the photoresist, and this is highly dependent on the presence on voids, cracks, defects, or sidewall access within the deposited film. Given that thicker or very conformal films are less likely to contain defects, and sonication of the sample could be potentially damaging to the desired structures, it is the latter parameter that we must focus on.

Ideally, a separation between the two layers of deposited films needs to be achieved to prevent the removed film tearing or pulling the desired film during lift-off. There are several methods that can be employed to achieve this. Firstly, an inwardly tapered sidewall, or undercut, creates a physical separation between the deposited film on the photoresist and the sample surface, thus allowing a gap for the stripper to access the photoresist and produce a clean lift-off. An inwardly tapered sidewall is aided by a direct (line-of-sight) deposition, i.e., a deposition orthogonal to the sample surface that prevents deposited material from penetrating the undercut. Thermal evaporation and sputtering, providing the pressures are kept low to reduce scattering of the deposition material, are ideal thin film deposition techniques for lift-off patterning as they produce less conformal thin films. Negative photoresists are typically more suited to lift-off patterning due to their natural inward tapering under exposure, which is a direct cause of proximity effect (section 5.1.4) in that minor over-exposure is more likely to occur at the top of the photoresist and cause an expansion, which in negative photoresists yields a natural undercut [147].
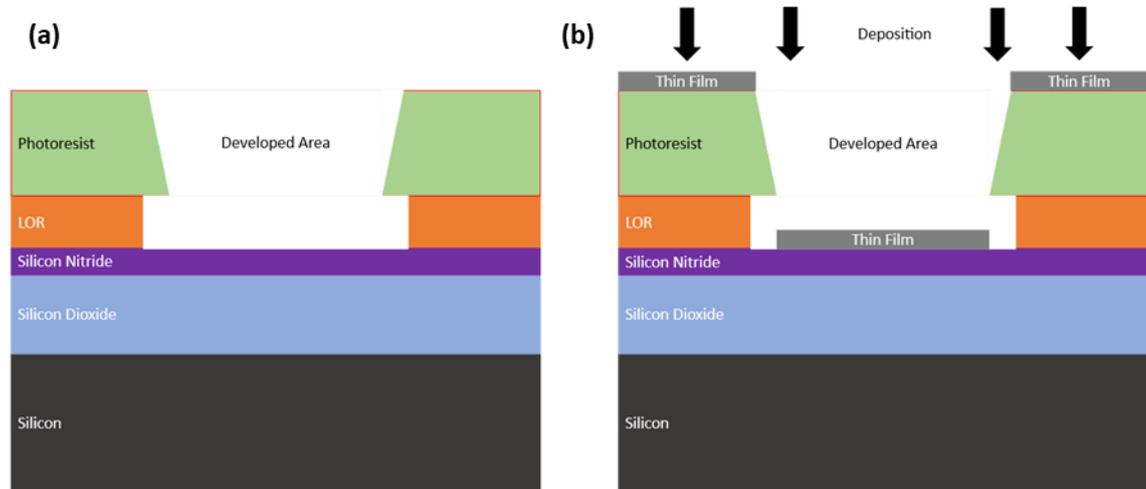
Figure 5.11: (a) Post UV exposure development with bilayer photoresist and (b) targeted thin film deposition.

Unfortunately, using negative photoresists increases exposure time for small structures, and are typically more difficult to remove entirely due to the nature of cross-linking in negative photoresists [148]. Instead, we can employ a positive photoresist and a lift-off resist (LOR), which contains resin with no photoactive compound, i.e., it does not react under exposure. The LOR is instead designed to be soluble isotropically in the photoresist developer and produces an undercut during the photoresist development phase [283]. The increase in undercut offered by using a LOR allows some tolerance in the orthogonality of the deposition direction due to the outward taper in the positive photoresist. The positive photoresist and corresponding lift-off resist used in this research are detailed in section 5.1.1. Figure 5.11 illustrates the (a) development and undercut and (b) orthogonal deposition patterning that can be achieved using a bilayer lift-off technique for deposited thin film pattering. In practice, the deposition is not orthogonal and a chamfered edge on the deposited thin film often remains.

5.4. Extra Fabrication Steps

There are extra fabrication steps to consider before I can detail the steps. Namely an adhesion layer required during the deposition of gold is required, but first I will discuss the thermal treatment of some of the layers in Figure 5.1(a) that have structural, optical, or electrical properties that can be developed or optimised under thermal treatment. SU-8 photoresist and ITO are two materials whose properties change when hard-baked and annealed, respectively. In

this section, I will detail the techniques used to hard-bake and thermally anneal these two materials, and the benefits therein.

5.4.1.   Annealing ITO

As discussed in chapter 4, the optical and electrical properties of ITO can be controlled using the exposure to oxygen at elevated temperature. To achieve this, I use an oxygen-environment annealing process in a tube furnace. The system available to my research group is shown in Figure 5.12 both with (a) a photograph and (b) a schematic of the internals. Figure 5.12(a) shows the (i) reactive gas flow controller, (ii) reactive gas inlet, (iii) furnace tube, (iv) heating filament, and (v) program controller that I will briefly describe in this section.
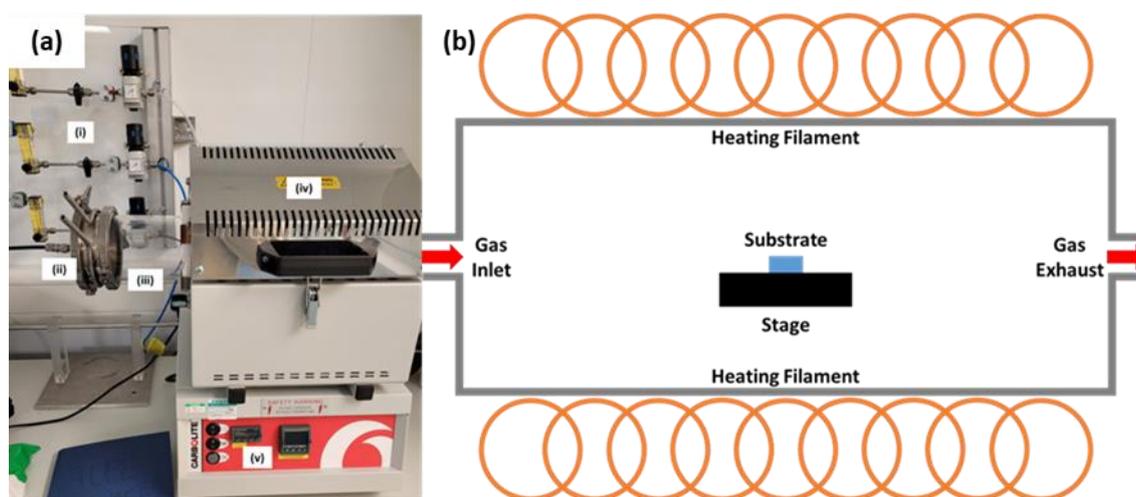
Figure 5.12: (a) Photograph and (b) schematic of the tube furnace available at York.

The annealing protocol I used is as follows. Reactive gas is allowed to flow into the tube furnace, in this research that is oxygen, at a rate of between 0-2 standard cubic foot per hour (SCFH), or 0-944 SCCM, 5 minutes prior to any heating and throughout the annealing process. The heating filaments are programmable, with a maximum ramp speed of 5°C/min and a maximum temperature of 1200°C. The heating controller allows for a series of steps, including ramps and dwells to be programmed in for a fully customisable process. I use a 500°C temperature for 1 hour, with a 5°C/min ramp, 1 SCFH oxygen flow, and I allow the system to cool unaided. The exhaust gas outlet is unassisted and vents into an extraction system only, therefore creating an oxygen environment in which ITO is annealed.

5.4.2.   Adhesion Layer

While gold is the ideal choice for an electrical contact material, due to its low resistivity and chemical inertness, it is the latter that introduces a further problem. Because gold is chemically unreactive, it adheres poorly to oxide surfaces and therefore requires an adhesion layer. The adhesion layer must be applied between the gold and insulator layer shown in Figure 5.1(a). Typical adhesion layers include nickel, titanium, and chromium, all of which are more chemically reactive than gold and thus have a higher adhesive property. All three adhesive metals are available in the e-beam evaporator available to me; therefore, the choice of adhesion layer should be carefully considered to ensure minimum effect on the final structure. There has been much investigation into metallic and non-metallic adhesion layers in the literature [284]–[287], primarily for nano-scale devices where the interference with periodic structures is not a concern for this work.

However, an in-depth analysis of titanium and chromium adhesion layers [288], shows that if interaction with the optical field is a concern, titanium is the adhesion layer of choice, and if mechanical stability, superior electrical and physical contact is required, chromium is the preferred choice. This is because chromium forms an oxide layer with the insulation layer and an alloy with the gold layer, where neither interface is diffuse, therefore allowing a strong bond for micro-scale bilayer lift-off (see section), whilst maintaining electrical compatibility and durability. I typically used an adhesion layer of 10% thickness with respect to the deposited gold film, although there are no hard rules to determine this thickness.

5.5.   Substrate Preparation

The substrate layers, as described in section 3.1.2, are commercially fabricated on a 4-inch wafer. The wafers are then diced into 15x15 mm samples, simply because this is the standard sample size now used in the research group and, as such, many of the fabrication steps have been optimised for these dimensions. Mechanical dicing, discussed here, and many of the other fabrication techniques used in this work can be aggressive and dirty, leaving a residue on the surface and/or damaging the sample surface or edges [289], so require substrate cleaning and preparation between each of the fabrication steps.

5.5.1.   Mechanical Dicing

Wafer dicing is common practice in the manufacturing of integrated circuits, and mechanical dicing is a subsection thereof, referring to the specific use of a mechanical saw (known as a dicing saw). Mechanical wafer dicing is the process of cutting through a wafer material to produce a specific size of sample, known as a die. The process is typically automated, and a sample size can be set prior to cutting. As mentioned in section 5.5, I chose a 15x15 mm sample size to easily assimilate with my research group. Given that the saw typically cuts through the entire thickness of the wafer, the dice quality is directly proportional to the blade speed, blade thickness, blade size, and cutting speed. The primary limitation of mechanical wafer dicing is that the bottom side of the wafer, with respect to the blade, may experience chipping during cutting. There are several techniques to minimise this effect, including bonding a glass layer to the bottom of the wafer, but this process is very labour intensive and has the potential to cause further wafer damage [290].
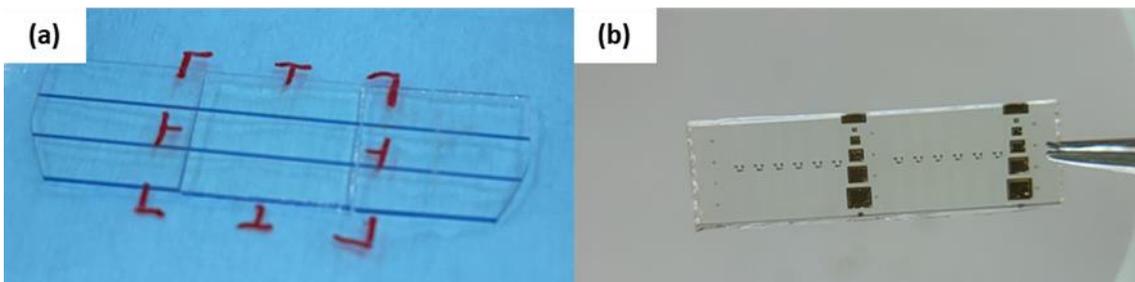
Figure 5.13: (a) Partially diced silicon nitride on borofloat 33 samples, and (b) cleaved chip from partially diced sample.

Initially, with Si_3N_4 on a borofloat 33 substrate, we found that cleaving the sample post-fabrication resulted in a significant drop in sample yield and often caused damage to the structures. We also found a significant amount of damage to the waveguide facet (see section) when purely cleaving the sample to produce the final chip. We used a technique, known as partial dicing, whereby the dicer was programmed to cut through the wafer only partially from the back, see Figure 5.13(a).

It is common practice to use a frame, clamp, or possibly a vacuum stage to maintain a stable cutting position of the wafer. We used a frame that was clamped into place with the wafer held in the frame using dicing tape (Figure 5.13(a)). To protect the silicon nitride surface, we place the nitride side of the wafer onto the tape, though this can have its own problem. At high cutting temperatures the tape can melt, leaving behind a residue on the silicon nitride surface of the wafer, which then requires an aggressive cleaning protocol to remove. To minimise the damage

to the wafer surface, the sample size needs to be larger than the desired circuit. Hence, while the sample size is arbitrary, it is important to have a larger sample size than necessary as edge damage does then not affect structures, providing they are fabricated away from the sample edges.

While a partial dice as in Figure 5.13(a) allows for a relatively good cleave in the case of an amorphous substrate, shown in Figure 5.13(b), the cleaving process required to remove the rough sample edges can leave excess material, leave a rough or distressed edge, or cause a sample to shatter. This problem was solved by changing the substrate from borosilicate to silicon, as a silicon substrate enables a good, smooth cleaved facet. We used a combination of the Disco DAD 320 semi-automatic dicing saw and the Loadpoint MicroAce 66 mechanical dicing saw throughout this work.

5.5.2. Piranha Cleaning Protocol

The aggressive cleaning protocol required to remove any residue left by mechanical dicing is a 3:1 mixture of sulfuric acid and hydrogen peroxide, commonly known as piranha solution. Piranha removes organic material, though it can also be used to remove certain trace metals [291]. Piranha cleaning is therefore the ideal cleaning protocol to use post-dicing. I use a 5-minute clean in piranha, without agitation or external heating, to remove any residue left by dicing the wafer. Samples are then submerged in deionised (DI) water for a further 10 minutes before drying with high-pressure nitrogen.

5.5.3. Acetone-Isopropyl Alcohol Cleaning Protocol

The acetone-isopropyl alcohol (ACE-IPA) cleaning protocol is common practice in pre-lithography sample preparation, and is performed to remove organic contamination including photoresist, dust, and other solutions/chemicals used during fabrication processes. Acetone is responsible for the removal of organic materials, softening or dissolving any residues, though this can often leave behind an organic film. Therefore, the additional use of IPA is required to remove the acetone before the sample can dry. IPA is hygroscopic and can therefore leave a surface dehydrated. As such, it is common practice to never allow the sample to dry naturally and instead to use high-pressure, dry nitrogen for sample drying.

Sonic cleaning is a useful addition to an ACE-IPA protocol and involves agitating the sample with high frequency acoustic waves on the order of 200kHz. Such a frequency can easily dislodge small particles (around 1μm in size) through the creation and collapse of partial vacuum bubbles around the sample. However, such a cleaning process can have an adverse effect on any structures, such as delamination, and should therefore be used with caution [292]. Fortunately, the structures in this work are large enough to avoid damage during sonic agitation.

The ACE-IPA protocol is used before each lithography step in conjunction with a sonication bath for added agitation, where the sample is allowed to sit for five minutes in acetone and IPA and then dried with nitrogen accordingly. The sample is never allowed to dry between either solvent and is never allowed to dry naturally to prevent any organic film growth or dehydration.

5.6.   Fabrication Steps

To this point, I have detailed all the fabrication techniques required to create the MOS EO modulator shown in Figure 5.1(a). Here I will summarise the layer development, fabrication techniques, and parameters used for each layer of the structure following the design parameters detailed in section 3.5.

5.6.1.   Choice of Photoresist

For photolithography, I chose the resist S1818 which is a member of the Microposit S1800 G2 photoresist series [233]. The S1800 photoresist series consists of positive photoresists, allowing for direct patterning, and residue-free removal techniques, and for complete removal in acetone. S1818 was specifically chosen to accommodate for the high-power UV-exposure technique discussed in section 5.1.4 and allows for a uniform coating when spinning over pre-existing structures due to its 1.5-2.5 μm thickness range as shown in Figure 5.2(b). S1818 photoresist serves as a patterning mask for bi-layer lift-off processing (section 5.3.3), which requires the use of a high yield undercut resist. Specifically, I used MicroChem LOR 7B [293] which is a simple bilayer resist with high thermal stability and thin film deposition compatibility. The final photoresist I used serves as the insulator layer in the MOS capacitor (Figure 5.1(a)), and is known as SU-8 [294]. SU-8 is a negative photoresist dissolved in cyclopentanone, where the concentration of cyclopentanone is inversely proportional to the final film thickness and directly

proportional to UV sensitivity [148]. SU-8 was chosen as the insulator layer due to its ability to become a structural film, due to its mechanical strength and chemical resistance.

For EBL, I chose to use AR-P 6200.13 (AR-P 13) from ALLRESIST GmbH [193], which is a high-contrast (high resolution) and highly stable positive electron-beam resist. AR-P 13 was chosen due to its resistance to plasma-etching relative to the silicon nitride platform, allowing for fine control over the etch depth (section 5.3.1). Due to the electrical properties of Si_3N_4, under electron-beam exposure, a significant charge can build up on the surface and a charge dissipation layer is required. I chose AR-PC 5091.01 (AR-PC) [295] from ALLRESIST GmbH. The advantage of using AR-PC is that it is directly compatible with AR-P 13 and soluble in water for easy, non-damaging removal post-exposure.

5.6.2.   Choice of Developers

Due to the multitude of photoresists, I used, I had to use a similarly extensive list of developers and will summarise them here with their corresponding resist layers. All developers were used at room temperature and with no external agitation. For EBL, AR-PC is easily removed with a deionised (DI) water wash, where AR-P 13 requires a xylene soak, followed by an IPA clean and dry with high-pressure nitrogen to remove any remaining developer. By removing the developer, the chance for over-development is reduced.

For UVL with S1818 and LOR 7B photoresists, I use MF-319 developer, which is a tetramethylammonium hydroxide (TMAH) based developer specifically designed for use with the S1800 photoresist series and provides a good undercut (section 5.3.3) when used with LOR 7B [296]. As is common with many photoresists, post-soaking in MF-319 the samples are rinsed with DI water to prevent further development and dried with high-pressure nitrogen. For UVL with SU-8, I use a 2-methoxy-1-methylethyl acetate, commonly known as EC solvent (from Microposit), soak with a DI rinse and high-pressure nitrogen drying step [297].

5.6.3.   Solvents for Bilayer Lift-Off

The removal of the lift-off resist requires the use of an appropriate solvent. The recommended solvent to use with LOR 7B is 1-methyl-2-pyrrolidinone solvent, specifically MICROPOSIT 1165 resist remover [298], which effectively dissolves LOR 7B and provides a clean lift-off procedure. Lift-off can be encouraged through gentle agitation, provided by a magnetic stir bar rotating at

300 rpm, and gentle heating at 50°C. Depending on the thickness, verticality, and possible overhang of the resist layer, lift-off can be a 1-hour to 5-hour process, where vigorous encouragement is not recommended due to the risk of damaging the lithographic structures. 1165 is removed using DI water.

5.6.4.   Alignment Marks

The multi-layer nature of the MOS EO structure discussed in this work requires multiple sequential lithography steps that have been discussed in this chapter. Maintaining precise layer alignment is a necessity to preserve the capacitive behaviour of the MOS structure and to reduce losses in the waveguide. To improve repeatability and increase fabrication speed, I developed a series of alignment marks that can be used in conjunction with a lithography tool known as "3-point alignment".

In practice, the alignment marks should allow for discrepancies in sample placement and orientation to be completely removed during any lithography step. Initially, I used a two-mark system where I designed a UVL mark (star) and an EBL mark (cross) to ensure that the two systems were aligned to each other.
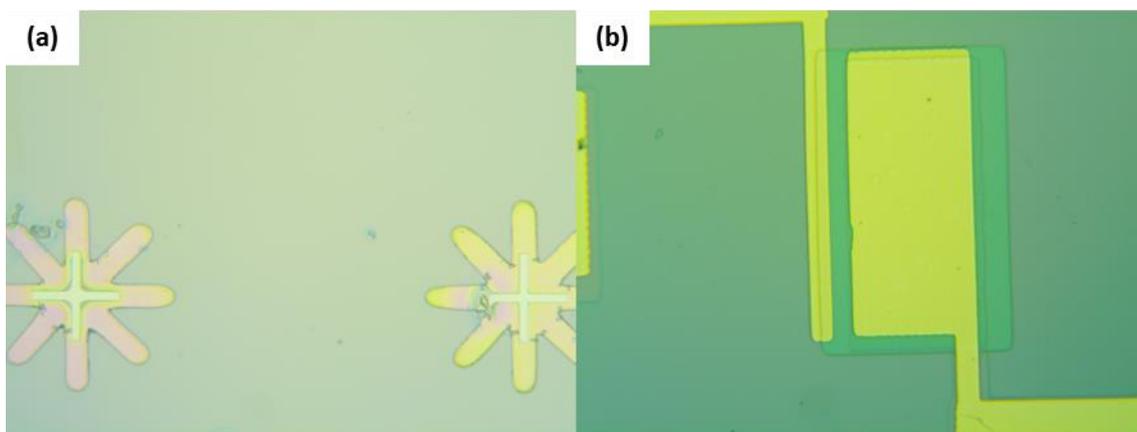
Figure 5.14: Micrographs of (a) UV and electron beam lithography alignment marks, and (b) misaligned MOS capacitor components.

An example of the two-mark system is shown in Figure 5.14(a), with the smaller EBL cross visible inside the larger UVL mark and showing that the inter-system alignment is good. I noticed, through the fabrication development portion of my research, that there can be drift between the UVL steps, shown in Figure 5.14(b). The chance for over-exposure created by the multiple

crossing lines, which leads to an increase in feature size, can diminish the effectiveness of the alignment marks as the centre point is more difficult to determine.
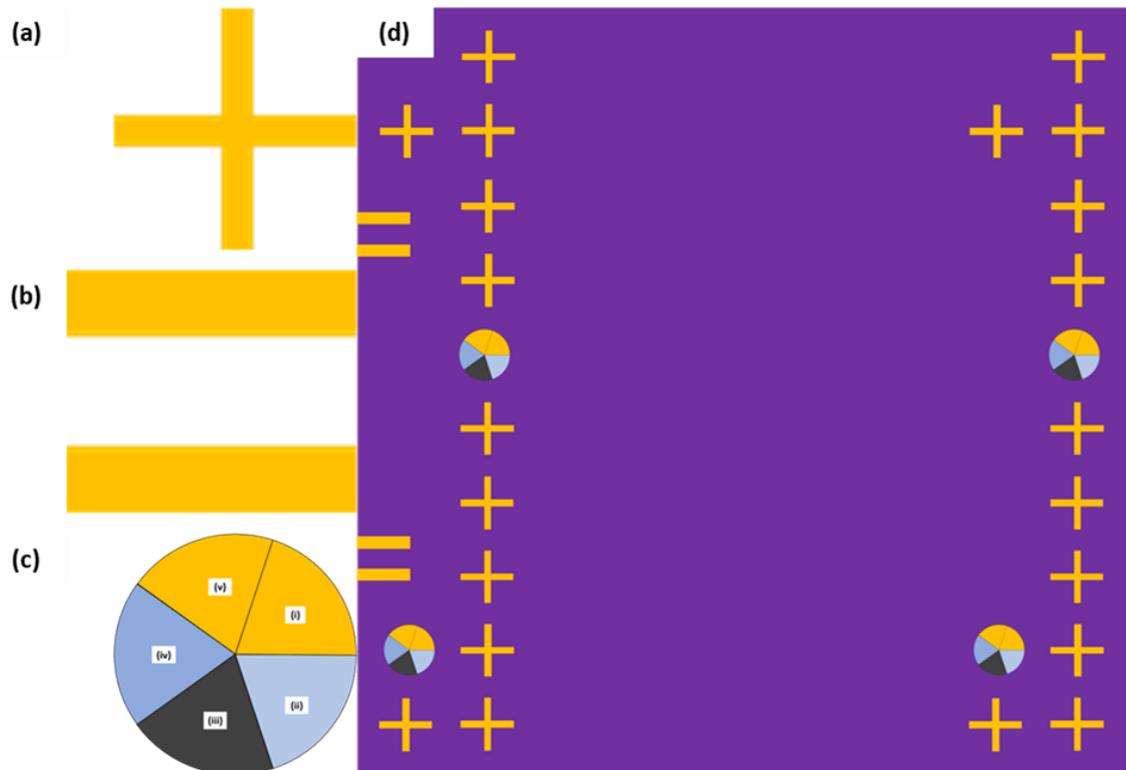
Figure 5.15: Alignment mark designs, showing (a) individual cross marker, (b) cleave marker, (c) layer alignment marker, and (d) sample layout (not to scale).

I chose to use a simple cross mark (Figure 5.15(a)) instead, and to repeat this mark across the sample (Figure 5.15(d)) in such a way that I am able to easily determine the orientation and position on the chip using any lithography or microscopy technique; essentially creating a coordinate system across the surface. The simple cross mark ensures minimal repeated writing areas, reducing over-exposure likelihood, whilst maintaining a clear centroid. To increase layer alignment precision, I included a "multi-layer alignment mark", illustrated in Figure 5.15(c), where each layer forms a part of the circle, allowing me to identify both the misaligned layer and the magnitude of the misalignment. The gold (i)(v), silicon nitride (ii), ITO (iii), and SU8 (iv) layers are illustrated respectively.

Finally, to ensure all cleaved chip sizes were the same, I included a cleave marker illustrated in Figure 5.15(b). these marks run down the sides of the sample and are spaced 5mm from top and bottom of the sample, thus leaving 5mm between them. By cleaving between the two lines, I can ensure that the resulting chip is approximately 5mm in length. Thus, the lengths of repeated

chips will be approximately the same, allowing for experimental comparison and ensuring the chips will fit on the optical stage (Figure 7.1).

Creating alignment marks is the first step in the multi-layer fabrication process and allows for precise alignment between the multiple lithography steps and both in EBL and UVL techniques. The samples are pre-diced to 15x15mm squares and are cleaned using the piranha and ACE-IPA protocols to produce a clean blank sample as illustrated in Figure 5.16(a). The alignment marks are patterned using UVL and a bilayer lift-off technique (sections 5.1.4 and 5.3.3 respectively), and therefore requires a layer of lift-off resist (LOR 7B) spun at 500 rpm for 20 seconds and 5000 rpm for 60 seconds to achieve an approximate thickness of 250 nm. LOR 7B is soft baked at 180°C for 10 minutes. The photoresist used for UVL is S1818, spun at 6000 rpm for 60 seconds to achieve a thickness of approximately 1.7 μm, followed by a soft bake at 115°C for 60 seconds. Temperatures are monitored throughout using a digital thermometer, and baking is performed on a hot plate. The photoresists are illustrated Figure 5.16(b).
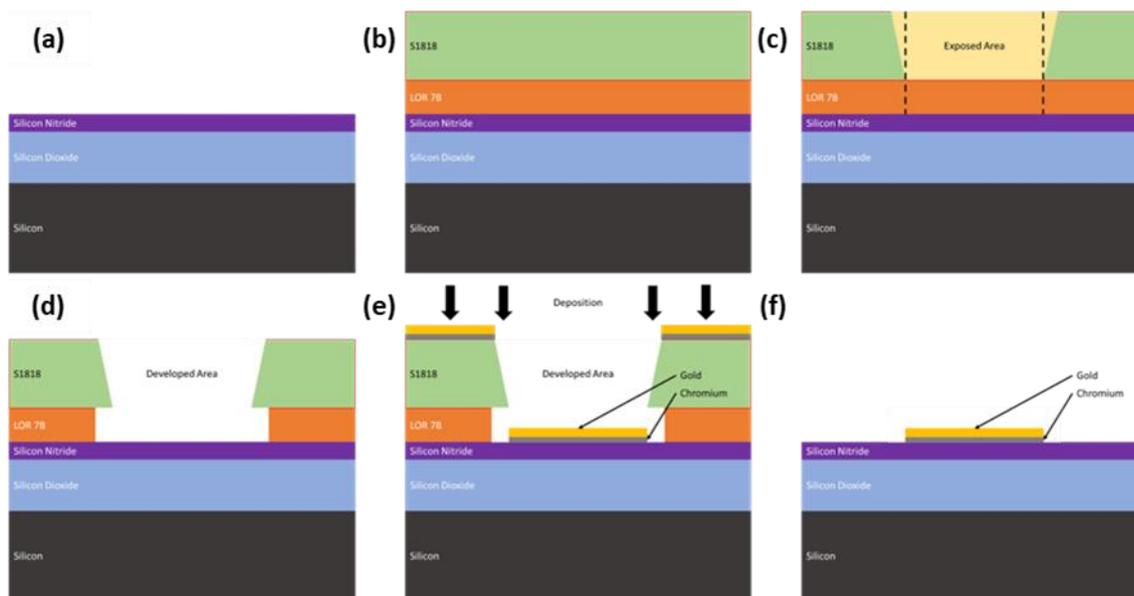
Figure 5.16: Bilayer lift-off for alignment mark patterning.

The pattern is exposed using UVL and developed for 1 minute in MF-319 developer using the protocol (section 5.6.2), illustrated in Figure 5.16(d). I then use an e-beam deposited chromium 20 nm and gold 80 nm protocol (illustrated in Figure 5.16(e)). I use gold because it can be clearly seen both optically and with electron microscopy (shown in Figure 5.17), allowing me to see the alignment marks clearly in both lithography steps. The patterning is completed using a bilayer lift-off by dissolving LOR 7B in 1165, as detailed in section 5.6.3 and illustrated in Figure 5.16(f).
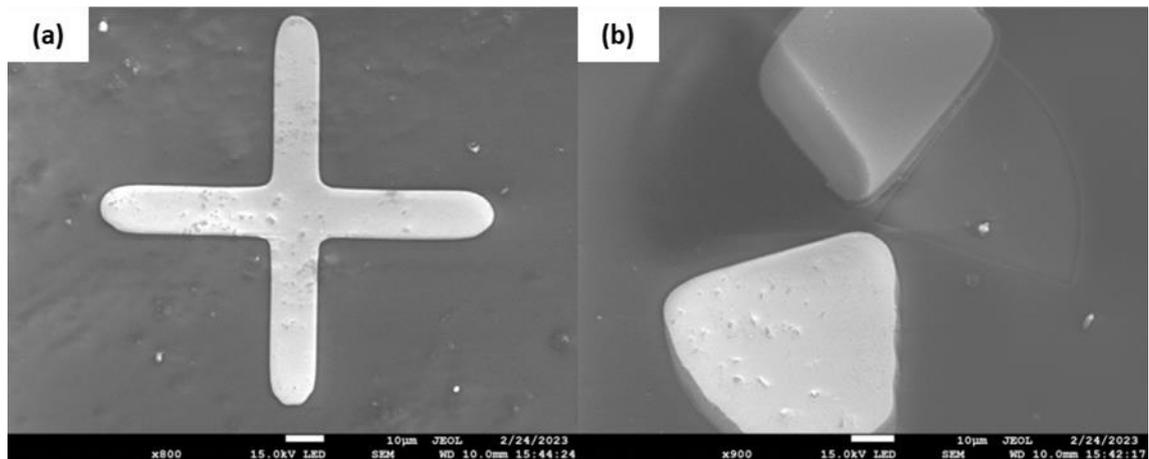
Figure 5.17: (a) individual alignment mark and (b) layer alignment marks.

The effectiveness of the multi-layer alignment mark can be seen in Figure 5.17(b). There is clearly 1 layer missing in the bottom right corner and is yet to be deposited, and the layer in the top left is poorly aligned; there are the gold electrical contact and silicon nitride etch layers respectively. Using an image like Figure 5.17(b) during the fabrication process, I can easily adjust the subsequent layers where necessary to realign, thus removing the need to restart the entire process.

5.6.5.   Waveguide

Samples are cleaned using the ACE-IPA protocol (section 5.5.3) to remove any remaining photoresist from the alignment mark layer, illustrated in Figure 5.18(a). The waveguides are fabricated using EBL and dry etching, therefore a layer of AR-P 13, spun at 5000 rpm for 60 seconds and soft-baked at 180°C for 5 minutes to achieve a thickness of 350 nm, followed by AR-PC spun at 2000 rpm for 60 seconds and soft-baked at 90°C for 2 minutes (illustrated in Figure 5.18(b)). The pattern is exposed using EBL (section 5.1.3) with a 40 µm aperture and beam dose of 130 µC/cm^2, the beam current is typically measured to be around 900 pA (Figure 5.18(c)).
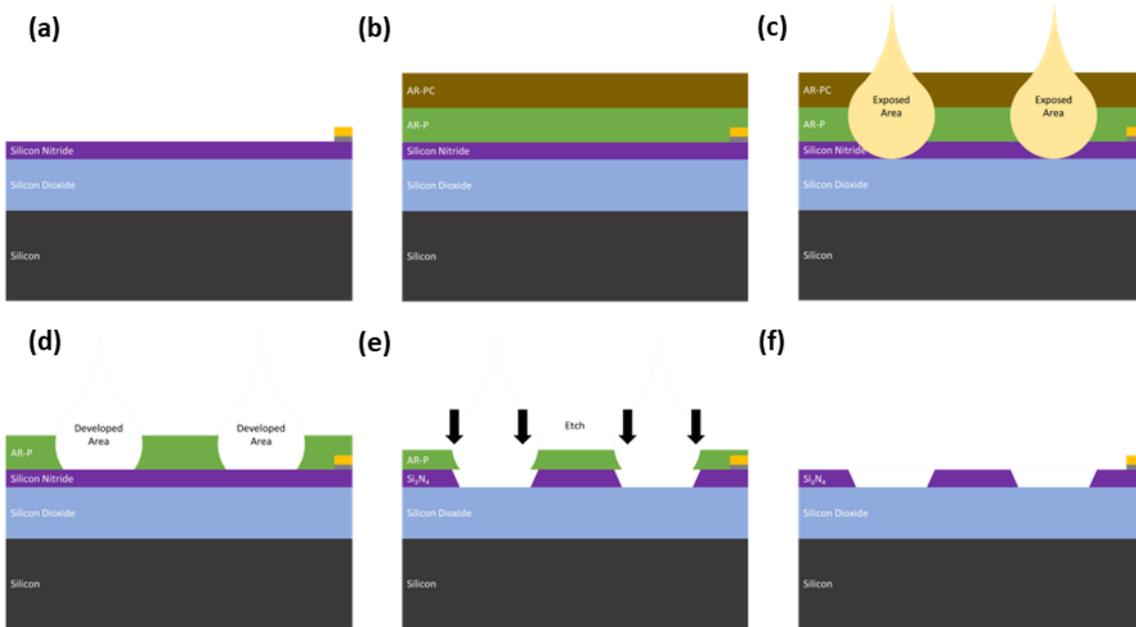
Figure 5.18: EBL and dry etching for waveguide patterning.

Post-exposure, the samples are developed using the xylene protocol describe in section 5.6.2, with 30 seconds in DI water to remove the charge dissipation layer and 2 minutes in xylene (Figure 5.18(d)). The patterning is completed using reactive ion etching (section 5.3.1) with a reactive gas of 2 SCCM O_2 and 58 SCCM CHF_3, with a DC volage maintained at 420 V for a consistent etch rate (Figure 5.18(e)).
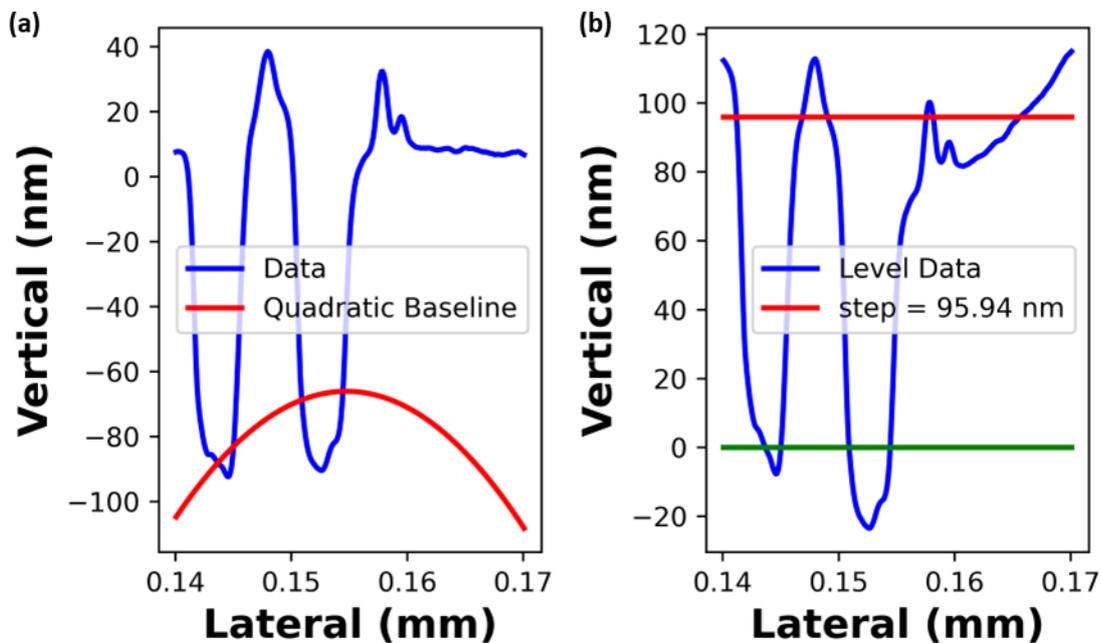
Figure 5.19: A (a) distorted and (b) levelled surface profile data for the silicon nitride waveguide. Data line marked in blue, quadratic line marked in red, zero marked in green.

The RIE etch of 2 minutes produces an etch depth of approximately 100nm, as shown in Figure 5.19(b). The long scans are taken to reduce the errors by increasing the distance over which an average step height can be calculated, with the Dektak moving across the waveguide to avoid damage when the tip touches down on the surface. The remaining resist is removed using 1165 resist remover and DI water, to produce the waveguide illustrated in Figure 5.18(f).
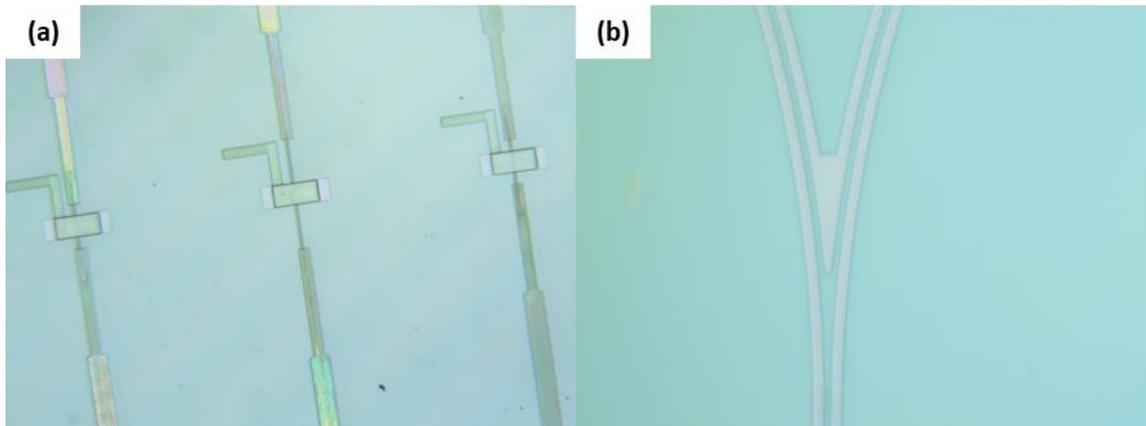
Figure 5.20: Micrographs of (a) tapered silicon nitride waveguide with SU-8 waveguides and (b) Y-junction splitter fabricated in silicon nitride.

Before moving to a silicon substrate, to compensate for the poor cleave of the borosilicate substrate, I used a tapered waveguide system. The silicon nitride waveguide tapered to a point and intersected with an SU-8 tapered waveguide, as shown in Figure 5.20(a). The overlap of the two tapers needs to be perfect to minimise optical losses. The change to a silicon substrate eliminated the problem of a poor waveguide facet, and a larger, SU-8, waveguide facet was no longer required. The SU-8 waveguide are exposed using UVL, which reduces the precision and increases the scale of the exposable patterns. Removing the need for SU-8 waveguides removed the need for UVL waveguide exposure. The multi-wavelength 1-to-2 splitter (section 3.6.2) is shown in Figure 5.20(b), where the gradual bend in the waveguide to reduce optical losses can be seen, along with the rounded edge on the split to reduce scattering losses and increase the bandwidth of accepted wavelengths.

5.6.6.   Semiconductor

Samples are cleaned using the ACE-IPA protocol (section 5.5.3). The semiconductor (ITO) layer is patterned using UVL and a bilayer lift-off technique, and therefore requires a layer of lift-off resist (LOR 7B) spun at 500 rpm for 20 seconds and 5000 rpm for 60 seconds, with a soft bake of 180°C for 10 minutes (Figure 5.21(b)). The photoresist used for UVL is S1818, spun at 6000 rpm for 60 seconds with a soft bake at 115°C for 60 seconds.
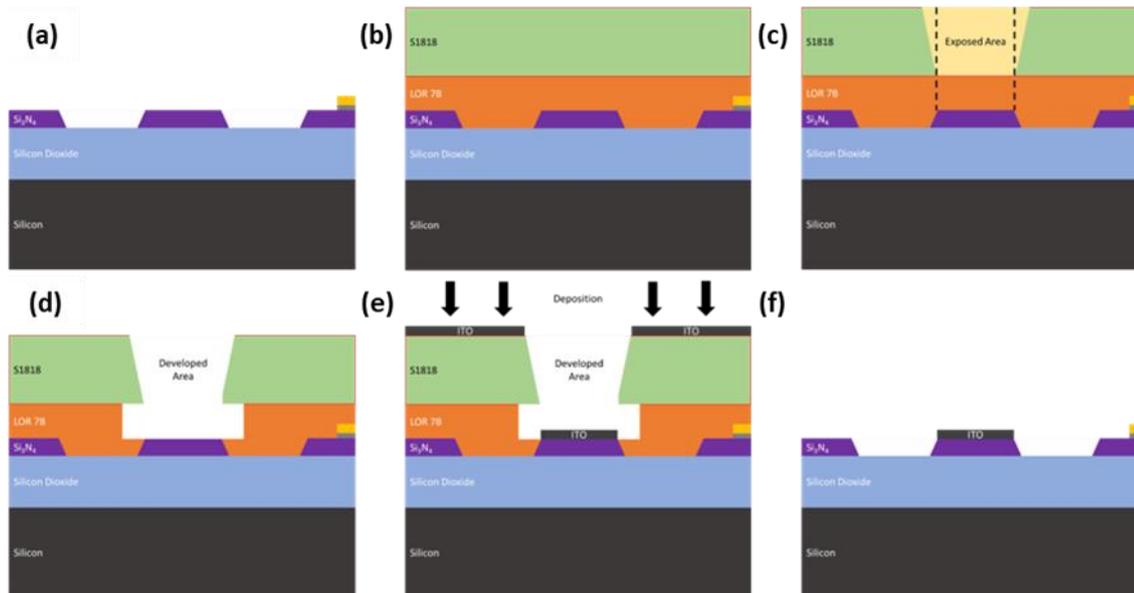
Figure 5.21: UVL and bilayer lift-off for ITO patterning.

The pattern is exposed using UVL (section 5.1.4) and developed for 1 minute in MF-319 using the protocol detailed in section 5.6.2 (illustrated in Figure 5.21(c) and (d)). I then deposit a thin film of ITO using a pulsed DC sputtering technique (section 5.2.2) with the following parameters: 20 SCCM argon, 7.5 SCCM oxygen, 100 W DC power, and with a pressure of 0.75 Pa. The patterning is completed using a bilayer lift-off in 1165 (section 5.6.3) as illustrated in Figure 5.21(f). Post-patterning, the ITO layer is annealed at 500°C for 1 hour in an oxygen flow of 1 SCFH, as detailed in section 5.4.1, and allowed to cool to room temperature overnight.
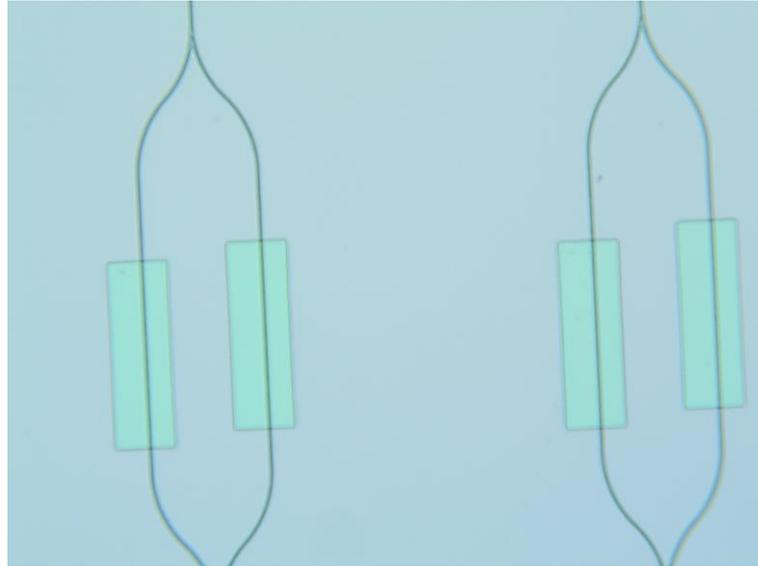
Figure 5.22: Micrograph of phase modulators with double capacitors, showing only the waveguide layer and ITO layer.

The characterisation of the ITO layer is discussed in section 5.7.2, but the result of the fabrication process can be seen in Figure 5.22, which shows two phase modulators with a capacitor on each arm of the Mach-Zehnder modulator.

5.6.7.   Insulator

The samples are cleaned using the ACE-IPA protocol (section 5.5.3). The insulator (SU-8) layer is patterned using UVL and hard-baking (section 5.1.7). First a layer of SU-8, either 2000.5 or a 7:2 2000.5:2050 mix (Appendix section 9.5), is spun at 5000 rpm for 60 seconds with a soft bake of 65°C, 95°C, 65°C for 2 minutes at each temperature to achieve a thickness between 200-800 nm (dependent on SU-8 mix), illustrated in Figure 5.23(b).
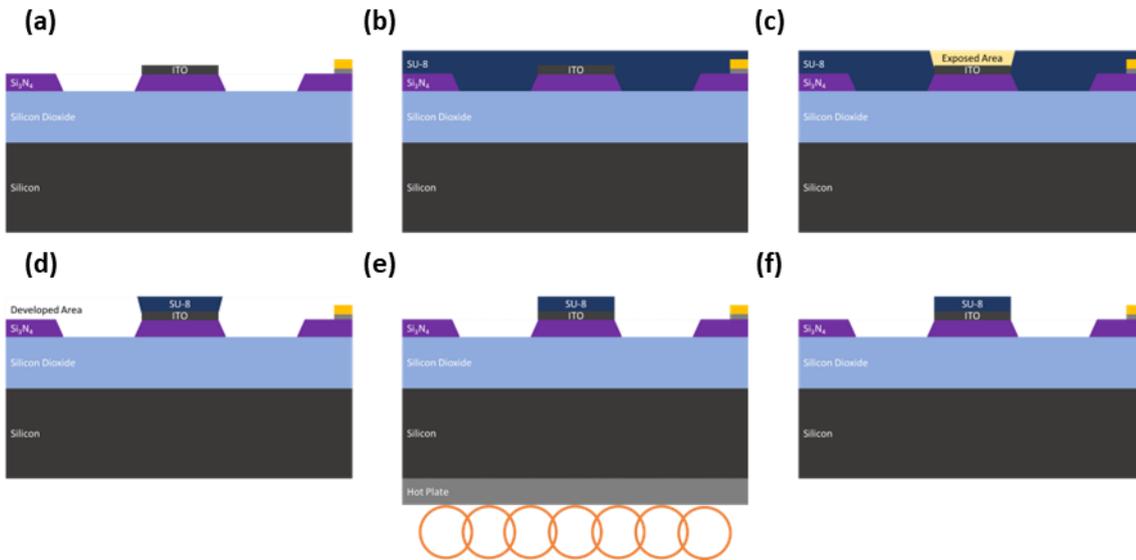
Figure 5.23: UVL and hard baking for SU-8 patterning.

