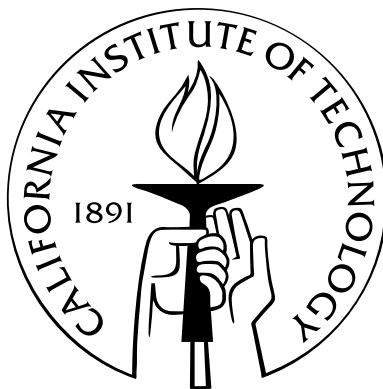# Electron Transport in Silicon Nanocrystal Devices: From Memory Applications to Silicon Photonics

Thesis by

Gerald M. Miller

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2012

(Defended September 16, 2011)

for my mother, Patti, and my wife, Kat

Acknowledgements

First off, I must thank Harry Atwater for allowing me to join his group. It has truly been a privilege to work here, and my experience in the Atwater group has been made memorable and enjoyable due to his scientific brilliance and strength of character. He has given me the freedom here to explore the topics that I wanted to explore, and has created a friendly and open environment here that not only fosters productivity, but is a pleasure to experience every day. Personally, Harry has an uncanny ability to turn frustration into opportunity, and I find that every time I meet with him his insights allow me to leave inspired and motivated, no matter how frustrated I was coming in. I'll miss that, and I hope I've learned enough to take some of it with me.

My thanks also go out to my defense committee, Professors Oskar Painter, Kerry Vahala, and Amnon Yariv, for their careful consideration of this thesis. My thanks as well to Professors Painter, Vahala, and Marc Bocrath for serving on my candidacy committee. It's been a humbling experience sharing Watson with all of them.

As for the Atwater group itself, a thousand thanks go first to Ryan Briggs. As my lab husband, his presence is entwined with all the erbium work in this thesis, and I shudder to think how it would have turned out without him. He set up the ion beam sputtering system used to grow the erbium films studied here, repeatedly helped me in the optics lab, and was a great help to me as I was starting to learn about erbium. He served as a daily source of motivation, inspiration, and encouragement during my time here. He's a brilliant scientist, a wonderful person, and I'm blessed to call him a friend. I must also thank Tao Feng, who mentored me when I first joined the group. He taught me about the UHV-AFM and got me started in the lab with the nanocrystal charging project. I'm also grateful to Robb Walters and Julie Biteen, who were great resources to me as I was beginning work with silicon nanocrystals. Domenico Pacifici was also of great help to me when I joined the erbium working group, and was an inspiration for the pulsed excitation model laid out in Chapter 4. Mike Kelzenberg has always been generous with his time and was a great help to me as I was learning to electrically characterize these films. I'd like to thank Greg Kimball, who also helped me with electrical characterization, as well as with using the Hall Probe stage. I spent a lot of time RF sputtering erbium films, and I'd like to thank Andrew Leenheer for his help with that system. Stan Burgos helped me prepare the TEM results of $LaAlO_3$ films, and I'm grateful for the

time he put in, as well as for having a fellow metalhead in the group.

A special thanks goes out to my office mates. Brendan Kayes, Jen Dionne, Krista Langeland, Dan Turner-Evans, Emily Warmann, and Carissa Eisler have made it fun to come to the office every day. I couldn't have asked for a nicer, more helpful set of people to share an office with. In grad school it was very easy for me to get wrapped up in my own work, but it was enlightening to see what my office mates were working on, and I appreciate the perspective they were able to give me on my own work.

Outside the Atwater group, I'd like to thank Carol Garland for her help obtaining TEM results and Ali Ghaffari for his help using the Watson cleanroom and SEM. Judy Lattimer in the Gray group trained me on the XPS that was used to determine silicon excess in my films.