The pattern is exposed using UVL (5.1.4), followed by a post-exposure soft bake of 65°C, 95°C, 65°C for 2 minutes at each temperature, and developed for 1 minute in EC solvent using the protocol recommended by the manufacturer and described in section 5.6.2 (illustrated in Figure 5.23(c) and (d)). As the SU-8 is the insulator layer, post-development, I hard bake the SU-8 at 180°C for 15 minutes. The sample can then be cleaned using the ACE-IPA protocol, illustrated in Figure 5.23(f).
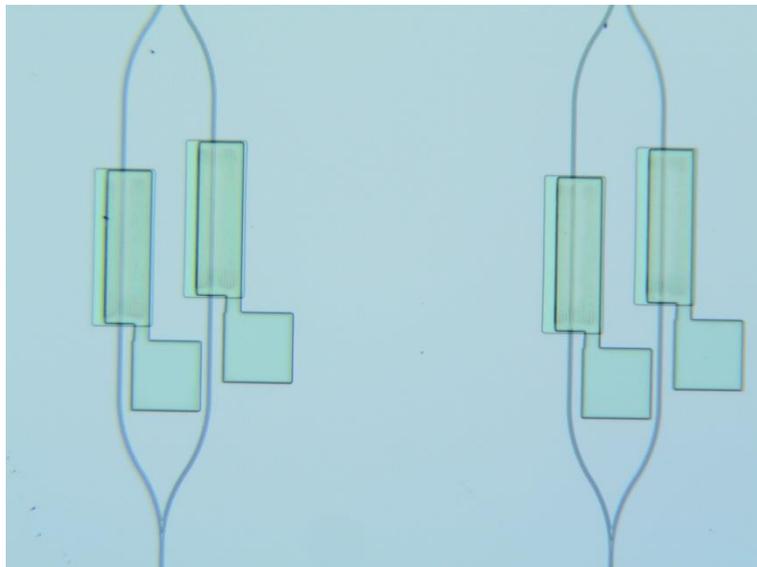
Figure 5.24: Micrograph of phase modulators with double capacitors, showing only the waveguide, ITO, and SU-8 layers.

At the start of my research, I found that I was overdosing the SU-8 pattern, which can result in a patch of SU-8 (not a pattern) exposed on the surface, which then becomes almost impossible

to remove. The problem lies in the consistency of the SU-8 across the sample and once optimised, I was able to precisely expose the insulator layer of the capacitor as shown in Figure 5.24.

5.6.8.   Electrical Contacts

Samples are cleaned using the ACE-IPA protocol (section 5.5.3). The electrical contact layer is a repeated process to the alignment mark layer with a different design (section 3.5.2). The electrical contacts are patterned using UVL and a bilayer lift-off technique, and therefore requires a layer of lift-off resist (LOR 7B) spun at 500 rpm for 20 seconds and 5000 rpm for 60 seconds with a soft bake at 180°C for 10 minutes. The photoresist used for UVL is S1818, spun at 6000 rpm for 60 seconds with a soft bake at 115°C for 60 seconds, illustrated in Figure 5.25(b).
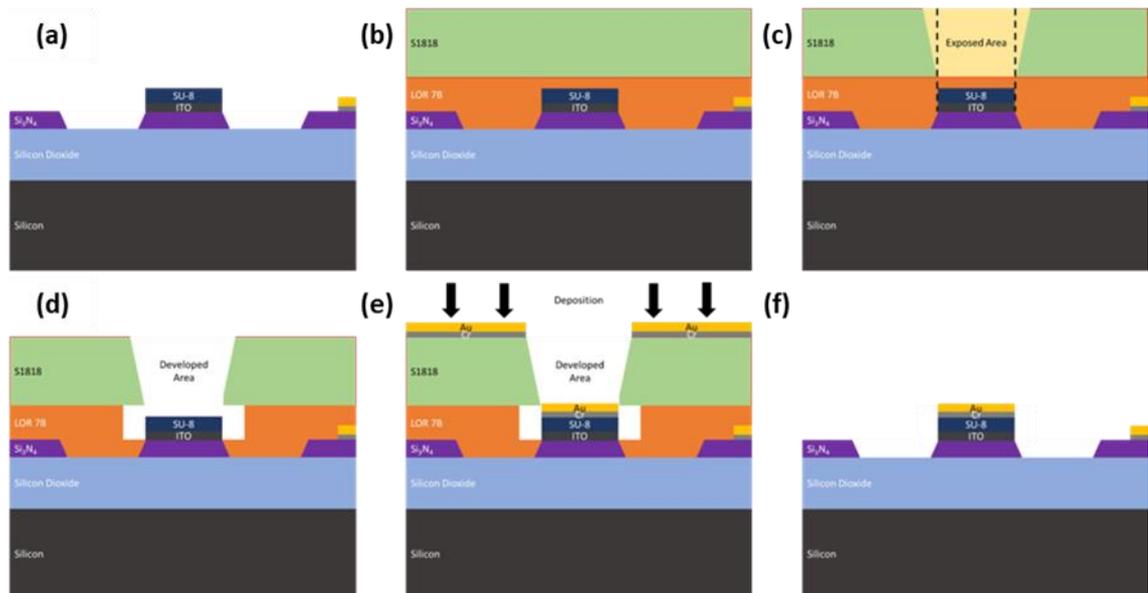
Figure 5.25: UVL and bilayer lift-off for electrical contact patterning.

The pattern is exposed using UVL and developed for 1 minute in MF-319 developer using the protocol (section 5.6.2), illustrated in Figure 5.25(c) and (d). I then use an e-beam deposited chromium 20nm and gold 80nm protocol (illustrated in Figure 5.25Figure 5.16(e)). The patterning is completed using a bilayer lift-off by dissolving LOR 7B in 1165, as detailed in section 5.6.3 and illustrated in Figure 5.25(f).

5.6.9.   Facet Cleave

Many of the techniques discussed in this chapter and the general handling of the sample throughout the fabrication sequence have the potential to be very damaging to the sample edges. Rough waveguide edges have the potential to scatter incident light and thus reduce the performance of the modulator. So, to create smooth edges on the waveguide facets, I use a technique known a cleaving. Cleaving is the process of encouraging a material to break down a crystal plane and is therefore far more compatible with crystalline materials than with amorphous materials. As discussed in section 5.5.1, this is why we introduced a partial dicing method for borosilicate substrates, but with silicon substrates I was able to perform a good cleave to obtain a smooth waveguide facet.
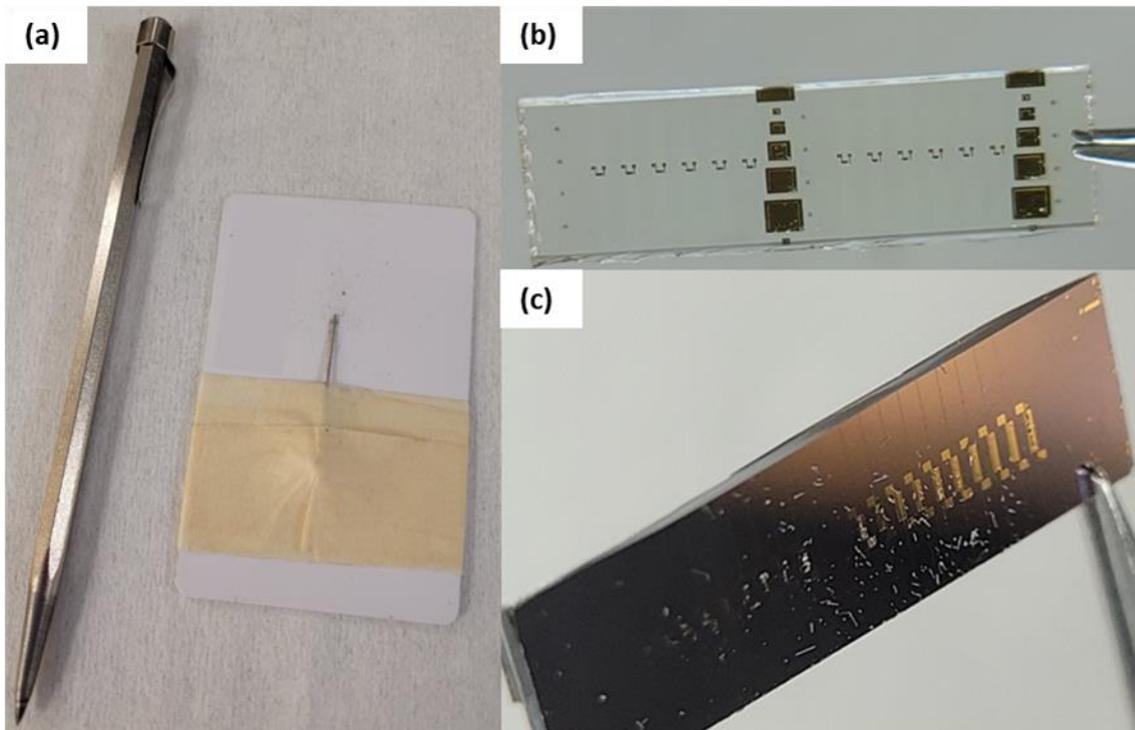
Figure 5.26: (a) Cleaving tool and scribe, (b) cleaved silicon nitride on borofloat 33 sample, and (c) cleaved silicon nitride on thermal oxide on silicon sample.

Cleaving requires a precise but shallow groove to be marked into the substrate edge and mechanical pressure to be applied around that groove. The groove can be applied using a diamond scribe, while the pressure can be applied using a similarly scaled tip to encourage breaking down the scribed crystal plane, where the tools I used are shown in Figure 5.26(a). Providing the material is crystalline, the cleaved facet follows the scribed mark to leave a smooth edge, Figure 5.26(c). However, cleaving in amorphous materials can leave torn substrate material behind on the sample edge due to an incomplete cleave, Figure 5.26(b). There is a clear

smooth facet edge visible in Figure 5.26(c), and the same edge in Figure 5.26(b) on the amorphous substrate is visibly rougher, even when observed with a standard camera. The difference in cleave shown in Figure 5.26(b) and (c) was the driving force in moving to a bulk silicon substrate.

5.7.   Fabrication Results

The culmination of optimising the fabrication process and developing the structure layout is shown in Figure 5.27. The precision alignment between layers, is visible in Figure 5.27(b), captured with high a scanning electron microscope (SEM). The two images show different structure layouts across two iterations of the design (section 3.5.2).
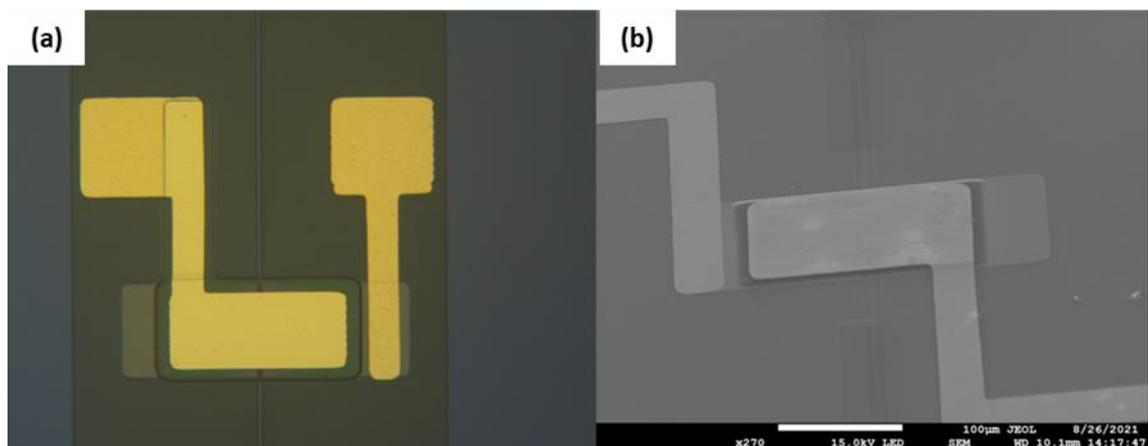
Figure 5.27: (a) Optical and (b) electron microscope image of fabricated MOS EO modulator showing waveguide, ITO, SU-8, and electrical contact layers.

The culmination of optimising the fabrication steps can be seen by the clearly distinguishable layers and good alignment in Figure 5.27(b). For the phase modulators, with a Mach-Zehnder architecture and double capacitor setup, the staggered electrical contacts and achieve a smaller system footprint can be seen in Figure 5.28.
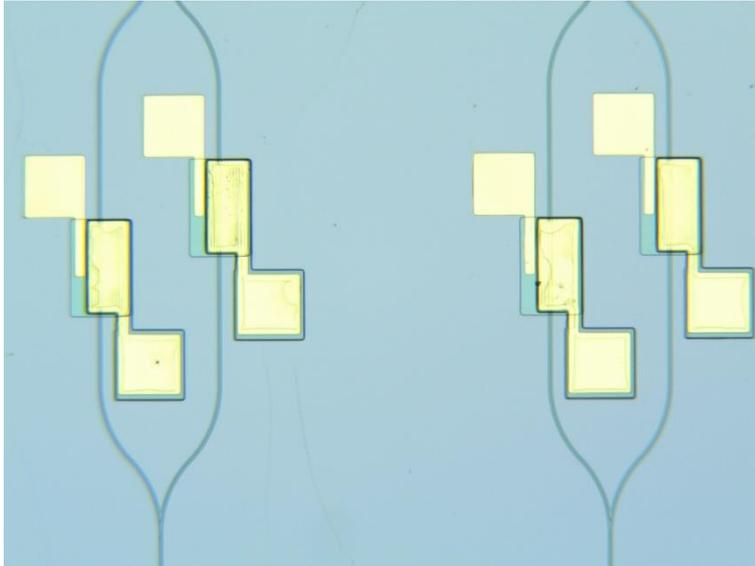
Figure 5.28: Micrograph of phase modulators with double capacitors, showing only the waveguide, ITO, SU-8, and gold layers.

The phase modulator layout is shown in Figure 5.29(a), observed with an SEM, further emphasising the need for gold alignment marks, as the gold layer is clearly visible with electron microscopy. The gradual bend and Y-junction splitter is shown in Figure 5.29(b). The gradual transition between the single input and double output (from left to right) is surrounded by the etched area of approximately 5 µm width. This width of air surrounding the waveguide is sufficient to maintain mode confinement and minimise exposure time.
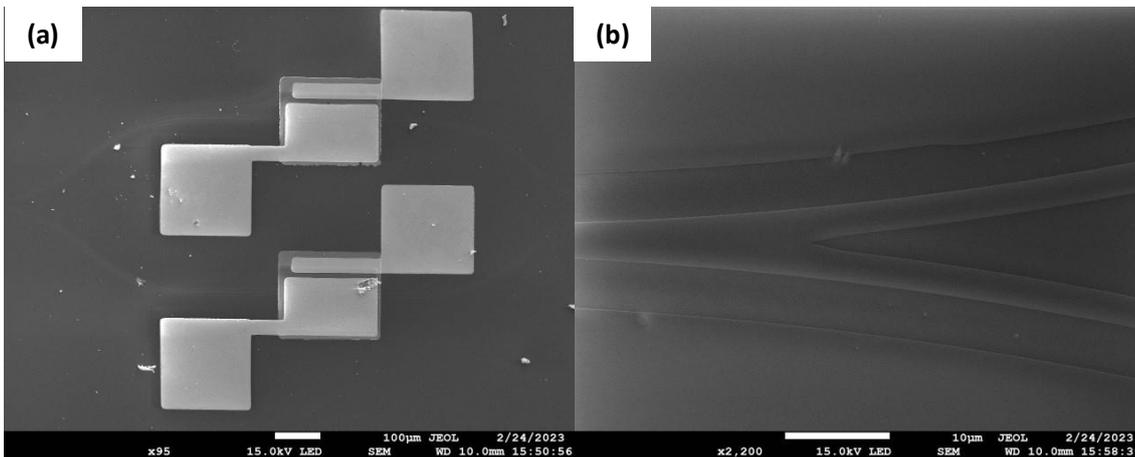
Figure 5.29: Scanning electron micrographs of (a) Y-junction splitter and (b) phase modulator with two capacitors.

5.7.1.   Waveguide Facet

A comparison between the cleaved facet between silicon nitride on (a) borofloat 33 and (b) thermal oxide on bulk silicon substrates is shown in Figure 5.30. The torn facet edge resulting from an amorphous substrate is easily visible, and a facet of this quality would cause significant scattering losses at the input facet.
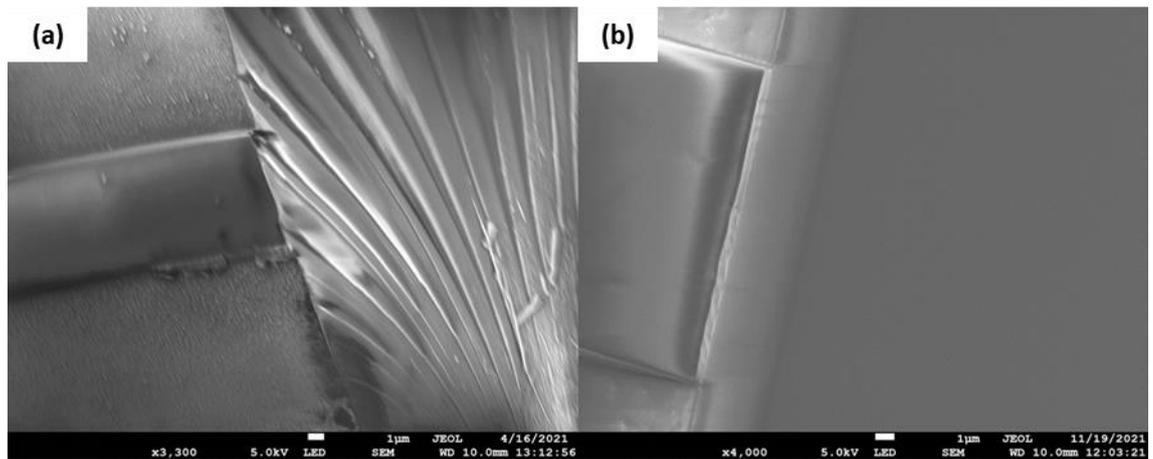
Figure 5.30: Scanning electron microscope images of cleaved facet from (a) silicon nitride on borofloat 33 substrate and (b) silicon nitride on thermal oxide on silicon substrate.

The scale of the damage that can be caused by a poor cleave on an amorphous substrate is not obvious in Figure 5.30(a). To understand the scale of substrate tearing and lack of smoothness to the facet edge that can occur, a wider field of view is required.
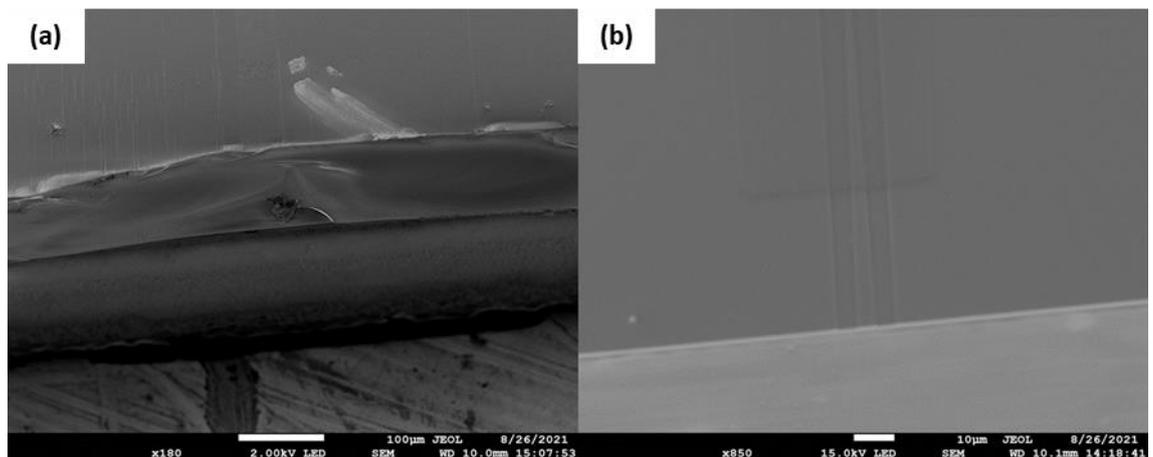
Figure 5.31: Scanning electron microscope images of cleaved facet from (a) silicon nitride on borofloat 33 substrate and (b) silicon nitride on thermal oxide on silicon substrate, captured with a low magnification.

The wider field of view, shown in Figure 5.31 for the (a) amorphous borosilicate and (b) crystalline silicon substrate, shows the scale of facet damage. The silicon nitride layer on the

borosilicate, in (a), is the middle horizontal layer appearing brighter in the image than the substrate. The silicon nitride layer appears rough and uneven. The contrasting quality of the silicon nitride and silicon dioxide layers, shown as the bottom layers in Figure 5.31(b), on the silicon substrate confirms the decision to use a silicon substrate was beneficial to the fabrication process.

5.7.2.   ITO Characteristics

The electrical and optical properties of the ITO used in this research was found using the novel technique discussed in chapter 4. The electrical properties were determined using surface profilometry (section 4.5.1), four-probe (section 4.5.2), and Hall probe (section 4.5.3), methods. These measurements yielded a film thickness of 197nm, a sheet resistance of (110±4) Ω/sq, a material conductivity of (460±20) S/cm, a carrier mobility of (4.67±0.06) cm^2/Vs, and a carrier density of (6.1±0.2) x10^{20} cm^{-3}, respectively. Using Equation 4.2, the electrical parameters give a plasma frequency of (2.1±0.2) x10^{15} rad/s. The sheet resistance and carrier mobility measurements were taken multiple times to obtain an average result with appropriate error (Appendix section 9.1.1), while the calculated values have an associated error calculated using quadrature (Appendix section 9.1.2). The carrier density of the ITO here is significantly higher than the ITO discussed in chapter 4. The higher carrier density is a result of the outgassing of the S1818 photoresist used for lithographic patterning, altering the sputter gas composition, which was not present during the deposition of the ITO films in chapter 4. Using the measured and calculated electrical properties, and our previous investigation, we determine an effective electron mass (m*) of 0.45m_e, where m_e is the standard electron mass.
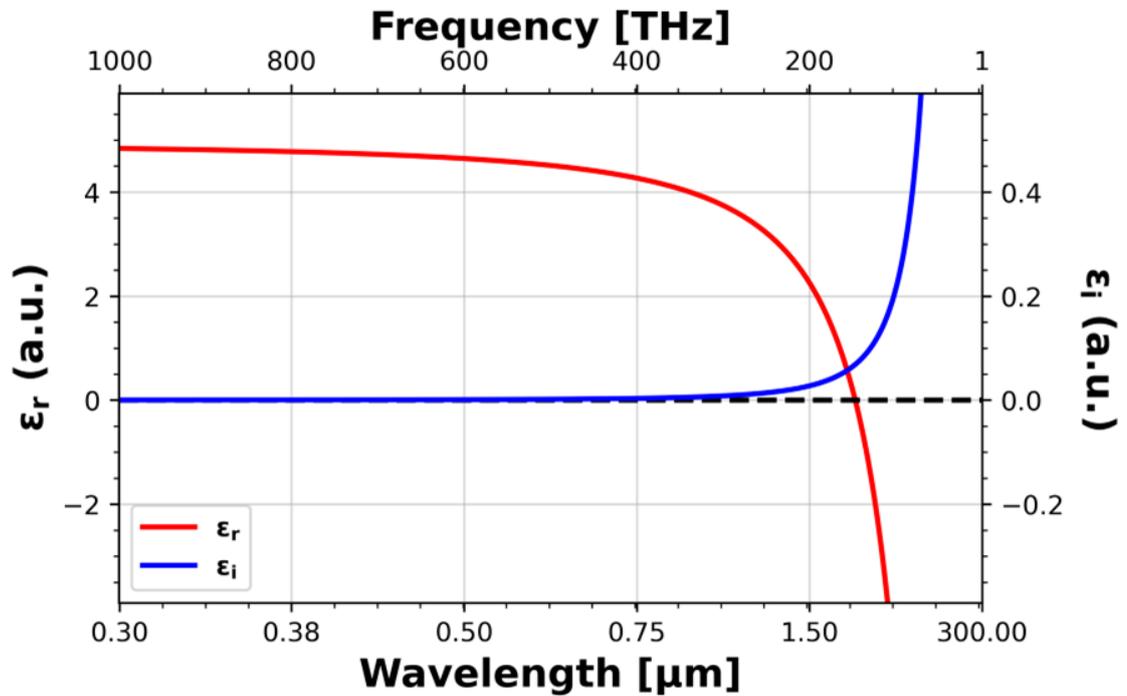
Figure 5.32: Real (red) and imaginary (blue) components of the Drude permittivity for sputter deposited ITO with a 27% oxygen flow concentration during deposition.

Using optical measurements, as detailed in chapter 4 and Appendix section 9.6, the high frequency permittivity (ε_∞) is measured as 4.95, and the collision frequency (Γ) is measured as 13THz. The measured m^*, ε_∞, and Γ parameters are indicative of a high carrier density, when compared to the ITO films shown in Table 4.4, which aligns with the electrical characterisation of this film. The optical and electrical parameters are used in the Drude equation (Equation 4.1) to plot the real (red) and imaginary (blue) components of the Drude dispersion shown in Figure 5.32, where the dispersion is plotted as a function of frequency with the corresponding wavelength values shown on the lower x-axis.
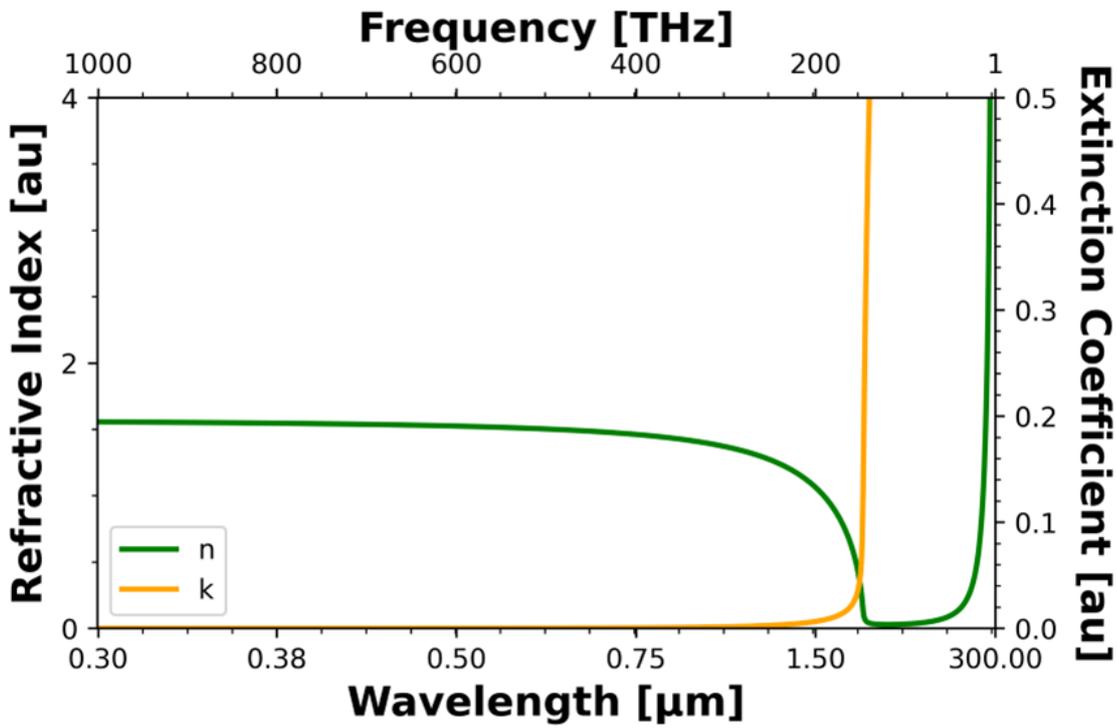
Figure 5.33: Refractive index (green) and extinction coefficient (orange) components of the complex refractive index of sputter deposited ITO with a 27% oxygen flow concentration during deposition.

The complex refractive index components are calculated from the dispersion curves in Figure 5.32 using Equation 4.4. The refractive index (n) and the extinction coefficient (k) are shown in green and orange lines, respectively, in Figure 5.33.

6. Metal Oxide Semiconductor Capacitor

Chapter Synopsis

This chapter presents a confirmation of suspected metal-insulator-semiconductor behaviour described in the literature as "needle-like" behaviour, where the metal and semiconductor materials penetrate the insulator in a MOS structure. I will also discuss the capacitance characterisation of the MOS structure used for optical modulation in this research. The capacitor I designed and fabricated had a capacitance too small to measure accurately, as a result I had to design and fabricate an array of capacitors in parallel to bring the total capacitance to a measurable value. The chapter concludes with some elemental analysis through the capacitor cross-section to confirm the literature predictions. Given further measurements, and time, the work discussed in this chapter could form an original contribution to the literature.

6.1.   Expected Capacitor Performance

The performance of the modulators discussed in my work relies heavily on the performance of the metal-oxide-semiconductor (MOS) capacitor, especially the magnitude of electron accumulation at the oxide/semiconductor interface (SU8/ITO).
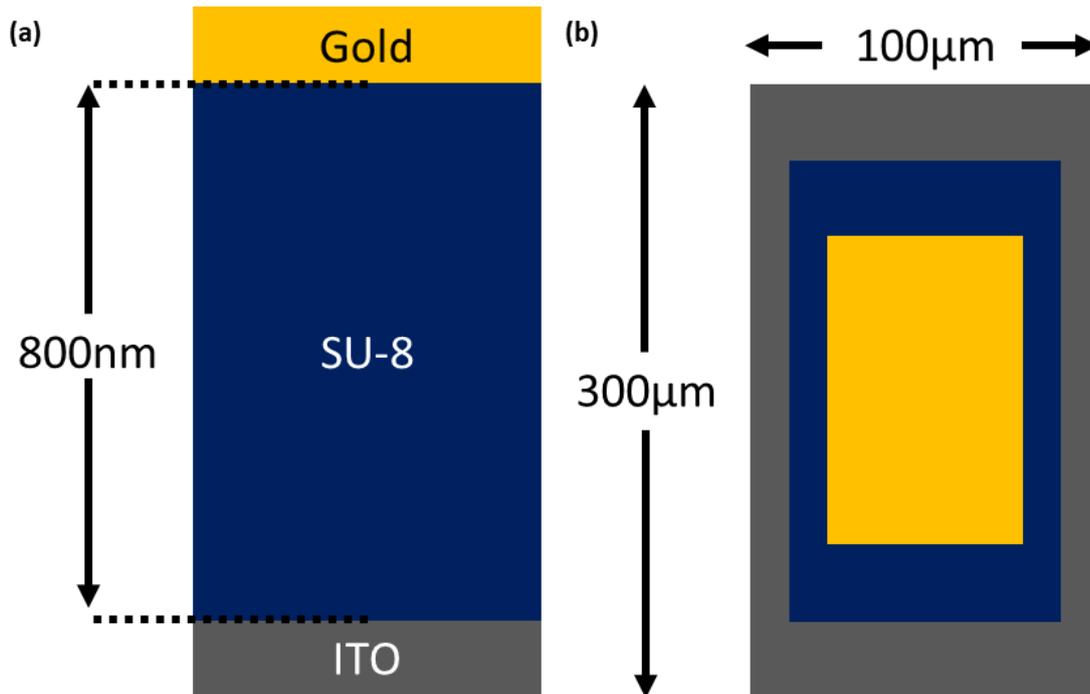
Figure 6.1: Schematic of the (a) cross-section and (b) top-down view of the capacitor with approximate layer thickness and dimensions.

Given the approximate dimensions of the capacitor, shown in Figure 6.1 with the (a) SU-8 thickness and (b) ITO capacitor plate dimensions, and the parallel plate capacitor equation (Equation 2.48), the predicted capacitance is approximately 1pF. In this chapter, I present a measurement technique used to verify the capacitance prediction and discuss the possible discrepancies between the physical and electrical thickness of the SU-8 layer and the impact on the modulator performance. The dimensions shown in Figure 6.1 are an approximation. To obtain a true estimate of the capacitance I measured the thickness of the SU-8 and the area of the ITO layer using surface profilometry (section 4.5.1).

6.1.1.   Capacitor Plate Area

In the ideal MOS capacitor theory, the semiconductor and metal plate areas are considered equal. However, due to fabrication constraints, I had to make the top gold contact and the bottom ITO contact different sizes to prevent the two layers touching in an over-exposure scenario (section 5.1.3). As such, it is more appropriate to consider the surface area of the ITO layer, given that the desired effect is the accumulation of electrons at the ITO/SU8 interface.
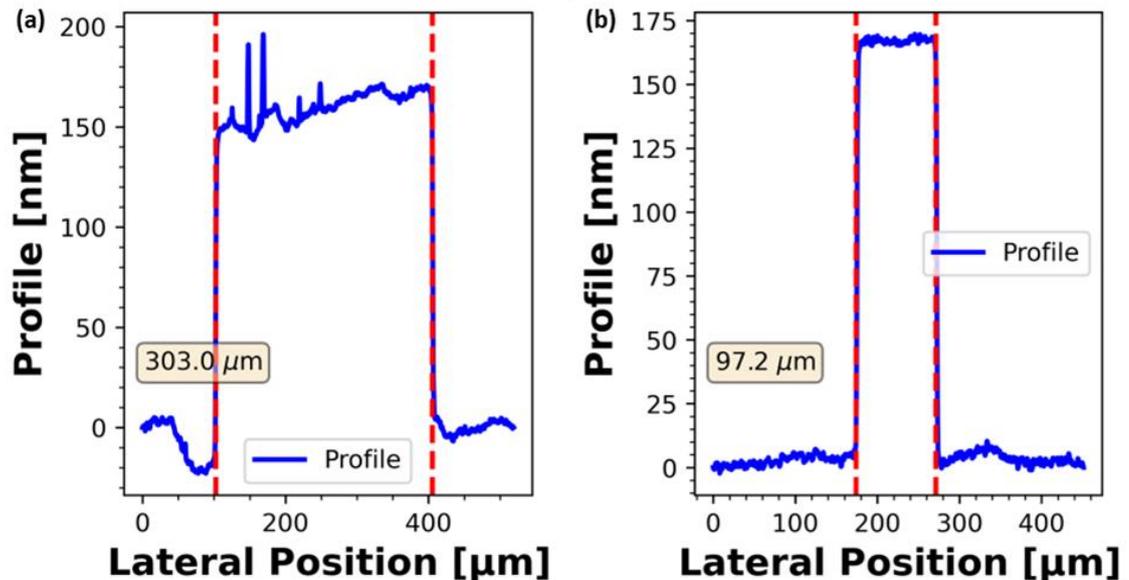
Figure 6.2: Surface profile analysis of indium tin oxide contact measuring (a) length and (b) width of the surface profile.

Surface profilometry was used to determine the dimensions of the example capacitor used in this chapter to measure the length (L) and width (w) of the ITO capacitor contact, with one of each of the measurements shown in Figure 6.2 (a) and (b) respectively. I took multiple measurements of the region, allowing for an average value and error to be calculated (Appendix section 9.1.1). This yielded an ITO capacitor contact with an area of (308±1) µm by (98.0±0.3) µm, so (30200±100) µm^2.

6.1.2.   Insulator Layer Thickness

Similarly, from Equation 2.48, the insulator layer thickness (d) and relative dielectric constant is required to fully characterise the MOS capacitor. I use an oxide layer of SU-8, which has a reported dielectric constant of 3.0 [299], and we can perform the same surface profile analysis as with the ITO layer to determine the SU8 layer thickness.
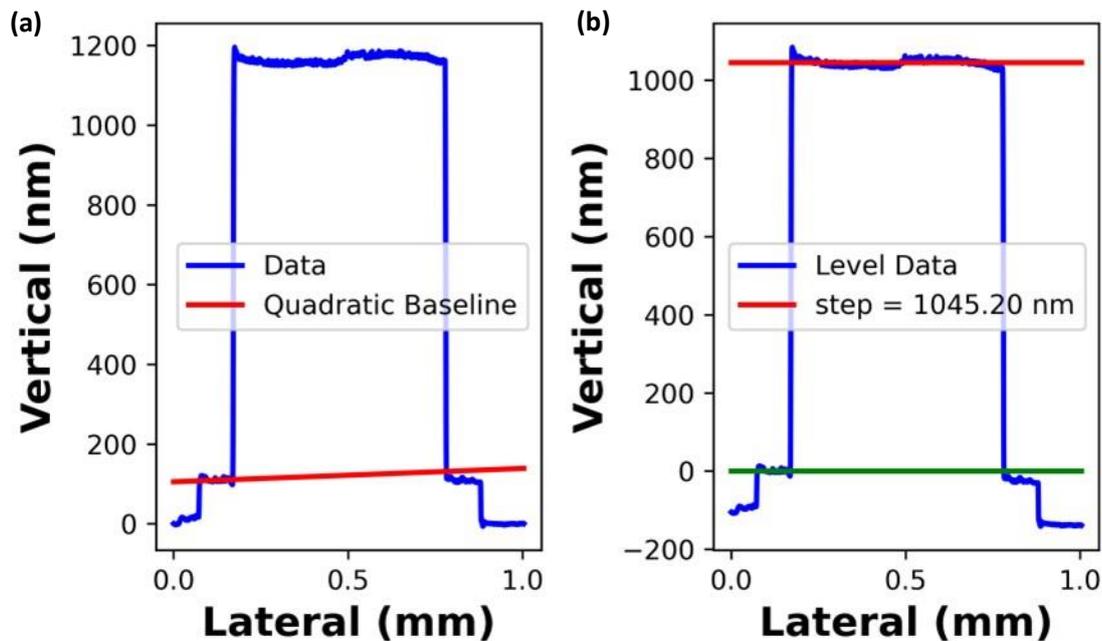
Figure 6.3: Surface profile analysis of SU-8 capacitor layer showing (a) distorted and (b) levelled surface profile.

An example of the surface profilometry on the SU-8 layer is shown in Figure 6.3, where the unlevelled (a) data and measured step height (b) data are shown in blue. The waveguide under the ITO and SU-8 layers is visible with a bump in the surface profile in the centre of the figure. I performed multiple measurements, took an average value, and calculated an error (as described in Appendix section 9.1.1) to measure a layer thickness of (760±10) nm for an SU-8 layer of 7:2 mix of SU-8 2000.5: 2050 (Appendix section 9.5).

6.1.3.   Capacitance Prediction

Thus, using the measured dimensions of the capacitor plate area and oxide layer thickness, and the literature values for the permittivity of free space and the relative dielectric constant, we

can use Equation 2.49 to calculate the expected capacitance. With appropriate errors, this can be calculated to be $(1.05 \pm 0.01) \times 10^{-12}$ F, or ~1 pF.

6.2.   Characterisation of the Capacitor

The expected capacitance of 1 pF, resulting from the small dimensions of the capacitor, is difficult to measure experimentally. Capacitors in parallel combine such that the total capacitance (C_T)

$$C_T = C_1 + C_2 + \cdots + C_{N-1} + C_N$$

Equation 6.1

Is the sum of the capacitances in the circuit. As such, the capacitance of a single modulator may be measured by increasing the number of capacitors in the circuit. I designed a measurement circuit with 32 capacitors in parallel, as shown in Figure 6.1(a).

(a)

(b)
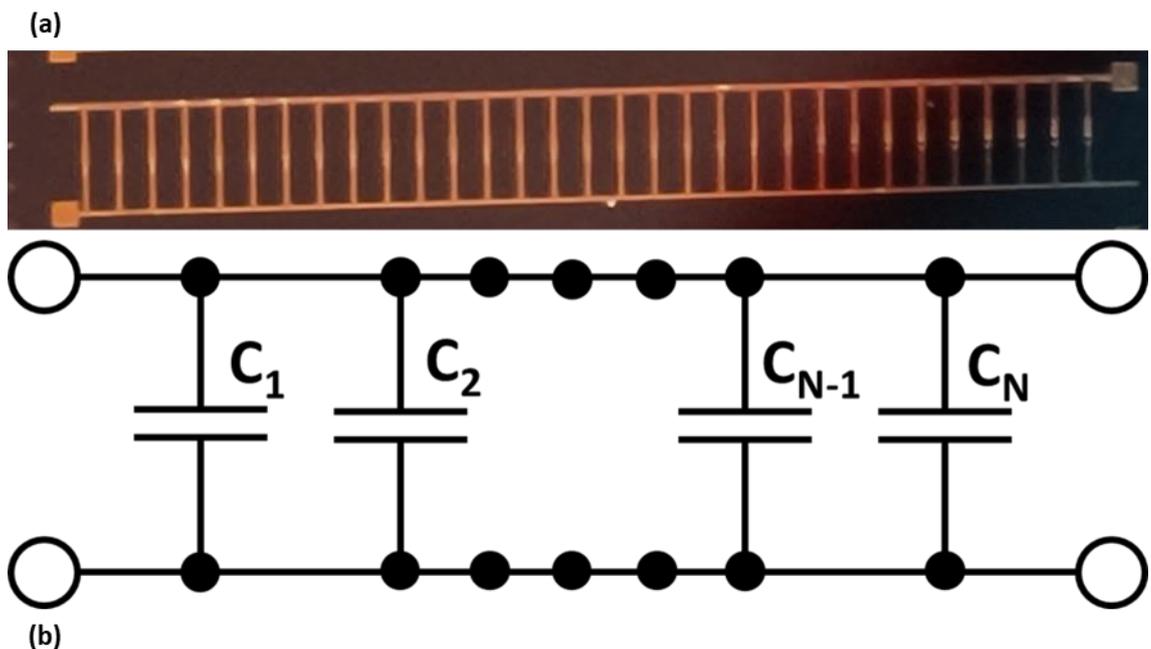
Figure 6.4: Metal-oxide-semiconductor capacitors in parallel, showing the (a) fabricated chip and (b) an illustration of the equivalent circuit.

The principle of the measurement is that many capacitors in parallel will sum up to a measurable capacitance. Thus, given that there are 32 equal capacitors, with an expected 1 pF capacitance, in the row in Figure 6.4, I expect a total capacitance of 32 pF from the measurement.

I used the same fabrication techniques discussed in section 5.6 to create rows of parallel capacitors, shown in Figure 6.4(a), with the row containing 32 capacitors. Here, I will discuss how the investigation was performed, and discuss the effect of the oxide layer on the capacitor performance.

6.2.1.   Capacitance Measurements

The charging of a capacitor is not instantaneous. Capacitors have I-V characteristics which depend exponentially on time, with a characteristic time constant (τ)

$$\tau = RC$$ Equation 6.2

and a voltage dependence across the capacitor (V_C) as a function of the supply voltage (V_S) described by:

$$V_C = V_S \left(1 - e^{-\frac{t}{RC}} \right)$$ Equation 6.3

Hence, we can use the time dependence of a simple circuit, shown in Figure 6.5, to measure the size of an unknown capacitor [300].
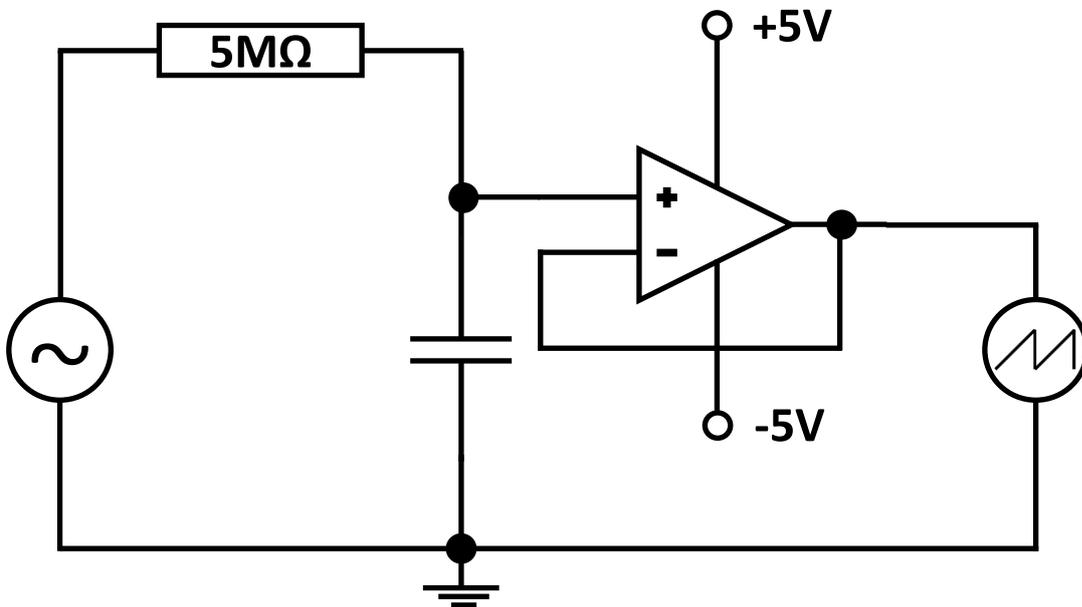
Figure 6.5: Illustration of an RC charging circuit with (from left to right) an input alternating current supply, a 5MΩ resistor, a capacitor, an operational amplifier, and an oscilloscope.

The input signal, on the far left, is an alternating current (AC) power supply, which feeds a 5MΩ resistor and the 32 capacitors indicated by the capacitor symbol. I chose a 5MΩ resistor to make

the time constant more easily measurable, i.e., around 150μs, given the ~32 pF capacitance of the array.

The output of the circuit is measured using a Multicomp MP720107 oscilloscope, capable of measuring in the GHz regime, but the signal is amplified by a 10 V operational amplifier (op-amp). Without the inclusion of the op-amp, the only measured signal would also be proportional to the internal impedance of the oscilloscope. The on-chip capacitors were connected to the circuit using electrical probes, which have their own impedance that can interfere with the measurement. This circuit capacitance was measured at 8pF and will be included as an error in the measurement.

The capacitance was calculated by taking 67% of the input voltage to obtain the time constant, and then using Equation 6.2 to determine the capacitance of the 32 parallel capacitors. The calculation is checked by fitting an exponential to the measured capacitor voltage signal using Equation 6.3. The result is shown in Figure 6.6.
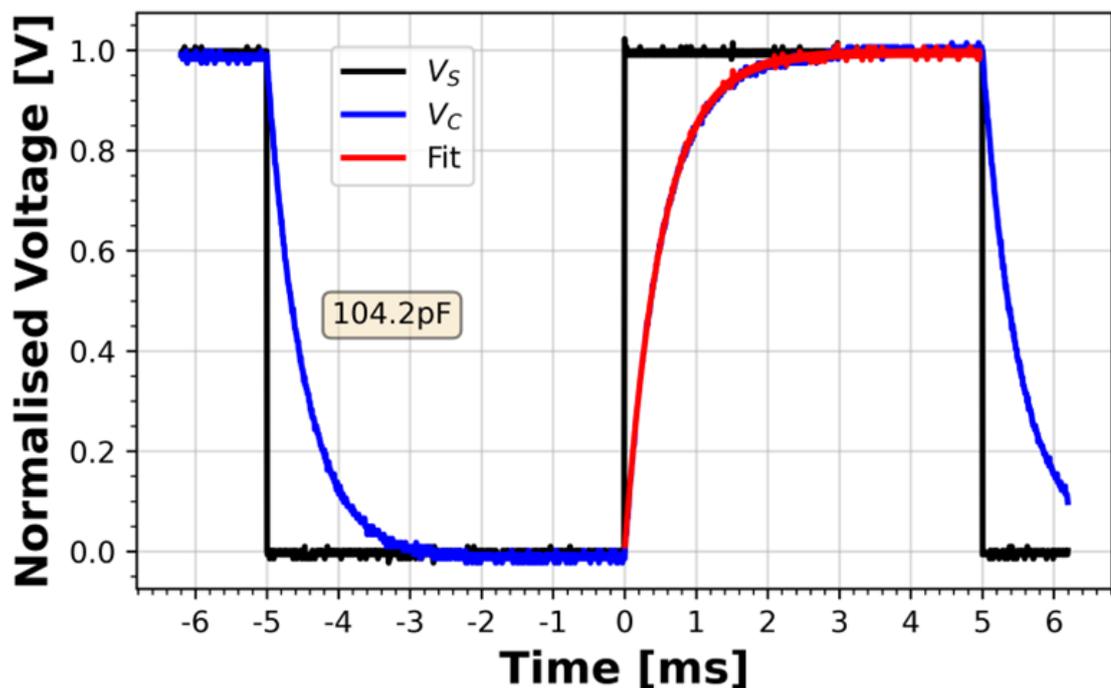
Figure 6.6: Capacitance measurement for 32 parallel capacitors in an RC charging circuit. Input voltage alternates between -0.5 and 0.5 V, normalised to 0-1 V, and is shown in black. The voltage across the capacitor is shown in blue, with the charging region between 0 and 0.005 s. The exponential fit using the calculated capacitance is shown in red.

The measured output, shown in Figure 6.6 includes a discharging region (-0.005-0s) and a charging region (0-0.005s), the input voltage supply (-0.5-0.5V, normalised to 0-1V) in black, the

voltage measured across the capacitor in blue, and the fitted exponential with the calculated capacitance in red. The exponential fit is good, and the measured capacitance for 32 parallel capacitors is shown as (104±8) pF. Using Equation 6.1, this yields an individual capacitance of (3.3±0.3) pF, approximately 3 times the expected capacitance. I will explore this discrepancy in the next section.

6.2.2.   Suggestions for Increased Capacitance

It is reported in the literature [80] that it is possible for the metal and semiconductor materials to extend into the oxide material, in a process described as "needle-like gold filaments", for glass-like materials. This would cause a decrease in effective the layer thickness (d) and thus increase the measured capacitance. Given that the dielectric constant is well documented for SU-8, and that the lateral dimensions of the capacitor are measured to a high level of precision, the "leaking" of the metal and semiconductor into the oxide layer is therefore the most plausible cause. I therefore decided to conduct further analysis to better understand whether the same effect may have happened during my fabrication process and whether metals may have leaked into the insulator layer thus effectively reducing its thickness and increasing the capacitance.

6.3.   Energy Dispersive X-ray Spectroscopy

Energy-dispersive x-ray spectroscopy (EDS) is a technique used to measure the elemental composition of a sample using high energy x-rays to eject core electrons from an atom. Removing these electrons leaves behind a hole that a higher energy electron can fill in a process that releases energy unique to each element. As such, the bombardment of a sample with high energy x-rays can determine the elemental composition of the material, and in what proportion they are present [301], [302]. The technique relies on Moseley's Law [303] and can be integrated into many electron microscopes. I used AZtecOne software, in conjunction with a scanning electron microscope, to measure the elemental composition of the cross-section of the MOS capacitors used in this research and in this section, I will discuss the relevant process and results.

6.3.1.   Sample Preparation

The EDS software is integrated into the JEOL JSM-7800F Prime scanning electron microscope (SEM) available at York, allowing precise alignment of the detector and imaging apparatus. I used the MOS capacitor described above and cleaved the substrate through the middle of the capacitor stack, as shown in Figure 6.7(b). This gave me a smooth facet edge to look down into with the EDS detector. The sample was mounted vertically into the SEM system due to detector alignment, as shown in Figure 6.7(a). The stage allows for minor angle adjustments up to 45°, and these have been accounted for in my analysis.
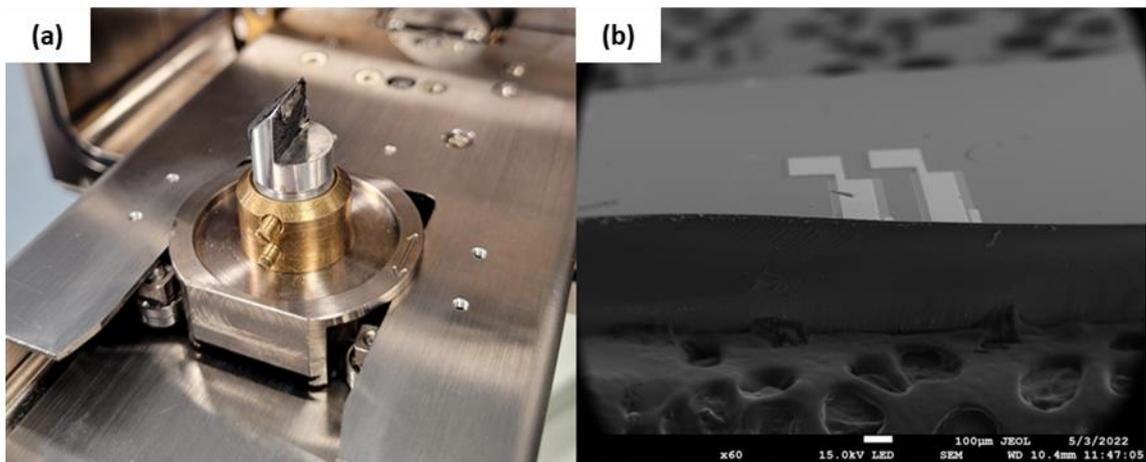
Figure 6.7: (a) Vertical sample mounting in the scanning electron microscope loading stage, (b) cleaved cross-section of the metal-oxide-semiconductor capacitor viewed under the scanning electron microscope at a 20° angle.

Despite the bulk metal stage (onto which the sample was mounted), the bulk silicon substrate, and metallic elements present in the capacitor, the oxide layer can still charge under examination and blur the resulting images. To prevent this, a thin (~5 nm) layer of platinum and palladium was deposited onto the facet edge to act as a charge dissipation layer. The EDS system detects these elements, but they can be ignored from the resulting analysis.

6.3.2.   Elemental Maps

The system can detect most of the elements of the periodic table, though I have narrowed the search parameters down to gold, tin, indium, silicon, oxygen, and carbon. The latter is the primary constituent of SU-8, a polymer, and will be treated as an indicator of the SU-8 layer in this analysis. It should be noted, however, that carbon may be present on the entire facet due

to handling and environmental factors. For the elemental maps shown in Figure 6.8, the sample was angled at 25° from vertical. For clarity, I have marked the regions of interest with dotted lines.
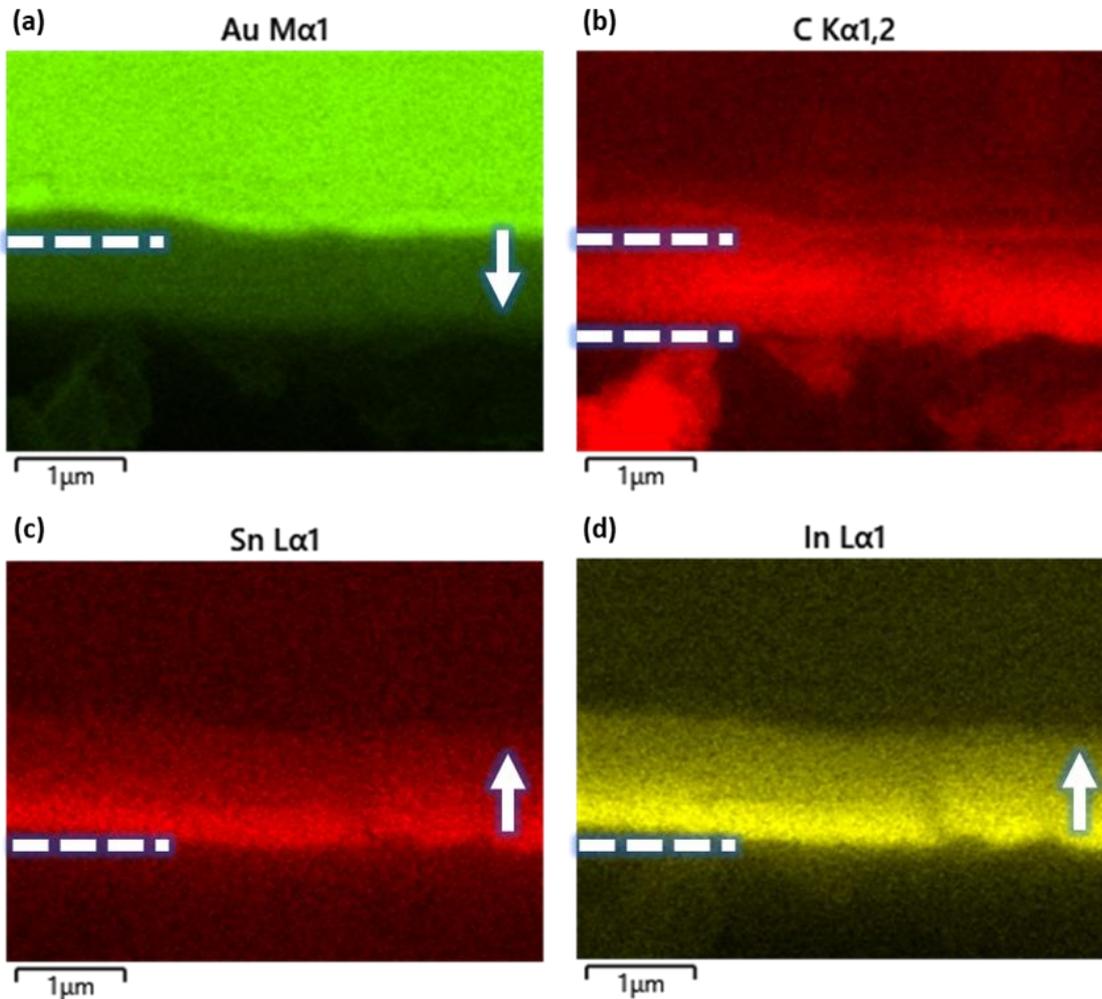
Figure 6.8: Energy-dispersive x-ray spectroscopy maps scanning for (a) gold, (b) carbon, (c) tin, and (d) indium elemental presence within a target region of the metal-oxide-semiconductor cross-section. The evidence of elemental leaking is indicated by an arrow.

The EDS detector scans across the image and measure the corresponding characteristic x-rays emitted by the sample to produce an elemental map of the area. These images are shown in Figure 6.8, for (a) gold, (b) carbon, (c) tin, and (d) indium. The top and bottom regions of the highest concentration of carbon is marked, and the corresponding regions in the gold, indium, and tin images are also marked.

It is clear from Figure 6.8(a) that there is some gold present in the carbon layer, with a faint green tint visible underneath the marked dotted line and indicated by an arrow. Considering that the images are shown with the underside of the gold angled away from the field of view,

this would indicate the presence of gold in the SU-8 layer, likely due to diffusion under the high temperatures of electron-beam evaporation (section 5.2.9). Equally I note some presence of tin and indium (Figure 6.8(c)(d)) in the carbon region. The heavily concentrated region of indium and tin, around the dotted line, is indicative of the desired ITO layer, and the slight offset in height is attributed to the angle of observation. However, the presence of both elements in the carbon is apparent when examining the less bright region in both images and shown by an arrow.

6.3.3.   Elemental Quantification

The elemental maps are only able to show a visual representation of the elemental concentration in the observed region. While it is possible to use the scale bar to determine the depth of penetration, it is far easier to use an element-wise counts per second (CPS) measurement as a function of the position in the image, which is a built-in software tool. This is done by taking a portion of the field of view and scanning along one axis only (Figure 6.9(a)). The counts per second are proportional to the detected x-rays per second and may be considered to quantify the concentration of that element in the region.
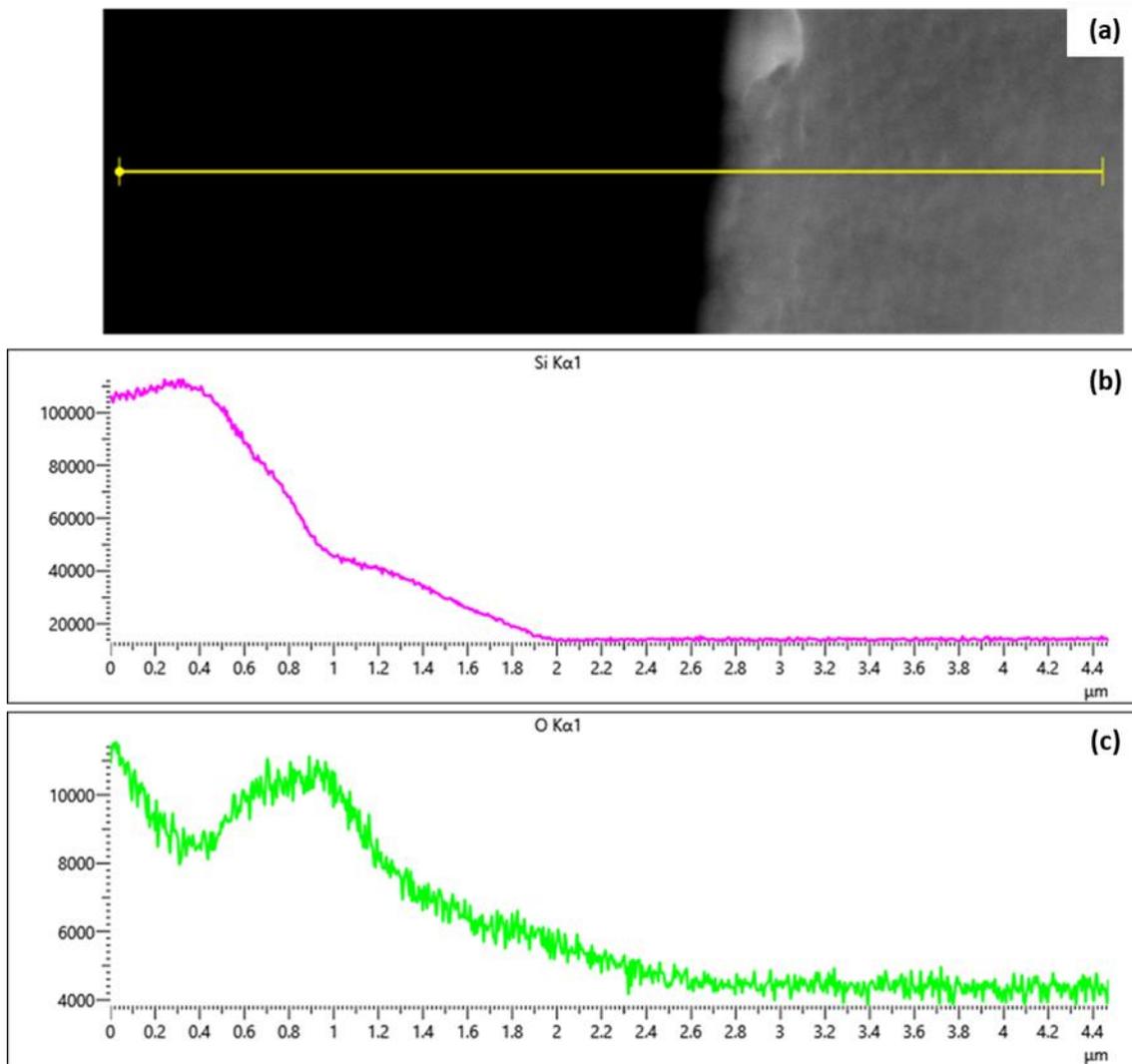
Figure 6.9: (a) Scanning electron micrograph of the region of interest and the corresponding elemental counts per second as a function of the distance across the cross-section of the region of interest for (b) silicon and (c) oxygen.

I measured along one axis of the same sample as shown in Figure 6.8, and measured the CPS of silicon, oxygen, gold, carbon and indium, with the assumption that the overlap between indium and tin would be equal, as indicated in Figure 6.8(c) and (d). I have divided the images for separate analysis. First, the silicon and oxygen elements are shown in Figure 6.9(b) and (c), which corresponds to the substrate layers of silicon and silicon dioxide, and for oxygen in the ITO layer. There is a high silicon presence in the substrate region, and the oxygen is high in both the substrate and ITO region, as expected.
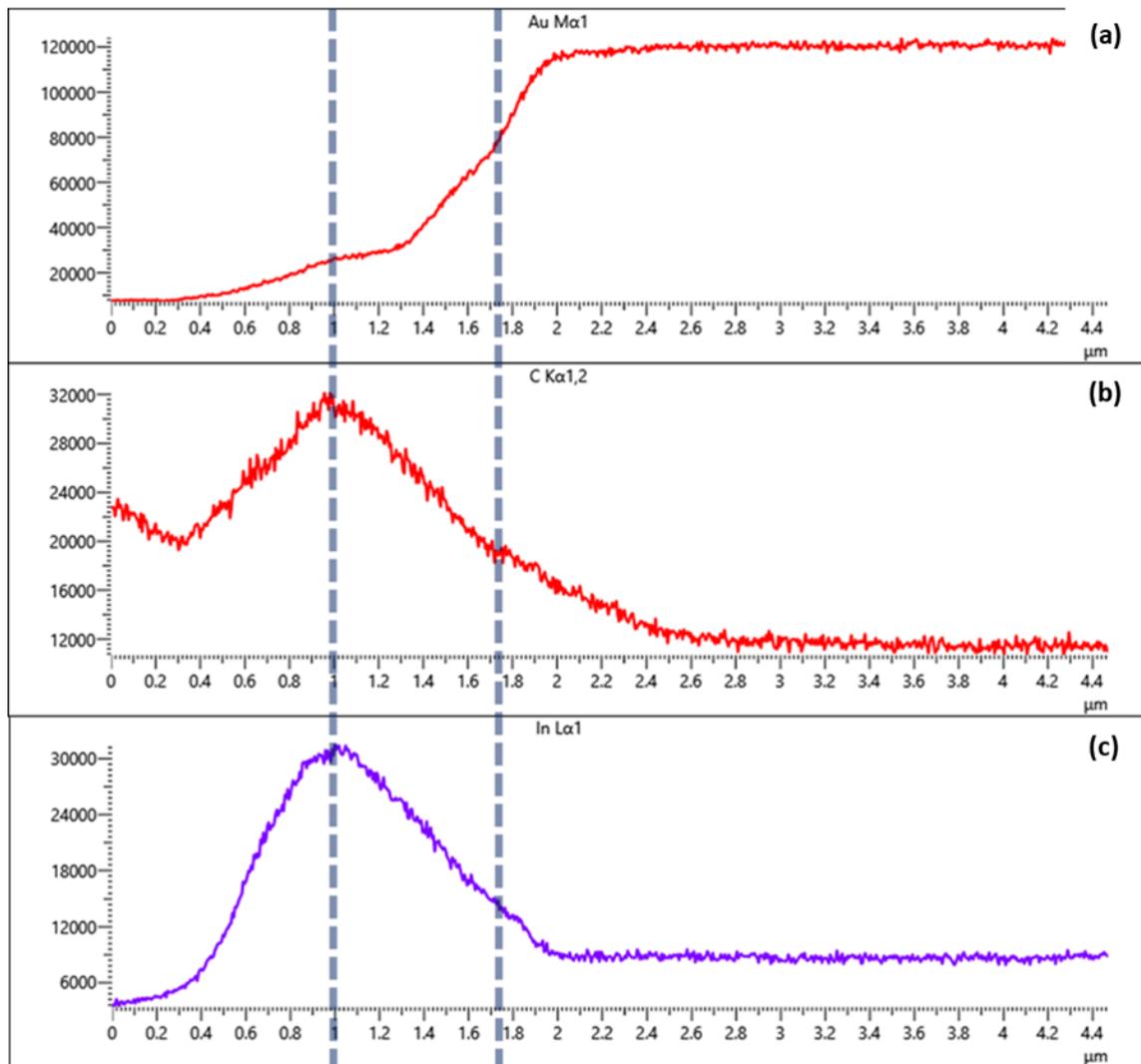
Figure 6.10: Elemental counts per second as a function of the distance across the cross-section of the rgion of interest for (a) gold, (b) carbon, and (c) indium, with the approximate insulator layer thickness marked with blue dotted lines.

As for the capacitor elements, shown in Figure 6.10 for (a) gold, (b) carbon, and (c) indium, there is a clear overlap across the oxide region of the MOS capacitor. I have marked the expected layer thickness with blue dotted lines, beginning at the end of the adhesion layer beneath the gold layer where the gold count drops steeply (~1.7 µm) and ending at the peak of the carbon count (~1µm). The measured SU8 thickness for the examined sample was approximately 750 nm, so this region is appropriate for the sample.

There is a strong presence of both gold and indium (and tin) in the oxide region of the capacitor. While the count diminishes quickly across the region, the depth of penetration of the electrons is sufficient to examine the depth of the capacitor (15 kV), and it is plausible to assume there are regions through the capacitor where gold or ITO extend significantly into the SU-8. The EDS

measurements therefore confirm my hypothesis in section 6.2.2 that the conductors "bleed" into the SU-8 layer, thus reducing the electrical thickness of the SU-8 compared to its physical thickness. While it is difficult to further quantify this statement, a reduction in electrical thickness by a factor 3 is entirely plausible.

6.4.   Capacitor Summary

We can draw some important conclusions from my very brief material analysis. Firstly, there is some elemental bleeding of the metal and semiconductor layers into the SU-8 layer. This is likely due to the diffusion of gold, indium, and tin during the electron-beam evaporation process used to deposit the gold layer. The temperature of evaporation of gold is approximately 1700-1800K (section 5.2.9), and there is significant radiative heat transfer towards the sample, which will therefore heat up to typically 500-600K, thereby causing some diffusion. Secondly, we have assumed, perhaps naively, that the SU-8 layer is pinhole-free. This may be possible, but it is known that many SU-8 layers exhibit pinholes.

From the measurements I have taken, I cannot directly quantify the depth of penetration of gold, indium, and tin at every point through the SU-8 layer, i.e., through the length of the modulator. To obtain these measurements a cut-back experiment, where the cross-section of the capacitor is measured at small increments through the modulator, would be a far better method. Because of this, it is not apparent if the counts per second values shown in Figure 6.10 are the result of surface element counts or a function of the x-ray penetration and reflection through the modulator. Using the calculation for the capacitance in section 6.2.1, which yielded an SU-8 electrical thickness of 300nm, the data shown in Figure 6.10 shows that this thickness is plausible and therefore the increase in measured capacitance compared to predicted capacitance is justified. Showing that the SU-8 is not a perfect insulator, despite its glass-like behaviour once hard-baked, agrees with the discussions amongst the literature about "needle-like gold filaments". I suggest further experimentation to determine the validity of this hypothesis.

7. Phase Modulator Measurements

Chapter Synopsis

This chapter presents the optical modulation results for the electro-optic (EO) modulators that I have designed and fabricated, as discussed in the previous chapters. The chapter discusses the optical setup and components required to perform the measurements, for which I acknowledge the influence and advice of Dr Christopher Reardon. From there I will discuss the results and a dual-mode phenomenon observed through an increased performance, which I have verified through finite-difference simulation. The chapter concludes with a comparison between my EO modulators and those in the literature, discussed in chapter 1. I acknowledge again the assistance from Dr Christopher Reardon for his assistance with the simulation results. The optical setup I used is shown in Figure 7.1(a) and illustrated in 2D in Figure 7.1(b).

7.1. Guiding Light
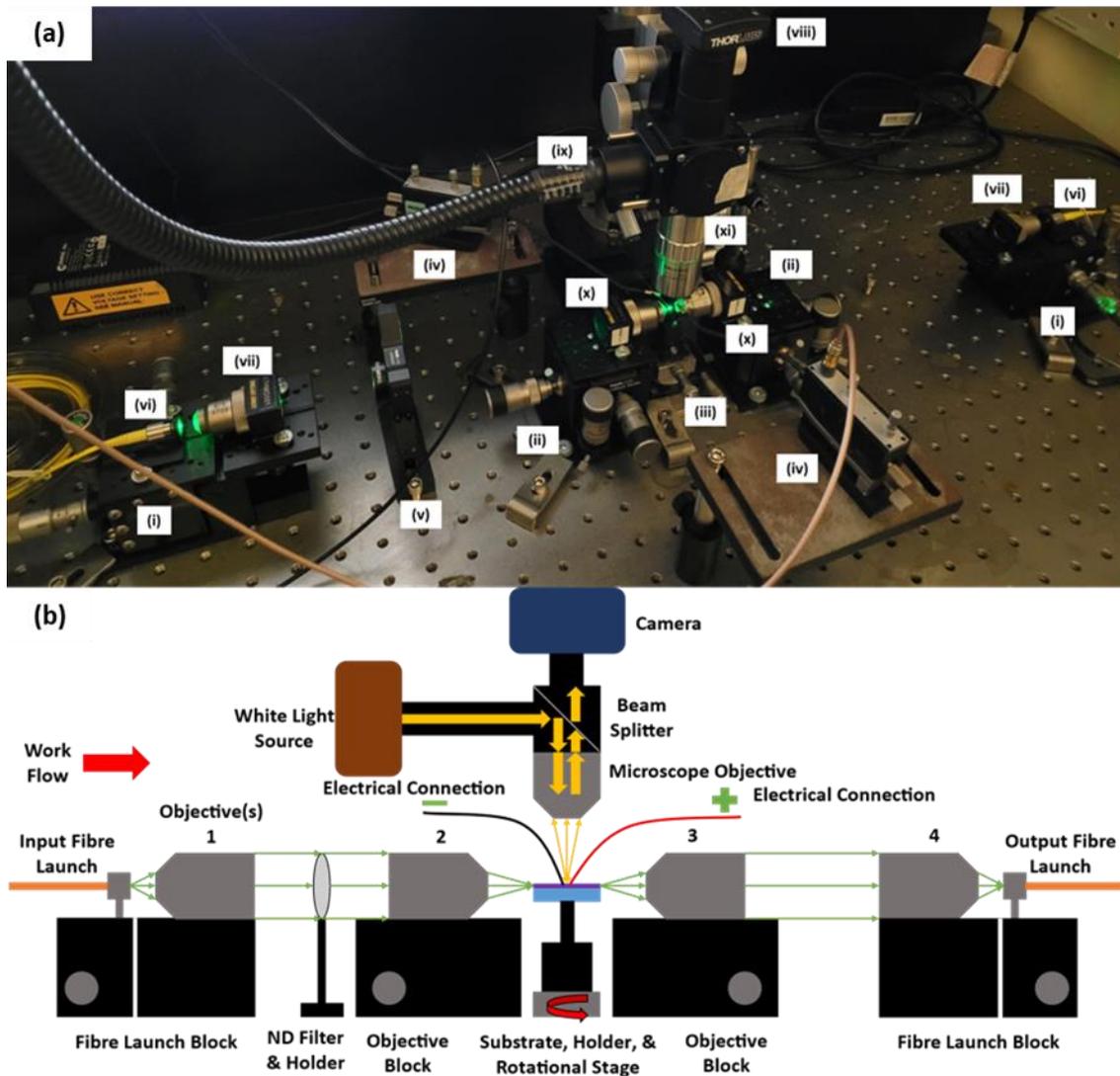
Figure 7.1: (a) Photograph and (b) illustrated schematic of the optical setup used to obtain amplitude and phase modulation results in this research. The individual components are discussed throughout this chapter.

While not directly apparent from Figure 7.1(a), the optical measurements in this research are conducted primarily using free-space optics, i.e., the light travels predominantly through air. The exception is that the input, output, and modulated signal are carried in optical fibres and the silicon nitride waveguide. Here I will discuss the components used to achieve free-space operation and, where appropriate, I will reference the components to Figure 7.1(a).

7.1.1.   Source of Light

The silicon nitride platform and the Y-junction splitters used on the Mach-Zehnder modulator (MZM) designs (section 2.5.1) were specifically chosen/designed to accept a wide range of wavelengths in the visible spectrum. In this chapter I will present results using a 520, 635, and 785 nm lasers. The 635 and 785 nm lasers are Thorlabs laser diodes, pigtailed with single mode optical fibres (LPS-635/785-FC), and the 520nm laser is a Thorlabs PL520 laser diode controlled by a Thorlabs CLD1010 temperature controller for stability of the source. The output powers were measured using a Thorlabs PM100D power meter and are shown in Table 7.1.

Table 7.1: Input power and output power of the three laser sources used in this research.

Wavelength	Drive Current	Output Power
[nm]	[mA]	[mW]
520	40	0.84
635	70	1.13
785	70	1.24

The 635nm and 785nm are powered by a 9V battery power supply and draw approximately 70mA consistently, while the 520nm source is powered by the temperature controller (mains electricity) and is controlled by an input current, as shown in Table 7.1. The lasers all output into an optical fibre, shown in Figure 7.1(a) at position (vi).

7.1.2.   Waveguide Coupling

To further emphasise the multi-wavelength capability of the EO modulators, I used a free-space coupling technique known as "end-fire", a technique that involves butting the input signal directly up to the waveguide facet. In the initial stages of my research, I did consider using a grating coupler, which is a periodic structure of alternating refractive indices (section 2.1.3) that operates in the diffraction regime to couple the fundamental mode into a neighbouring waveguide structure [304]. However, as the diffraction regime requires the wavelength of the light inside the grating material being smaller than the grating period, use of a grating coupler limits the accepted wavelengths of the modulator and can introduce excess loss into the system.

Figure 7.2: Fibre launch block with input/output optical fibre (left) a 16x magnification (11mm focal length) objective on a three-axis adjustable optical block.

To achieve end-fire coupling, a collimated beam is focused to a spot matching the dimensions of the waveguide facet, thus reducing scattering loss and encouraging maximum coupling of the input signal. Naturally, a high index contrast between the waveguide and surrounding materials aids in the coupling process, as described in section 2.1.2. A collimated beam of light is a beam propagating in a homogenous medium with low divergence, i.e., a laser propagating in air with a beam radius that does not experience significant change in a reasonable propagation distance. It is typical for this distance, known as the Rayleigh length, to be several orders of magnitude larger than the required propagation distance [305]. Collimation is achieved using an objective lens in the beam path at an appropriate distance, focal length, and magnification for the wavelength and fibre used. I used a fixed 16x magnification objective and adjusted the distance between the lens and the fibre using a three-axis stage to achieve collimation, as pictured in Figure 7.2. The same system may be used to focus a collimated beam into a fibre, and I used the same objective, fibre, and stage to couple the output signal into the measurement apparatus, shown in Figure 7.1(a) at position (i) stage, (vi) input/output fibre, and (viii) objective lens.

The collimated signal is then focused on to the waveguide facet using a 60x magnification 2.8mm focal length objective where the beam spot size (W) may be calculated using:

$$W = \frac{4M^2\lambda f}{\pi D}$$

<div align="right">*Equation 7.1*</div>

Assuming the input is collimated and may be considered a Gaussian Beam (beam quality M=1) with diameter (D) no greater than 1mm, the beam spot size is determined by the input wavelength (λ) and the input objective focal length (f).
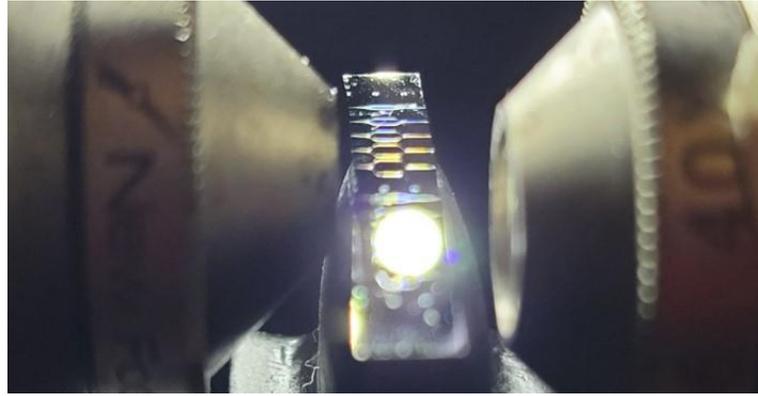
Figure 7.3: Facet coupled input (left) and output (right) objectives, with magnification and focal lengths 60x and 40x, 2.8mm and 4.5mm, respectively.

The objective at the input facet, pictured in Figure 7.3 (left), has a focal length of 2.8 mm, yielding the spot sizes given in Table 7.2. Similarly, the output objective, with a focal length of 4.5 mm, the acceptance range of the output facet is shown in Table 7.2. The output objective was chosen for its wide acceptance range to collect as much light as possible from the output facet of the waveguide [306].

Table 7.2: Input spot size and output acceptance spot size for input wavelengths used in this research.

Wavelength	Input Spot Size	Output Acceptance Spot Size
[nm]	[µm]	[µm]
520	1.66	3.18
635	2.02	3.88
785	2.50	4.80

The waveguide facets, as discussed in section 3.4.2, were 5µm in width and were chosen to accommodate minor variations in beam diameter and coupling angle, and to allow minor imperfections on the waveguide facet to be avoided. The input and output objectives, shown in Figure 7.1(a) at position (x) and are mounted to a three-axis stage at position (ii).

To further improve coupling into the waveguide and allow for tolerances, the chip is mounted to a holder on a two-axis (rotation and one positional) stage. The stage is custom printed using a Formlabs Form 3 printer and coated in an anti-reflective paint to prevent scattering from the holder. The stage can rotate about the centre and can move in the axis perpendicular to beam propagation in Figure 7.3. This allows for a full coverage of a 15 mm length chip in conjunction with the three-axis stage of the objectives, i.e., both objectives and stage may move in the same direction, while the objectives also move in the vertical direction and in the direction of beam propagation. The rotational stage is shown in Figure 7.1(a) at position (iii).

7.1.3.   Filtering

The silicon nitride waveguide was chosen to provide a loss-less platform for wavelengths in the visible spectrum and larger [307] and, given the magnitude of the input power of the lasers used in this research, it follows that it could be possible to oversaturate the measurement instruments and thus introduce inaccuracies into the measurements. To obtain a finer control of the output intensity of the modulator, a neutral density (ND) filter may be used to reduce the intensity of the input signal. An ND filter reduces the intensity of all wavelengths equally, thus providing a linear decrease in intensity across the measured wavelengths in this research. The fractional transmittance (T_F)

$$T_F = \frac{I}{I_0} = 10^{-d}$$

Equation 7.2

of an ND filter is the measure of the post-filter intensity (I) and the input intensity (I_0) and is often expressed on the ND filter as an optical density (d), as expressed in Equation 7.2. In my work, I used a ND 0.6 filter, which reduces the input signal by 25%, and this may then be accounted for in further calculations [308]. To prevent oversaturation of the camera in the optical setup, I placed the ND filter in the path of the input collimated beam, shown in Figure 7.1(a) at position (v).

7.2.   Modulation Technique

Indium tin oxide (ITO), as I have discussed chapter 4, has a permittivity described by the Drude model which is governed by the movement of free carriers (electrons) within the material. Modulation of the input signal is achieved through the control of the permittivity of ITO through an applied gate voltage across a metal-oxide-semiconductor (MOS) capacitor as described in section 2.6.3. In this section I will discuss how I measured amplitude and phase modulation using the electro-optic (EO) modulators in this research, using the MOS capacitor structure describe in chapter 6.

7.2.1.   Applying a Gate Voltage

To apply the gate voltage across the MOS capacitor, two gold contact patches were included in the modulator design (section 3.5.2), where each arm of the capacitor was insulated from each

other to encourage the electric field to be concentrated above the waveguide. Electrical contact was made using two 4-axis electrical probes, as shown in Figure 7.4(a).
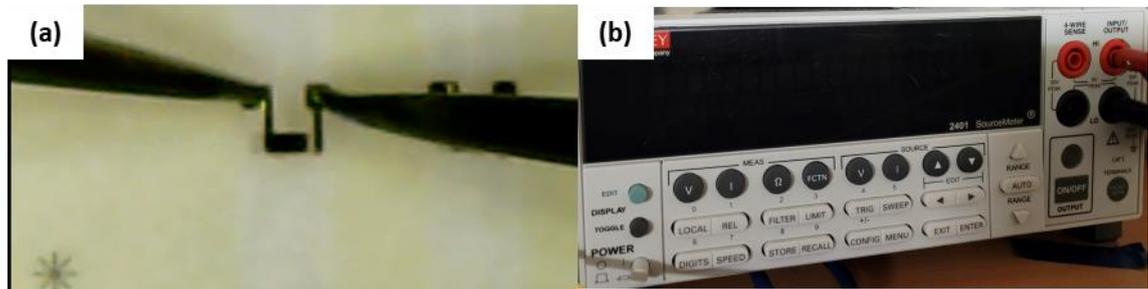
Figure 7.4: (a) Electrical contact made to the metal-oxide-semiconductor capacitor arms with 4-axis probes. (b) Keithley 2440 Bench Top DC Power Supply

The electrical probes are relatively large but allowed for a 200x200 µm contact patch and minimal resistance in applying the gate voltage. I used a Keithley 2440 DC Power Supply (Figure 7.4(b)). The gate voltage is considered positive, in this research, when the positive terminal is connected to the gold plate of the capacitor, thus allowing electrons to accumulate in the ITO layer.

The electrical probes have a magnetic base to ensure their position is fixed and are placed either side of the chip, as shown in Figure 7.1(a) at position (iv). I found separating the probes in this way meant that there was very little chance of the tips contacting each other and potentially causing damage to the capacitor or waveguide, or cross contact allowing a current to flow between the probes and thus neutralising the capacitor.

7.3.   Measuring Light

I employed two methods of measuring the output intensity of the modulator in this research, a Thorlabs DCC1545M CMOS camera and a Thorlabs PDF10A silicon detector. The camera was used to observe the chip through a microscope objective with 20x magnification and serves two purposes; firstly, to align the electrical contacts for the modulator and the input/output objectives, and secondly it can be used to measure the intensity of the scattered light from the chip. The camera requires illumination during observation/alignment, in the same fashion as an optical microscope requires illumination, as such a white light source is included and passes through a beam splitter as illustrated in Figure 7.1(b). The silicon detector is fibre coupled into a Tektronix TDS 1012B oscilloscope and serves to quantify the intensity of the out-coupled light

as a voltage on the detector. The camera, white light source, and microscope objective are shown in Figure 7.1(a) at positions (viii), (ix), and (xi) respectively, and the silicon detector is not shown. I will now briefly describe how the camera and silicon detector are used to measure the output intensity and how the data is processed.

7.3.1.   Camera Intensity

The camera has several properties that need to be manipulated to obtain an accurate measure of the output intensity. The first is the camera exposure time, the second is digitisation; the camera operates is monochrome and provides a pixel intensity 0-255, where 255 is fully saturated. The intensity measurements are taken with the white light source switched off, and the exposure time is adjusted under the same circumstances, i.e., the only light source is the input laser. The exposure time is adjusted so that the camera is not saturated, allowing for increases in intensity to be measured. While not a direct camera property, the ND filter strength, if used, is also noted and accounted for in the intensity measurements with the exposure time, such that the measured intensity (I)

$$I_r = I t_e T_F$$ *Equation 7.3*

is multiplied by the exposure time (t_e) and the fractional transmittance (T_F) to yield a raw intensity (I_r), as shown in Equation 7.3.
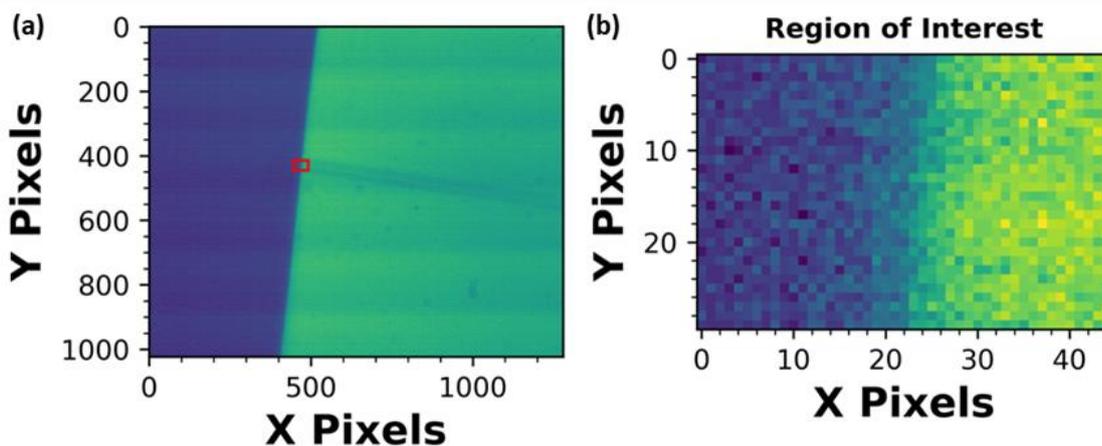
Figure 7.5: (a) Raw image from the camera showing the output waveguide facet from the top-down, the desired region of interest is marked with a red box and shown in (b).

The camera's field of view is significantly larger than the waveguide width, as shown in Figure 7.5(a), and this can affect the accuracy of the intensity measurement if scattered light is incident

on the camera lens from elsewhere on the chip. To reduce the impact this has on the measurement, I trim the image down to a selected region of interest (ROI), indicated by the red box in Figure 7.5(a). The resulting ROI can be seen in Figure 7.5(b), and it is within this region where the intensity is measured.

Multiple images are taken for each voltage increment to allow for average results and standard error on the mean errors to be calculated (Appendix section 9.1.1, which removes minor variations and serves to reduce experimental errors overall. To reduce external (background) noise from the measurements, the intensity of the room light is measured, i.e., a measurement without an input signal is measured. This value is then subtracted from the measurements as an offset. The error in the background corrected intensity is then calculated using standard quadrature methods [215] (Appendix section 9.1.2). The intensity is then calculated as a function of the ROI area, where again errors are calculated using the quadrature formulation. The processing of the measured images is analogous to the image processing software "ImageJ" [309], recreated in Python to allow for live processing and monitoring of the images.

7.3.2.   Oscilloscope Intensity

The oscilloscope intensity is measured as a voltage output of the silicon detector. The specification for the silicon detector states that the voltage out (V_{out})

$$V_{out} = P_{out}R(\lambda)G$$

Equation 7.4

is a function of the incident light power (P_{out}), the wavelength dependent responsivity ($R(\lambda)$), and the transimpedance gain (G), as described in Equation 7.4. The specification gives the transimpedance gain as 1×10^{12} V/A (±10%), and the responsivity for 520, 635, and 785 nm as approximately 0.35, 0.45, and 0.5 VW^{-1}, respectively. Recording a measurement using the oscilloscope, however, is significantly simpler than with camera intensity measurements.
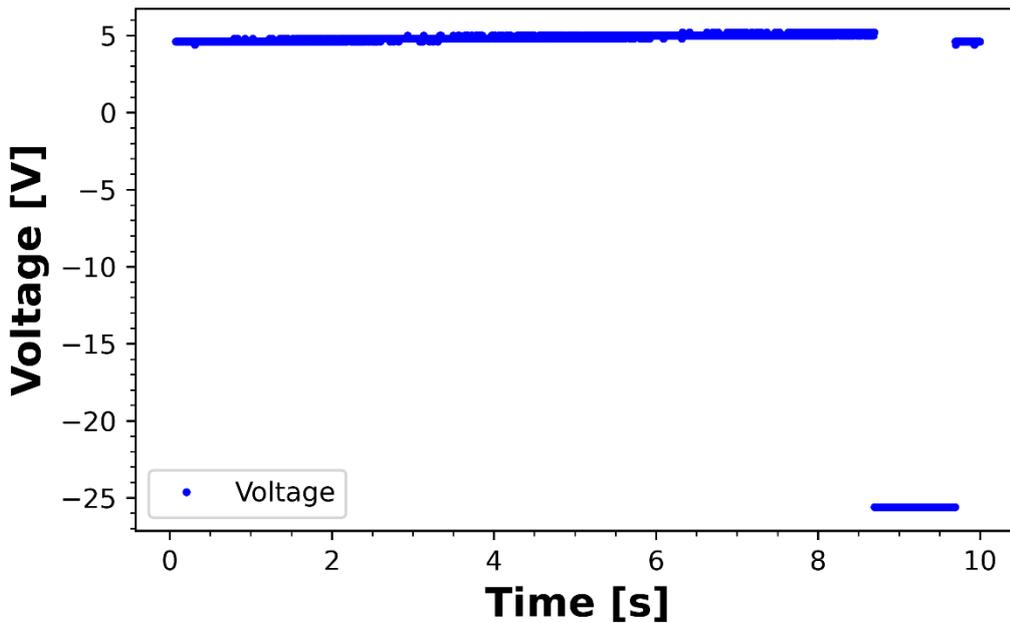
Figure 7.6: Oscilloscope voltage measurement.

An example of the voltage output measured by the oscilloscope is shown in Figure 7.6. The oscilloscope records continuously with a repeating time base. When a scan is saved, the voltage recorded drops below 0, which can be seen in Figure 7.6 between 8.5-9.5 s. As this region is clearly not accurate, it is removed from the data and further processing.