I've also had the privilege of working with scientists outside of Caltech. I'd like to thank Doug Bell at JPL for all of his help on the layered tunnel barrier project. He did all the IPE measurements for that work and was very helpful as I was learning about that project. The erbium work was funded by the AFOSR under the Silicon Laser MURI project, and I'd like to thank my collaborators on the project, who have done great work to push the project forward. Specifically, the discussions I had with Rohan Kekatpure in Marc Brongersma's lab at Stanford were very helpful when I was studying carrier absorption. I'd also like to thank Seljuk Yerci for discussions on sputtering when he was working on erbium doped nitride films in Luca Dal Negro's group at Boston University.

I also owe a debt of gratitude to the assistants of the Atwater group. April Neidholdt, Lyra Haas, Tiffany Kimoto, and Jennifer Blankenship are the grease in the wheels of this group. They are all helpful and professional while still being a lot of fun. They're people who get things done, and this lab would grind to a halt without them. In the Applied Physics department I'd like to thank Connie Rodriguez, Rosalie Rowe, Eleanora Vorobieff, and Christy Jenstad for all their help throughout the years.

Finally, I'd like to thank all my friends and family for always being there, and providing a crucial sense of work-life balance. To my mom, dad, and stepmom — you have been an endless source of encouragement for me, and I love you all very much. And to my wife, Kat, thank you for your patience during my seven years of grad school, this degree is yours too. I love you and can't wait to see what the future brings us.

Gerald Miller
September, 2011
Pasadena, CA

Abstract

The push to integrate the realms of microelectronics and photonics on the silicon platform is currently lacking an efficient, electrically pumped silicon light source. One promising material system for photonics on the silicon platform is erbium-doped silicon nanoclusters (Er:Si-nc), which uses silicon nanoclusters to sensitize erbium ions in a SiO_2 matrix. This medium can be pumped electrically, and this thesis focuses primarily on the electrical properties of Er:Si-nc films and their possible development as a silicon light source in the erbium emission band around 1.5 μm. Silicon nanocrystals can also be used as the floating gate in a flash memory device, and work is also presented examining charge transport in novel systems for flash memory applications.

To explore silicon nanocrystals as a potential replacement for metallic floating gates in flash memory, the charging dynamics in silicon nanocrystal films are first studied using UHV-AFM. This approach uses a non-contact AFM tip to locally charge a layer of nanocrystals. Subsequent imaging allows the injected charge to be observed in real time as it moves through the layer. Simulation of this interaction allows the quantification of the charge in the layer, where we find that each nanocrystal is only singly charged after injection, while holes are retained in the film for hours.

Work towards developing a dielectric stack with a voltage-tunable barrier is presented, with applications for flash memory and hyperspectral imaging. For hyperspectral imaging applications, film stacks containing various dielectrics are studied using I-V, TEM, and internal photoemission, with barrier tunability demonstrated in the Sc_2O_3/SiO_2 system.

To study Er:Si-nc as a potential lasing medium for silicon photonics, a theoretical approach is presented where Er:Si-nc is the gain medium in a silicon slot waveguide. By accounting for the local density of optical states effect on the emitters, and carrier absorption due to electrical pumping, it is shown that a pulsed excitation method is needed to achieve gain in this system. A gain of up to 2 db/cm is predicted for an electrically pumped gain medium 50 nm thick. To test these predictions Er:Si-nc LEDs were fabricated and studied. Reactive oxygen sputtering is found to produce more robust films, and an electrical excitation cross section of 2.5×10^{-13} cm^2 is found, which is two orders of magnitude larger than the optical cross section. The fabricated devices exhibited low lifetimes and low current densities which prevent observation of gain, and the modeling is used to predict how the films must be improved to achieve gain and lasing in this system.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.0.1 Moore's Law and Silicon Microelectronics