To introduce the average measurement, and standard error on the mean, the oscilloscope is set to record for 25 seconds. The average output voltage is recorded over this period and averaged over the length of the data set. As with the camera, a zero-input measurement is recorded to measure the background intensity of light in the room during experiments. This value is, of course, treated as an offset and is subtracted from the measured oscilloscope responses, with errors calculated using the quadrature formulation (Appendix section 9.1.2). As the output voltage is directly proportional to the intensity of the light incident on the silicon detector, I have treated the output voltage as a measure of intensity.

7.4.   Interferometry Phase Modulation

The phase modulation measurements reported here are for both single- and double-modulator in a Mach-Zehnder Interferometer (MZI) architecture, where the single/double refers to the number of capacitors in the system, as shown in Figure 7.7(a)/(b).
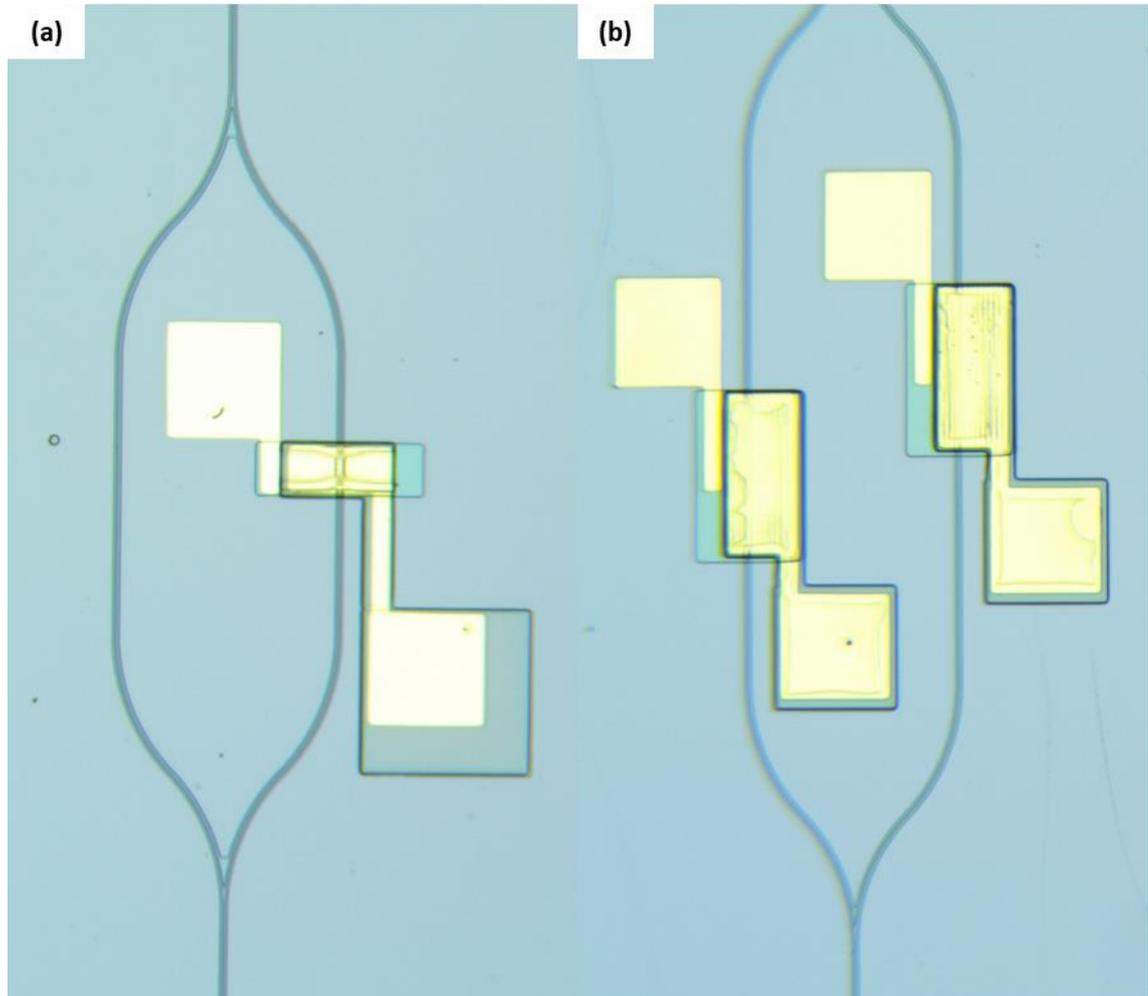
Figure 7.7: (a) Single and (b) double capacitor phase modulators.

The single modulator (Figure 7.7(a)) has the advantage of lower optical losses due to one arm not intersecting any other material and potentially scattering at the interface. However, the single capacitor has an asymmetry to the interference of the two arms in the measurement, i.e., there is an inherent phase difference between the two sides prior to the application of a gate voltage. The double capacitor (Figure 7.7(b)) has overall higher optical losses, but a symmetrical phase difference between the two arms. The phase modulation measurements shown in this section were measured using the 635nm laser, with details shown in Table 7.1.

The MZM output, is described by Equation 2.42 and is a function of the superposition of the two arms of the system. Thus, the time-dependent transfer function (Equation 2.43) describes the system as a slowly varying envelope approximation, one in which the interference of the two arms creates a sine wave, and the loss created by modulating the complex refractive index of the ITO describes a loss envelope. It should be noted that not all losses in the system are a direct cause of modulation, some may arise from scattering and intrinsic material loss.

As the material loss of ITO is not sufficient to entirely absorb the input signal, a bias is only applied to one arm at a time, and we are not operating near the epsilon-near-zero point of ITO, there will always be a remaining signal at the interferometer output. As such, it is appropriate to use a damped sine wave

$$y(x) = y_0 + Ae^{-\frac{x}{t_0}}\sin\left(\frac{x - x_c}{2w}\right)$$

Equation 7.5

fitted to the data. Here, the maximum amplitude of the sine wave is given by A, with the decay, or loss, of the system described by the exponent, and the period given by 2w. I used non-linear regression (section 4.7.1) to fit a damped sine wave (Equation 7.5) to the measured oscilloscope voltage. To calculate the figure of merit discussed in section 1.4, the half-wave voltage length ($V_\pi L$), from Equation 7.5, we take the product of the half-wave period (w) and the modulator length (l):

$$V_\pi L = wl$$

Equation 7.6

The damped sine wave is shown as a black dashed line in the optical measurement figures in the following sections. To verify the non-linear regression fits and create a parameter space in which the parameters in Equation 7.6 can vary, I checked the data against a commercial software tool. The results of this are presented in appendix section 9.6.

7.4.1.   Single-Capacitor Phase Modulator

Figure 7.8 shows the oscilloscope-measured voltage (representing light intensity) as a function of the gate voltage across the capacitor for a single-capacitor modulator of length 60μm.
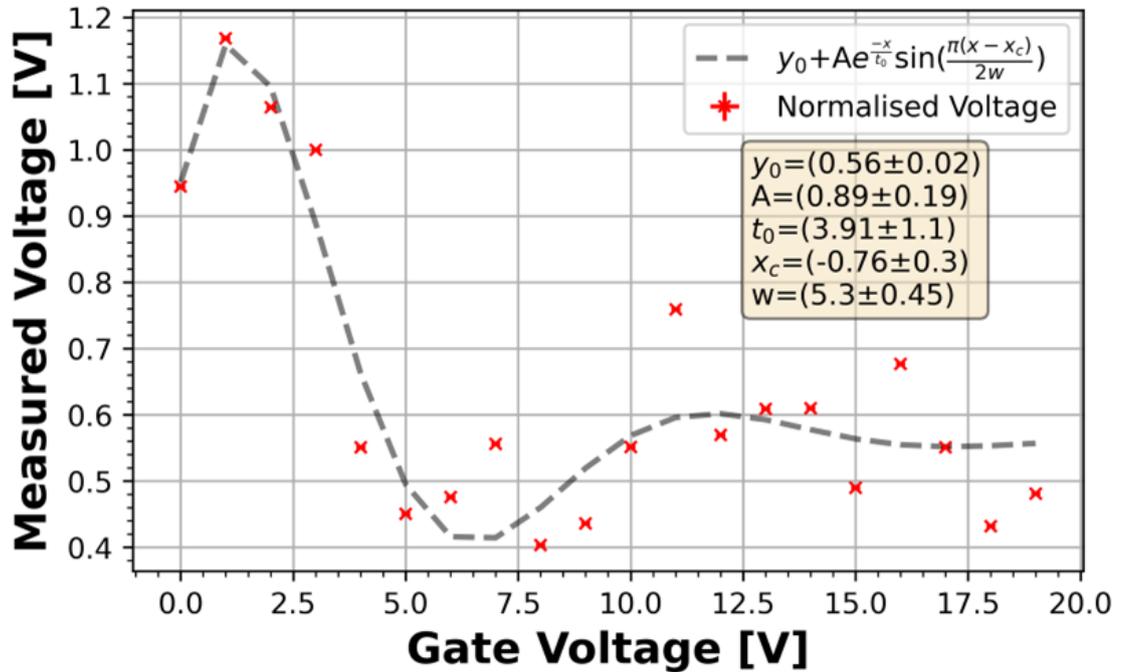
Figure 7.8: Oscilloscope measured voltage as a function of capacitor gate voltage for a 60 μm single capacitor modulator.

Using Equation 7.6, with a modulator length of (60±10) μm and a half-period of (5.3±0.5) V, which is, for the modulator shown in Figure 7.8, is calculated as (0.3±0.1) Vmm. The error is calculated using the standard quadrature technique (Appendix section 9.1.2).
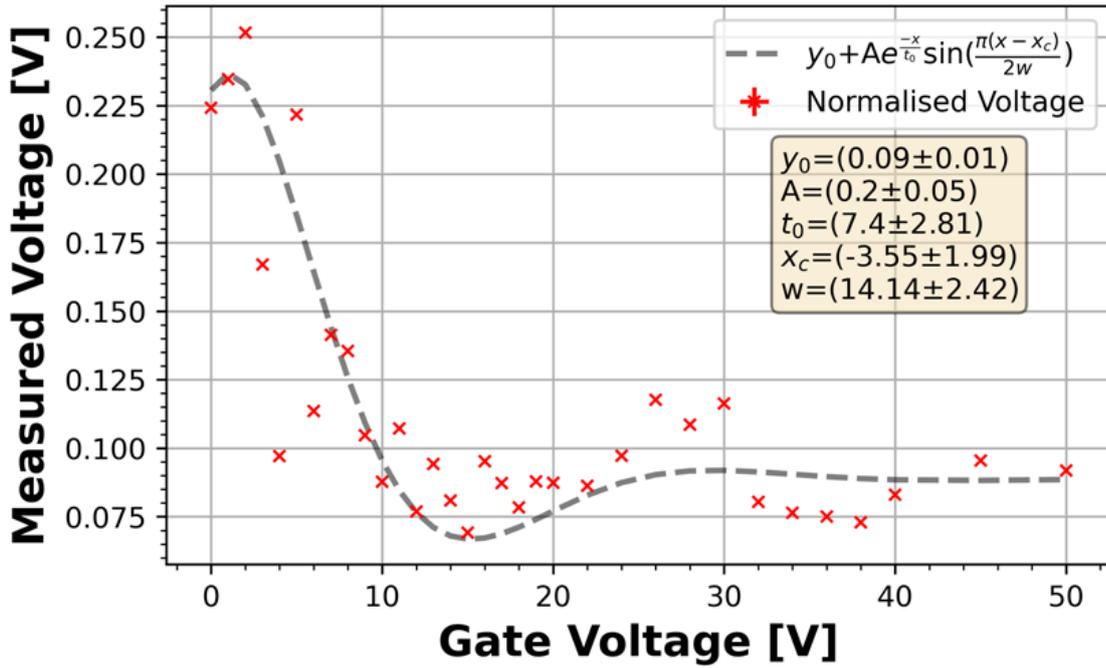
Figure 7.9: Oscilloscope measured voltage as a function of capacitor gate voltage for a 70 μm single capacitor modulator.

Figure 7.9 show the oscilloscope-measured voltage as a function of the gate voltage for a single-capacitor modulator of length 70 μm. Using Equation 7.6, with a modulator length of (70±10) μm and a half-period of (14±2) V, which is, for the modulator shown in Figure 7.9, is calculated as (1.0±0.2)Vmm.

7.4.2.   Double-Capacitor Phase Modulator

Figure 7.10 shows the oscilloscope-measured voltage as a function of the gate voltage for a double-capacitor modulator of length 100 μm. Using Equation 7.6, with a modulator length of (100±10) μm and a half-period of (13±2) V, which is, for the modulator shown in Figure 7.10, is calculated as (1.3±0.2) Vmm.
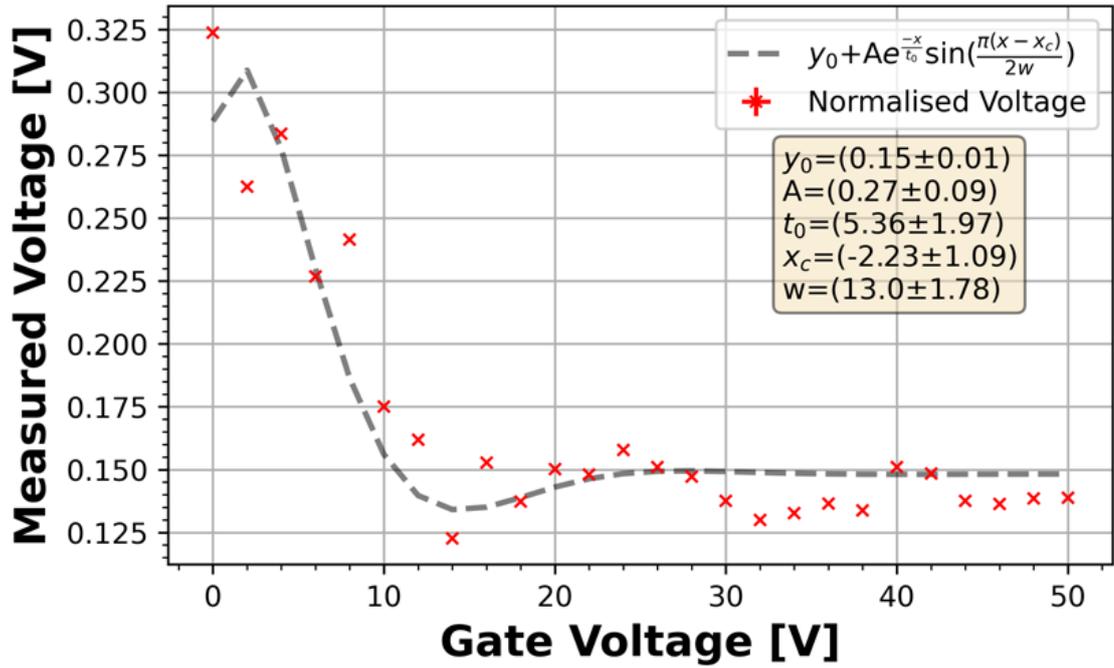
Figure 7.10: Oscilloscope measured voltage as a function of capacitor gate voltage for a 100 µm double capacitor modulator.

7.4.3.   Increased Modulation Efficiency

Phase is given as a function of the change in optical path length between the two sides of the modulator (OPD) and the wavelength (λ), given by Equation 2.41. For a phase change of 2π, the equation can be rearranged to express the change in effective index (n_{eff})

$$\Delta\Phi = 2\pi\left(\frac{\Delta OPD}{\lambda}\right) = 2\pi\left(\frac{\Delta n_{eff}L}{\lambda}\right), \qquad \Delta n_{eff} = \frac{\lambda}{L} \, for \, \Delta\Phi = 2\pi \qquad \text{\textit{Equation 7.7}}$$

as a function of the wavelength and modulator length (L). Here I have chosen 2π to reduce the equation to an effective index relationship. Using Equation 7.7 for a wavelength of 635nm, the change in effective index for a 60, 70, and 100 µm modulator is 0.01, 0.009, and 0.006, respectively. Using the mode overlap formulation in section 2.4.1, the accumulation layer of ITO accounts for 1% of the effective index of the modulator, and thus a change in the refractive index of the accumulation layer of 0.1, 0.09, and 0.06 would be required to achieve a 2π phase change. As shown in Figure 9.21, this is higher than expected and so there must be another phenomenon contributing to the phase modulation. Moreover, the $V_{\pi}L$ parameter is expected to be broadly independent of modulator length on this scale, which is clearly not the case and it a further indicator that another effect is contributing to the modulation.

7.5.   Secondary Mode Interference

As observed in the phase modulation measurements, there is clearly another effect contributing to the performance of the modulator that has not yet been considered. To investigate the cause of the increased modulator performance, we used a combination of simulated structures and single waveguide (amplitude) modulators, which I will discuss here.

7.5.1.   Dual-Mode Effect

We used Lumerical (3.2.2) to simulate a reduced-scale modulator and monitor the waveguide mode from a side profile. The simulated structure included a silicon nitride waveguide on a glass substrate, a thin ITO layer, an SU-8 layer, and a gold contact layer. The silicon nitride waveguide, with refractive index of 2.0 [84], has a length of 100 µm, height of 100 nm, and width of 3 µm. The modulator length is 1.5 µm. The ITO layer is 100 nm thick and has a refractive index of 1.9-2.2 [167]. The SU8 layer is 250 nm thick and has a refractive index of 1.6 [96]. The gold layer, although it does not interact with the mode, has a dispersive refractive index [101] which allows for the correct boundary reflections in the simulated environment, and is 100 nm thick. The wavelength of the incident light is 550 nm, with a 300 nm bandwidth (400-700 nm). The modulator and top surface of the waveguide is surrounded by air.

The side profile of the structure was monitored using a movie monitor within Lumerical, which monitored the propagation of the fundamental waveguide mode. The fundamental waveguide mode profile through the modulator was measured as a function of modulator length (distance) by merging several images of the side profile as described in appendix section 9.7. I have included the mode intensity, sine wave, and material names/colours used throughout this thesis for simplicity.
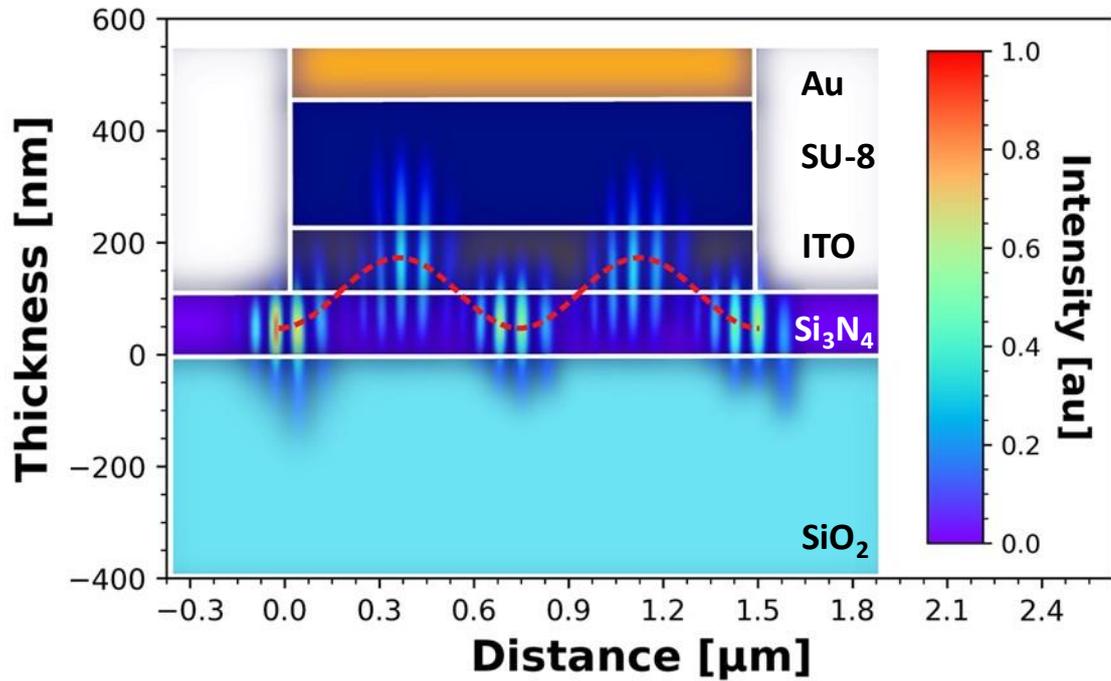
Figure 7.11: Fundamental waveguide mode in a 1.5µm gold, SU-8, ITO, silicon nitride on glass modulator side-profile. The waveguide mode profile is imposed onto an illustration of the modulator with thicknesses and lengths shown. Intensity is shown by the colour scale on the right of the figure. The ITO has a refractive index of 1.9.

The mode profiles for ITO with refractive indices 1.9 and 2.2 are shown in Figure 7.11 and Figure 7.12, respectively. WE observe, in the simulation, the excitation of a second optical mode in the ITO layer upon the mode reaching the ITO/air interface at distance 0 (x=0) in Figure 7.11, caused by the small change in refractive index between the silicon nitride and ITO layers. The result is "dual mode" behaviour, in much the same way as an MMI (section 3.6.1) which causes the superposition of the two modes to form a vertically oscillating mode propagating through the structure. This effect can be seen in Figure 7.11 and Figure 7.12 by the mode travelling into the ITO layer and has been highlighted by a sinusoidal wave drawn with a dotted red line.

While a full parameter sweep including a layer thicknesses, waveguide lengths, mode profile calculations, and multiple wavelengths have not been considered here, as these results serve to illustrate a phenomenon rather than characterise it, what we can observe is the effect a change in optical path length has on the waveguide transmission, and thus the effect it has on the modulation results.
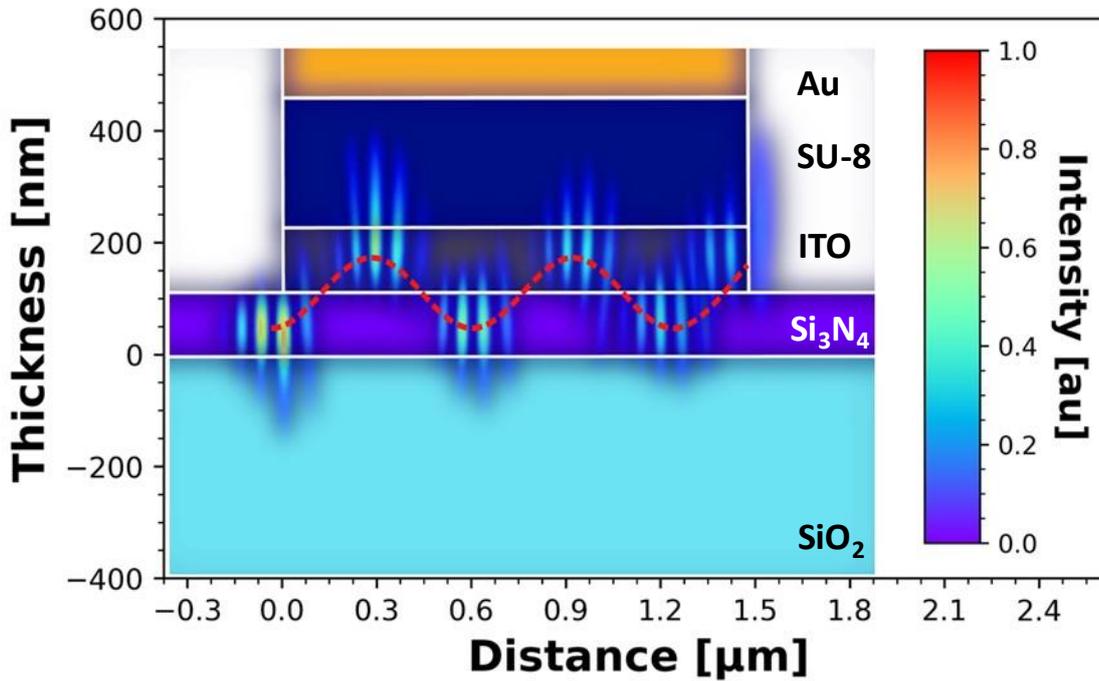
Figure 7.12: Fundamental waveguide mode in a 1.5μm gold, SU-8, ITO, silicon nitride on glass modulator side-profile. The waveguide mode profile is imposed onto an illustration of the modulator with thicknesses and lengths shown. Intensity is shown by the colour scale on the right of the figure. The ITO has a refractive index of 2.2.

A change in refractive index, of the ITO, from 1.9 to 2.2 is shown in Figure 7.11 and Figure 7.12. Using the illustrative sine wave (dotted red line), the change in transmission at the waveguide exit (far right of the x-axis, x > 1.9) is almost 100%. Most of the waveguide mode is scattered into the air for the higher index ITO, while most is coupled into the waveguide for the lower refractive index ITO. The change in refractive index, essentially optical path length of the mode, serves to increase the oscillation speed of the signal too, with the lower refractive index ITO showing an approximate period of 0.8 μm and the higher refractive index ITO showing an approximate period of 0.6 μm.

The implications of this simulation are that it may be possible to reduce the physical length of the modulator to a few micrometres, and that the minimum length is dependent on the degree of index modulation that can be achieved in the ITO layer, i.e., a more dramatic refractive index shift allows for a shorter device length. Similarly, it shows that there are more losses to consider in the phase modulation measurements than just material loss. To further investigate the dual mode behaviour of the modulator shown in Figure 7.11 and Figure 7.12, we used a transmittance monitor on the waveguide exit in Lumerical, which I will discuss in the next section.

7.5.2.   Effective Modulator Length

The waveguide transmission is affected by the optical path length of the waveguide mode, which is a product of the effective index and the modulator length. Thus, the effective optical path length experienced by the waveguide mode through the modulator section (the MOS capacitor stack) I will refer to as the effective modulator length. We used the same simulation model as described in section 7.5.1, varying the modulator length between 0.5-3 μm, the equivalent of varying the effective index of the modulator due to the optical path length formulation shown in Equation 2.41. The transmittance was measured as a percentage, and the simulation repeated for a bulk ITO index of 1.9-2.2 [167]. The effect of the modulator length on transmittance into the waveguide on the modulator exit, for the different ITO indices (shown as effective index of the waveguide mode) is shown in Figure 7.13.
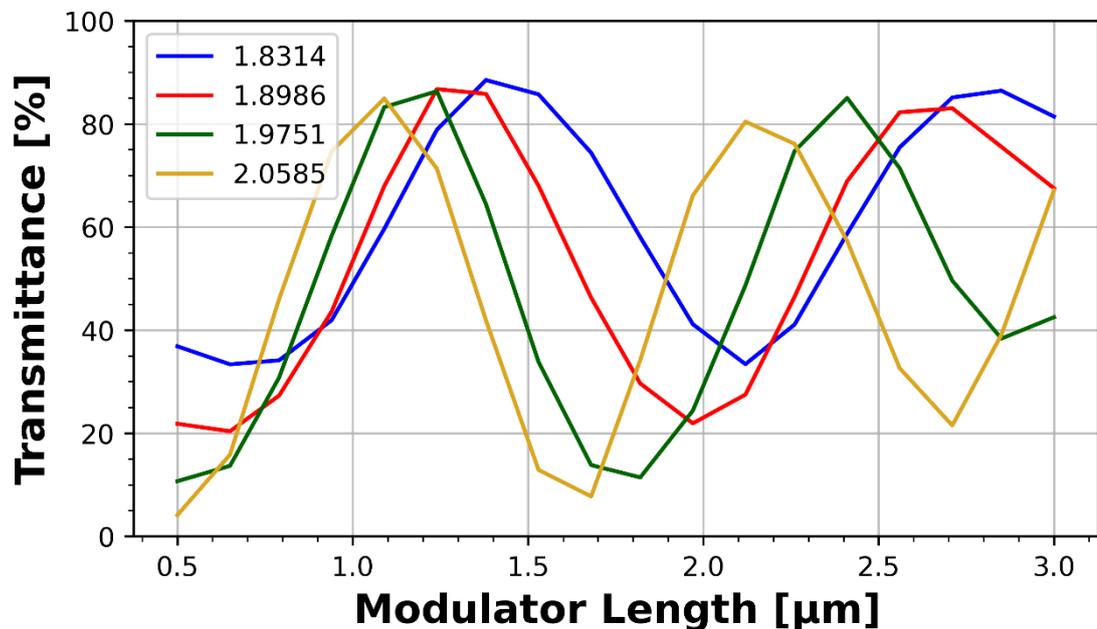
Figure 7.13: Intensity transmitted into the waveguide after passing through the MOS capacitor modulator stack as a function of effective waveguide mode index.

It is clear, from Figure 7.13, that the effective mode index has a significant effect on the transmittance of the mode into the waveguide at the exit of the capacitor. Here it is assumed that the percentage of missing light is scattered above the waveguide at the exit of the ITO layer, as seen in Figure 7.12. Using Equation 7.6, I measured the period of the transmittance, which is summarised in Table 7.3, where the errors are taken from the regression of the equation on to the data (section 4.7.1).

Table 7.3: Summary of period of oscillation of transmittance into the waveguide exit of the modulator for the different indices of ITO.

Bulk ITO Refractive Index [RIU]	Effective Index of Waveguide Mode [RIU]	Period of Oscillation of Transmittance [μm]
1.9	1.8314	1.40±0.02
2.0	1.8986	1.38±0.02
2.1	1.9751	1.34±0.02
2.2	2.0585	1.06±0.02

The range of refractive indices chosen give a quasi-linear relationship between the transmittance period and the effective index of the waveguide mode (shown in Figure 7.14), though it is not expected that this trend continues. As the effective index of the waveguide mode approaches the refractive index of ITO, the transmittance into the waveguide will approach an asymptote close to 0% i.e., most of the waveguide mode will be in the ITO layer and scatter at the end of the modulator. Similarly, as the ITO index drops significantly, most of the waveguide mode will be in the waveguide and transmittance will approach 100%.
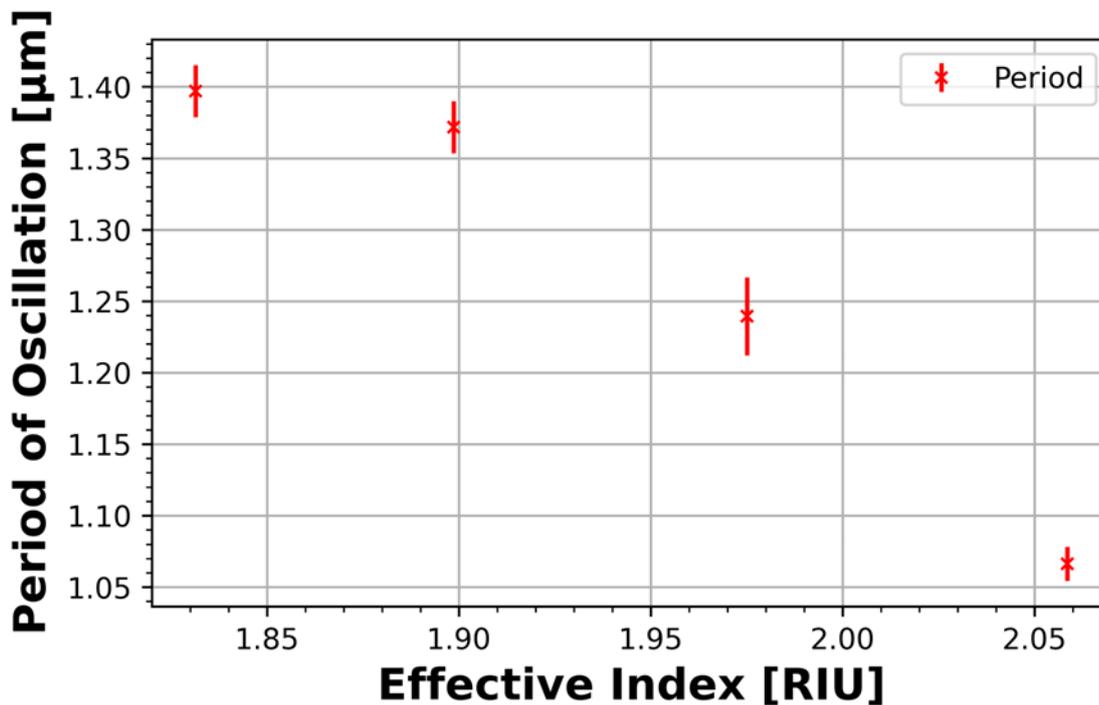
Figure 7.14: Period of transmittance oscillation as a function of waveguide mode effective index.

It is likely the relationship shown in Figure 7.14 is parabolic, as in the instance of total transmittance into the waveguide and 0 transmittance into the waveguide, the period of oscillation will be 0. Though a larger data set is required to draw that conclusion. However, the

conclusion I can draw from Figure 7.14, is that the minimum modulator length for a 2π phase shift is on the order of 1-3 µm, if a change in effective index on the order of 0.2-0.5 RIU can be achieved (calculated using Equation 7.7).

7.6.   Straight-Through Phase Modulation

Given the oscillation of the waveguide mode propagation in the modulator, and the resulting transmittance response (Figure 7.13), a MZM architecture is not required to achieve a phase and amplitude modulation. A straight-through modulator is sufficient to provide a phase modulation of the transmitted signal. The phase measurements reported in this section are for a straight-through, single-capacitor, modulator fabricated as described in section 5.6 with a 200 nm SU-8 layer in the capacitor. The phase modulation measurements are reported for 520, 635, and 785 nm sources, as described in Table 7.1. Camera intensity and oscilloscope voltage measurements are presented, with measurement techniques as discussed in sections 7.3.1 and 7.3.2, respectively.

Again, to verify the non-linear regression fits, and create a parameter space in which the parameters in Equation 7.6 can vary, I checked the data against a commercial software tool. The results of this are presented in appendix section 9.6.

7.6.1.   520nm Measurement

Figure 7.15 shows the oscilloscope-measured voltage (representing light intensity) as a function of the gate voltage across the capacitor for a single-capacitor single-waveguide modulator of length 600 µm with a 520 nm laser input.

The plot legend shows:
$$y_0 + Ae^{\frac{-x}{t_0}}\sin\left(\frac{\pi(x - x_c)}{2w}\right)$$

Normalised Voltage

$y_0 = (0.6 \pm 0.06)$
$A = (0.14 \pm 0.15)$
$t_0 = (20.0 \pm 34.5)$
$x_c = (1.45 \pm 1.48)$
$w = (10.0 \pm 4.24)$
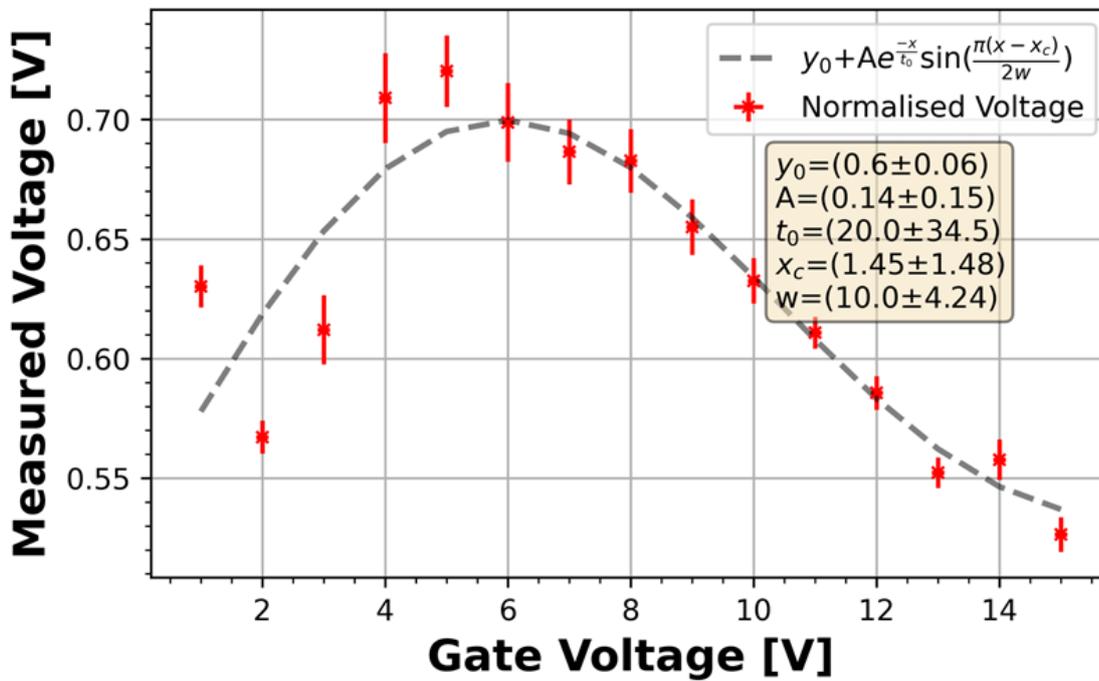
Figure 7.15: Oscilloscope measured voltage as a function of capacitor gate voltage for a 600 µm single capacitor, single waveguide, modulator for a 520 nm input.

Using Equation 7.6, with a modulator length of (600±10) µm and a half-period of (10±4) V, which is, for the modulator shown in Figure 7.15, is calculated as (6±3) Vmm. The error is calculated using the standard quadrature technique (Appendix section 9.1.2).

Figure 7.16 shows the camera-measured intensity (normalised and background corrected) as a function of the gate voltage across the capacitor for a single-capacitor single-waveguide modulator of length 600 μm with a 520 nm laser input, measured at the output of the capacitor.
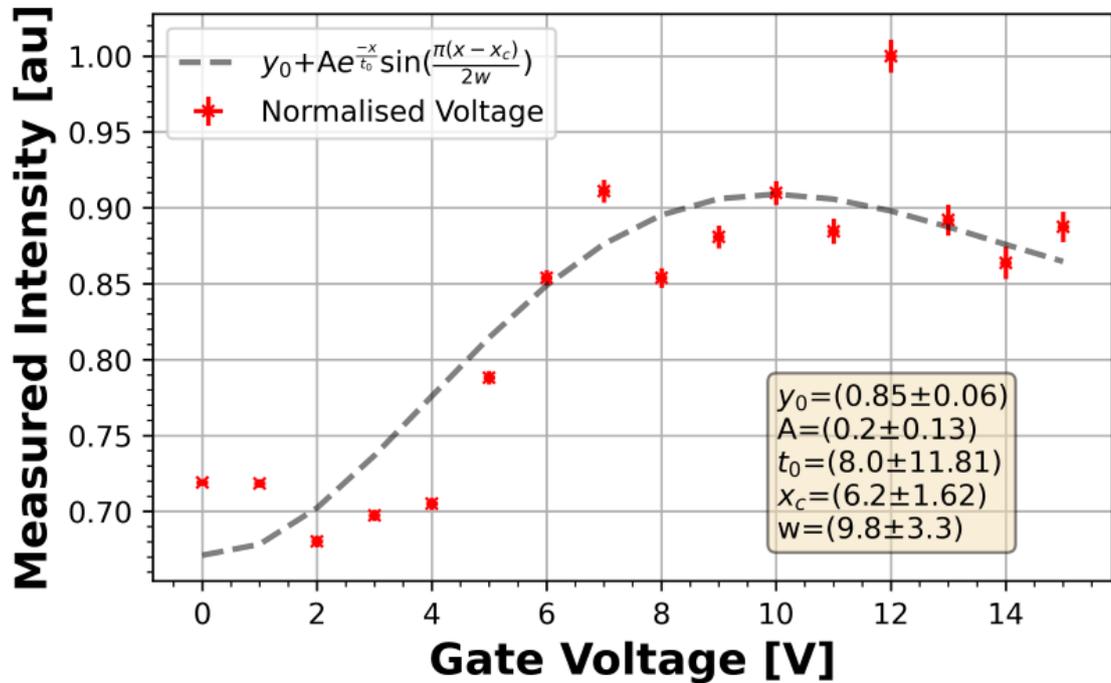
Figure 7.16: Normalised camera intensity at capacitor exit as a function of capacitor gate voltage for a 600 μm single capacitor, single waveguide, modulator for a 520 nm input.

The measured half-period in Figure 7.16 is equal, and within error, to that in Figure 7.15, verifying the phase modulation and figure of merit. The amplitude of the measured camera intensity at the modulator exit (Figure 7.16) is approximately the opposite of that measured with the oscilloscope at the waveguide exit (Figure 7.15), thus confirming the simulated results in Figure 7.11 and Figure 7.13, i.e., light is either coupled into the waveguide or scattered at the capacitor exit.

231

7.6.2.   635nm Measurement

Figure 7.17 shows the oscilloscope-measured voltage (representing light intensity) as a function of the gate voltage across the capacitor for a single-capacitor single-waveguide modulator of length 50 μm with a 635 nm laser input.
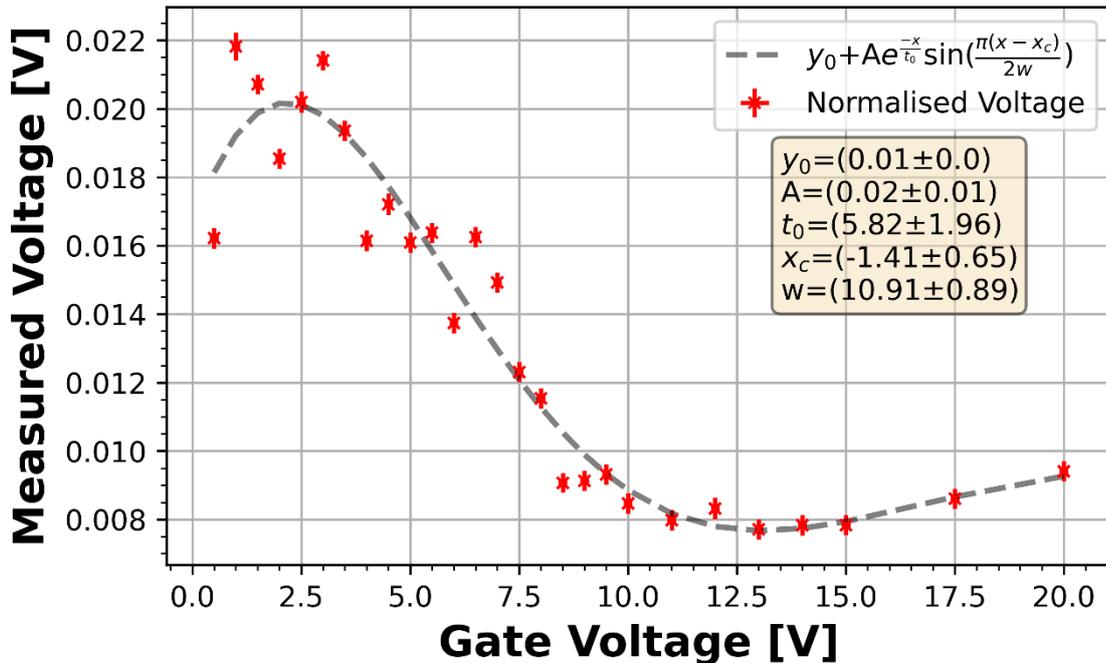
Figure 7.17: Oscilloscope measured voltage as a function of capacitor gate voltage for a 50 μm single capacitor, single waveguide, modulator for a 635 nm input.

Using Equation 7.6, with a modulator length of (50±10) μm and a half-period of (10.9±0.9) V, which is, for the modulator shown in Figure 7.17, is calculated as (0.5±0.1) Vmm. Figure 7.18 shows the camera-measured intensity (normalised and background corrected) as a function of the gate voltage across the capacitor for a single-capacitor single-waveguide modulator of length 50 μm with a 635 nm laser input, measured at the output of the capacitor.
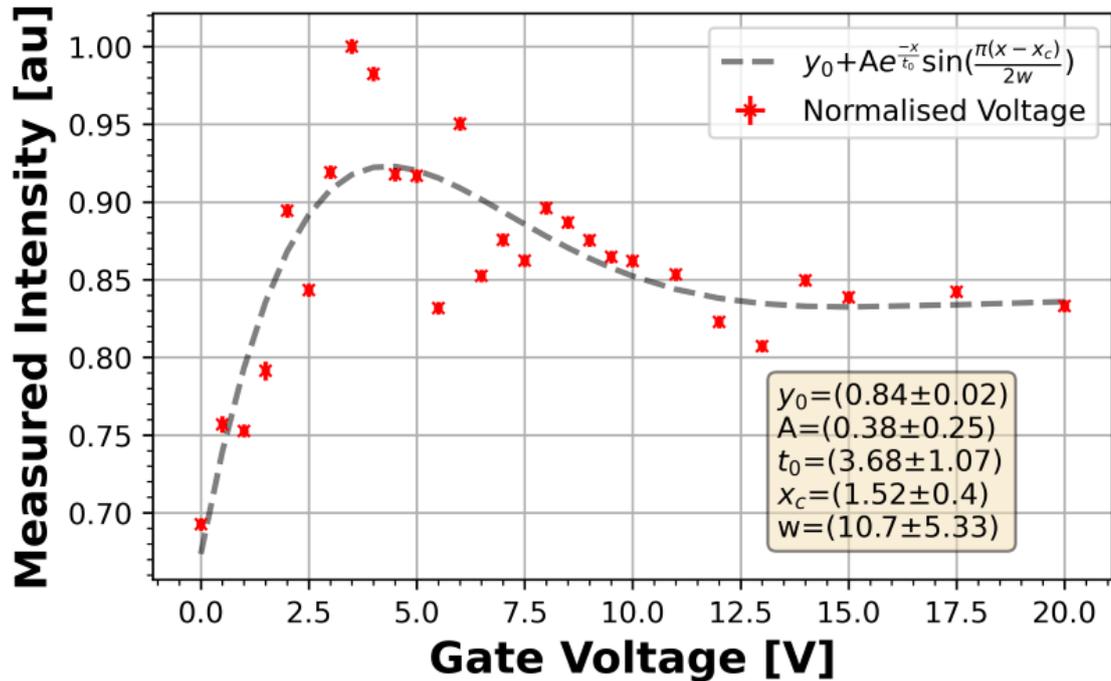
Figure 7.18: Normalised camera intensity at capacitor exit as a function of capacitor gate voltage for a 50 μm single capacitor, single waveguide, modulator for a 635 nm input.

The measured half-period in Figure 7.18 is equal, and within error, to that in Figure 7.17, verifying the phase modulation and figure of merit. The amplitude of the measured camera intensity at the modulator exit (Figure 7.18) is low in the high oscilloscope intensity region (Figure 7.17) at small voltages. As the oscilloscope voltage decreases, the camera intensity increases and maintains a high level in the high gate voltage region. The oscillation of the waveguide mode, while not as clearly demonstrated as with 520nm input, is observed at 635nm.

7.6.3.   785nm Measurement

Figure 7.19 shows the oscilloscope-measured voltage (representing light intensity) as a function of the gate voltage across the capacitor for a single-capacitor single-waveguide modulator of length 100 μm with a 785 nm laser input.
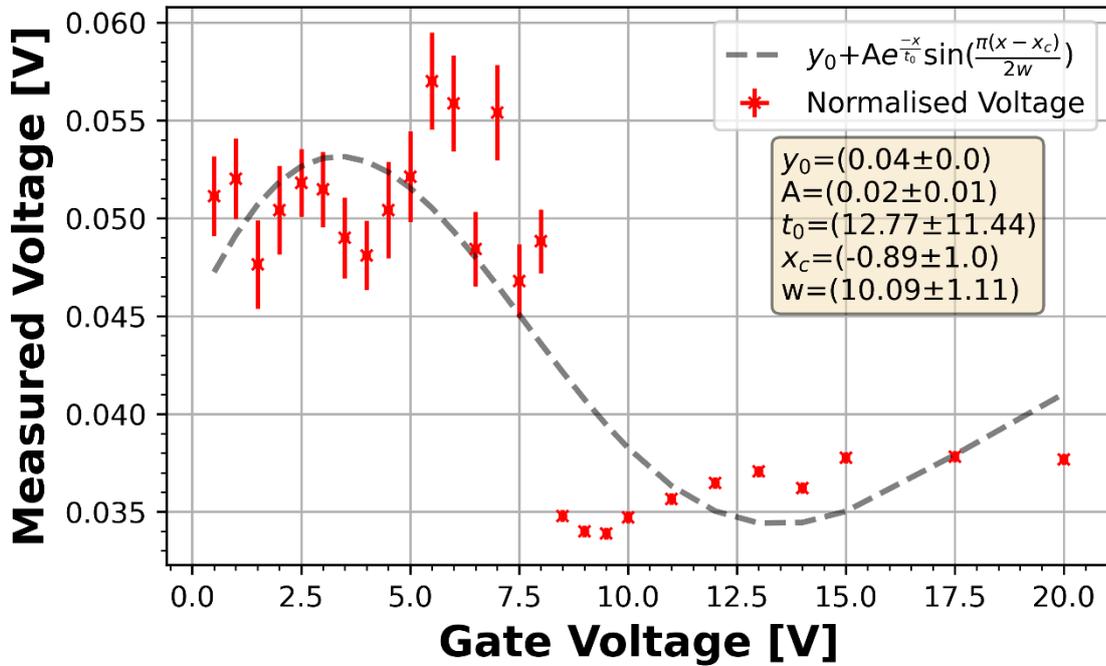
Figure 7.19: Oscilloscope measured voltage as a function of capacitor gate voltage for a 100 μm single capacitor, single waveguide, modulator for a 785 nm input.

Using Equation 7.6, with a modulator length of (100±10) μm and a half-period of (10±1) V, which is, for the modulator shown in Figure 7.19, is calculated as (1.0±0.2) Vmm.

Figure 7.20 shows the camera-measured intensity (normalised and background corrected) as a function of the gate voltage across the capacitor for a single-capacitor single-waveguide modulator of length 100 μm with a 785 nm laser input, measured at the output of the capacitor.
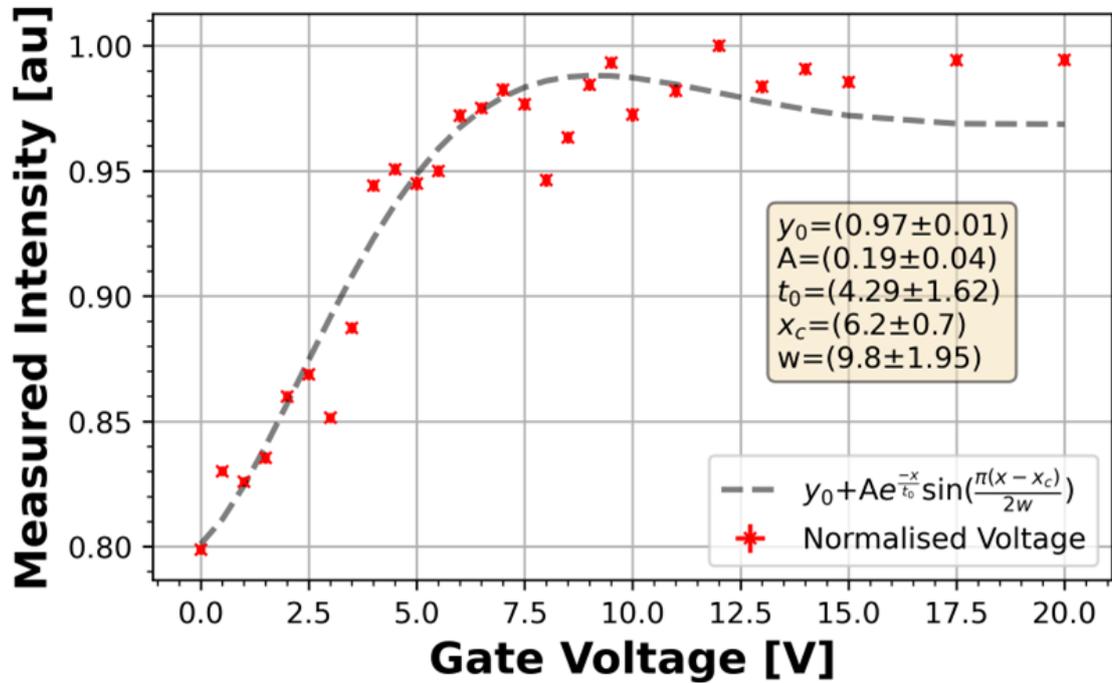
Figure 7.20: Normalised camera intensity at capacitor exit as a function of capacitor gate voltage for a 100 μm single capacitor, single waveguide, modulator for a 785 nm input.

The measured half-period in Figure 7.20 is equal, and within error, to that in Figure 7.19, verifying the intensity modulation and, as described in section 2.5, phase modulation, as well as the figure of merit. The amplitude of the measured camera intensity at the modulator exit (Figure 7.20) is approximately opposite of that measured with the oscilloscope (Figure 7.19) like the 520 nm result. Thus, confirming the oscillation of the waveguide mode as it propagates through the modulator.

7.7.   Summary

In this chapter, I have presented the optical measurement technique and resulting phase modulation results using an ITO-based MOS capacitor modulator on a silicon nitride waveguide platform for a Mach-Zehnder Interferometer (MZI) and a straight-through modulator design. The results of the MZI phase modulators are summarised in Table 7.4, with values of 0.32 Vmm, 1.0 Vmm, and 1.3 Vmm. The MZI phase modulators are competitive within the literature, with notable performance metrics of 0.52 Vmm from Amin et al. [77] and their ITO-based waveguide modulator, 0.6 Vmm from Dogru and Dagli [37] with a multiple-quantum-well system, and an order of magnitude better than Wang et al. [67] with 36 Vmm from a lithium niobate modulator.

Table 7.4: Summary of MZI architecture phase modulator performance metrics.

Modulator Length [mm]	Modulation Half-Period [V]	Half-Wave Voltage-Length (FOM $V_\pi L$) [Vmm]
(0.06±0.01)	(5.3±0.5)	(0.32±0.06)
(0.07±0.01)	(14±2)	(1.0±0.2)
(0.10±0.01)	(13±2)	(1.3±0.2)

Similarly, the straight-through architecture phase modulators, summarised in Table 7.5, show a performance of 6, 0.5, and 1 Vmm and are competitive within the literature.

Table 7.5: Summary of straight-through architecture phase modulator performance metrics.

Modulator Length [mm]	Modulation Half-Period [V]	Half-Wave Voltage-Length (FOM $V_\pi L$) [Vmm]
(0.60±0.01)	(10±4)	(6±3)
(0.05±0.01)	(10.9±0.9)	(0.5±0.1)
(0.10±0.01)	(10±1)	(1.0±0.2)

The extinction ratio (ER) is a representation of the ratio between the output optical power (P) and the input optical power (P_0) of a system and may be expressed as:

$$ER = 10\log\left(\frac{P}{P_0}\right)$$

Equation 7.8

I have calculated the ER and expressed in decibels, as is common practice, for the MZI architecture (Table 7.6) and straight-through (Table 7.7) modulators, with values ranging from 0.9-3.9 dB, within the expected literature range.

Table 7.6: Summary of MZI architecture loss metrics.

Modulator Length [mm]	Effective Extinction Coefficient (k_{eff})	Extinction Ratio [dB]
(0.06±0.01)	(0.006±0.002)	(3.9±0.4)
(0.07±0.01)	(0.011±0.004)	(2.9±0.3)
(0.10±0.01)	(0.005±0.002)	(3.7±0.4)

Using (appendix equation) I have also calculated the effective extinction coefficient (k_{eff}) using the loss profile of the enveloping exponential function in Equation 7.5(t_0) and the modulator dimensions. The extinction coefficients are within the expected range of the Drude profile shown in Figure 5.32.

Table 7.7: Summary of the straight-through architecture loss metrics.

Modulator Length [mm]	Effective Extinction Coefficient (k_{eff})	Extinction Ratio [dB]
(0.60±0.01)	(0.003±0.001)	(0.9±0.1)
(0.05±0.01)	(0.012±0.004)	(2.4±0.2)
(0.10±0.01)	(0.016±0.004)	(1.3±0.1)

In this chapter, I noted that the refractive index modulation of the accumulation layer within the ITO layer of the MOS capacitor for the MZI architecture phase modulators needed to be significantly higher than achievable with the capacitor architecture used. We conducted a simulation of the structure and determined that the waveguide mode excites a dual-mode interference within the capacitor structure, and the superposition of the two modes creates an oscillating mode between the silicon nitride and ITO layers as it propagates through the system. The modulator length governs the portion of the mode that is transmitted in the waveguide (and measured), and the portion which scatters from the edge of the capacitor. We determined that, with sufficient refractive index modulation, a minimum modulator of 1-3 μm in length could achieve a 2π phase modulation. This would significantly reduce the footprint of the modulator, reduce losses, and increase the potential array size for a larger system.

8. Conclusions and Outlook

Chapter Synopsis

This chapter presents the results and conclusions of my research project and an outline of where the research is going and could be taken in the future. Albert Einstein is quoted as saying "If we knew what we were doing, it would not be called research", and when it comes to a research project such as a PhD thesis, this can be very apt. In my introduction I outlined a motivation and overall goal of my research, and now I will discuss what conclusions can be made, what improvements could be made, how I have developed as a researcher, and look at what impact my research could have on the future of the field.

8.1.   Conclusions

The core research question of this thesis was if we could develop a new, highly efficient, carrier-based electro-optic modulator platform suitable for large-scale coherent optical processors. The thesis concludes with the answer to that question, a "yes". The silicon nitride waveguide platform, combined with an indium tin oxide (ITO) -based metal-oxide-semiconductor (MOS) modulator has been demonstrated to perform to a competitive level in the field. There are, however, some caveats and further research questions that were discovered and answered along the way, namely how the modulator gains that performance considering the wavelength range over which it has been demonstrated and how the MOS structure described in this work has aided in the increased performance.

An electro-optic modulators performance is characterised by the product of the half-wave voltage (V_π) and the length of the modulator (L). In other words, the performance is a balance between the efficiency of the modulation (i.e., the performance of the modulating material and mechanism) and the footprint of the device. Ideally the voltage and length would both be a minimum, but as I will discuss here, it is not that simple.

8.1.1.   ITO Conclusions

The performance of ITO as a modulating material is governed by the electro-optic properties of the material, which we have shown to be highly dependent on the deposition parameters, and this too is a common theme in the literature. The permittivity of ITO is described by the Drude-Lorentz model, which describes the optical properties of a material in terms of the electrical properties of the material. Hence the dependence on the deposition conditions for ITO, as the electrical properties are determined by the oxygen and tin content of the material.

To determine the electrical properties of the sputter-deposited ITO used in this research, we found many inconsistencies in the literature. Many authors have published literature without accounting for the deposition conditions for which the optical/electrical properties were obtained. Given the limitations of spectroscopic ellipsometry/reflectometry available to me, we circumvented the problem by using a resonant structure in the material to determine the optical properties and techniques like the Hall probe and four-probe methods to determine the electrical properties. The investigation was a success, and we were able to publish the results

and I was then able to apply the same technique to the ITO used in modulators discussed in this research.

The key parameters of the Drude-Lorentz model, the high frequency permittivity (ε_∞) , the collision frequency (Γ), the effective electron mass (m*), and the carrier density (n), were measured to be 4.95, 13 THz, 0.45, and 6.1×10^{20} cm^{-3}, respectively. These were calculated from a measured sheet resistance, conductivity, and carrier mobility of 110 Ω/sq, 460 S/cm, and 4.67 cm^2/Vs, respectively, where the associated errors and discussions may be found in section 7.7.

With the sensitivity of the deposition conditions so high, literature values for the sputter-deposited ITO in this research are non-existent. However, the values obtained are well within common literature values for such a high carrier density. Thus, I can determine that the characterisation technique that combined resonant structures and electrical measurements is an accurate and acceptable alternative to common spectroscopic ellipsometry techniques that require a deep understanding of the technique and a high level of skill to obtain the information.

8.1.2.   MOS Capacitor Conclusions

The performance of a carrier-based modulator is naturally determined by the quantity of carriers that can be accumulated in the Drude-material (ITO). The operation of a MOS capacitor is quite simple; two contacts are spaced by a non-conducting material, and under an applied electric field charges (electron-hole pairs) accumulate on the two surfaces at the insulator interfaces (metal-insulator and semiconductor-insulator). While, for ITO, this is somewhat dependent on the bulk carrier concentration in the semiconductor, it is also largely dependent on the properties of the insulating material, namely thickness and dielectric constant. Using the standard capacitor equation, the capacitance of the gold-SU-8-ITO capacitor used in this research was predicted to be on the order of 1 pF.

However, under measurement it was found to be 3.3 pF, 3 times the expected capacitance. Thus, given that the capacitance is directly proportional to the charge that can be accumulated at the insulator interface with the semiconductor, yielding a 3-fold increase in optical modulation. It is detailed in the literature that, on occasion, the metal and semiconductor materials extend into the insulator layer, thus decreasing the plate separation distance and increasing the capacitance. There is, however, little evidence to suggest that this is the case.

Using energy dispersive x-ray spectroscopy (EDS) to examine the cross-section of the gold-SU-8-ITO capacitor, I determined that there was gold, indium, and tin elemental presence within the SU-8 layer of the capacitor, thus agreeing with the literature. The depth of the penetration of these materials into the SU-8 layer has not been quantified, but calculations indicate a 3-fold reduction in the distance between the metal and semiconductor materials. The diffusion of gold and ITO into the SU-8 layer is likely due to the temperatures associated with the deposition of these materials and that, perhaps naively, we assumed the SU-8 layer was pinhole-free.

However, given that the ideal situation for an electro-optic modulator is for the gold (highly absorptive) layer to be far away from the optical mode whilst maintaining a high capacitance, the diffusion of the metal and semiconductor materials into the insulator layer may have benefitted the overall performance of the modulator. The indication from the EDS measurements was not that a bulk gold layer had diffused into the SU-8 but suggested that it may be more like hotspots of the material through the length of the modulator. This would imply a decrease in electrical thickness of the SU-8, but not physical thickness, and therefore an increase in electrical performance without the deficit of losing more optical power due to closer proximity of the gold to the optical mode.

8.1.3.   Dual-Mode Conclusions

Finally, the performance of the modulators. Two modulator architectures were investigated during my research: a Mach-Zehnder Interferometer (MZI) architecture, in which the interference of a modulated and unmodulated arm induce a phase change, and a straight-through architecture in which the amplitude of the input signal is the primary modulation effect. Typically, the figure of merit ($V_\pi L$) is given in Vmm, with state-of-the-art modulators quoting 0.5-0.6 Vmm using a similar architecture. Measuring the superposition of the modulated and unmodulated arms with a fibre-coupled oscilloscope on the exit facet of the waveguide, I measured a $V_\pi L$ for three modulators of 0.32, 1.0, and 1.3 Vmm.

The phase measurements should have produced a figure of merit (FOM) approximately of the same order regardless of length, since the length of the modulator is proportional to the required gate voltage, i.e., modulators of larger length require larger voltage to achieve a π-phase change, for the same ITO and SU-8 properties. A simulation of the modulator, examining the side-profile and the propagation of the waveguide mode through the modulator, revealed the excitation of a dual-mode. The cause of the dual-mode is a matching of refractive indices

between the waveguide, ITO, and SU-8 layers creating a substantial optical thickness in which the waveguide mode can propagate, with secondary modes excited when the thickness of the guiding material suddenly increases at the input of the modulator.

The simulated structure, despite having an ITO index closer to the silicon nitride index than the experimental structure, showed that, with a strong enough effective index change, a modulator on the order of 1-3 μm is sufficient in the visible spectrum. A modulator of this length would require an effective index change of 0.18-0.5 RIU, which is beyond the capabilities of the ITO presented in this work, but with further optimisation it could be possible to reduce the modulator length down to ~10 μm.

The dual-mode behaviour of the waveguide means that the straight-through architecture produces intensity modulation of similar magnitude to the MZI architecture. Using the same measurement technique, I was able to measure a $V_\pi L$ of 6, 0.5, and 1 Vmm, for 520, 635, and 785 nm input sources, thus demonstrating the capabilities of the modulator across the visible spectrum.

In conclusion, I have demonstrated that an ITO-based electro-optic modulator on a silicon nitride platform, which allows for application in the visible spectrum and with higher input powers than the commonly used silicon platform, is competitive with state-of-the-art electro-optic modulators in the literature. Further improvements to the structure are required to tune the dual-mode transmittance such that a smaller (<50 μm), as well as determining the elemental composition of the capacitor to further understand the electrical capabilities of the modulator, I will discuss these in the next section.

8.2. Developments and Improvements

Here, I would like to reflect on my research, my development, and offer suggestions to how I could improve on the project if I were to have the knowledge I do now that I am at the end of my research.

8.2.1.   Developments

It goes without saying that a large portion of the early part of my project was impacted by the Covid-19 pandemic, a time of familiarising oneself with the project, the fabrication, and perhaps designing an experimental setup, lost. During this time, I focussed by efforts on simulation techniques and environments, as well as familiarising myself with the literature and wider field of study. This is evident from my introduction and variety of simulation techniques presented throughout this thesis. Fortunately, I was comfortable with nanofabrication to quickly develop and optimise the extensive fabrication process required to achieve the multi-layer modulator and focus most of my efforts in characterisation of the ITO and capacitor.

A large amount of time could have been saved if the realisation of using a silicon substrate had been made earlier in my research. I used many different tricks to circumvent the rough facet created by cleaving a borosilicate substrate, including partial dicing and SU-8 tapered waveguides to increase the facet size with the intention of minimising the effect of surface roughness. The switch to a silicon substrate with an oxide layer to prevent coupling into the silicon drastically improved the speed and reliability of the fabrication process and allowed optical experiments to become the focus.

While I have not discussed spectroscopic ellipsometry in detail in this thesis, for a long time it was the primary technique I used to determine the optical properties of the sputter-deposited ITO. It became clear that, to accurately determine the optical properties of the material, a more sophisticated analytical system was required, allowing multiple models and angles to be examined simultaneously, which meant outsourcing the process. The resonant structure analysis in ITO, and the collaboration with a colleague, helped in obtaining a detailed understanding of the material in-house, thus reducing the analysis time.

The optical setup is highly sensitive to external noise, despite being insulated from minor vibrations by the optical bench on which it sits. I found that many of the micro positioners that hold the objectives still during measurement had begun to drift out of alignment during an experiment. Determining which objectives were creating the problem and replacing them and realigning the setup certainly delayed the process of obtaining experimental data, and in hindsight this is a process that should have started much earlier than it did.

Overall, the biggest development on a personal level, throughout my research, has been to manage a project of this size. The simulation and fabrication techniques were largely familiar or easily translated from other environments or protocols in the past. The knowledge of how to

deal with a problem and find the answer or solution is certainly the one of the highlights of my personal development during this time. Now that I have discussed my own personal developments and pitfalls of the project, I would like to suggest how the results could be improved if I had started the project with the knowledge I have just described.

8.2.2.   Improvements

The main improvement to the experimental data that I can suggest is simply more of it. The optical experiments took a significant amount of time to obtain, and the delays caused by the setup issues described above, meant that repeated wavelengths and modulator lengths, as well as more wavelengths in the visible spectrum, were simply not possible to obtain before the end of my research. We had planned, initially, to measure the FOM as a function of input wavelength, and this would certainly have required a more predictable FOM, and more wavelengths measured.

Similarly, with the capacitor characterisation, a more detailed analysis of the cross-section to examine the elemental composition, and measurements of the elemental composition for capacitors treated with different temperatures during fabrication, would have allowed for further discussions on the composition and how to reduce elemental leaking into the insulator layer. A "cut-back" analysis of the cross section, in which sections of the capacitor are cleaved away and examined to determine the lateral elemental composition of the capacitor would indicate whether it is indeed "fingers" or "hotspots" of metal/semiconductor material that extends into the insulator layer. An investigation into different insulating materials in the capacitor, and the optical/electrical benefits or deficits thereof would allow for a full characterisation of the capacitor and a potential publication in the field, considering the interest in such a structure and the lack of information surrounding the phenomenon.

Finally, I would like to have achieved smaller modulators, on the order of 10µm or shorter in length. The simulated results showing that perhaps 1-3µm is sufficient presented an interesting experiment, but one that would have required a complete redesign of the fabrication process too late in my research. A modulator of this scale would have required a switch to electron-beam lithography techniques for all layers, which I did not have a starting protocol for from which I could have found an optimum process.

8.3.   Outlook

The last thing I would like to discuss, is what I think ITO-based modulators, and specifically the modulators discussed in this thesis, could achieve in the field of electro-optic modulators, and the wider field of photonics-based neuromorphic computing.

Planar modulators, such as the waveguide modulators presented in this thesis, will always occupy more space than out-of-plane modulators. It follows, therefore, that the next generation of ITO-based electro-optic modulators would be out-of-plane modulators that utilise the local permittivity modulation achievable in electrically-gated ITO. My concept design of the planar modulator is shown in Figure 8.1(a), with an exaggerated waveguide input and output for visual effect only. Shown in Figure 8.1(b), is a similar concept design for an out-of-plane ITO-based modulator that uses a resonant structure to reflect an input signal out of the top of the ITO layer, in which the resonant structure is fabricated, thus allowing for a smaller overall system, as shown in Figure 8.1(c). Considering the 4f-coherent optical processor, a planar system would require significant free-space optics at the input and output of the waveguide, but an out-of-plane solution can operate with flat optics in reflection, thus minimising the overall footprint.
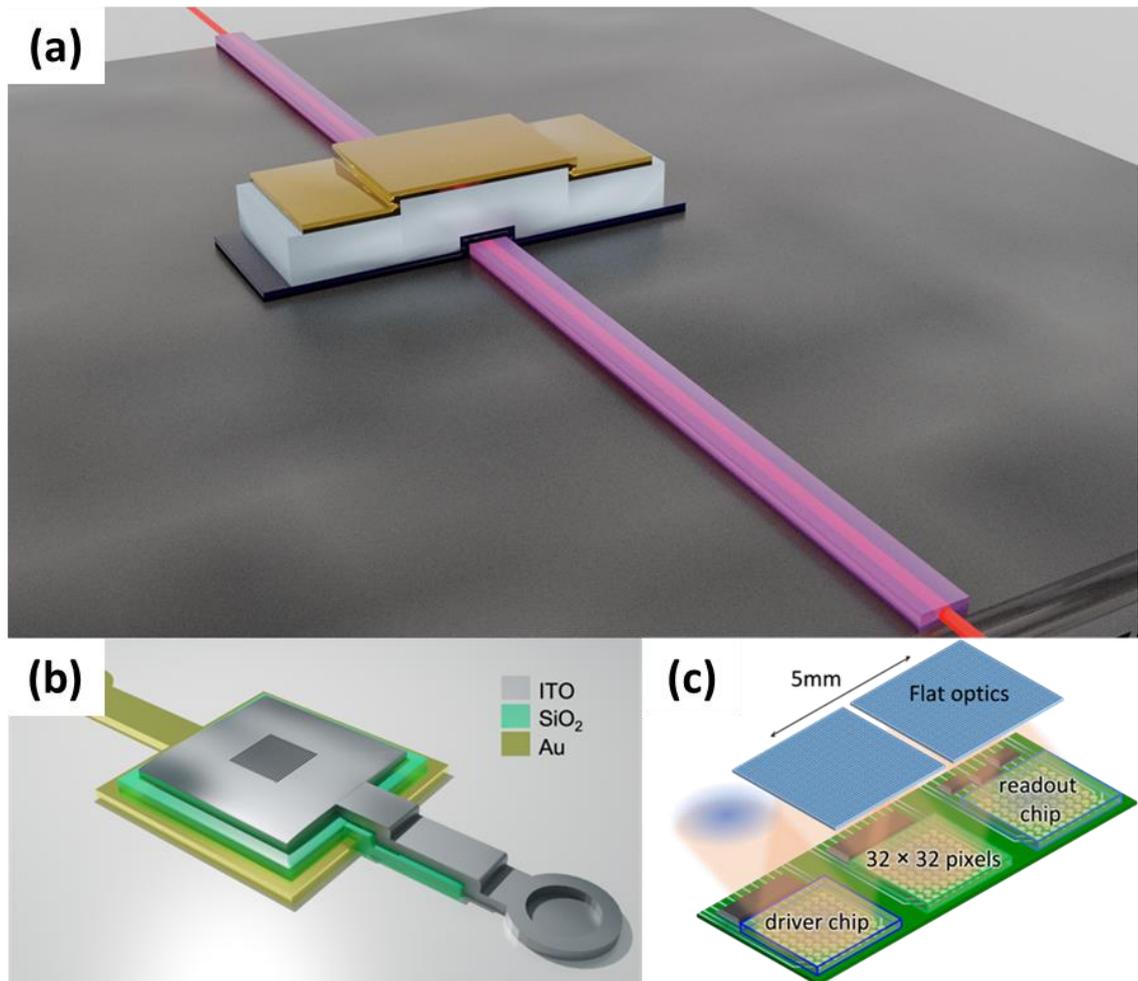
Figure 8.1: Blender designed renders of (a) a planar (waveguide) modulator and (b) an out-of-plane (resonant) modulator. (c) A concept of an out-of-plane 4f coherent optical processor. Credit: (b) & (c) S. Blair 2023.

Out-of-plane modulators are something I have spent a small amount of time working on throughout my research. Specifically, using distributed Bragg reflector (DBR), a reflector formed from multiple layers of alternating materials with different refractive indices. The repeated distribution of a refractive index pattern creates partial reflections at the interface, resulting in a specific reflectivity spectrum with a centre wavelength range known as the flatband. The introduction of a cavity or defect in the repeating material structure determines the wavelength of the flatband, and electrically gating the cavity (thus altering the refractive index) can tune the reflectivity in this region as shown in Figure 8.2.
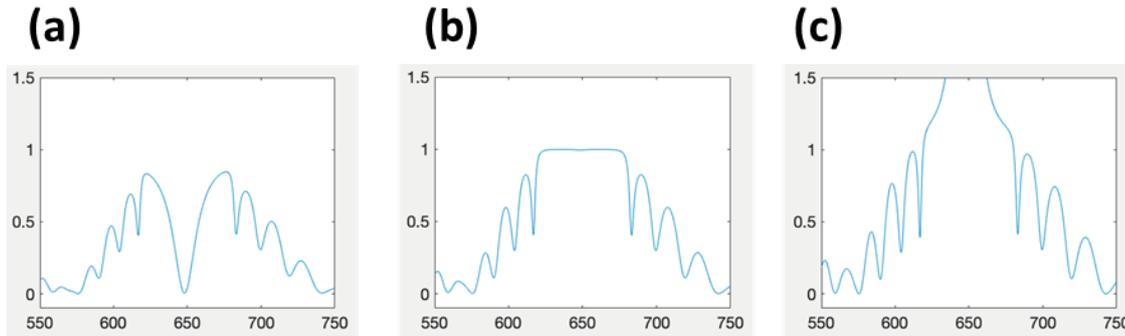
Figure 8.2: Reflectivity spectrum for a distributed Bragg reflector showing a (a) dip in reflection, (b) a flatband response, and (c) a peak in reflection.

The reflectivity of the central region, observed in Figure 8.2 as a dip in (a), a flat region in (b), and a peak in (c), serves as the optical signal, like the intensity of the output signal of the waveguide modulators discussed in this thesis. The DBR system would provide a minor improvement to speed over a waveguide modulator but is predicted to have slower operating speeds than the resonant system shown in Figure 8.2(b).

In the wider field of neuromorphic computing, electro-optic modulators and flat optics present the next generation of Fourier-space computation. The most taxing aspect of computing can, and I think will, be replaced by optical processing systems. It is proposed that a Fourier engine, combining several optical modulators for different processing techniques, could be implements in current computing systems to improve processing speed and efficiency.

9. Appendix

Chapter Synopsis

This chapter presets the aspects of the research project that could not be discussed in the main body of a thesis whilst having the ideas flow in a nice order. For those extra bits, I include an appendix. Here I will discuss all the extra parts of my research that didn't make it into the main body of text. In this chapter there are discussions of the error formulation used throughout the results presented in this thesis, there is a discussion into the degenerate semiconductor analysis, as well as references to the ellipsometry and reflectometry techniques used. I have also used this appendix to include the step-analysis of final measurements presented in the main body text.

9.1.   Error Formulation

Quantifying experimental errors is an important aspect of any research project. There are two concepts used throughout this research that I will discuss here. For experimental results, it is often better to take multiple measurements and calculate an average value to reduce minor experimental errors that occur either through variation in the measurement equipment or human error. It is also often the case that experimentally measured values are used in calculations, perhaps in a complex function with multiple variables. In the case of a multivariable calculation, the partial differential of the function can be used to quantify errors.

9.1.1.   Standard Error on the Mean Formulation

A common method to reduce variations in measurements, and help improve precision, is to take repeat measurements and calculate an average result (Equation 9.1). Taking average results allows the use of a standard error on the mean (SEOM) technique to calculate an error value. The SEOM is calculated from the standard deviation in the measurement (STDV) (Equation 9.2) and the number of measurements taken (N) (Equation 9.3) [310].

$$\bar{x} = \frac{\sum x_i}{N}$$

Equation 9.1

$$STDV = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N - 1}}$$

Equation 9.2

$$SEOM = \frac{STDV}{\sqrt{N}}$$

Equation 9.3

I used the average and SEOM techniques in most of this work to reduce experimental errors.

9.1.2.   Quadrature

The quadrature formulation is another technique for quantifying errors in calculated values [215]. The formulation may be used to quantify propagated errors in calculation. Consider a parameter Z given as a function of variables A, B, …, Y, shown in Equation 9.4, where the variables A, B, …, Y have associated errors ΔA, ΔB, …, ΔY.

$$Z = f(A, B, \ldots, Y), \qquad \frac{\Delta Z}{Z} = \sqrt{\left(\frac{\Delta A^2}{A}\right) + \cdots + \left(\frac{\Delta Y^2}{Y}\right)}$$

The error in Z is given by the partial differentiation of the function, such that the error calculation is given as the square root of the sum of the square of the fractional variable errors, as shown in Equation 9.4.

9.2. Degenerate Semiconductor Analysis

The space-charge region of the metal-oxide-semiconductor (MOS) diode is a key aspect of my research. Charge accumulation in the space-charge region of indium tin oxide (ITO) is the modulation mechanism used in the amplitude and phase modulators in this work. It follows that a detailed understanding of the charge accumulation is required to predict and quantify the modulation efficiency. In chapter 2 I discussed the formulation for calculating the charge accumulation from the capacitor dimensions and materials. To verify this simplistic formulation, we investigated the MOS theory for a strongly degenerate semiconductor with E. Avrutin in preparation [311], that I will discuss here.

9.2.1. General MOS Considerations

As with any MOS analysis, there are three elements to consider: the metal, an ideal insulator in which we can neglect charge accumulation and tunnelling, and a semiconductor. Using numerical solutions of the steady-state drift-diffusion problem, with numerical values taken from Lesina et. al [312], the potential across the oxide and ITO layer may be calculated as shown in Figure 9.1.
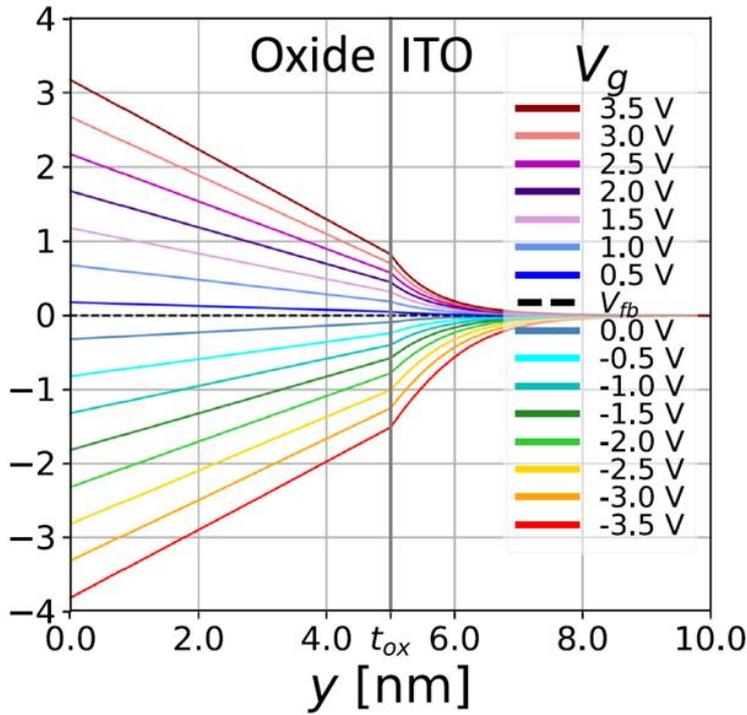
Figure 9.1: The potential profile across the oxide and ITO layers from numerical simulations of [312].

The potential in the ITO layer is described by the one-dimensional Poisson equation, where the potential (φ)

$$\frac{d^2\varphi}{dy^2} = -\frac{\rho(y)}{\varepsilon_{ITO}}$$

Equation 9.5

is given as a function of the space charge ($\rho(y)$) and the dielectric permittivity of ITO (ε_{ITO}), as shown in Equation 9.5. Applying appropriate boundary conditions and taking the oxide-ITO interface at y=0, the built-in flatband voltage (V_{FB})

$$V_{FB} = \Phi_m - \Phi_s = \Phi_m - \chi_s + E_F$$

Equation 9.6

of the metal-semiconductor junction may be estimated from the work functions of the metal and semiconductor (Φ_m and Φ_s, respectively) in the MOS structure, as given by Equation 9.6. Here the work function of the semiconductor may be calculated from the electron affinity (χ_s) and the Fermi level (E_F) of the semiconductor. Here we have assumed that there is no charge accumulation in the semiconductor [135].

9.2.2. Classical Analysis

Using the nondegenerate semiconductor analysis in [135] with the consideration of charge accumulation (positive bias), we consider that local equilibrium will result in a local carrier density in the semiconductor, such that $n(y) = n(\varphi(y))$, and the Poisson equation becomes:

$$\frac{d^2\varphi}{dy^2} = \frac{e}{\varepsilon_{ITO}}(n(\varphi) - N_D^+)$$

<div align="right">*Equation 9.7*</div>

Where $n(\varphi)$ and N_D^+ are the local electron density and the density of electrons in the bulk ITO material. Using the differential formulation for the Poisson equation, we find that the local potential may be expressed as:

$$\frac{d\varphi}{dy} = -\sqrt{2\left(\frac{eN_D^+ k_B T}{\varepsilon_{ITO}}\right)} F\left(\frac{e\varphi}{k_B T}\right)$$

<div align="right">*Equation 9.8*</div>

Where the electron charge (e), Boltzmann constant (k_B), and temperature (T) are given their usual notations. The dimensionless positive function (F(y)) is defined as:

$$F^2(y) = \int_0^y \left(\frac{1}{N_D^+} n\left(\frac{k_B T}{e} y'\right) - 1\right) dy'$$

<div align="right">*Equation 9.9*</div>

At the insulator-ITO interface, the electric field strength E_s is found as a function of the local potential (φ_s):

$$E_s = \sqrt{2\left(\frac{eN_D^+ k_B T}{\varepsilon_{ITO}}\right)} F(y)$$

<div align="right">*Equation 9.10*</div>

Which may be used to estimate the characteristic scale of potential variation in ITO at a given voltage, which is essentially the effective space-charge region width:

$$l_D(\varphi_s) = \frac{\varphi_s}{E_s} = \frac{y_s}{\sqrt{2}F(y)} l_{D0}, \qquad l_{D0} = \sqrt{\frac{\varepsilon_{ITO} k_B T}{e^2 N_D^+}}$$

<div align="right">*Equation 9.11*</div>

Where l_{D0} is the Debye length as normally defined for nondegenerate semiconductors. Equation 9.11 may be used to determine the potential in the ITO layer as a function of the gate voltage (V_g) using the solution as a boundary condition for Equation 9.8. The dependence of the space-charge accumulation on the gate voltage is determined from the condition of thermodynamic equilibrium, whereby the absolute position of the Fermi level is the same for all values of y whereas the conduction band edge at each point in the semiconductor is located at $-\varphi(y)$ at each point y (downward bending). This gives the dependence $n(\varphi)$ in the form of an exponential for degenerate semiconductors. The derivation in [135] gives an approximation for l_D as:

Equation 9.12

$$l_D \approx \frac{\pi^{\frac{2}{3}}\hbar}{e}\sqrt{\frac{\varepsilon_{ITO}}{m_e^*}}(3N_D^+)^{-\frac{1}{6}}$$

Where m_e^* is the effective electron mass. Using literature values [312] for the effective electron mass, and the charge density in the bulk ITO, a space layer width of $l_D \approx 0.74nm$ is obtained, which is in reasonable agreement with the literature assumption of 1nm. The dependence on the space-charge region width with gate voltage is shown as the blue dotted line in Figure 9.2. The approximation in Equation 9.12 assumes that the value of the carrier density in the material (N_D^+) only has a modest effect on the width of the space-charge region. Rather than using the potential and its derivative at the interface, we can use an alternative method to define the effective space-charge region width. Instead, we find the integral of the excess carrier density and divide it by the excess carrier density at the interface:

$$l_a(\varphi_s) = \frac{\int_0^\infty (n(\varphi(y)) - N_D^+)dy}{n(\varphi_s) - N_D^+}$$

Equation 9.13

Again, using the literature values in [312], we can use Equation 9.13 to predict the space-charge region width as a function of the gate voltage, shown as the blue solid line in Figure 9.2.
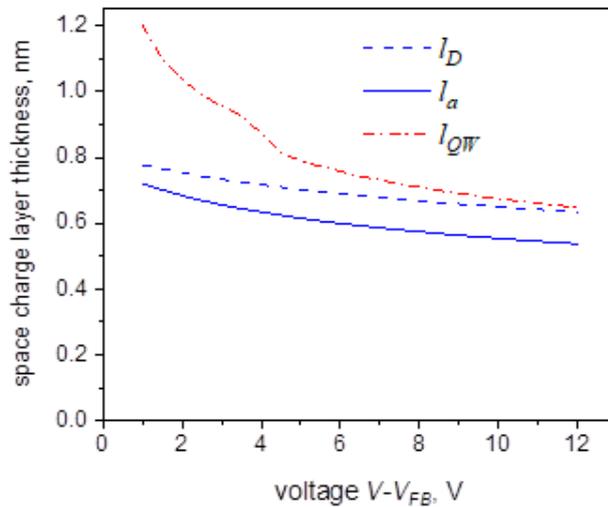
Figure 9.2: Effective space-charge region width, estimated using different analytical methods [311].

The result for l_a, shown as the solid blue line in Figure 9.2, is very close to l_D, with the lower values a result of the linear dependence of $n(\varphi)$. The two measures would be identical if both the carrier density and potential decayed exponentially and at the same rate with distance into the ITO. The similarity between the two means the dependences are close to exponential, as confirmed in the numerical simulations in Figure 9.1. The results in Figure 9.2 show that we can approximate the space-charge region width to 1nm as detailed in chapter 2.

Further considerations were made towards electron band nonparabolicity and considering the quantum behaviour of the electrons, shown as the red dotted line in Figure 9.2, though this yielded similar space-charge region widths and would require further computational analysis to confirm, which we decided was beyond the scope of this thesis.

9.3.   Capacitor Design

In chapter 3 I performed a mode overlap investigation looking at how the layer thicknesses and waveguide dimensions affect the waveguide mode overlap with each layer. In determining the capacitor layout around the waveguide, i.e., fully surrounded, partially surrounded, or open on both side of the waveguide, I performed the same mode overlap investigation using the same parameters as in section 3.4. Here, I will present the figures and discuss their similarities, thus justifying my design choices.
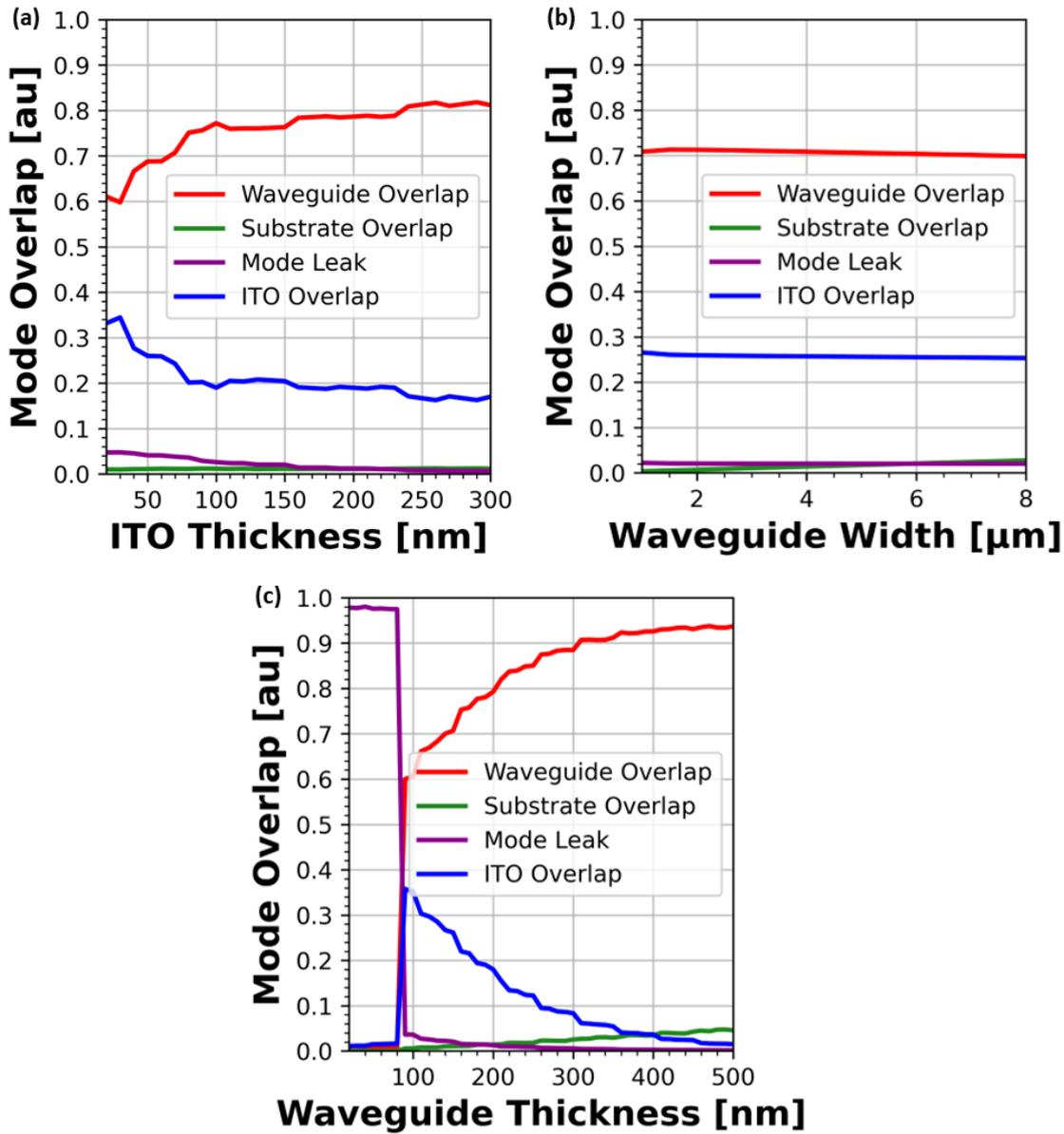
Figure 9.3: Mode overlap as a function of (a) indium tin oxide layer thickness, (b) waveguide width, and (c) waveguide layer thickness for a slab structure.

The mode overlaps for the waveguide layer (red), substrate layer (green), and ITO layer (blue) for a slab waveguide/capacitor structure are shown in Figure 9.3 as a function of (a) ITO layer thickness, (b) waveguide width, and (c) waveguide layer thickness. As expected, these trends are identical to the original simulations in section 3.4.2.
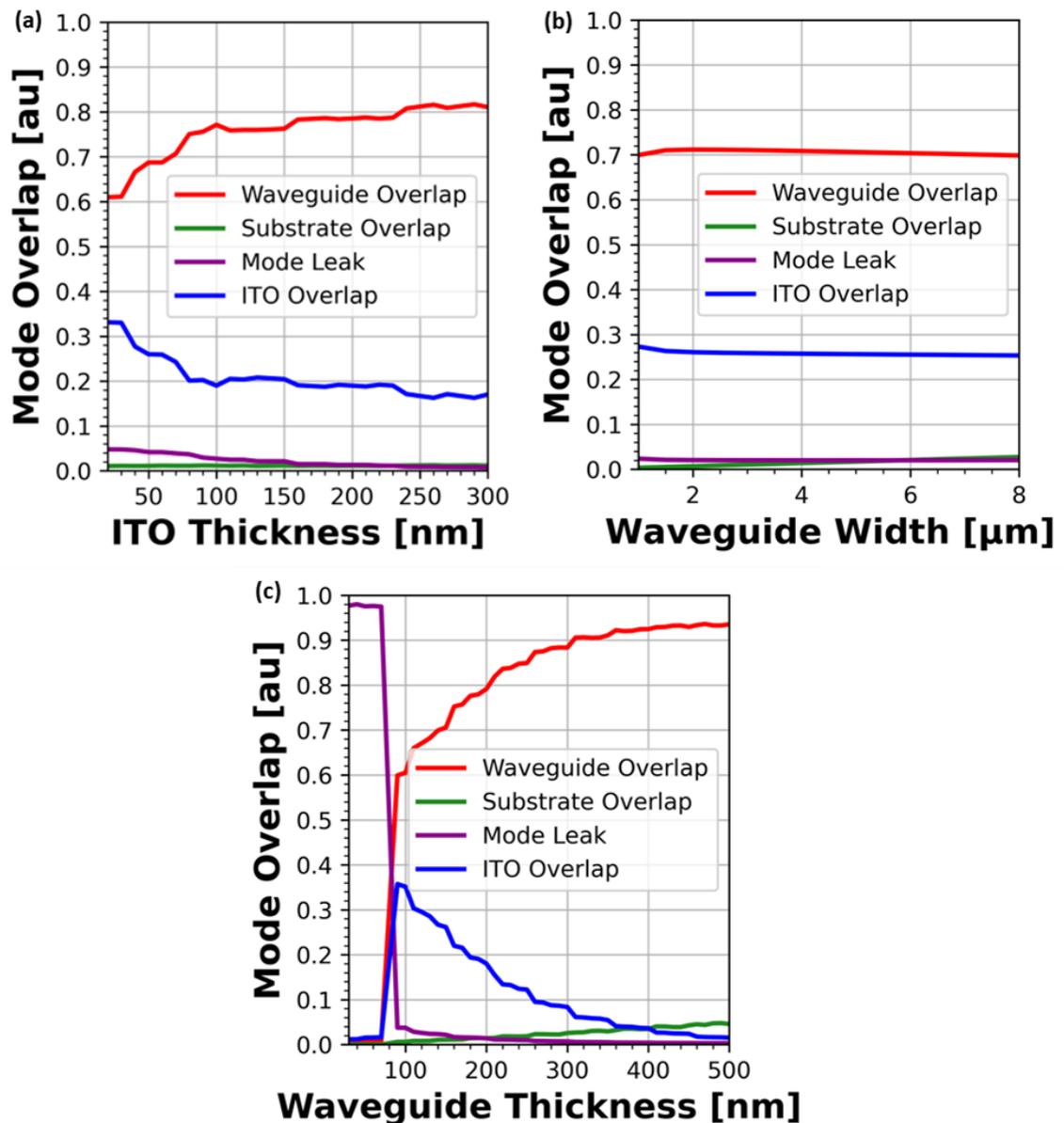
Figure 9.4: Mode overlap as a function of (a) indium tin oxide layer thickness, (b) waveguide width, and (c) waveguide layer thickness for a fully surrounded waveguide structure.