The exponential development of the silicon microelectronic industry is responsible for the profound societal change experienced in recent times. Ever since the first transistor was fabricated at Bell labs by Shockley, Bardeen, and Brittain in 1947, and the first integrated circuit was introduced by Kirby at Texas instruments in 1957, the computing power achieved by silicon microelectronics has increased exponentially. This behavior has faithfully obeyed Moore's Law by, as he so eloquently put it, "Cramming more components onto integrated circuits" [1]. The exponential increase in the number of transistors fabricated on silicon chips and the resulting increase in computing power has revolutionized every industry on the planet. For comparison, Fig. 1.1a shows the very first transistor fabricated by Bardeen in 1947. Modern advances in fabrication have increased the number of transistors integrated on a chip into the billions. For example, the Intel x7 microprocessor is shown in Fig. 1.1b, and has 1.17 billion transistors. Concurrent with the exponential increase in transistor density is an exponential decrease in the cost of computing. According to the International Technology Roadmap for Semiconductors (ITRS), the cost per transistor has dropped from $5.54 in 1952 to a nano-dollar (a billionth of a dollar) in 2004 [2]. This incredible behavior has put a little computer in the pocket of almost every American, in the form of cell phones, digital cameras, and tablets.

For most of the past 40 years, Moore's law has been followed by device scaling. To 'cram' more components onto chips, each of these components must be made smaller and smaller. Increasing the number of transistors has lead to an increase in computing power by increasing the number of computations that can be done per second. In addition, miniaturizing these devices has lead to increased speeds and lower power consumption. These improvements come from the basic structure of the transistor, which at its core is a metal-oxide-semiconductor (MOS) capacitor. To keep the field constant in the device as it is scaled, both the gate length and the oxide thickness are shrunk. This decreases the voltage needed to charge the capacitor and create the inversion channel basic to
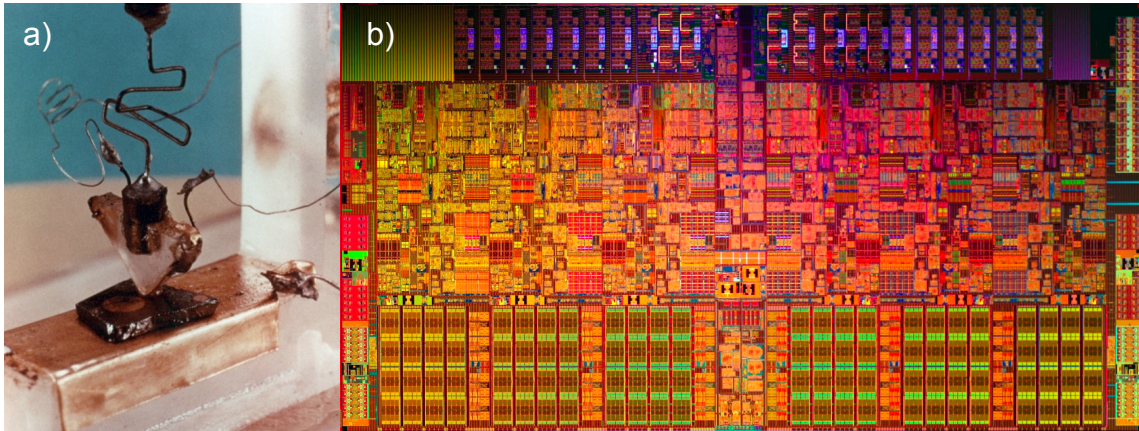
Figure 1.1: Moore's law has described the microelectronics industry since the introduction of the integrated circuit. Shown in a) is the very first transistor fabricated by Shockley. Today, silicon chips hold billions of transistors. Shown in b) is the Intel x7 processor, which contains 1.17 billion transistors.

MOSFET operation.

However, this device scaling limit has been reached for devices using SiO_2 as the gate oxide, as device scaling has shrunk the devices so far that the thickness of the oxide is only 1.2 nm for the 65 nm node. At this thickness the tunneling current across the thin oxide greatly increases the power consumption of the device. Furthermore, this thickness represents only about 4 atomic layers of SiO_2, as can be seen from the cross section in Fig. 1.2. At this point, further device scaling would require thinning this film further, and at this point the SiO_2 is less a homogenous film than it is a series of Si and O atoms. For these reasons, the industry had to break with SiO_2 as the gate dielectric. The solution to this problem lay in using high-κ, hafnium-based dielectrics, which would provide the same capacitance as SiO_2, but in a thicker film that exhibited much less leakage. This technology has enabled further device scaling into the 45 nm, 32 nm, and 22 nm nodes.