The mode overlaps for the waveguide layer (red), substrate layer (green), and ITO layer (blue) for a fully surrounded waveguide structure are shown in Figure 9.4 as a function of (a) ITO layer thickness, (b) waveguide width, and (c) waveguide layer thickness. The overlaps are nearly identical to the slab waveguide/capacitor structure shown in Figure 9.3, and as such the design choice of a fully surrounded waveguide structure will make little difference to the modulator performance.
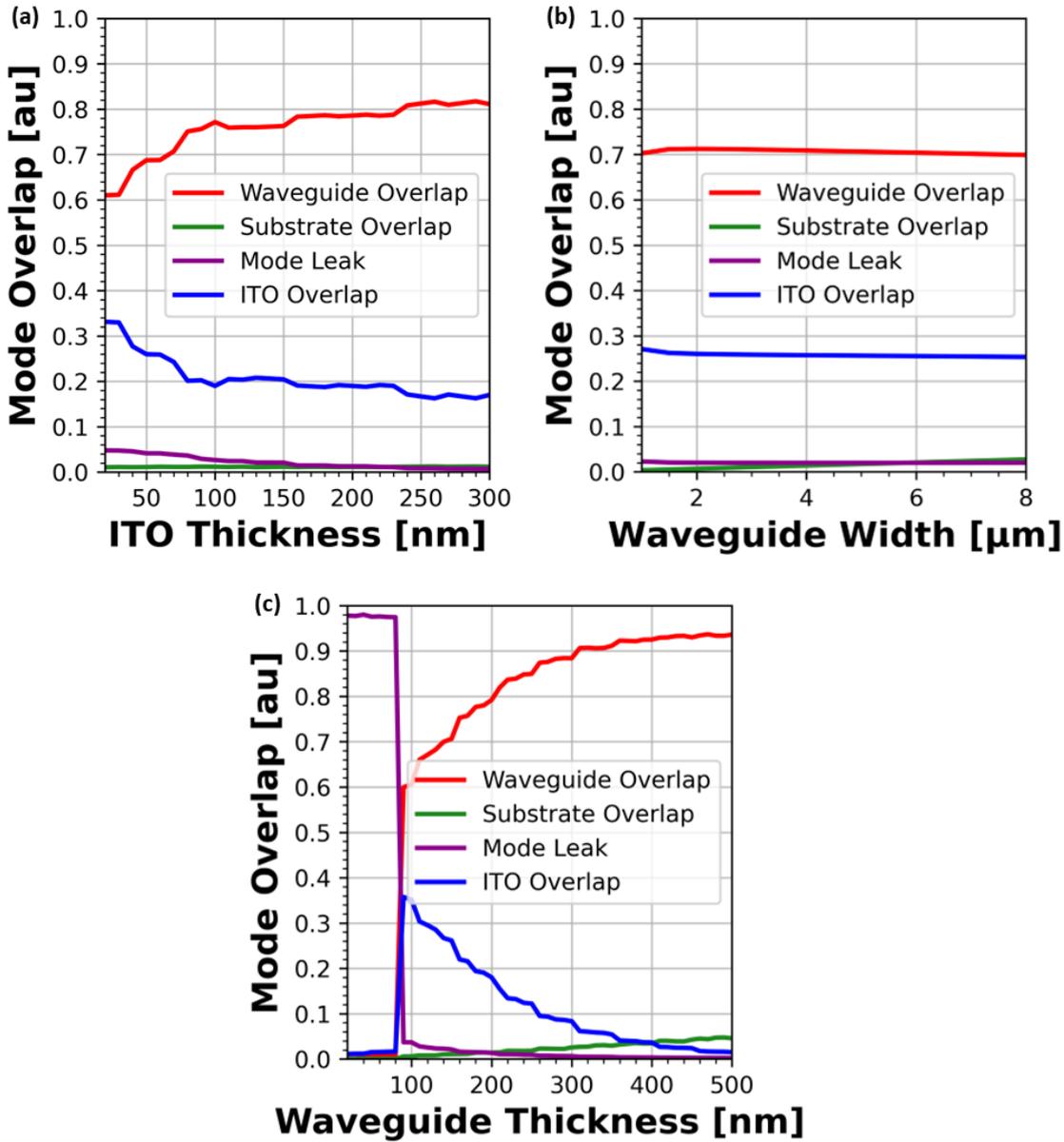
Figure 9.5: Mode overlap as a function of (a) indium tin oxide layer thickness, (b) waveguide width, and (c) waveguide layer thickness for a partially surrounded waveguide structure.

The mode overlaps for the waveguide layer (red), substrate layer (green), and ITO layer (blue) for a partially surrounded waveguide structure are shown in Figure 9.5 as a function of (a) ITO layer thickness, (b) waveguide width, and (c) waveguide layer thickness. The overlaps are nearly identical to the slab and fully surrounded structures shown in Figure 9.3 and Figure 9.4, respectively, and as such the design choice of a partially surrounded waveguide structure will make little difference to the modulator performance.

9.4. Ellipsometry and Reflectometry

9.4.1. Spectroscopic Ellipsometry

Spectroscopic ellipsometry is a common technique used to determine the optical properties of a thin film material. Ellipsometry provides a contactless, non-invasive, and non-destructive optical technique to determine the thin film thickness and optical properties through the change in polarisation state of light reflected from a thin film [313]. The polarisation change is represented as the complex Fresnel reflection coefficients, amplitude change (Ψ), and phase difference (Δ). Ellipsometry can determine the refractive index and extinction coefficient for a known thickness of material, or vice versa. Spectroscopic ellipsometry is a technique that was available to me during my research but was not used to characterise the sputter-deposited ITO thin films, here I will detail why that is the case.

9.4.1.1. Principle of Operation

Ellipsometry detects the change in p- (parallel) and s- (perpendicular) polarisation components of incident light upon reflection on a thin film material, with respect to each other. The change in polarisation is used to determine the material's thickness and optical constants.
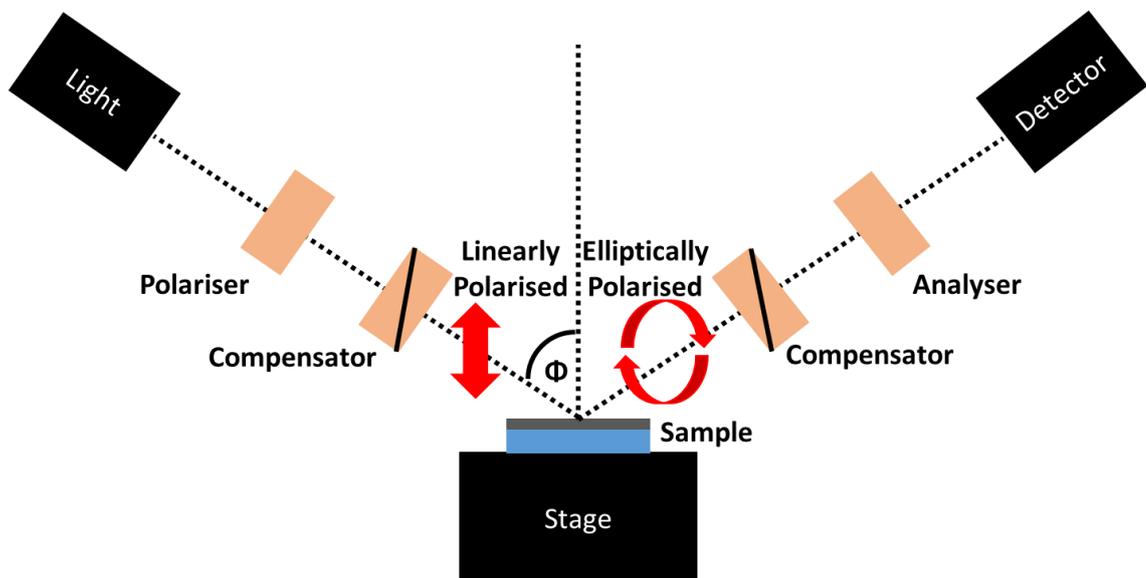
Figure 9.6: Optical elements required to determine the change in polarisation of the light upon reflection at the sample, (Spectroscopic Ellipsometer setup) [314].

Figure 9.6 illustrates the optical components required to collect spectroscopic ellipsometry data (an ellipsometer). A broadband light source (left) produces an unpolarised light wave that passes through a polariser. The wave becomes linearly polarised and is incident on the sample surface at an angle (Φ). Post-reflection, the light is elliptically polarised and passes through an analyser, which allows an amount of light to pass through depending on the orientation of the polariser with respect to the electric field coming from the sample. The reflected polarisation is determined by the detector and the change in polarisation (ρ) is

$$\rho = \tan(\Psi)\,e^{i\Delta}$$

Equation 9.14

calculated using Equation 9.14. The polariser and analyser rotate independently, determined by the controlling software, to obtain the desired angles. The compensators on either side of the sample are optional and are to alter the beam size for varying sample dimensions.

9.4.1.2. Results

The optical components, or the film thickness, are obtained in spectroscopic ellipsometry by fitting experimental data to a known oscillator model, such as the Cauchy model or Lorentz model [99], [315]. The input parameters are varied until a good match between the experimental data and the model data is achieved. Unfortunately, the system available to me at the University of York used open-source software with a limited number of models available. The simplicity of such a system was not suitable to measure a complex material such as ITO.
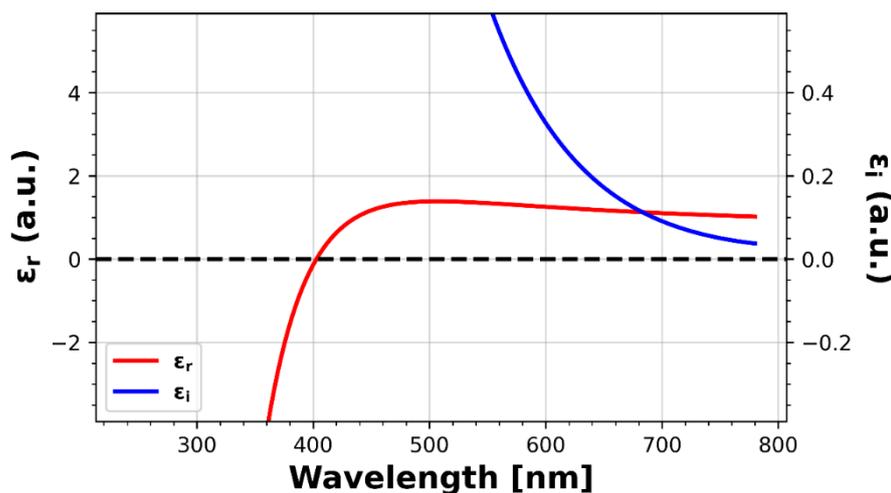
Figure 9.7: Real (red) and imaginary (blue) components of the complex permittivity of a sputter deposited ITO thin film with 1% oxygen flow concentration.

The simplicity of the available oscillator models is evident in the real (red) and imaginary (blue) components of the complex material permittivity shown in Figure 9.7, and the corresponding refractive index (green) and extinction coefficient (orange) components of the complex refractive index shown in Figure 9.8 and calculated using Equation 2.41. The Drude model, which describes ITO, does not produce dispersion curves that show the relationship observed in Figure 9.7, which is more common of a dielectric material.
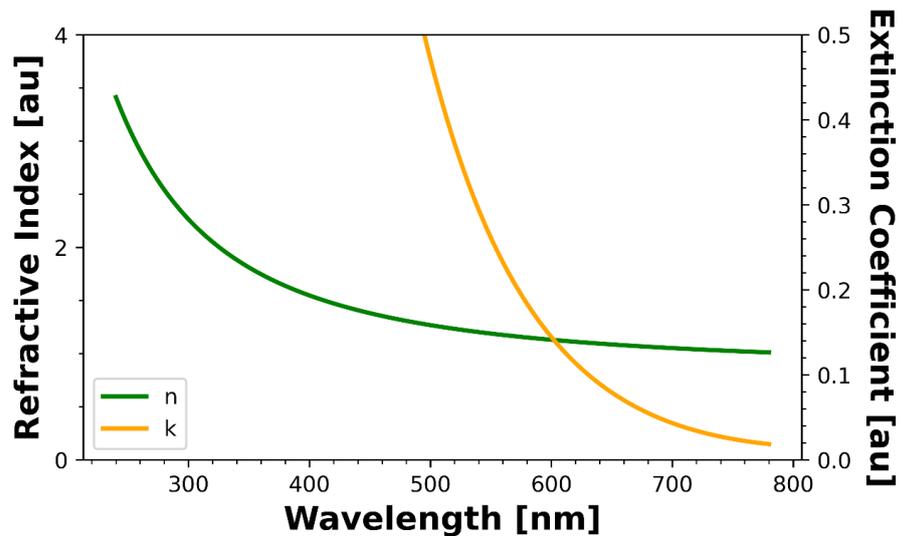
Figure 9.8: Real (green) and imaginary (orange) components of the complex refractive index of a sputter deposited ITO thin film with 1% oxygen flow concentration.

Given the results of the spectroscopic ellipsometry measurement, we decided to use a spectral reflectometry technique to determine the optical properties of the sputter deposited ITO, which I will discuss in the next section.

9.4.2.   Spectroscopic Reflectometry

Spectroscopic reflectometry is another common technique used to determine the optical properties, or film thickness, of a thin film material. Reflectometry is a non-destructive analytical technique that uses the reflection of light from a sample to measure the film thickness and complex refractive index components.

9.4.2.1. Principle of Operation

Unlike spectroscopic ellipsometry, reflectometry uses an intensity ratio of light to determine the optical properties of the material. Incident light with intensity (I_0) is incident orthogonally on the sample which produces multiple reflections inside the thin film. The interactions between these reflections produces interference fringes in reflection, with intensity (I_R), and in transmission, with intensity (I_T), as illustrated in Figure 9.9.
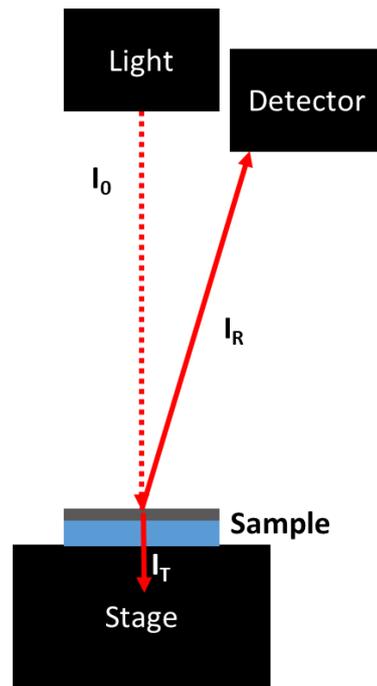
Figure 9.9: Illustration of a spectroscopic reflectometer.:

A model of a known material dispersion is then fitted to the measured interference pattern to determine the optical properties and material thickness, i.e., the components of the optical path length that determines the interference pattern.

9.4.2.2. Results

We used an F20 spectroscopic reflectometer from Filmetrics to measure the optical properties of sputter deposited ITO. As with ellipsometry, the simplicity of a purely optical measurement is not sufficient to fully characterise a complex material such as ITO. Figure 9.10 show the real (red) and imaginary (blue) components of the complex material permittivity for a sputter deposited thin film of ITO with a 27% oxygen flow concentration during the deposition. For comparison, I have plotted the dispersion on the same axis scale. We know from the Drude model that as the

wavelength of incident light increases, the real component of the permittivity tends to 0 and the imaginary component increases exponentially.
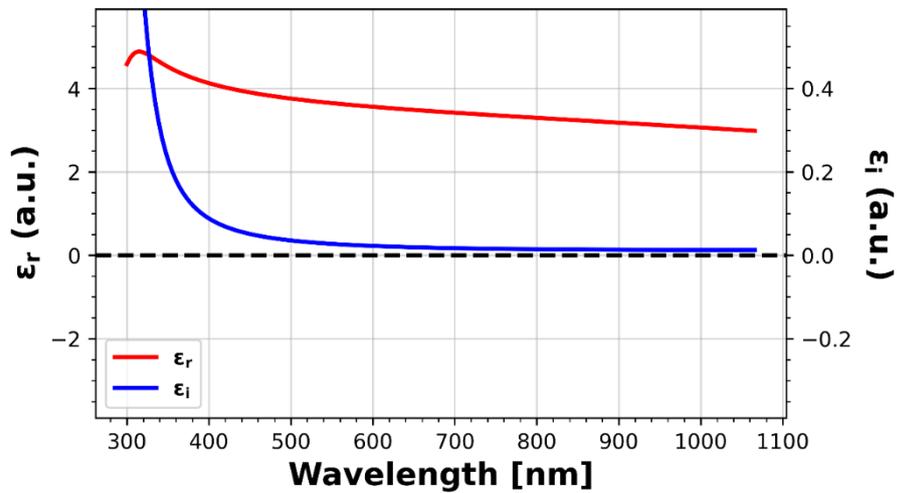
Figure 9.10: Real (red) and imaginary (blue) components of the complex permittivity of a sputter deposited ITO thin film with 17% oxygen flow concentration.

This is not the case in the measurement shown in Figure 9.10, nor with the real (green) and imaginary (orange) components of the complex refractive index shown in Figure 9.11. We expect the extinction coefficient to increase with wavelength, not exhibit the behaviour of a dielectric material.
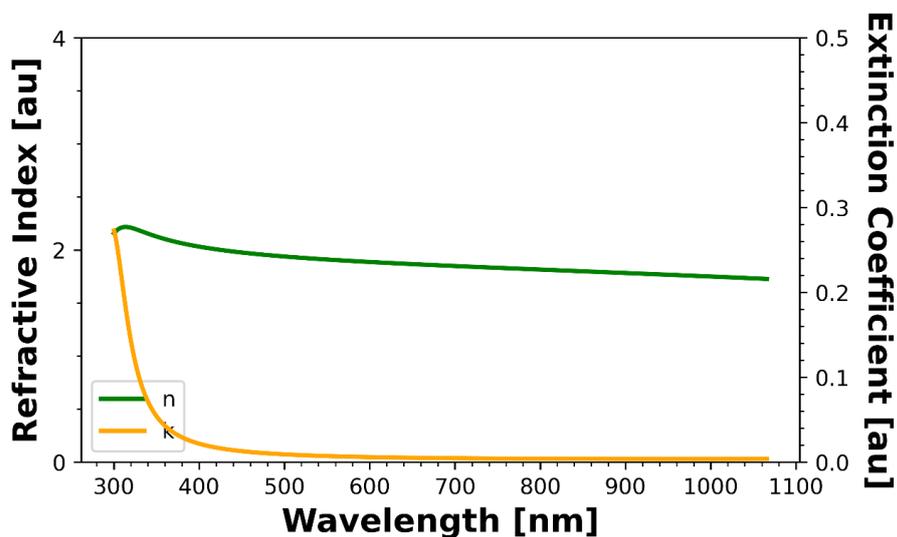
Figure 9.11: Real (green) and imaginary (orange) components of the complex refractive index of a sputter deposited ITO thin film with 1% oxygen flow concentration.

The simplicity of an all-optical measurement and the inconsistent results it produced, is what lead us to the measurement technique presented in chapter 4.

9.5.   Additional Photoresists

In section 0, I detailed the application of photoresists and referenced the film thickness as a function of spin speed for the S1800 photoresist series. S1818 is not the only photoresist used in my research, I used a plethora of resists for different purposes throughout (5). Here, I will detail the other resists and show their spin curves.

9.5.1.   Comment on Spinning

Most photoresists are designed with a range of film thicknesses that can be achieved by adjusting the spin coating spin at which the photoresist is applied. It is common for the photoresist to have a minimum achievable thickness (an asymptote of the relationship). Due to the stability of the material in this region, I use a spin speed in the asymptotic region of the spin curve and choose the corresponding photoresist for a desired thickness.

9.5.2.   Ultraviolet Lithography Resits

MicroChem's LOR series is a simple bilayer resist with high thermal stability and thin film deposition compatibility; I use LOR 7B as a bilayer lift-off resist. The resist is spun using a 500 rpm for 20 seconds and 5000 rpm for 60 seconds protocol, and the final film thickness may be estimated in the 500nm region from the spin curve in Figure 9.12.
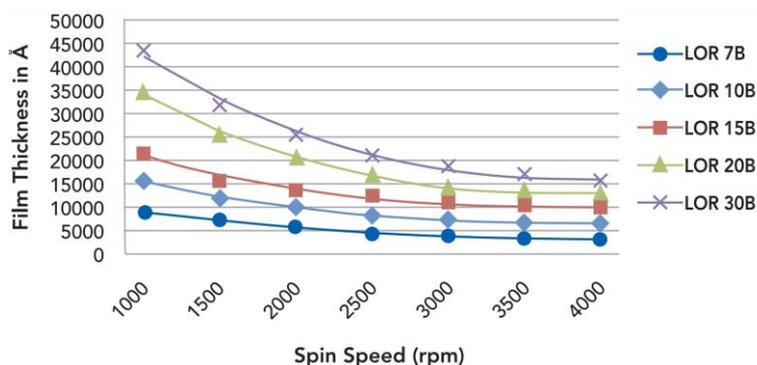
Figure 9.12: Photoresist thickness as a function of spin speed for lift-off resist (LOR) series [293].

In chapter 5, I detailed that I used two different concentrations of SU-8 as the insulator layer in the modulator. The first was a mix with volume ration 7:2 of SU-8 2050: 2000.5, where the

2000.5 resist serves to thin the final mixture and target a thinner final layer. The second was a layer of SU-8 2000.5.
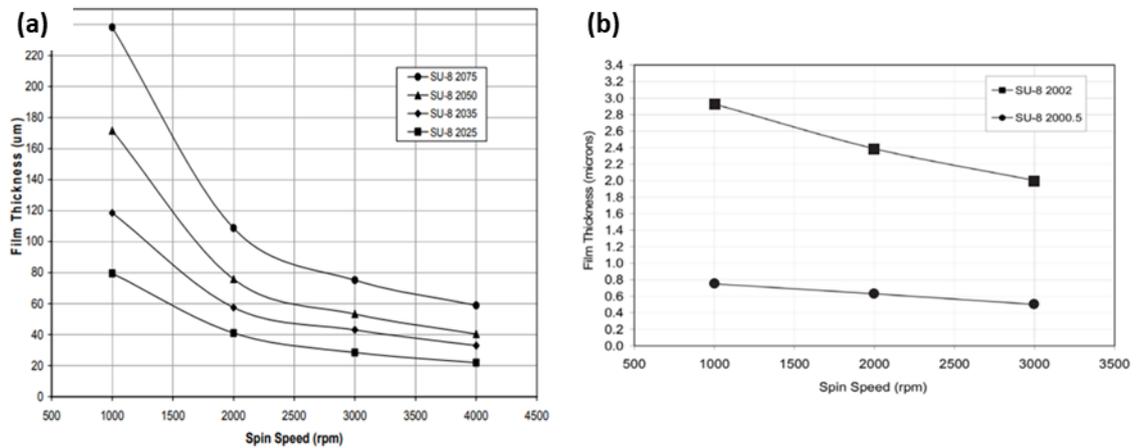
Figure 9.13: Thickness as a function of spin speed for SU-8 2000 series [294] showing the (a) thicker and (b) thinner films in the range.

The spin curves for these two SU-8 photoresists are shown in Figure 9.13, though neither shows the spin speed of 5000 rpm that I used. The thickness of the 7:2 mixture has been calculated as (760±10) nm in chapter 6 using surface profilometry. I performed the same repeat measurements using the 2000.5 SU-8 film, where an example of the surface profilometry data is shown in Figure 9.14, and measured a film thickness of (210±20) nm.
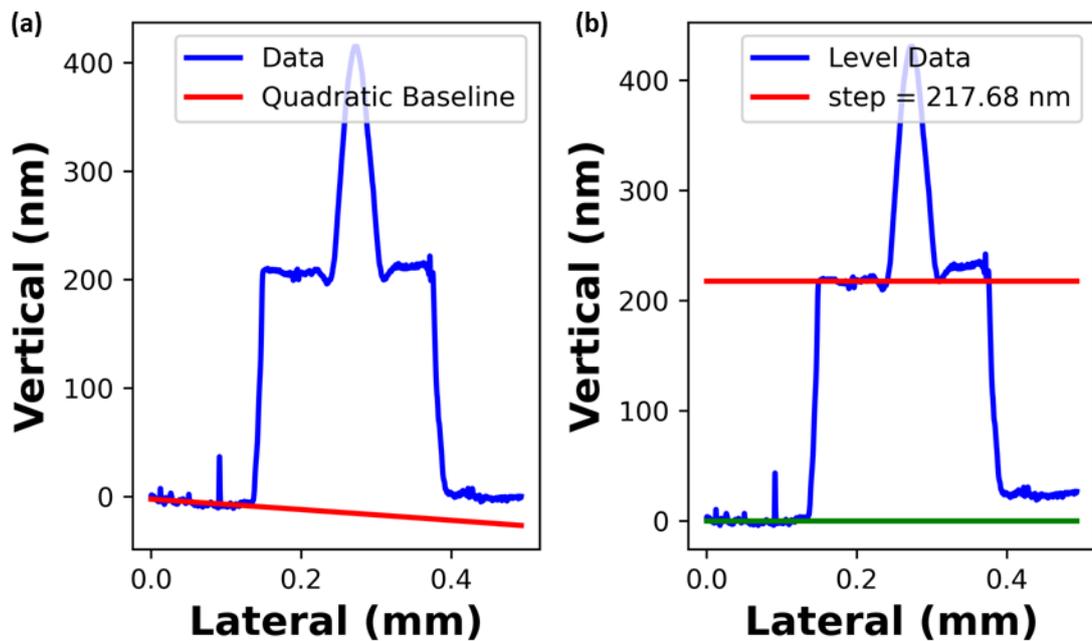
Figure 9.14: Surface profile analysis of SU-8 layer showing (a) distorted and (b) levelled surface profile.

9.5.3.   Electron-Beam Lithography Resists

Finally, the electron-beam lithography resists I use are ALLRESIST's AR-P 6200.13 and AR-PC 5091.01. The resist layer (AR-P) is spun at 5000 rpm for 60 seconds to achieve a thickness of approximately 350 nm, shown by the green line in Figure 9.15(a). The charge dissipation layer (AR-PC) is spun at 2000rpm for 60 seconds, to achieve a thickness of approximately 3 μm, shown by the orange line in Figure 9.15(b).
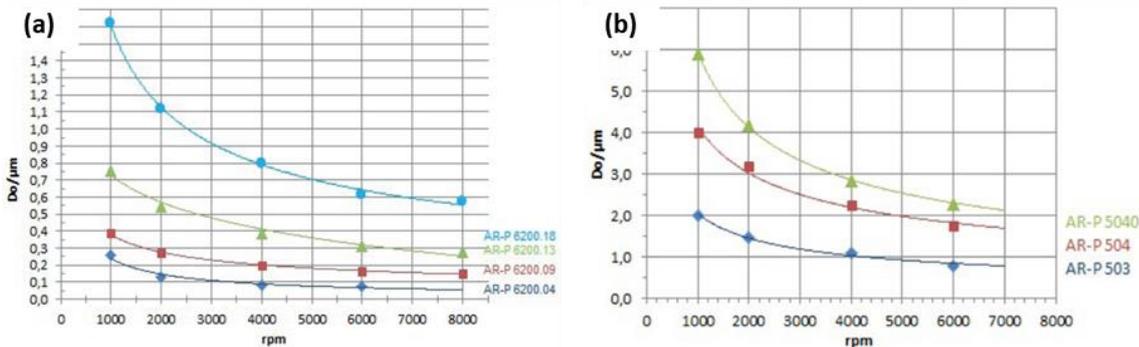
Figure 9.15: Film thickness as a function of spin speed for (a) AR-P 6200 and (b) AR-PC 5000 series [193], [295].

9.6.   Dual-Mode Effect

The dual mode simulations discussed in section 7.5.1 feature two figures (Figure 7.11 and Figure 7.12) that are constructed from several images taken from the simulation. Here, I will detail how those figures come together and show the individual images for data integrity. In the simulations, a short pulse of light is used to reduce the required computational power. As such, it is only possible to measure the position of the pulse through the waveguide by combining multiple images. The output of the simulation is a video of the waveguide mode travelling through the structure, and I use time-stamped screenshots to capture the position of the mode.
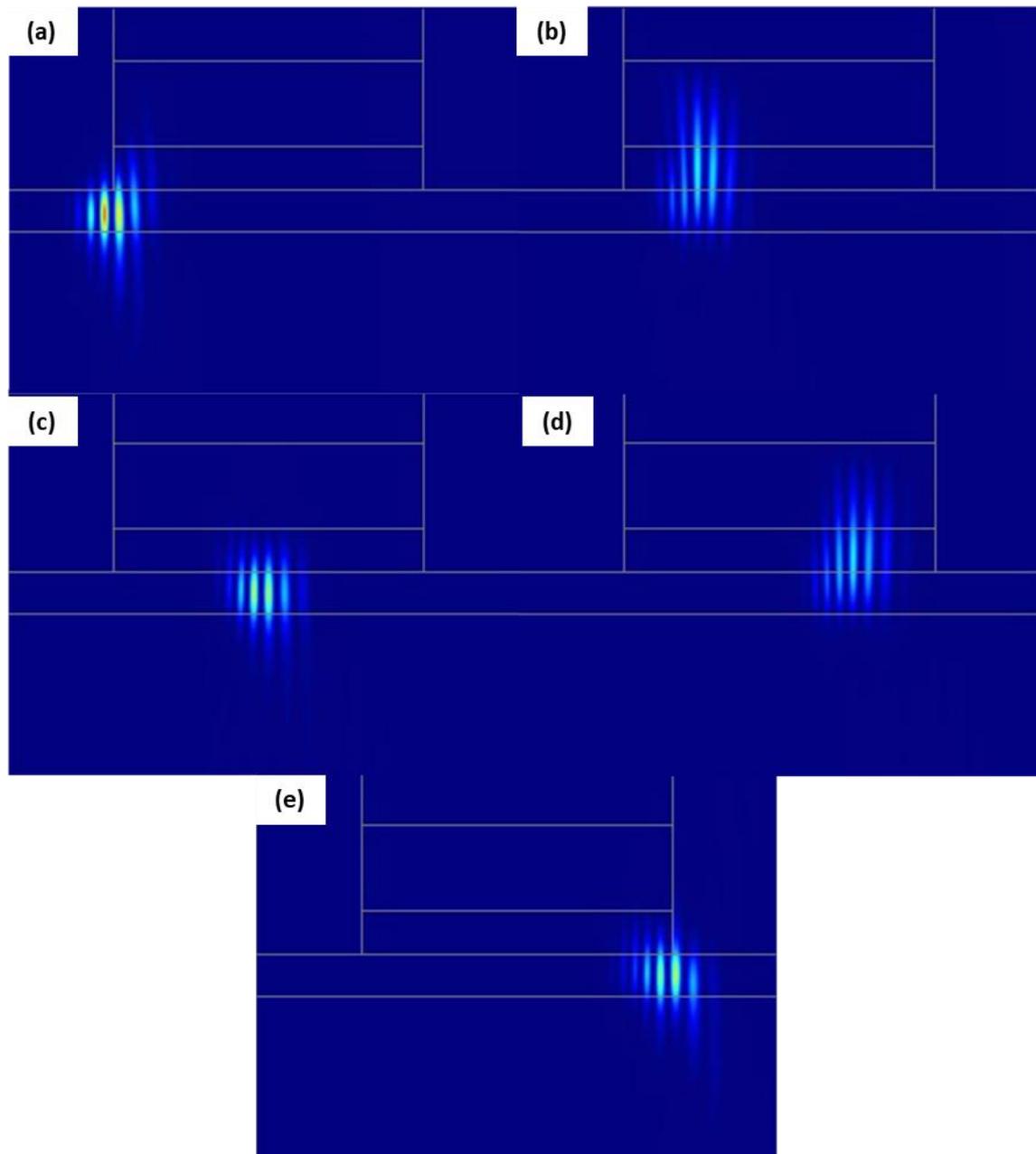
Figure 9.16: Fundamental waveguide mode in a 1.5µm gold, SU-8, ITO, silicon nitride on glass modulator side-profile. The waveguide mode profile is shown at (a) the waveguide entry (b)-(c) travelling through the waveguide, and (e) exiting the waveguide. The ITO has a refractive index of 1.9.

The individual images are shown in Figure 9.16(a)-(e) for the mode at the waveguide entrance (a), through the modulator (b)-(d) and the waveguide exit (e) for the simulation with ITO refractive index of 1.9 RIU. I import the images into an image processing software (Gimp) and remove the blue background using a filter. This allows the waveguide modes to remain in each image, and the images may then be merged into one. Then I use the line tool and background tool to replace the simulated structure and background respectively to produce the image shown in Figure 9.17.
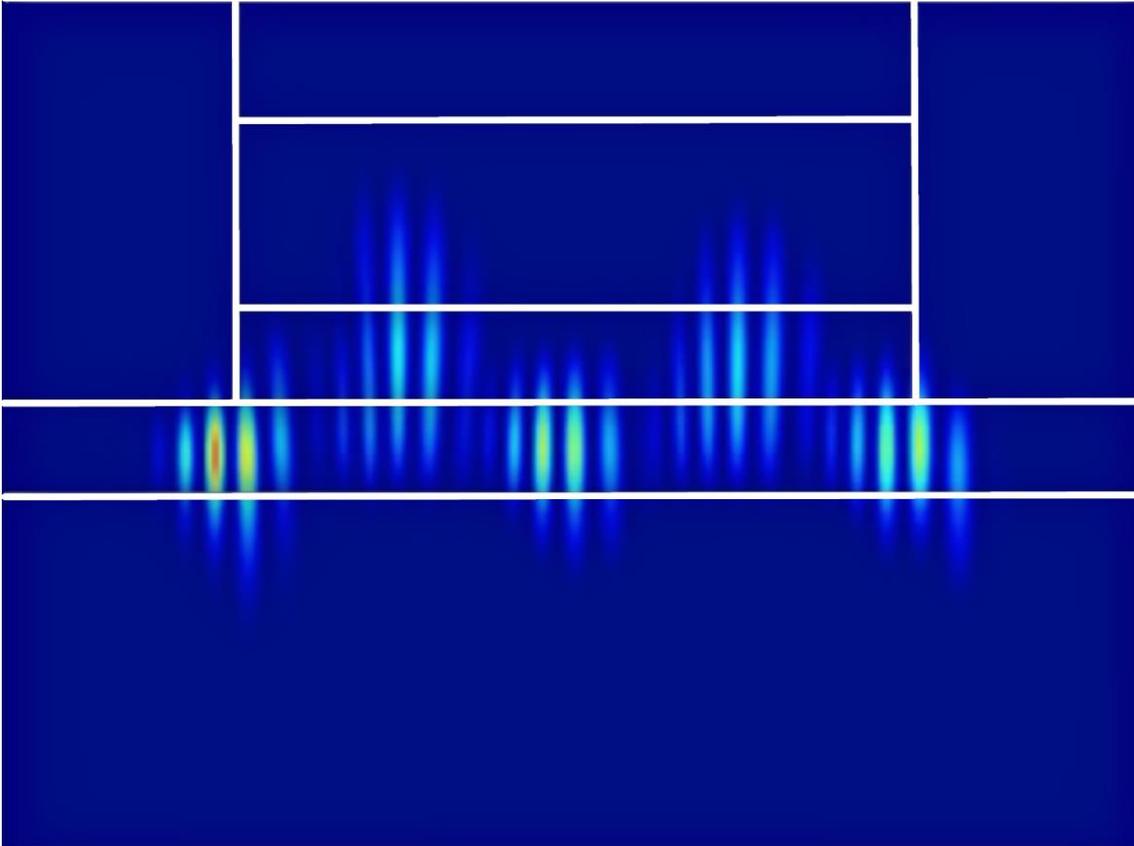
Figure 9.17: Fundamental waveguide mode in a 1.5 µm gold, SU-8, ITO, silicon nitride on glass modulator side-profile. The waveguide mode profile is shown imposed on top of each other from 5 images. The ITO has a refractive index of 1.9.

Using the rectangle filling tool, I can overlay the material colours and labels used throughout this thesis onto Figure 9.17 and give the intensity scale bar and axes to produce the result shown in Figure 7.11. The process is repeated for Figure 7.12.

9.7.   Damped Sine Wave Fits

To verify the non-linear regression used to fit a damped sine wave to the optical measurements in section 7.4 onwards, I used Origin Pro's waveform damped sine wave fit with some of the measurements to create a parameter space in which I could vary the non-linear regression. The fitting tool within Origin Pro uses a Lavenberg-Marquardt algorithm, which is a damped least-squares (DLS) method used to solve non-linear least squares problems [316]. The fits resemble those presented in the main thesis, thus confirming the non-linear regression method used in section 4.7.1.

9.7.1. Interferometry Phase Modulation
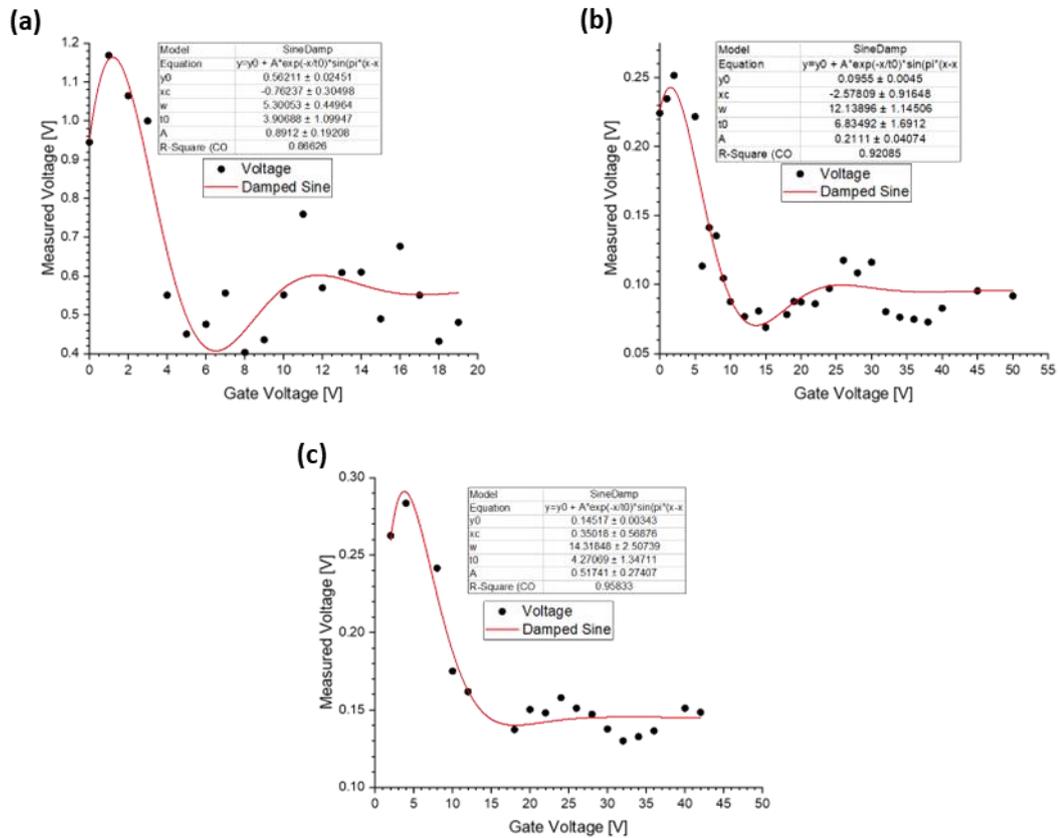
Figure 9.18: Oscilloscope measured voltage as a function of capacitor gate voltage for (a) 60 µm, (b) 70 µm, and (c) 100 µm (a)(b) single and (c) double capacitor modulator.

For the modulator of 60 µm length, presented in Figure 7.8, the DLS method calculates a half oscillation period of 5.3, which matches that calculated with the Rosenbrock function. For the 70 µm length modulator, presented in Figure 7.9, the DLS method calculates a half oscillation period of 12.1, which is in reasonable agreement with the Rosenbrock function calculated value. Finally, for the double modulator of length 100 µm, presented in Figure 7.10, the DLS method estimates a half oscillation period of 14.3, which is also in reasonable agreement with the Rosenbrock function. To obtain the DLS method results, some data points were excluded.

9.7.2.   Straight-Through Phase Modulation
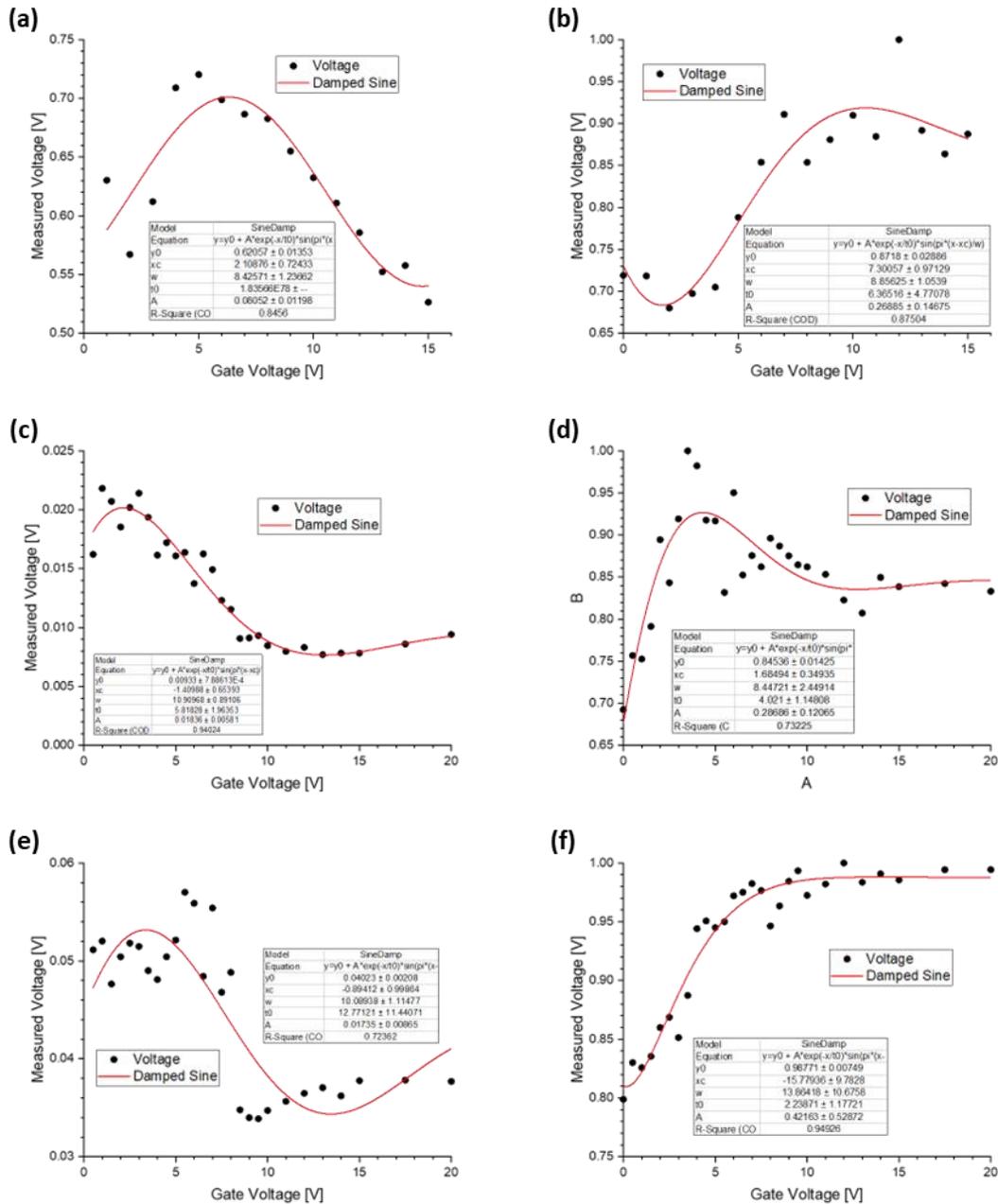
Figure 9.19: Oscilloscope measured voltage as a function of capacitor gate voltage for (a)(b) 600 μm, (c)(d) 50 μm, and (e)(f) 100 μm straight-through modulator at (a)(b) 520 nm, (c)(d) 635 nm, and (e)(f) 785 nm input wavelength.

For the 600 μm straight through modulator in Figure 9.19(a) and (b), presented in Figure 7.15 and Figure 7.16, the half oscillation period of 8.4 and 8.6 is in good agreement with the Rosenbrock formulation-fitted data. Similarly the 50 μm and 100 μm straight through modulators are in good agreement with the Rosenbrock-fitted data. Hence, I can assume that

the non-linear regression performed is accurate, and any higher order frequencies that may be observed are anomalous results.

9.8.   Verifying Collision Frequency

The collision frequency parameter in the Drude model (Equation 4.1) has been discussed throughout this thesis. The collision frequency (Γ) is a measure of the time between electron collisions in a Drude material, and it follows that a material with high collision frequency has a large density of electrons and a high optical absorption. In section 4.7 we introduced a method of measuring the collision frequency of sputter deposited ITO using a guided mode resonance technique. The ITO used in the modulators discussed in this thesis, however, did not use this technique. Instead, I used the optical measurements for a straight-through modulator to determine the absorption and calculate the collision frequency. Here I will discuss how the collision frequency may be obtained using this method.

9.8.1.   Optical Absorption

The optical absorption of a material (α)

$$\alpha = \frac{4\pi k}{\lambda}$$

<div align="right">*Equation 9.15*</div>

is related to the extinction coefficient (k) and the incident wavelength (λ) using Lambert's law [123], as shown in Equation 9.15. The fraction of incident power (P(x))

$$P(x) = P(0)e^{-\alpha x}$$

<div align="right">*Equation 9.16*</div>

that has propagated from a position (x=0) to a distance (x) through a material with an absorption given by Equation 9.15, may be expressed as a fraction of the incident power (P(0)), as shown in Equation 9.16. To determine the collision frequency, I first needed to find the extinction coefficient of the material using the optical measurement technique discussed in section 7.3.

9.8.2.　　　Optical Measurements

Using the optical measurement setup (Figure 7.1) and the oscilloscope voltage measurement technique discussed in section 7.3.2, for a straight-through modulator, I took the 0V gate voltage measurement for modulators of different length. Rearranging Equation 9.16, and taking the natural logarithm of the function

$$\ln\big(P(x)\big) = -\alpha x + \ln\big(P(0)\big)$$

Equation 9.17

a linear relationship is expected between the power (P(x)) and the modulator length (x), as shown in Equation 9.17. The results are shown in Figure 9.20 where the incident wavelengths are 520nm (black), 635nm (blue), and 785nm (red). Linear fits are plotted with weighted errors, with errors calculated using the technique discussed in Appendix section 9.1.2, for the measured power, and half the UV lithography spot size on either end of the modulator length (10 µm). I have plotted the gradients as positive (not as indicated in Equation 9.17) to easily translate the absorption into extinction coefficient using Equation 9.15.

Equation	y = a + b*x		
Plot	520nm	635nm	785nm
Intercept	-1.82111 ± 0.05686		
Slope	0.00657 ± 8.45035E-5	0.00656 ± 1.5594E-4	0.00843 ± 3.65057E-4
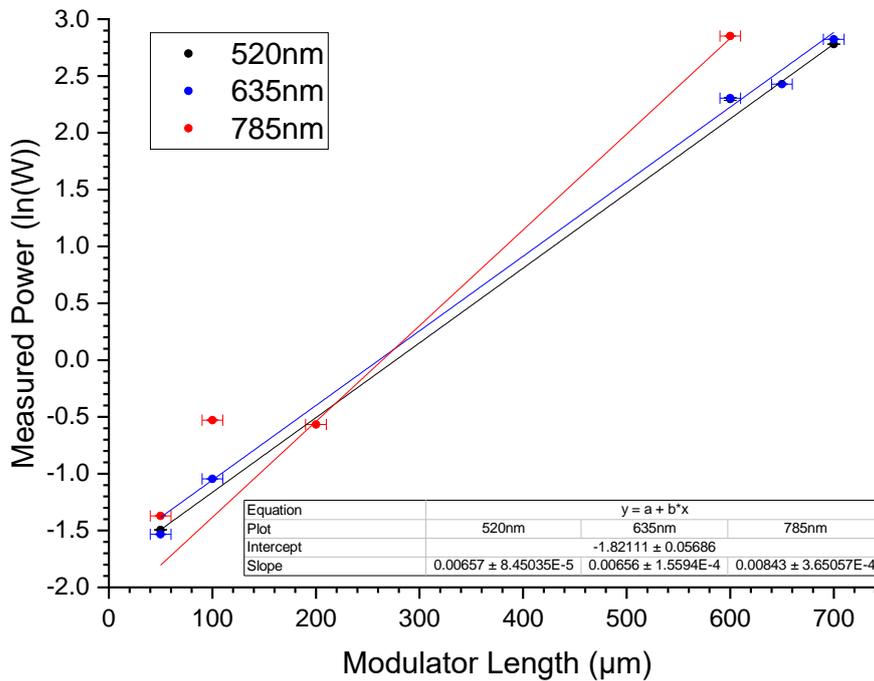
Figure 9.20: Measured incident power on the oscilloscope as a function of modulator length for incident wavelengths of 520nm (black), 635nm (blue), and 785nm (red).

The error-weighted gradients for the incident wavelengths are shown in the box in Figure 9.20, and are shown in Table 9.1. Using Equation 9.15, I calculated the wavelength dependent extinction coefficients, which are shown in Table 9.1.

Table 9.1: Absorption and extinction coefficient of sputter deposited ITO with 27% oxygen flow concentration during deposition as a function of incident wavelength.

Wavelength [nm]	Absorption $[\mu m^{-1}]$ (x10^{-3})	Extinction Coefficient [au] (x10^{-4})
520	(6.6±0.1)	(2.72±0.03)
635	(6.6±0.2)	(3.31±0.08)
785	(8.4±0.4)	(5.3±0.2)

9.8.3.   Collision Frequency

To calculate the collision frequency, I used the regression algorithm discussed in section 4.7.1 with the measured electrical parameters and the extinction coefficients in Table 9.1 as known data points. This technique yielded a collision frequency of 13 THz, which was used to plot the dispersion curves shown in Figure 5.32.

9.9.   Expected Modulator Performance

Using the analytical model (2.4.1) and the electrical properties discussed in chapter 6, I can now obtain an estimate of how the modulator will perform. I used an input wavelength of 635nm, with a silicon nitride waveguide on a glass substrate with refractive indices (n=2.01 and n=1.44, respectively) {7, 8}, and nitride thickness of 150nm. I used the Drude model optical properties (section 5.7.2) and a layer thickness of 197nm for ITO, as discussed in chapter 6. I modelled the accumulation layer with 5nm thickness, and optical properties calculated using the Drude model, where the free carrier density is modelled using the parallel plate capacitor formulation (section 2.6.3). For the SU-8 layer, I used a refractive index (n=1.6) from the literature {9}, and infinite thickness, i.e., the SU-8 layer serves as the cladding layer in the optical model.
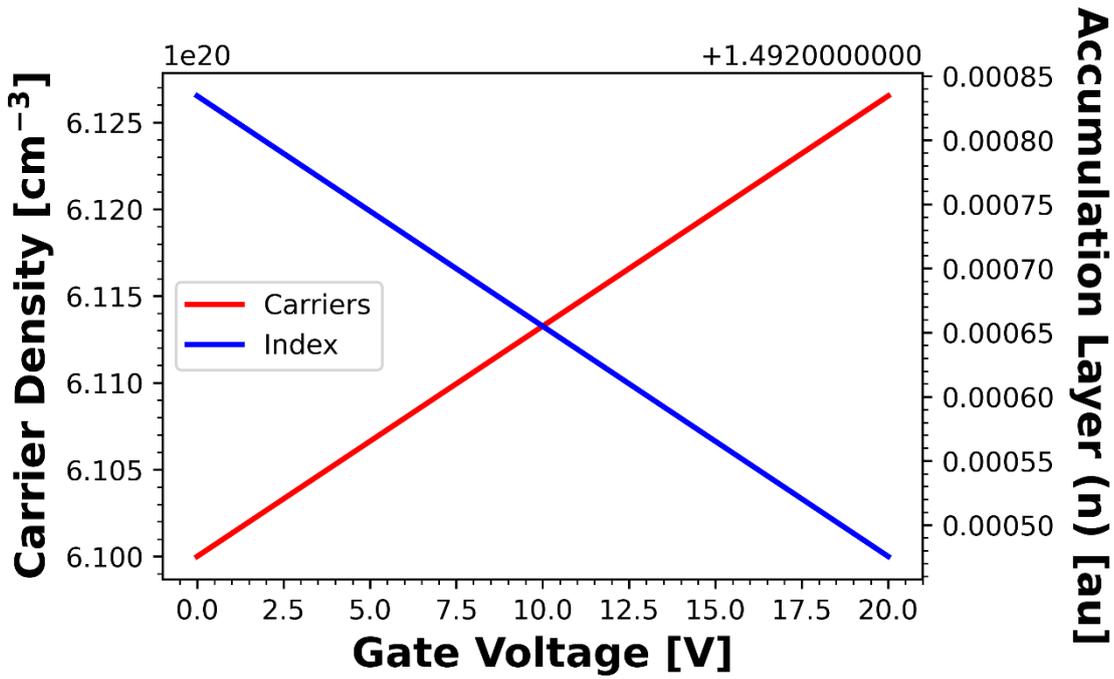
Figure 9.21: Predicted local carrier density (red) and refractive index (blue) within a 5nm accumulation layer of ITO. The letter has a linear relationship.

The free carrier density in the accumulation layer is shown by the red line in Figure 9.21, with the corresponding refractive index shown in blue. As expected, the relationship between gate voltage and carrier density/refractive index is linear. I then used the accumulation layer refractive index in the optical model to determine the effective mode index, shown in Figure 9.22.
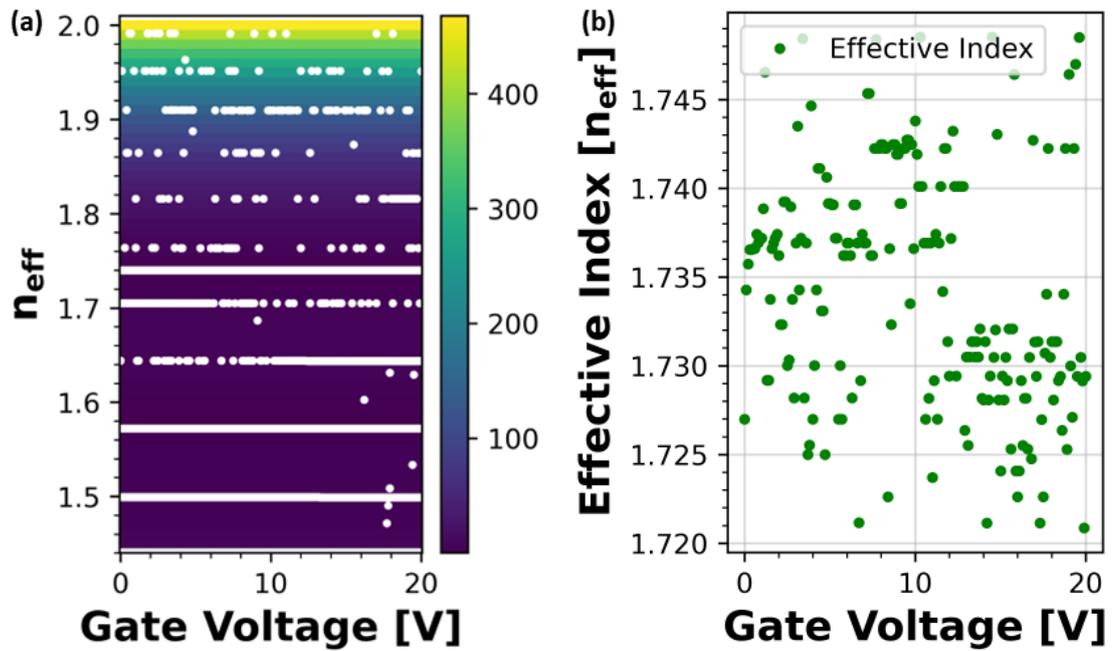
Figure 9.22: Effective index as a function of gate voltage for an ITO based modulator plotted as (a) a surface plot with the propagation constant and (b) the gate voltage.

As shown in Figure 9.22(a), the number of guided modes with a smaller wavelength and lower waveguide refractive index than the literature modulators yield a significantly higher number of guided modes. The fundamental TE mode is shown around n_{eff}=1.73, as verified using other optical modelling techniques (section 3.2.2) and has been plotted a function of the gate voltage in Figure 9.22(b). The nonlinear trend observed in Figure 9.22(b) is a result of the small difference of indices in the structure and is discussed further in chapter 7.

Bibliography

[1]   S. F. J. Blair, J. S. Male, S. A. Cavill, C. P. Reardon, and T. F. Krauss, 'Photonic Characterisation of Indium Tin Oxide as a Function of Deposition Conditions', *Nanomaterials*, vol. 13, no. 13, Art. no. 13, Jan. 2023, doi: 10.3390/nano13131990.

[2]   G. E. Moore, 'Cramming More Components Onto Integrated Circuits', *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, Jan. 1998, doi: 10.1109/JPROC.1998.658762.

[3]   T. N. Theis and H.-S. P. Wong, 'The End of Moore's Law: A New Beginning for Information Technology', *Computing in Science Engineering*, vol. 19, no. 2, pp. 41–50, Mar. 2017, doi: 10.1109/MCSE.2017.29.

[4]   K. Rupp, '48 Years of Microprocessor Trend Data'. Zenodo, Jul. 2020. doi: 10.5281/zenodo.3947824.

[5]   R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, 'Design of ion-implanted MOSFET's with very small physical dimensions', *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, Oct. 1974, doi: 10.1109/JSSC.1974.1050511.

[6]   N. S. Kim *et al.*, 'Leakage current: Moore's law meets static power', *Computer*, vol. 36, no. 12, pp. 68–75, Dec. 2003, doi: 10.1109/MC.2003.1250885.

[7]   P. Bose, 'Power Wall', in *Encyclopedia of Parallel Computing*, D. Padua, Ed., Boston, MA: Springer US, 2011, pp. 1593–1608. doi: 10.1007/978-0-387-09766-4_499.

[8]   H. Esmaeilzadeh, E. Blem, R. S. Amant, K. Sankaralingam, and D. Burger, 'Dark silicon and the end of multicore scaling', *Annual International Symposium on Computer Architecture*, Jul. 2011, doi: 10.1145/2000064.2000108.

[9]   S. Borkar, 'Designing reliable systems from unreliable components: the challenges of transistor variability and degradation', *IEEE Micro*, vol. 25, no. 6, pp. 10–16, Nov. 2005, doi: 10.1109/MM.2005.110.

[10]   T. F. de Lima, B. J. Shastri, A. N. Tait, M. A. Nahmias, and P. R. Prucnal, 'Progress in neuromorphic photonics', *Nanophotonics*, vol. 6, no. 3, pp. 577–599, May 2017, doi: 10.1515/nanoph-2016-0139.

[11]   '1.3: Von Neumann and Harvard Architectures', Engineering LibreTexts. Accessed: Mar. 24, 2023. [Online]. Available: https://eng.libretexts.org/Bookshelves/Electrical_Engineering/Electronics/Implementin g_a_One_Address_CPU_in_Logisim_(Kann)/01%3A_Introduction/1.03%3A_Von_Neum ann_and_Harvard_Architectures

[12]   J. Edwards and S. O'Keefe, 'Eager recirculating memory to alleviate the von Neumann Bottleneck', in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, Dec. 2016, pp. 1–5. doi: 10.1109/SSCI.2016.7850155.

[13]   S. S. Sahoo, B. Ranjbar, and A. Kumar, 'Reliability-Aware Resource Management in Multi-/Many-Core Systems: A Perspective Paper', *Journal of Low Power Electronics and Applications*, vol. 11, p. 7, Jan. 2021, doi: 10.3390/jlpea11010007.

[14]   P. J. Denning and T. G. Lewis, 'Exponential laws of computing growth', *Commun. ACM*, vol. 60, no. 1, pp. 54–65, Dec. 2016, doi: 10.1145/2976758.

[15]   'How Fast Is Technology Advancing in 2023?', Techjury. Accessed: Mar. 25, 2023. [Online]. Available: https://techjury.net/blog/how-fast-is-technology-growing/

[16]   A. S. G. Andrae and T. Edler, 'On Global Electricity Usage of Communication Technology: Trends to 2030', *Challenges*, vol. 6, no. 1, Art. no. 1, Jun. 2015, doi: 10.3390/challe6010117.

Bibliography

[17]   K. Liu, C. R. Ye, S. Khan, and V. J. Sorger, 'Review and perspective on ultrafast wavelength-size electro-optic modulators', *Laser & Photonics Reviews*, vol. 9, no. 2, pp. 172–194, Mar. 2015, doi: 10.1002/lpor.201400219.

[18]   B. E. Floren, 'Optical interconnects in the Touchstone supercomputer program', in *Integrated Optoelectronics for Communication and Processing*, SPIE, Dec. 1992, p. 46. doi: 10.1117/12.2321818.

[19]   A. Louri and H. Sung, 'An optical multi-mesh hypercube: a scalable optical interconnection network for massively parallel computing', *Journal of Lightwave Technology*, vol. 12, no. 4, pp. 704–716, Apr. 1994, doi: 10.1109/50.285368.

[20]   A. F. Benner, M. Ignatowski, J. A. Kash, D. M. Kuchta, and M. B. Ritter, 'Exploitation of optical interconnects in future server architectures', *IBM Journal of Research and Development*, vol. 49, no. 4.5, pp. 755–775, Jul. 2005, doi: 10.1147/rd.494.0755.

[21]   J. A. Kash, 'Leveraging Optical Interconnects in Future Supercomputers and Servers', in *2008 16th IEEE Symposium on High Performance Interconnects*, Aug. 2008, pp. 190–194. doi: 10.1109/HOTI.2008.29.

[22]   A. Imamura, A. Matsutani, and F. Koyama, 'Multi-wavelength 1060nm-band VCSEL array with tapered hollow waveguide multiplexer', in *22nd IEEE International Semiconductor Laser Conference*, Sep. 2010, pp. 7–8. doi: 10.1109/ISLC.2010.5642780.

[23]   T. Saeki *et al.*, '100Gbit/s compact transmitter module integrated with optical multiplexer', in *2013 IEEE Photonics Conference*, Sep. 2013, pp. 307–308. doi: 10.1109/IPCon.2013.6656558.

[24]   D. Nikolova *et al.*, 'Scaling silicon photonic switch fabrics for data center interconnection networks', *Opt. Express, OE*, vol. 23, no. 2, pp. 1159–1175, Jan. 2015, doi: 10.1364/OE.23.001159.

[25]   J. E. Cunningham *et al.*, 'Highly-efficient thermally-tuned resonant optical filters', *Opt. Express, OE*, vol. 18, no. 18, pp. 19055–19063, Aug. 2010, doi: 10.1364/OE.18.019055.

[26]   R. Soref, 'The Past, Present, and Future of Silicon Photonics', *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 6, pp. 1678–1687, Nov. 2006, doi: 10.1109/JSTQE.2006.883151.

[27]   R. Ding *et al.*, 'A Compact Low-Power 320-Gb/s WDM Transmitter Based on Silicon Microrings', *IEEE Photonics Journal*, vol. 6, no. 3, pp. 1–8, Jun. 2014, doi: 10.1109/JPHOT.2014.2326656.

[28]   Y. Bengio, A. Courville, and P. Vincent, 'Representation Learning: A Review and New Perspectives', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013, doi: 10.1109/TPAMI.2013.50.

[29]   J. Hasler and H. Marr, 'Finding a roadmap to achieve large neuromorphic hardware systems', *Frontiers in Neuroscience*, vol. 7, 2013, Accessed: Mar. 25, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2013.00118

[30]   'What is the relationship between GMACs and GFLOPs? · Issue #16 · sovrasov/flops-counter.pytorch', GitHub. Accessed: Apr. 10, 2023. [Online]. Available: https://github.com/sovrasov/flops-counter.pytorch/issues/16

[31]   'AMD-Powered Frontier Supercomputer Tops Top500 At 1.1 Exaflops, Tops Green500 Too - Phoronix'. Accessed: Apr. 10, 2023. [Online]. Available: https://www.phoronix.com/news/Top500-Green500-Frontier

[32]   S.-C. Liu, T. Delbruck, G. Indiveri, A. Whatley, and R. Douglas, *Event-Based Neuromorphic Systems*. Wiley, 2015.

[33]   P. A. Merolla *et al.*, 'A million spiking-neuron integrated circuit with a scalable communication network and interface', *Science*, vol. 345, no. 6197, pp. 668–673, Aug. 2014, doi: 10.1126/science.1254642.

[34]   S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, 'The SpiNNaker project', *Institute of Electrical and Electronics Engineers. Proceedings*, vol. 102, no. 5, pp. 652–665, Feb. 2014, doi: 10.1109/JPROC.2014.2304638.

[35]   B. Shastri, A. N. Tait, T. Ferreira de Lima, M. Nahmias, H.-T. Peng, and P. Prucnal, 'Principles of Neuromorphic Photonics: A Volume in the Encyclopedia of Complexity and Systems Science, Second Edition', 2018, pp. 83–118. doi: 10.1007/978-1-4939-6883-1_702.

[36]   D. Liang, G. Roelkens, R. Baets, and J. E. Bowers, 'Hybrid Integrated Platforms for Silicon Photonics', *Materials*, vol. 3, no. 3, Art. no. 3, Mar. 2010, doi: 10.3390/ma3031782.

[37]   S. Dogru and N. Dagli, '0.2 V Drive Voltage Substrate Removed Electro-Optic Mach–Zehnder Modulators With MQW Cores at 1.55 μm', *J. Lightwave Technol., JLT*, vol. 32, no. 3, pp. 435–439, Feb. 2014.

[38]   R. Amin, J. George, J. Khurgin, T. El-Ghazawi, P. R. Prucnal, and V. J. Sorger, 'Attojoule Modulators for Photonic Neuromorphic Computing', in *2018 Conference on Lasers and Electro-Optics (CLEO)*, May 2018, pp. 1–2.

[39]   'Scaling up silicon photonic-based accelerators: Challenges and opportunities: APL Photonics: Vol 7, No 2'. Accessed: Mar. 25, 2023. [Online]. Available: https://aip.scitation.org/doi/full/10.1063/5.0070992

[40]   Y. LeCun, Y. Bengio, and G. Hinton, 'Deep learning', *Nature*, vol. 521, no. 7553, Art. no. 7553, May 2015, doi: 10.1038/nature14539.

[41]   A. Krizhevsky, I. Sutskever, and G. E. Hinton, 'ImageNet Classification with Deep Convolutional Neural Networks', in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2012, pp. 1097–1105. Accessed: Jun. 17, 2020. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[42]   M. Davies *et al.*, 'Loihi: A Neuromorphic Manycore Processor with On-Chip Learning', *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan. 2018, doi: 10.1109/MM.2018.112130359.

[43]   T. Mikolov, A. Deoras, D. Povey, L. Burget, and J. Cernocky, 'Strategies for training large scale neural network language models', *2011 IEEE Workshop on Automatic Speech Recognition & Understanding*, pp. 196–201, Dec. 2011, doi: 10.1109/ASRU.2011.6163930.

[44]   G. Hinton *et al.*, 'Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups', *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, Nov. 2012, doi: 10.1109/MSP.2012.2205597.

[45]   Y. LeCun *et al.*, 'Handwritten Digit Recognition with a Back-Propagation Network', in *Advances in Neural Information Processing Systems*, Morgan-Kaufmann, 1989. Accessed: Mar. 25, 2023. [Online]. Available: https://proceedings.neurips.cc/paper/1989/hash/53c3bce66e43be4f209556518c2fcb54-Abstract.html

[46]   R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, 'Convolutional neural networks: an overview and application in radiology', *Insights Imaging*, vol. 9, no. 4, Art. no. 4, Aug. 2018, doi: 10.1007/s13244-018-0639-9.

[47]   P. Merolla, J. Arthur, F. Akopyan, N. Imam, R. Manohar, and D. Modha, *A digital neurosynaptic core using embedded crossbar memory with 45pJ per spike in 45nm*. 2011, p. 4. doi: 10.1109/CICC.2011.6055294.

[48]   F. Akopyan *et al.*, 'TrueNorth: Design and Tool Flow of a 65 mW 1 Million Neuron Programmable Neurosynaptic Chip', *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 10, pp. 1537–1557, Oct. 2015, doi: 10.1109/TCAD.2015.2474396.

[49]   B. V. Benjamin *et al.*, 'Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations', *Proceedings of the IEEE*, vol. 102, no. 5, pp. 699–716, May 2014, doi: 10.1109/JPROC.2014.2313565.

[50] C. Frenkel, M. Lefebvre, J. Legat, and D. Bol, 'A 0.086-mm2 12.7-pJ/SOP 64k-synapse 256-neuron online-learning digital spiking neuromorphic processor in 28-nm CMOS', *IEEE Trans. Biomed. Cirtcuits Syst.*, vol. 13, no. no.1, pp. 145–158, 2019.

[51] S. Furber *et al.*, 'Overview of the SpiNNaker System Architecture', *Computers, IEEE Transactions on*, vol. 62, pp. 2454–2467, Dec. 2013, doi: 10.1109/TC.2012.142.

[52] E. Painkras *et al.*, 'SpiNNaker: A 1-W 18-Core System-on-Chip for Massively-Parallel Neural Network Simulation', *Solid-State Circuits, IEEE Journal of*, vol. 48, pp. 1943–1953, Aug. 2013, doi: 10.1109/JSSC.2013.2259038.

[53] L. Plana *et al.*, 'SpiNNaker: Design and Implementation of a GALS Multicore System-on-Chip', *JETC*, vol. 7, p. 17, Dec. 2011, doi: 10.1145/2043643.2043647.

[54] K. Meier, 'A mixed-signal universal neuromorphic computing system', in *2015 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2015, p. 4.6.1-4.6.4.

[55] J. Schemmel, S. Billaudelle, P. Dauer, and J. Weis, 'Accelerated Analog Neuromorphic Computing'. arXiv, Mar. 26, 2020. doi: 10.48550/arXiv.2003.11996.

[56] S. Schmitt *et al.*, 'Neuromorphic hardware in the loop: Training a deep spiking network on the BrainScaleS wafer-scale system', in *2017 International Joint Conference on Neural Networks (IJCNN)*, May 2017, pp. 2227–2234. doi: 10.1109/IJCNN.2017.7966125.

[57] J. W. Goodman, 'Introduction to Fourier optics', *Introduction to Fourier optics, 3rd ed., by JW Goodman. Englewood, CO: Roberts & Co. Publishers, 2005*, vol. 1, 2005.

[58] L. Bottou and O. Bousquet, 'The Tradeoffs of Large Scale Learning', in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2007. Accessed: Mar. 25, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2007/hash/0d3180d672e08b4c5312 dcdafdf6ef36-Abstract.html

[59] M. Mathieu, M. Henaff, and Y. LeCun, 'Fast Training of Convolutional Networks through FFTs', *arXiv:1312.5851 [cs]*, Mar. 2014, Accessed: Jun. 17, 2020. [Online]. Available: http://arxiv.org/abs/1312.5851

[60] 'Intel® Core™ i9-10980XE Extreme Edition Processor 198017'. Accessed: Jun. 17, 2020. [Online]. Available: https://www.intel.co.uk/content/www/uk/en/products/processors/core/x-series/i9-10980xe.html

[61] 'Optalysys launches FT:X 2000 - The world's first commercial optical processing system', insideHPC. Accessed: Jun. 17, 2020. [Online]. Available: https://insidehpc.com/2019/03/optalysys-launches-ftx-2000-the-worlds-first-commercial-optical-processing-system/

[62] N. Savage, 'Digital spatial light modulators', *Nature Photonics*, vol. 3, no. 3, Art. no. 3, Mar. 2009, doi: 10.1038/nphoton.2009.18.

[63] D. Wang, Y. Zhang, C. Yang, and K. Wang, 'Study on processing synthetic aperture radar data based on an optical 4f system for fast imaging', *Opt. Express, OE*, vol. 30, no. 25, pp. 44408–44419, Dec. 2022, doi: 10.1364/OE.471716.

[64] 'Optical phase modulators for MHz and GHz modulation in silicon-on-insulator (SOI) - University of Surrey'. Accessed: Sep. 15, 2023. [Online]. Available: https://openresearch.surrey.ac.uk/esploro/outputs/journalArticle/Optical-phase-modulators-for-MHz-and/99514050502346

[65] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, 'Silicon optical modulators', *Nature Photonics*, vol. 4, no. 8, Art. no. 8, Aug. 2010, doi: 10.1038/nphoton.2010.179.

[66] D. Janner, D. Tulli, M. García-Granda, M. Belmonte, and V. Pruneri, 'Micro-structured integrated electro-optic LiNbO3 modulators', *Laser & Photonics Reviews*, vol. 3, no. 3, pp. 301–313, 2009, doi: 10.1002/lpor.200810073.

[67] C. Wang, M. Zhang, B. Stern, M. Lipson, and M. Lončar, 'Nanophotonic lithium niobate electro-optic modulators', *Opt. Express, OE*, vol. 26, no. 2, pp. 1547–1555, Jan. 2018, doi: 10.1364/OE.26.001547.

[68] X. Huang, Y. Liu, H. Guan, Z. Yu, M. Tan, and Z. Li, 'High-Efficiency, Slow-Light Modulator on Hybrid Thin-Film Lithium Niobate Platform', *IEEE Photonics Technology Letters*, vol. 33, no. 19, pp. 1093–1096, Oct. 2021, doi: 10.1109/LPT.2021.3106642.

[69] R. Soref and B. Bennett, 'Electrooptical effects in silicon', *IEEE Journal of Quantum Electronics*, vol. 23, no. 1, pp. 123–129, Jan. 1987, doi: 10.1109/JQE.1987.1073206.

[70] A. Liu *et al.*, 'A high-speed silicon optical modulator based on a metal–oxide–semiconductor capacitor', *Nature*, vol. 427, no. 6975, Art. no. 6975, Feb. 2004, doi: 10.1038/nature02310.

[71] W. M. J. Green, M. J. Rooks, L. Sekaric, and Y. A. Vlasov, 'Ultra-compact, low RF power, 10 Gb/s silicon Mach-Zehnder modulator', *Opt. Express, OE*, vol. 15, no. 25, pp. 17106–17113, Dec. 2007, doi: 10.1364/OE.15.017106.

[72] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, and M. Lipson, '12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators', *Opt Express*, vol. 15, no. 2, pp. 430–436, Jan. 2007, doi: 10.1364/oe.15.000430.

[73] D. Kharas *et al.*, 'High-Power (>300 mW) On-Chip Laser With Passively Aligned Silicon-Nitride Waveguide DBR Cavity', *IEEE Photonics Journal*, vol. 12, no. 6, pp. 1–12, Dec. 2020, doi: 10.1109/JPHOT.2020.3037834.

[74] R. Amin *et al.*, 'A Lateral MOS-Capacitor-Enabled ITO Mach–Zehnder Modulator for Beam Steering', *Journal of Lightwave Technology*, vol. 38, no. 2, pp. 282–290, Jan. 2020, doi: 10.1109/JLT.2019.2956719.

[75] R. Amin *et al.*, 'Heterogeneously integrated ITO plasmonic Mach–Zehnder interferometric modulator on SOI', *Scientific Reports*, vol. 11, no. 1, Art. no. 1, Jan. 2021, doi: 10.1038/s41598-020-80381-3.

[76] J. Witzens, 'High-Speed Silicon Photonics Modulators', *Proceedings of the IEEE*, vol. 106, no. 12, pp. 2158–2182, Dec. 2018, doi: 10.1109/JPROC.2018.2877636.

[77] R. Amin *et al.*, '0.52 V mm ITO-based Mach-Zehnder modulator in silicon photonics', *APL Photonics*, vol. 3, no. 12, p. 126104, Dec. 2018, doi: 10.1063/1.5052635.

[78] R. Amin *et al.*, 'ITO-based electro-absorption modulator for photonic neural activation function', *APL Materials*, vol. 7, no. 8, p. 081112, Aug. 2019, doi: 10.1063/1.5109039.

[79] Z. Ma, Z. Li, K. Liu, C. Ye, and V. Sorger, 'Indium-Tin-Oxide for High-performance Electro-optic Modulation', *Nanophotonics*, vol. 4, pp. 198–213, Jun. 2015, doi: 10.1515/nanoph-2015-0006.

[80] E. Feigenbaum, K. Diest, and H. A. Atwater, 'Unity-Order Index Change in Transparent Conducting Oxides at Visible Frequencies', *Nano Lett.*, vol. 10, no. 6, pp. 2111–2116, Jun. 2010, doi: 10.1021/nl1006307.

[81] Y. Shen *et al.*, 'Deep learning with coherent nanophotonic circuits', *Nature Photon*, vol. 11, no. 7, Art. no. 7, Jul. 2017, doi: 10.1038/nphoton.2017.93.

[82] J. Bueno *et al.*, 'Reinforcement learning in a large-scale photonic recurrent neural network', *Optica, OPTICA*, vol. 5, no. 6, pp. 756–760, Jun. 2018, doi: 10.1364/OPTICA.5.000756.

[83] T.-Y. Cheng, D.-Y. Chou, C.-C. Liu, Y.-J. Chang, and C.-C. Chen, 'Optical neural networks based on optical fiber-communication system', *Neurocomputing*, vol. 364, pp. 239–244, Oct. 2019, doi: 10.1016/j.neucom.2019.07.051.

[84] H. R. Philipp, 'Optical Properties of Silicon Nitride', *J. Electrochem. Soc.*, vol. 120, no. 2, pp. 295–300, Jan. 1973, doi: 10.1149/1.2403440.

[85] J. D. Joannopoulos, Ed., *Photonic crystals: molding the flow of light*, 2nd ed. Princeton: Princeton University Press, 2008.

[86] A. I. Arbab and F. A. Yassein, 'A New Formulation of Electrodynamics', *JEMAA*, vol. 02, no. 08, pp. 457–461, 2010, doi: 10.4236/jemaa.2010.28060.

[87] G. J. Triggs, 'Resonant Grating Surfaces for Biosensing', p. 142.

[88] K. Okamoto, *Fundamentals of Optical Waveguides*. Elsevier, 2021.

[89]   M. Moskowitz, *A Course In Complex Analysis In One Variable*. World Scientific Publishing Company, 2002.

[90]   T. Tamir, Ed., *Integrated Optics*, vol. 7. in Topics in Applied Physics, vol. 7. Berlin, Heidelberg: Springer, 1975. doi: 10.1007/978-3-662-43208-2.

[91]   D. Marcuse, *Light Transmission Optics*. Van Nostrand Reinhold, 1982.

[92]   F. Bloch, 'Über die Quantenmechanik der Elektronen in Kristallgittern', *Z. Physik*, vol. 52, no. 7, pp. 555–600, Jul. 1929, doi: 10.1007/BF01339455.

[93]   C. Kittel, *Introduction to Solid State Physics*, 8th ed. Wiley, 2004. [Online]. Available: http://www.amazon.com/Introduction-Solid-Physics-Charles-Kittel/dp/047141526X/ref=dp_ob_title_bk

[94]   A. Drayton, 'Autonomous photonic biosensor', phd, University of York, 2021. Accessed: Dec. 11, 2022. [Online]. Available: https://etheses.whiterose.ac.uk/29096/

[95]   I. H. Malitson, 'Interspecimen Comparison of the Refractive Index of Fused Silica$\ast$,†', *J. Opt. Soc. Am.*, vol. 55, no. 10, pp. 1205–1209, Oct. 1965, doi: 10.1364/JOSA.55.001205.

[96]   'Refractive Index of SU-8 for Thin Film Thickness Measurement'. Accessed: May 02, 2023. [Online]. Available: https://www.filmetrics.com/refractive-index-database/SU-8

[97]   H.-Y. Li *et al.*, 'Analysis of the Drude Model in Metallic Films', *Applied optics*, vol. 40, pp. 6307–11, Jan. 2002, doi: 10.1364/AO.40.006307.

[98]   P. Drude, 'Zur Elektronentheorie der Metalle; II. Teil. Galvanomagnetische und thermomagnetische Effecte', *Annalen der Physik*, vol. 308, no. 11, pp. 369–402, 1900, doi: 10.1002/andp.19003081102.

[99]   M. Eldlio, F. Che, and M. Cada, 'Drude-Lorentz Model of Semiconductor Optical Plasmons', in *IAENG Transactions on Engineering Technologies: Special Issue of the World Congress on Engineering and Computer Science 2012*, H. K. Kim, S.-I. Ao, M. A. Amouzegar, and B. B. Rieger, Eds., in Lecture Notes in Electrical Engineering. , Dordrecht: Springer Netherlands, 2014, pp. 41–49. doi: 10.1007/978-94-007-6818-5_4.

[100]   G. L. Zeng and M. Zeng, 'Ohm's Law', in *Electric Circuits: A Concise, Conceptual Tutorial*, G. L. Zeng and M. Zeng, Eds., Cham: Springer International Publishing, 2021, pp. 17–21. doi: 10.1007/978-3-030-60515-5_3.

[101]   P. B. Johnson and R. W. Christy, 'Optical Constants of the Noble Metals', *Phys. Rev. B*, vol. 6, no. 12, pp. 4370–4379, Dec. 1972, doi: 10.1103/PhysRevB.6.4370.

[102]   G. Duffett, 'Metal-insulator-metal nanoresonators for refractive index sensing', phd, University of York, 2021. Accessed: Feb. 01, 2023. [Online]. Available: https://etheses.whiterose.ac.uk/29493/

[103]   C. J. Chang-Hasnain and W. Yang, 'High-contrast gratings for integrated optoelectronics', *Adv. Opt. Photon., AOP*, vol. 4, no. 3, pp. 379–440, Sep. 2012, doi: 10.1364/AOP.4.000379.

[104]   M. G. Moharam and T. K. Gaylord, 'Rigorous coupled-wave analysis of planar-grating diffraction', *J. Opt. Soc. Am., JOSA*, vol. 71, no. 7, pp. 811–818, Jul. 1981, doi: 10.1364/JOSA.71.000811.

[105]   Y. Ding and R. Magnusson, 'Band gaps and leaky-wave effects in resonant photonic-crystal waveguides', *Opt. Express, OE*, vol. 15, no. 2, pp. 680–694, Jan. 2007, doi: 10.1364/OE.15.000680.

[106]   M. Shokooh-Saremi and R. Magnusson, 'Wideband leaky-mode resonance reflectors: Influence of grating profile and sublayers', *Optics express*, vol. 16, pp. 18249–63, Nov. 2008, doi: 10.1364/OE.16.018249.

[107]   I. D. Block *et al.*, 'A detection instrument for enhanced-fluorescence and label-free imaging on photonic crystal surfaces', *Opt. Express, OE*, vol. 17, no. 15, pp. 13222–13235, Jul. 2009, doi: 10.1364/OE.17.013222.

[108]   yuke ding and R. Magnusson, 'Resonant leaky-mode spectral-band engineering and device applications', *Optics express*, vol. 12, pp. 5661–74, Dec. 2004, doi: 10.1364/OPEX.12.005661.

[109] C. Clark, 'Effects of Configuration Interaction on Intensities and Phase Shifts', 2001, pp. 116–119.

[110] S. Thorpe, 'Electrochemical Guided Mode Resonance Biosensors', phd, University of York, 2019. Accessed: Mar. 13, 2023. [Online]. Available: https://etheses.whiterose.ac.uk/26224/

[111] Y. Wang, C. P. Reardon, N. Read, G. J. Triggs, and T. F. Krauss, 'Guided mode resonance imaging — A novel sensing technique to study bacterial biofilm antibiotic resistance', in *2017 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*, Jun. 2017, pp. 1–1. doi: 10.1109/CLEOE-EQEC.2017.8087191.

[112] G. J. Triggs, Y. Wang, C. P. Reardon, M. Fischer, G. J. O. Evans, and T. F. Krauss, 'Chirped guided-mode resonance biosensor', *Optica, OPTICA*, vol. 4, no. 2, pp. 229–234, Feb. 2017, doi: 10.1364/OPTICA.4.000229.

[113] J. S. Male, M. Simmons, C. P. Reardon, and T. F. Krauss, 'Increased Biological Sensitivity and Throughput through Application of Lateral Flow Device Integrated Microheaters'. Aug. 2018.

[114] J. S. Male, L. Grant, M. Simmons, C. P. Reardon, and T. F. Krauss, 'Thermally-Induced Pre-Concentration of Analytes in Paper Microfluidics'. Mar. 2018.

[115] H. Benisty *et al.*, 'Optical and confinement properties of two-dimensional photonic crystals', *Journal of Lightwave Technology*, vol. 17, no. 11, pp. 2063–2077, Nov. 1999, doi: 10.1109/50.802996.

[116] D. Brunner, A. Marandi, W. Bogaerts, and A. Ozcan, 'Photonics for computing and computing for photonics', *Nanophotonics*, vol. 9, no. 13, pp. 4053–4054, Sep. 2020, doi: 10.1515/nanoph-2020-0470.

[117] J. Leuthold *et al.*, 'Signal Processing with Silicon-Organic Hybrid Waveguides', in *Integrated Photonics Research, Silicon and Nanophotonics and Photonics in Switching*, Monterey, California: OSA, 2010, p. IWC1. doi: 10.1364/IPRSN.2010.IWC1.

[118] C. Koos *et al.*, 'All-optical high-speed signal processing with silicon–organic hybrid slot waveguides', *Nature Photon*, vol. 3, no. 4, pp. 216–219, Apr. 2009, doi: 10.1038/nphoton.2009.25.

[119] K. P. Zetie, S. F. Adams, and R. M. Tocknell, 'How does a Mach-Zehnder interferometer work?', *Phys. Educ.*, vol. 35, no. 1, p. 46, Jan. 2000, doi: 10.1088/0031-9120/35/1/308.

[120] M. Streshinsky *et al.*, 'Low power 50 Gb/s silicon traveling wave Mach-Zehnder modulator near 1300 nm', *Opt. Express, OE*, vol. 21, no. 25, pp. 30350–30357, Dec. 2013, doi: 10.1364/OE.21.030350.

[121] J. Fujita, M. Levy, R. M. Osgood, L. Wilkens, and H. Dötsch, 'Waveguide optical isolator based on Mach–Zehnder interferometer', *Appl. Phys. Lett.*, vol. 76, no. 16, pp. 2158–2160, Apr. 2000, doi: 10.1063/1.126284.

[122] H. Saghaei, P. Elyasi, and R. Karimzadeh, 'Design, fabrication, and characterization of Mach–Zehnder interferometers', *Photonics and Nanostructures - Fundamentals and Applications*, vol. 37, p. 100733, Dec. 2019, doi: 10.1016/j.photonics.2019.100733.

[123] J. H. Lambert and E. Anding, *Lamberts Photometrie. (Photometria, sive De mensura et gradibus luminis, colorum et umbrae) (1760)*. in Ostwalds Klassiker der exakten Wissenschaften. Leipzig: W. Engelmann, 1892.

[124] B. J. Baliga, *Advanced Power MOSFET Concepts*. Springer Science & Business Media, 2010.

[125] J. R. Brews, 'A charge-sheet model of the MOSFET', *Solid-State Electronics*, vol. 21, no. 2, pp. 345–355, Feb. 1978, doi: 10.1016/0038-1101(78)90264-2.

[126] A. Ortiz-Conde, F. J. García Sánchez, J. J. Liou, A. Cerdeira, M. Estrada, and Y. Yue, 'A review of recent MOSFET threshold voltage extraction methods', *Microelectronics Reliability*, vol. 42, no. 4, pp. 583–596, Apr. 2002, doi: 10.1016/S0026-2714(02)00027-6.

[127] T. Sakurai and A. R. Newton, 'A simple MOSFET model for circuit analysis', *IEEE Transactions on Electron Devices*, vol. 38, no. 4, pp. 887–894, Apr. 1991, doi: 10.1109/16.75219.

[128] C. H. Wann, K. Noda, T. Tanaka, M. Yoshida, and C. Hu, 'A comparative study of advanced MOSFET concepts', *IEEE Transactions on Electron Devices*, vol. 43, no. 10, pp. 1742–1753, Oct. 1996, doi: 10.1109/16.536820.

[129] J. R. Janesick, T. Elliott, S. Collins, M. M. Blouke, and J. Freeman, 'Scientific Charge-Coupled Devices', *OE*, vol. 26, no. 8, pp. 692–714, Aug. 1987, doi: 10.1117/12.7974139.

[130] G. F. Amelio, M. F. Tompsett, and G. E. Smith, 'Experimental Verification of the Charge Coupled Device Concept', *Bell System Technical Journal*, vol. 49, no. 4, pp. 593–600, 1970, doi: 10.1002/j.1538-7305.1970.tb01791.x.

[131] S. M. Gruner, M. W. Tate, and E. F. Eikenberry, 'Charge-coupled device area x-ray detectors', *Review of Scientific Instruments*, vol. 73, no. 8, pp. 2815–2842, Jul. 2002, doi: 10.1063/1.1488674.

[132] Y. Hiraoka, J. W. Sedat, and D. A. Agard, 'The Use of a Charge-Coupled Device for Quantitative Optical Microscopy of Biological Structures', *Science*, vol. 238, no. 4823, pp. 36–41, Oct. 1987, doi: 10.1126/science.3116667.

[133] A. Forouzmand and H. Mosallaei, 'Electro-optical Amplitude and Phase Modulators Based on Tunable Guided-Mode Resonance Effect', *ACS Photonics*, vol. 6, no. 11, pp. 2860–2869, Nov. 2019, doi: 10.1021/acsphotonics.9b00950.

[134] X. Qiu, X. Ruan, Y. Li, and F. Zhang, 'Multi-layer MOS capacitor based polarization insensitive electro-optic intensity modulator', *Opt. Express*, vol. 26, no. 11, p. 13902, May 2018, doi: 10.1364/OE.26.013902.

[135] 'Metal-Insulator-Semiconductor Capacitors', in *Physics of Semiconductor Devices*, John Wiley & Sons, Ltd, 2006, pp. 197–240. doi: 10.1002/9780470068328.ch4.

[136] S. M. Sze, *Semiconductor devices: physics and technology / S.M. Sze.*, 2nd ed. New York: Wiley, 2002.

[137] G. Giusti, 'Deposition and characterisation of functional ITO thin films', Jul. 2011. Accessed: Apr. 21, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Deposition-and-characterisation-of-functional-ITO-Giusti/f9c378be5bb251f9b39792c02a1b741c076828fc

[138] W. D. Greason, 'Electrostatic discharge: a charge driven phenomenon', *Journal of Electrostatics - J ELECTROSTAT*, vol. 28, pp. 199–218, Sep. 1992, doi: 10.1016/0304-3886(92)90073-3.

[139] K. Aoki *et al.*, 'Single-drive X-cut thin-sheet LiNbO/sub 3/ optical modulator with chirp adjusted using asymmetric CPW electrode', *Journal of Lightwave Technology*, vol. 24, no. 5, pp. 2233–2237, May 2006, doi: 10.1109/JLT.2006.872303.

[140] P. Rabiei and W. H. Steier, 'Lithium niobate ridge waveguides and modulators fabricated using smart guide', *Applied Physics Letters*, vol. 86, no. 16, p. 161115, Apr. 2005, doi: 10.1063/1.1906311.

[141] F. Lucchi *et al.*, 'Very low voltage single drive domain inverted $LiNbO_3$ integrated electro-optic modulator', *Opt. Express, OE*, vol. 15, no. 17, pp. 10739–10743, Aug. 2007, doi: 10.1364/OE.15.010739.

[142] N. Anwar, S. S. A. Obayya, S. Haxha, C. Thernistos, B. M. A. Rahman, and K. T. V. Grattan, 'The effect of fabrication parameters on a ridge Mach-Zehnder interferometric (MZI) modulator', *Journal of Lightwave Technology*, vol. 20, no. 5, pp. 854–861, May 2002, doi: 10.1109/JLT.2002.1007940.

[143] D. E. Aspnes and A. A. Studna, 'Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV', *Phys. Rev. B*, vol. 27, no. 2, pp. 985–1009, Jan. 1983, doi: 10.1103/PhysRevB.27.985.

[144] H.-J. Hagemann, W. Gudat, and C. Kunz, 'Optical constants from the far infrared to the x-ray region: Mg, Al, Cu, Ag, Au, Bi, C, and Al_2O_3', *J. Opt. Soc. Am., JOSA*, vol. 65, no. 6, pp. 742–744, Jun. 1975, doi: 10.1364/JOSA.65.000742.

[145] K. Luke, Y. Okawachi, M. R. E. Lamont, A. L. Gaeta, and M. Lipson, 'Broadband mid-infrared frequency comb generation in a Si_3N_4 microresonator', *Opt. Lett., OL*, vol. 40, no. 21, pp. 4823–4826, Nov. 2015, doi: 10.1364/OL.40.004823.

[146] G. J. Triggs, M. Fischer, D. Stellinga, M. G. Scullion, G. J. O. Evans, and T. F. Krauss, 'Spatial Resolution and Refractive Index Contrast of Resonant Photonic Crystal Surfaces for Biosensing', *IEEE Photonics Journal*, vol. 7, no. 3, pp. 1–10, Jun. 2015, doi: 10.1109/JPHOT.2015.2435699.

[147] A. Sarangan, *Nanofabrication: Principles to Laboratory Practice*. Boca Raton: CRC Press, 2016. doi: 10.1201/9781315370514.

[148] A. del Campo and C. Greiner, 'SU-8: a photoresist for high-aspect-ratio and 3D submicron lithography', *J. Micromech. Microeng.*, vol. 17, no. 6, p. R81, May 2007, doi: 10.1088/0960-1317/17/6/R01.

[149] B. Hammer and J. K. Norskov, 'Why gold is the noblest of all the metals', *Nature*, vol. 376, no. 6537, Art. no. 6537, Jul. 1995, doi: 10.1038/376238a0.

[150] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*. Artech House, 2005.

[151] G. Veronis, R. W. Dutton, and S. Fan, 'Method for sensitivity analysis of photonic crystal devices', *Opt. Lett., OL*, vol. 29, no. 19, pp. 2288–2290, Oct. 2004, doi: 10.1364/OL.29.002288.

[152] 'The Finite Element Method in Electromagnetics, 3rd Edition | Wiley', Wiley.com. Accessed: Apr. 01, 2023. [Online]. Available: https://www.wiley.com/en-gb/The+Finite+Element+Method+in+Electromagnetics%2C+3rd+Edition-p-9781118571361

[153] L. Li, 'New formulation of the Fourier modal method for crossed surface-relief gratings', *J. Opt. Soc. Am. A, JOSAA*, vol. 14, no. 10, pp. 2758–2767, Oct. 1997, doi: 10.1364/JOSAA.14.002758.

[154] R. Bräuer and O. Bryngdahl, 'Electromagnetic diffraction analysis of two-dimensional gratings', *Optics Communications*, vol. 100, no. 1, pp. 1–5, Jul. 1993, doi: 10.1016/0030-4018(93)90546-H.

[155] F. G. Kaspar, 'Diffraction by thick, periodically stratified gratings with complex dielectric constant', *J. Opt. Soc. Am., JOSA*, vol. 63, no. 1, pp. 37–45, Jan. 1973, doi: 10.1364/JOSA.63.000037.

[156] M. Liscidini, D. Gerace, L. C. Andreani, and J. E. Sipe, 'Scattering-matrix analysis of periodically patterned multilayers with asymmetric unit cells and birefringent media', *Phys. Rev. B*, vol. 77, no. 3, p. 035324, Jan. 2008, doi: 10.1103/PhysRevB.77.035324.

[157] D. M. Whittaker and I. S. Culshaw, 'Scattering-matrix treatment of patterned multilayer photonic structures', *Phys. Rev. B*, vol. 60, no. 4, pp. 2610–2618, Jul. 1999, doi: 10.1103/PhysRevB.60.2610.

[158] G. Yoon and J. Rho, 'MAXIM: Metasurfaces-oriented electromagnetic wave simulation software with intuitive graphical user interfaces', *Computer Physics Communications*, vol. 264, p. 107846, Jul. 2021, doi: 10.1016/j.cpc.2021.107846.

[159] 'Classical Electrodynamics, 3rd Edition | Wiley', Wiley.com. Accessed: Apr. 01, 2023. [Online]. Available: https://www.wiley.com/en-us/Classical+Electrodynamics%2C+3rd+Edition-p-9780471309321

[160] M. G. Moharam, D. A. Pommet, E. B. Grann, and T. K. Gaylord, 'Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach', *J. Opt. Soc. Am. A, JOSAA*, vol. 12, no. 5, pp. 1077–1086, May 1995, doi: 10.1364/JOSAA.12.001077.

[161] 'Ansys Lumerical FDTD | Simulation for Photonic Components'. Accessed: Apr. 29, 2023. [Online]. Available: https://www.ansys.com/products/photonics/fdtd

[162] 'COMSOL: Multiphysics Software for Optimizing Designs', COMSOL. Accessed: May 18, 2023. [Online]. Available: https://www.comsol.com/

[163] K. Yee, 'Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media', *IEEE Transactions on Antennas and Propagation*, vol. 14, no. 3, pp. 302–307, May 1966, doi: 10.1109/TAP.1966.1138693.

[164] W. Yu and R. Mittra, 'A conformal finite difference time domain technique for modeling curved dielectric surfaces', *IEEE Microwave and Wireless Components Letters*, vol. 11, no. 1, pp. 25–27, Jan. 2001, doi: 10.1109/7260.905957.

[165] A. Taflove and M. E. Brodwin, 'Numerical Solution of Steady-State Electromagnetic Scattering Problems Using the Time-Dependent Maxwell's Equations', *IEEE Transactions on Microwave Theory and Techniques*, vol. 23, no. 8, pp. 623–630, Aug. 1975, doi: 10.1109/TMTT.1975.1128640.

[166] 'Handbook of Optical Constants of Solids - 1st Edition'. Accessed: May 02, 2023. [Online]. Available: https://www.elsevier.com/books/handbook-of-optical-constants-of-solids/palik/978-0-08-054721-3

[167] T. A. F. König *et al.*, 'Electrically Tunable Plasmonic Behavior of Nanocube–Polymer Nanomaterials Induced by a Redox-Active Electrochromic Polymer', *ACS Nano*, vol. 8, no. 6, pp. 6182–6192, Jun. 2014, doi: 10.1021/nn501601e.

[168] R. Jenkins, R. Devereux, and J. Heaton, 'Waveguide beam splitters and recombiners based on multimode propagation phenomena', *Optics letters*, vol. 17, pp. 991–3, Jul. 1992, doi: 10.1364/OL.17.000991.

[169] G. Singh, A. Sirohi, and S. Verma, 'Estimation of the performance of a 3-dB Y-junction optical coupler with a channel profile of proton-exchanged lithium niobate', *Physics of Wave Phenomena*, vol. 21, pp. 201–206, Jul. 2013, doi: 10.3103/S1541308X13030059.

[170] A. Melloni, P. Monguzzi, R. Costa, and M. Martinelli, 'Design of curved waveguides: The matched bend', *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 20, pp. 130–7, Feb. 2003, doi: 10.1364/JOSAA.20.000130.

[171] T. W. H. Ward, 'Conductive coating for glass and method of application', US2280135A, Apr. 21, 1942 Accessed: Apr. 21, 2023. [Online]. Available: https://patents.google.com/patent/US2280135/en

[172] H. Köstlin, R. Jost, and W. Lems, 'Optical and electrical properties of doped In2O3 films', *Phys. Stat. Sol. (a)*, vol. 29, no. 1, pp. 87–93, May 1975, doi: 10.1002/pssa.2210290110.

[173] D. L. White and M. Feldman, 'Liquid-crystal light valves', *Electronics Letters*, vol. 6, no. 26, pp. 837–839, Dec. 1970, doi: 10.1049/el:19700578.

[174] U. Betz, M. Olsson, J. Marthy, M. F. Escola, and F. Atamny, 'Thin Film Engineering of Indium Tin Oxide: Large area flat panel displays application', *Surface & Coatings Technology*, vol. 200, no. 20–21, pp. 5751–5759, 2006, doi: 10.1016/j.surfcoat.2005.08.144.

[175] M. Marezio, 'Refinement of the crystal structure of In2O3 at two wavelengths', *Acta Cryst*, vol. 20, no. 6, Art. no. 6, Jun. 1966, doi: 10.1107/S0365110X66001749.

[176] I. Rauf and J. Yuan, 'Effects of Microstructure on the Optical Properties of Tin-Doped Indium Oxide Thin Films Studies by Electron Energy Loss Spectroscopy', *Materials Letters - MATER LETT*, vol. 25, pp. 217–222, Dec. 1995, doi: 10.1016/0167-577X(95)00189-1.

[177] J. C. C. Fan and J. B. Goodenough, 'X-ray photoemission spectroscopy studies of Sn-doped indium-oxide films', *Journal of Applied Physics*, vol. 48, no. 8, pp. 3524–3531, Aug. 2008, doi: 10.1063/1.324149.

[178] N. F. Mott and L. Friedman, 'Metal-insulator transitions in VO$_2$, Ti$_2$O$_3$ and Ti$_{2-x}$V$_x$O$_3$', *Philosophical Magazine*, vol. 30, no. 2, pp. 389–402, Aug. 1974, doi: 10.1080/14786439808206565.

[179] D. Slocombe, A. Porch, M. Pepper, and P. P. Edwards, 'The Mott transition and optimal performance of transparent conducting oxides in thin-film solar cells', *Energy Environ. Sci.*, vol. 5, no. 1, pp. 5387–5391, Jan. 2012, doi: 10.1039/C1EE02585F.

[180] M. Bender, W. Seelig, C. Daube, H. Frankenberger, B. Ocker, and J. Stollenwerk, 'Dependence of oxygen flow on optical and electrical properties of DC-magnetron sputtered ITO films', *Thin Solid Films*, vol. 326, no. 1, pp. 72–77, Aug. 1998, doi: 10.1016/S0040-6090(98)00521-5.

[181] K. Suzuki, N. Hashimoto, T. Oyama, J. Shimizu, Y. Akao, and H. Kojima, 'Large scale and low resistance ITO films formed at high deposition rates', *Thin Solid Films*, vol. 226, no. 1, pp. 104–109, Apr. 1993, doi: 10.1016/0040-6090(93)90213-9.

[182] A. B. Chebotareva, G. G. Untila, T. N. Kost, S. Jorgensen, and A. G. Ulyashin, 'ITO deposited by pyrosol for photovoltaic applications', *Thin Solid Films*, vol. 515, no. 24, pp. 8505–8510, Oct. 2007, doi: 10.1016/j.tsf.2007.03.097.

[183] D. U. Yildirim *et al.*, 'Disordered and Densely Packed ITO Nanorods as an Excellent Lithography-Free Optical Solar Reflector Metasurface', *ACS Photonics*, vol. 6, no. 7, pp. 1812–1822, Jul. 2019, doi: 10.1021/acsphotonics.9b00636.

[184] '2022 Index IEEE Photonics Journal Vol. 14', *IEEE Photonics Journal*, vol. 14, no. 6, pp. 1–142, Dec. 2022, doi: 10.1109/JPHOT.2023.3234505.

[185] M. Shak Sadi and E. Kumpikaitė, 'Advances in the Robustness of Wearable Electronic Textiles: Strategies, Stability, Washability and Perspective', *Nanomaterials (Basel)*, vol. 12, no. 12, p. 2039, Jun. 2022, doi: 10.3390/nano12122039.

[186] J. Park, J.-H. Kang, X. Liu, and M. L. Brongersma, 'Electrically Tunable Epsilon-Near-Zero (ENZ) Metafilm Absorbers', *Sci Rep*, vol. 5, p. 15754, Nov. 2015, doi: 10.1038/srep15754.

[187] V. J. Sorger, N. D. Lanzillotti-Kimura, R.-M. Ma, and X. Zhang, 'Ultra-compact silicon nanophotonic modulator with broadband response', *Nanophotonics*, vol. 1, no. 1, pp. 17–22, Jul. 2012, doi: 10.1515/nanoph-2012-0009.

[188] Y.-W. Huang *et al.*, 'Gate-Tunable Conducting Oxide Metasurfaces', *Nano Lett.*, vol. 16, no. 9, pp. 5319–5325, Sep. 2016, doi: 10.1021/acs.nanolett.6b00555.

[189] F. Wooten, 'Chapter 3 - ABSORPTION AND DISPERSION', in *Optical Properties of Solids*, F. Wooten, Ed., Academic Press, 1972, pp. 42–84. doi: 10.1016/B978-0-12-763450-0.50008-8.

[190] A. Kriesch, S. P. Burgos, D. Ploss, H. Pfeifer, H. A. Atwater, and U. Peschel, 'Functional Plasmonic Nanocircuits with Low Insertion and Propagation Losses', *Nano Lett.*, vol. 13, no. 9, pp. 4539–4545, Sep. 2013, doi: 10.1021/nl402580c.

[191] N. M. Ahmed, F. A. Sabah, H. I. Abdulgafour, A. Alsadig, A. Sulieman, and M. Alkhoaryef, 'The effect of post annealing temperature on grain size of indium-tin-oxide for optical and electrical properties improvement', *Results in Physics*, vol. 13, p. 102159, Jun. 2019, doi: 10.1016/j.rinp.2019.102159.

[192] C. Nunes de Carvalho, A. M. Botelho do Rego, A. Amaral, P. Brogueira, and G. Lavareda, 'Effect of substrate temperature on the surface structure, composition and morphology of indium–tin oxide films', *Surface and Coatings Technology*, vol. 124, no. 1, pp. 70–75, Feb. 2000, doi: 10.1016/S0257-8972(99)00619-2.

[193] 'E-Beam Resist AR-P 6200 series (CSAR 62)', Allresist EN. Accessed: Dec. 11, 2022. [Online]. Available: https://www.allresist.com/portfolio-item/e-beam-resist-ar-p-6200-series-csar-62/

[194] K. I. Park, 'Probability Theory', in *Fundamentals of Probability and Stochastic Processes with Applications to Communications*, K. I. Park, Ed., Cham: Springer International Publishing, 2018, pp. 51–71. doi: 10.1007/978-3-319-68075-0_3.

[195] R. A John, *Mathematical statistics and data analysis*, 3rd ed. Belmont, CA: Thomson/Brooks/Cole, 2007.

[196] 'DektakXT - Overview | Bruker Stylus Profiler | Bruker'. Accessed: Aug. 05, 2019. [Online]. Available: https://www.bruker.com/products/surface-and-dimensional-analysis/stylus-profilometers/dektak-xt/overview.html

[197] D.-H. Lee, '3-Dimensional profile distortion measured by stylus type surface profilometer', *Measurement*, vol. 46, pp. 803–814, Jan. 2013, doi: 10.1016/j.measurement.2012.09.022.

[198] A. Charnes, E. L. Frome, and P. L. Yu, 'The Equivalence of Generalized Least Squares and Maximum Likelihood Estimates in the Exponential Family', *Journal of the American Statistical Association*, vol. 71, no. 353, pp. 169–171, Mar. 1976, doi: 10.1080/01621459.1976.10481508.

[199] 'jm1261/SurfaceProfileAnalysis: Step height finder for AFM/Dektak', GitHub. Accessed: Jan. 15, 2023. [Online]. Available: https://github.com/jm1261/SurfaceProfileAnalysis

[200] I. Manyamwa, B. Mburunge, A. Lukamba, and P. Lambin, 'Loi de Pouillet généralisée et applications', *Cahiers du Ceruki : nouvelle série*, vol. 64, pp. 351–364, 2021.

[201] F. M. Smits, 'Measurement of sheet resistivities with the four-point probe', *The Bell System Technical Journal*, vol. 37, no. 3, pp. 711–718, May 1958, doi: 10.1002/j.1538-7305.1958.tb03883.x.

[202] L. Valdes, 'Resistivity Measurements on Germanium for Transistors', *Proc. IRE*, vol. 42, no. 2, pp. 420–427, Feb. 1954, doi: 10.1109/JRPROC.1954.274680.

[203] I. Miccoli, F. Edler, H. Pfnür, and C. Tegenkamp, 'The 100th anniversary of the four-point probe technique: the role of probe geometries in isotropic and anisotropic systems', *J. Phys.: Condens. Matter*, vol. 27, no. 22, p. 223201, May 2015, doi: 10.1088/0953-8984/27/22/223201.

[204] K. Hauffe, 'The Hall Effect and Related Phenomena. Von E. H. Putley, Semi-Conductor Monography, Butterworths, London 1960. 263 Seiten. Preis: 50 s.', *Zeitschrift für Elektrochemie, Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 65, no. 4, pp. 398–398, 1961, doi: 10.1002/bbpc.19610650426.

[205] R. Chwang, B. J. Smith, and C. R. Crowell, 'Contact size effects on the van der Pauw method for resistivity and Hall coefficient measurement', *Solid-State Electronics*, vol. 17, no. 12, pp. 1217–1227, Dec. 1974, doi: 10.1016/0038-1101(74)90001-X.

[206] C. F. Holder and R. E. Schaak, 'Tutorial on Powder X-ray Diffraction for Characterizing Nanoscale Materials', *ACS Nano*, vol. 13, no. 7, pp. 7359–7365, Jul. 2019, doi: 10.1021/acsnano.9b05157.

[207] Y. J. Kim, S. B. Jin, S. I. Kim, Y. S. Choi, I. S. Choi, and J. G. Han, 'Study on the electrical properties of ITO films deposited by facing target sputter deposition', *J. Phys. D: Appl. Phys.*, vol. 42, no. 7, p. 075412, Mar. 2009, doi: 10.1088/0022-3727/42/7/075412.

[208] S. Złotnik *et al.*, 'Experimental Verification of Single-Type Electron Population in Indium Tin Oxide Layers', *physica status solidi (RRL) – Rapid Research Letters*, vol. 16, no. 8, p. 2200170, 2022, doi: 10.1002/pssr.202200170.

[209] C.-L. Tien, H.-Y. Lin, C.-K. Chang, and C.-J. Tang, 'Effect of Oxygen Flow Rate on the Optical, Electrical, and Mechanical Properties of DC Sputtering ITO Thin Films', *Advances in Condensed Matter Physics*, vol. 2018, p. e2647282, Feb. 2018, doi: 10.1155/2018/2647282.

[210] S. Song, T. Yang, J. Liu, Y. Xin, Y. Li, and S. Han, 'Rapid Thermal Annealing of ITO Films', *Applied Surface Science - APPL SURF SCI*, vol. 257, pp. 7061–7064, Jun. 2011, doi: 10.1016/j.apsusc.2011.03.009.

[211] X. Chai *et al.*, 'Modulation of photoelectric properties of indium tin oxide thin films via oxygen control, and its application to epsilon-near-zero properties for an infrared absorber', *Journal of Applied Physics*, vol. 128, no. 18, p. 185301, Nov. 2020, doi: 10.1063/5.0025997.

[212] D. C. Paine, T. Whitson, D. Janiac, R. Beresford, C. O. Yang, and B. Lewis, 'A study of low temperature crystallization of amorphous thin film indium–tin–oxide', *Journal of Applied Physics*, vol. 85, no. 12, pp. 8445–8450, Jun. 1999, doi: 10.1063/1.370695.

[213] Y. Hu, X. Diao, C. Wang, W. Hao, and T. Wang, 'Effects of Heat Treatment on Properties of ITO Films Prepared by RF Magnetron Sputtering', *Vacuum*, vol. 75, pp. 183–188, Jul. 2004, doi: 10.1016/j.vacuum.2004.01.081.

[214] F. J. Giessibl, 'Advances in atomic force microscopy', *Rev. Mod. Phys.*, vol. 75, no. 3, pp. 949–983, Jul. 2003, doi: 10.1103/RevModPhys.75.949.

[215] 'Quadrature Error'. Accessed: Jan. 15, 2023. [Online]. Available: https://faraday.physics.utoronto.ca/PVB/Harrison/ErrorAnalysis/Propagation.html

[216] 'Principles of SEM', in *Principles and Practice of Variable Pressure/Environmental Scanning Electron Microscopy (VP-ESEM)*, John Wiley & Sons, Ltd, 2008, pp. 17–62. doi: 10.1002/9780470758731.ch2.

[217] 'jm1261/MicroscopyPeriodAnalysis: Grating Image Analysis'. Accessed: Jan. 15, 2023. [Online]. Available: https://github.com/jm1261/MicroscopyPeriodAnalysis

[218] M. F. Limonov, M. V. Rybin, A. N. Poddubny, and Y. S. Kivshar, 'Fano resonances in photonics', *Nature Photon*, vol. 11, no. 9, Art. no. 9, Sep. 2017, doi: 10.1038/nphoton.2017.142.

[219] J. Male, 'Resonant Wavelength Analysis'. Jan. 23, 2023. Accessed: Apr. 20, 2023. [Online]. Available: https://github.com/jm1261/PeakFinder

[220] H. H. Rosenbrock, 'An Automatic Method for Finding the Greatest or Least Value of a Function', *The Computer Journal*, vol. 3, no. 3, pp. 175–184, Jan. 1960, doi: 10.1093/comjnl/3.3.175.

[221] 'Practical Methods of Optimization, 2nd Edition | Wiley', Wiley.com. Accessed: Apr. 20, 2023. [Online]. Available: https://www.wiley.com/en-gb/Practical+Methods+of+Optimization%2C+2nd+Edition-p-9780471494638

[222] J. Davies and K. Binmore, *Calculus, concepts and methods*. Cambridge, UK: Cambridge University Press, 2001. Accessed: Apr. 20, 2023. [Online]. Available: http://www.cambridge.org/catalogue/catalogue.asp?isbn=0521775418

[223] Y. Maeda, J. Kobashi, H. Yoshida, and M. Ozaki, 'Diffusion-based liquid crystal substitution for the improvement of electro-optic properties in polymer/cholesteric liquid crystal composites', *Opt. Mater. Express, OME*, vol. 7, no. 1, pp. 85–92, Jan. 2017, doi: 10.1364/OME.7.000085.

[224] G. Hulley, E. Pavlis, and V. Mendes, 'Validation of improved atmospheric refraction models for Satellite Laser Ranging (SLR)', vol. 130, Jan. 2007, doi: 10.1007/978-3-540-49350-1_119.

[225] J. Male, 'jm1261/S4GratingOptimizer'. Dec. 05, 2022. Accessed: Jan. 15, 2023. [Online]. Available: https://github.com/jm1261/S4GratingOptimizer

[226] 'jm1261/DrudeModulators: Code for calculating optical properties of Drude-Model governed transparent conductive oxide-controlled optical modulators.', GitHub. Accessed: Jan. 15, 2023. [Online]. Available: https://github.com/jm1261/DrudeModulators

[227] J. A. Woollam, W. A. McGaham, and B. Johs, 'Spectroscopic ellipsometry studies of indium tin oxide and other flat panel display multilayer materials', *Thin Solid Films*, vol. 241, no. 1, pp. 44–46, Apr. 1994, doi: 10.1016/0040-6090(94)90393-X.

[228] R. A. Synowicki, 'Spectroscopic ellipsometry characterization of indium tin oxide film microstructure and optical constants', *Thin Solid Films*, vol. 313–314, pp. 394–397, Feb. 1998, doi: 10.1016/S0040-6090(97)00853-5.

[229] C. Bohórquez, H. Bakkali, J. J. Delgado, E. Blanco, M. Herrera, and M. Domínguez, 'Spectroscopic Ellipsometry Study on Tuning the Electrical and Optical Properties of Zr-Doped ZnO Thin Films Grown by Atomic Layer Deposition', *ACS Appl. Electron. Mater.*, vol. 4, no. 3, pp. 925–935, Mar. 2022, doi: 10.1021/acsaelm.1c01026.

[230] S.-H. Jeong, B.-S. Kim, and B.-T. Lee, 'Photoluminescence dependence of ZnO films grown on Si(100) by radio-frequency magnetron sputtering on the growth ambient', *Applied Physics Letters*, vol. 82, p. 2625, Apr. 2003, doi: 10.1063/1.1568543.

[231] J. Kim *et al.*, 'Optical Properties of Gallium-Doped Zinc Oxide---A Low-Loss Plasmonic Material: First-Principles Theory and Experiment', *Phys. Rev. X*, vol. 3, no. 4, p. 041037, Dec. 2013, doi: 10.1103/PhysRevX.3.041037.

[232] 'Photoresist', Semiconductor Engineering. Accessed: Dec. 11, 2022. [Online]. Available: https://semiengineering.com/knowledge_centers/manufacturing/lithography/photoresist/

[233] 'MICROPOSIT^{TM} S1800^{TM} G2 series – Microresist'. Accessed: Dec. 11, 2022. [Online]. Available: https://www.microresist.de/en/produkt/microposit-s1800-g2-series/

[234] N. Pham, J. Burghartz, and P. Sarro, 'Spray coating of photoresist for pattern transfer on high topography surfaces', *Journal of Micromechanics and Microengineering*, vol. 15, no. 4, pp. 691–697, 2005, doi: 10.1088/0960-1317/15/4/003.

[235] L. Yu, Y. Y. Lee, F. E. H. Tay, and C. Iliescu, 'Spray Coating of Photoresist for 3D Microstructures with Different Geometries', *J. Phys.: Conf. Ser.*, vol. 34, no. 1, p. 937, Apr. 2006, doi: 10.1088/1742-6596/34/1/155.

[236] Y. Chen, 'Nanofabrication by electron beam lithography and its applications: A review', *Microelectronic Engineering*, vol. 135, pp. 57–72, Mar. 2015, doi: 10.1016/j.mee.2015.02.042.

[237] adminNEA, 'Dilase 650 - Multifunction maskless lithography equipment', KLOE. Accessed: Jan. 30, 2023. [Online]. Available: https://www.kloe-france.com/en/photolithography-equipment/direct-laser-writing/direct-laser-writer-dilase-650

[238] C. J. R. Sheppard, 'MICROSCOPY | Overview', in *Encyclopedia of Modern Optics*, R. D. Guenther, Ed., Oxford: Elsevier, 2005, pp. 61–69. doi: 10.1016/B0-12-369395-0/00823-X.

[239] 'Field Guide to Geometrical Optics | (2004) | Greivenkamp | Publications | Spie'. Accessed: Mar. 02, 2023. [Online]. Available: https://spie.org/Publications/Book/547461?SSO=1

[240] 'Proximity effect in electron-beam lithography: Journal of Vacuum Science and Technology: Vol 12, No 6'. Accessed: Mar. 02, 2023. [Online]. Available: https://avs.scitation.org/doi/abs/10.1116/1.568515

[241] G. Owen and P. Rissman, 'Proximity effect correction for electron beam lithography by equalization of background dose', *Journal of Applied Physics*, vol. 54, no. 6, pp. 3573–3581, Jun. 1983, doi: 10.1063/1.332426.

[242] R. DellaGuardia, W.-S. Huang, K. R. Chen, and D. Kang, 'Correlations between dissolution data and lithography of various resists', in *Advances in Resist Technology and Processing XVI*, SPIE, Jun. 1999, pp. 316–328. doi: 10.1117/12.350215.

[243] K. Y. Lee *et al.*, 'Micromachining applications of a high resolution ultrathick photoresist', *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 13, no. 6, pp. 3012–3016, Nov. 1995, doi: 10.1116/1.588297.

[244] S. D'Elia, N. Scaramuzza, F. Ciuchi, C. Versace, G. Strangi, and R. Bartolino., 'Ellipsometry investigation of the effects of annealing temperature on the optical properties of indium tin oxide thin films studied by Drude–Lorentz model', *Applied Surface Science*, vol. 255, no. 16, pp. 7203–7211, May 2009, doi: 10.1016/j.apsusc.2009.03.064.

[245] P. J. Kelly and R. D. Arnell, 'Magnetron sputtering: a review of recent developments and applications', *Vacuum*, vol. 56, no. 3, pp. 159–172, Mar. 2000, doi: 10.1016/S0042-207X(99)00189-X.

[246] I. Safi, 'Recent aspects concerning DC reactive magnetron sputtering of thin films: a review', *Surface and Coatings Technology*, vol. 127, no. 2, pp. 203–218, May 2000, doi: 10.1016/S0257-8972(00)00566-1.

[247] Y. Yamamura and H. Tawara, 'ENERGY DEPENDENCE OF ION-INDUCED SPUTTERING YIELDS FROM MONATOMIC SOLIDS AT NORMAL INCIDENCE', *Atomic Data and Nuclear Data Tables*, vol. 62, no. 2, pp. 149–253, Mar. 1996, doi: 10.1006/adnd.1996.0005.

[248] P. Sigmund, 'Theory of Sputtering. I. Sputtering Yield of Amorphous and Polycrystalline Targets', *Phys. Rev.*, vol. 184, no. 2, pp. 383–416, Aug. 1969, doi: 10.1103/PhysRev.184.383.

[249] J. Bohdansky, 'A universal relation for the sputtering yield of monatomic solids at normal ion incidence', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 2, no. 1, pp. 587–591, Mar. 1984, doi: 10.1016/0168-583X(84)90271-4.

[250] A. Anders, 'Tutorial: Reactive high power impulse magnetron sputtering (R-HiPIMS)', *Journal of Applied Physics*, vol. 121, p. 171101, Mar. 2017, doi: 10.1063/1.4978350.

[251] S. Swann, 'Magnetron sputtering', *Physics in Technology*, vol. 19, no. 2, p. 67, Mar. 1988, doi: 10.1088/0305-4624/19/2/304.

[252] J.-C. Pivin, 'An overview of ion sputtering physics and practical implications', *Journal of Materials Science*, vol. 18, pp. 1267–1290, May 1983, doi: 10.1007/BF01111944.

[253] S. Xian *et al.*, 'Effect of oxygen stoichiometry on the structure, optical and epsilon-near-zero properties of indium tin oxide films', *Opt. Express, OE*, vol. 27, no. 20, pp. 28618–28628, Sep. 2019, doi: 10.1364/OE.27.028618.

[254] 'Energy distribution of O− ions during reactive magnetron sputtering: Applied Physics Letters: Vol 89, No 5'. Accessed: Feb. 11, 2023. [Online]. Available: https://aip.scitation.org/doi/10.1063/1.2266888

[255] T. Jäger, Y. Romanyuk, A. Tiwari, and A. Anders, 'Controlling ion fluxes during reactive sputter-deposition of SnO2:F', *Journal of Applied Physics*, vol. 116, pp. 033301–033301, Jul. 2014, doi: 10.1063/1.4887119.

[256] T. Welzel and K. Ellmer, 'Negative oxygen ion formation in reactive magnetron sputtering processes for transparent conductive oxides', *Journal of Vacuum Science & Technology A*, vol. 30, no. 6, p. 061306, Nov. 2012, doi: 10.1116/1.4762815.

[257] H. Kupfer and F. Richter, 'Reactive Magnetron Sputtering of Indium Tin Oxide Thin Films: The Cross-Corner and Cross-Magnetron Effect', in *Reactive Sputter Deposition*, D. Depla and S. Mahieu, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 337–366. doi: 10.1007/978-3-540-76664-3_10.

[258] D. Depla and R. Gryse, 'Target poisoning during reactive magnetron sputtering: Part II: The influence of chemisorption and gettering', *Surface & Coatings Technology - SURF COAT TECH*, vol. 183, pp. 190–195, May 2004, doi: 10.1016/j.surfcoat.2003.10.007.

[259] D. Depla and R. De Gryse, 'Target poisoning during reactive magnetron sputtering: Part I: the influence of ion implantation', *Surface and Coatings Technology*, vol. 183, no. 2, pp. 184–189, May 2004, doi: 10.1016/j.surfcoat.2003.10.006.

[260] Shive John Northrup and R. L. Weber, 'Similarities in physics / John N. Shive, Robert L. Weber'. Adam Hilger Bristol, 1982.

[261] A. Anders, 'Physics of arcing, and implications to sputter deposition', *Thin Solid Films*, vol. 502, no. 1, pp. 22–28, Apr. 2006, doi: 10.1016/j.tsf.2005.07.228.

[262] S. Karthikeyan, A. E. Hill, J. S. Cowpe, and R. D. Pilkington, 'The influence of operating parameters on pulsed D.C. magnetron sputtering plasma', *Vacuum*, vol. 85, no. 5, pp. 634–638, Nov. 2010, doi: 10.1016/j.vacuum.2010.09.007.

[263] K. Ver Heyen, E. Cielen, A. Tahri, A. Saleh, N. Boens, and G. J. Hoornaert, 'Synthesis and characterisation of new fluorescent Na+ and K+ indicators', *Tetrahedron*, vol. 55, no. 16, pp. 5207–5226, Apr. 1999, doi: 10.1016/S0040-4020(99)00167-2.

[264] 'Vacuum Coating - MBRAUN'. Accessed: Mar. 03, 2023. [Online]. Available: https://www.mbraun.com/us/products/coating/vacuum-coating.html

[265] 'Ceramic coating deposition by electron beam evaporation - ScienceDirect'. Accessed: Mar. 03, 2023. [Online]. Available:

https://www.sciencedirect.com/science/article/pii/S025789721730645X?casa_token=d
PUFHQtJaQ0AAAAA:CYmYbAhPSxk1coZTQX2mSqr0LjCvGPNxrw2e7tdcA-
efjfZCJFboRl9lyneXgKkaTXuxRiH1

[266] F. Nazeer, Z. Ma, L. Gao, F. Wang, M. A. Khan, and A. Malik, 'Thermal and mechanical properties of copper-graphite and copper-reduced graphene oxide composites', *Composites Part B: Engineering*, vol. 163, pp. 77–85, Apr. 2019, doi: 10.1016/j.compositesb.2018.11.004.

[267] B. Moller, J. Rarey, and D. Ramjugernath, 'Estimation of the vapour pressure of non-electrolyte organic compounds via group contributions and group interactions', *Journal of Molecular Liquids*, vol. 143, no. 1, pp. 52–63, Sep. 2008, doi: 10.1016/j.molliq.2008.04.020.

[268] N. Inogamov, V. Zhakhovsky, and V. Khokhlov, 'Dynamics of Gold Ablation into Water', *Journal of Experimental and Theoretical Physics*, vol. 127, pp. 79–106, Jul. 2018, doi: 10.1134/S1063776118070075.

[269] '(13) Thermodynamic Properties of Gold'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.researchgate.net/publication/292074975_Thermodynamic_Properties_of_Gold

[270] 'WebElements Periodic Table » Gold » thermochemistry and thermodynamics'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.webelements.com/gold/thermochemistry.html

[271] 'WebElements Periodic Table » Chromium » thermochemistry and thermodynamics'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.webelements.com/chromium/thermochemistry.html

[272] 'WebElements Periodic Table » Indium » thermochemistry and thermodynamics'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.webelements.com/indium/thermochemistry.html

[273] 'WebElements Periodic Table » Nickel » thermochemistry and thermodynamics'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.webelements.com/nickel/thermochemistry.html

[274] 'WebElements Periodic Table » Tin » thermochemistry and thermodynamics'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.webelements.com/tin/thermochemistry.html

[275] 'WebElements Periodic Table » Titanium » thermochemistry and thermodynamics'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.webelements.com/titanium/thermochemistry.html

[276] 'Universal energy dependence of physical and ion-enhanced chemical etch yields at low ion energy: Applied Physics Letters: Vol 55, No 19'. Accessed: Mar. 13, 2023. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.102336?casa_token=pUzoX067Q1EAAAAA: pKpt4b3_pYskJoIeSEmBJJci5XEiPq5Fs0BpSJ9Vb-krHm-- XhZ1wIzPrK04eLJDy6WUKrpAA2T6

[277] H. Li, K. Karahashi, M. Fukasawa, K. Nagahata, T. Tatsumi, and S. Hamaguchi, 'Sputtering yields and surface chemical modification of tin-doped indium oxide in hydrocarbon-based plasma etching', *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 33, p. 060606, Nov. 2015, doi: 10.1116/1.4927125.

[278] 'Sputtering Yields of Metals for Ar+ and Ne+ Ions with Energies from 50 to 600 ev: Journal of Applied Physics: Vol 32, No 3'. Accessed: Mar. 13, 2023. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.1736012?casa_token=kLvFxP_nDFAAAAAA: hc_W7g8sJOT- aNHdRkn6MOCiWSNeQNQH3NLSPG85Hc78Gp_Cl5CUD5BcFEuhFRBseizOlLEw-tYH

[279] 'Principles of Plasma Discharges and Materials Processing, 2nd Edition | Wiley'. Accessed: Mar. 13, 2023. [Online]. Available: https://www.wiley.com/en-

us/Principles+of+Plasma+Discharges+and+Materials+Processing%2C+2nd+Edition-p-9780471720010

[280] D. Manos and D. Flamm, 'Plasma etching : an introduction', 1989. Accessed: Mar. 13, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Plasma-etching-%3A-an-introduction-Manos-Flamm/a283d4068b03e47e0358e5d58167eda858452f52

[281] R. Legtenberg, H. Jansen, M. de Boer, and M. Elwenspoek, 'Anisotropic Reactive Ion Etching of Silicon Using SF 6 / O 2 / CHF 3 Gas Mixtures', *J. Electrochem. Soc.*, vol. 142, no. 6, p. 2020, Jun. 1995, doi: 10.1149/1.2044234.

[282] V. Bhalla, S. Carrara, C. Stagni, and B. Samorì, 'Chip cleaning and regeneration for electrochemical sensor arrays', *Thin Solid Films*, vol. 518, no. 12, pp. 3360–3366, Apr. 2010, doi: 10.1016/j.tsf.2009.10.022.

[283] Y. Chen, 'A lift-off process for high resolution patterns using PMMA/LOR resist stack', *Microelectronic Engineering*, vol. 73–74, pp. 278–281, Jun. 2004, doi: 10.1016/S0167-9317(04)00111-X.

[284] D. M. Mattox, 'Influence of Oxygen on the Adherence of Gold Films to Oxide Substrates', *Journal of Applied Physics*, vol. 37, no. 9, pp. 3613–3615, Aug. 1966, doi: 10.1063/1.1708913.

[285] S. Tosoni and G. Pacchioni, 'Oxide-Supported Gold Clusters and Nanoparticles in Catalysis: A Computational Chemistry Perspective', *ChemCatChem*, vol. 11, no. 1, pp. 73–89, 2019, doi: 10.1002/cctc.201801082.

[286] H. Dallaporta and A. Cros, 'Influence of low-energy electron irradiation on the adhesion of gold films on a silicon substrate', *Appl. Phys. Lett.*, vol. 48, no. 20, pp. 1357–1359, May 1986, doi: 10.1063/1.96908.

[287] K. W. Vogt, P. A. Kohl, W. B. Carter, R. A. Bell, and L. A. Bottomley, 'Characterization of thin titanium oxide adhesion layers on gold: resistivity, morphology, and composition', *Surface Science*, vol. 301, no. 1, pp. 203–213, Jan. 1994, doi: 10.1016/0039-6028(94)91300-5.

[288] M. Todeschini, A. Bastos da Silva Fanta, F. Jensen, J. B. Wagner, and A. Han, 'Influence of Ti and Cr Adhesion Layers on Ultrathin Au Films', *ACS applied materials & interfaces*, vol. 9, no. 42, pp. 37374–37385, 2017, doi: 10.1021/acsami.7b10136.

[289] S. Takyu, T. Kurosawa, N. Shimizu, and S. Harada, 'Novel Wafer Dicing and Chip Thinning Technologies Realizing High Chip Strength', *MRS Online Proceedings Library (OPL)*, vol. 970, p. 0970, ed 2006, doi: 10.1557/PROC-0970-Y06-05.

[290] 'All About Wafer Dicing in Semiconductor/IC Manufacturing'. Accessed: Feb. 28, 2023. [Online]. Available: https://www.thomasnet.com/articles/custom-manufacturing-fabricating/wafer-dicing/

[291] 'Piranha Solutions'. Accessed: Dec. 11, 2022. [Online]. Available: https://ehs.princeton.edu/book/export/html/513

[292] W. Kim, T.-H. Kim, J. Choi, and H.-Y. Kim, 'Mechanism of particle removal by megasonic waves', *Applied Physics Letters*, vol. 94, pp. 081908–081908, Mar. 2009, doi: 10.1063/1.3089820.

[293] 'LOR / PMGI series – Microresist'. Accessed: Dec. 11, 2022. [Online]. Available: https://www.microresist.de/en/produkt/lor-pmgi-series/

[294] 'SU-8', Kayaku Advanced Materials, Inc. Accessed: Dec. 11, 2022. [Online]. Available: https://kayakuam.com/products/su-8-photoresists/

[295] 'Protective Coating AR-PC 5091.02 (Electra 92)', Allresist EN. Accessed: Dec. 11, 2022. [Online]. Available: https://www.allresist.com/portfolio-item/protective-coating-ar-pc-5091-02-electra-92/

[296] 'Microposit MF-319 - nanoFAB Knowledge Base - nanoFAB Confluence'. Accessed: Jan. 31, 2023. [Online]. Available: https://confluence.nanofab.ualberta.ca/display/NFCORE/Microposit+MF-319

[297] 'EC Solvent, EC Solvent 11 – Microresist'. Accessed: Jan. 31, 2023. [Online]. Available: https://www.microresist.de/en/produkt/ec-solvent-ec-solvent-11/

[298] 'MICROPOSIT™ REMOVER 1165', *Semiconductor Technologies*, no. 889, p. 3.

[299] 'SU-8 for Dielectrics in Organic TFT Back-Planes', Kayaku Advanced Materials, Inc. Accessed: May 24, 2023. [Online]. Available: https://kayakuam.com/products/display-dielectric-layers/

[300] D. Sun, S. Tian, X. Jiao, and Y. Jiang, 'Discussions for the charging efficiency of the resistor–capacitor and resistor–inductor–capacitor series circuits under different excitation sources', *Energy Reports*, vol. 8, pp. 300–306, Nov. 2022, doi: 10.1016/j.egyr.2022.09.163.

[301] D. Shindo and T. Oikawa, 'Energy Dispersive X-ray Spectroscopy', in *Analytical Electron Microscopy for Materials Science*, D. Shindo and T. Oikawa, Eds., Tokyo: Springer Japan, 2002, pp. 81–102. doi: 10.1007/978-4-431-66988-3_4.

[302] S. Wenner, L. Jones, C. D. Marioara, and R. Holmestad, 'Atomic-resolution chemical mapping of ordered precipitates in Al alloys using energy-dispersive X-ray spectroscopy', *Micron*, vol. 96, pp. 103–111, May 2017, doi: 10.1016/j.micron.2017.02.007.

[303] *The London, Edinburgh and Dublin philosophical magazine and journal of science*. London : Taylor & Francis, 1840. Accessed: May 24, 2023. [Online]. Available: http://archive.org/details/londonedinburg6271914lond

[304] L. Cheng, S. Mao, Z. Li, Y. Han, and H. Y. Fu, 'Grating Couplers on Silicon Photonics: Design Principles, Emerging Trends and Practical Issues', *Micromachines*, vol. 11, no. 7, Art. no. 7, Jul. 2020, doi: 10.3390/mi11070666.

[305] D. R. Paschotta, 'Collimated beams'. Accessed: May 29, 2023. [Online]. Available: https://www.rp-photonics.com/collimated_beams.html

[306] J. S. Weaver, J. C. Heigel, and B. M. Lane, 'Laser spot size and scaling laws for laser beam additive manufacturing', *J Mater Process Technol*, vol. 73, p. 10.1016/j.jmapro.2021.10.053, Jan. 2022, doi: 10.1016/j.jmapro.2021.10.053.

[307] L. Yu. Beliaev, E. Shkondin, A. V. Lavrinenko, and O. Takayama, 'Optical, structural and composition properties of silicon nitride films deposited by reactive radio-frequency sputtering, low pressure and plasma-enhanced chemical vapor deposition', *Thin Solid Films*, vol. 763, p. 139568, Dec. 2022, doi: 10.1016/j.tsf.2022.139568.

[308] R. Hanke, *Filter-Fascinatie [Filter-Faszination, niederländ.] Alles over filters en hun toepassing*. 1981.

[309] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, 'NIH Image to ImageJ: 25 years of Image Analysis', *Nat Methods*, vol. 9, no. 7, pp. 671–675, Jul. 2012.

[310] D. G. Altman and J. M. Bland, 'Standard deviations and standard errors', *BMJ*, vol. 331, no. 7521, p. 903, Oct. 2005, doi: 10.1136/bmj.331.7521.903.

[311] E. A. Avrutin, 'Numerical and semi-analytical analysis of effective modulation layer properties in ITO modulators', presented at the Numerical Simulation of Optoelectronic Devices, Turin, Italy, Sep. 2023.

[312] A. C. Lesina, D. Goodwill, E. Bernier, L. Ramunno, and P. Berini, 'On the performance of optical phased array technology for beam steering: effect of pixel limitations', *Opt. Express, OE*, vol. 28, no. 21, pp. 31637–31657, Oct. 2020, doi: 10.1364/OE.402894.

[313] W. M. Duncan and S. A. Henck, 'Insitu spectral ellipsometry for real-time measurement and control', *Applied Surface Science*, vol. 63, no. 1, pp. 9–16, Jan. 1993, doi: 10.1016/0169-4332(93)90056-H.

[314] J. Male, 'Thermally Enhanced Paper Microfluidics for Photonic Crystal Biosensors in Amorphous Silicon', mscresearch, University of York, 2019. Accessed: Jun. 22, 2023. [Online]. Available: https://etheses.whiterose.ac.uk/25801/

[315] J. Li and S.-T. Wu, 'Two-coefficient Cauchy model for low birefringence liquid crystals', *Journal of Applied Physics*, vol. 96, no. 1, pp. 170–174, Jun. 2004, doi: 10.1063/1.1738526.

[316] K. Levenberg, 'A method for the solution of certain non-linear problems in least squares', *Quart. Appl. Math.*, vol. 2, no. 2, pp. 164–168, 1944, doi: 10.1090/qam/10666.