This is important not because it represents recent events — the Hf-based 45 nm node debuted in 2008 — but because it represents a significant point in the history of the transistor. Silicon was chosen as the semiconductor of choice primarily because of the very high quality of the interface between Si and SiO_2. It is stable, has a very low defect density, is grown from Si rather than deposited, and is an excellent insulator. The robust quality of this interface drove device scaling for over thirty years, and SiO_2 performed its duties admirably even as its thickness was shrunk to atomic scales. This thesis begins with these properties in mind, as we explore the properties of silicon nanocrystals. When embedded in a host medium of SiO_2, nanocrystals can act as charge storage nodes for non-volatile memory. To program and erase these nanocrystals, charge must flow in and out of them through the oxide matrix on a length scale comparable to the one shown in Fig. 1.2, while still retaining their charge for nonvolatile memory purposes.
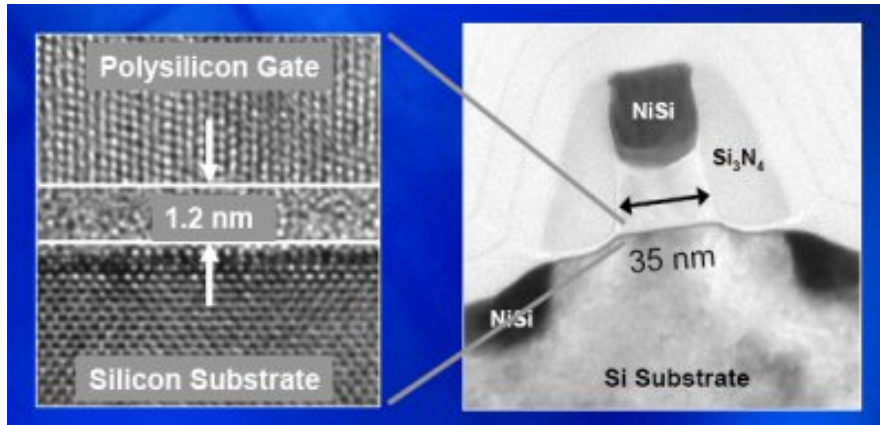
Figure 1.2: Shown at the right is a cross section of a MOSFET from Intel's 65 node, the last to use SiO_2 as the tunnel oxide. The thickness of the tunnel oxide was scaled down to 1.2 nm, as shown by the inset on the left. Further device scaling required the replacement of SiO_2 with other dielectrics for subsequent nodes.

1.0.2   Silicon Photonics

Complimentary to the advances in the silicon microelectronics industry are advances in the field of photonics. The distinction between these fields is primarily one of scale, with microelectronics focused on high-speed electrical computation and data transfer over a single Si chip, while photonics is focused on high-speed optical data transfer over a distance of kilometers. Photonics has become the long-haul communication method of choice over electronics because the low cost and power necessary to transmit data large distances. Using light allows the very fastest of transit speeds — the speed of light — and does not suffer from electromagnetic interference like electronics do. The data fidelity of long-range photonics was made possible by the development of low-loss silica optical fiber [3]. These fibers exhibit loss below 1 dB/km for wavelengths of 1.3 and 1.6 μm, which is why most modern telecommunications operate in these bands.

The ability to quickly and easily transmit data all over the world has also had a profound effect on modern society, but advances in the realm of photonics must also be matched by advances in microelectronics. The two have been inextricably linked by the need to process data electronically when it comes out of the fiber, and as the world consumes more and more bandwidth, research has focused on merging these two worlds by integrating optical components on silicon chips.

The field of silicon microphotonics has sought to develop silicon as an optical material, and has thus far been successful developing passive silicon devices such as waveguides [4], and active devices like switches and modulators [5, 6, 7]. However, developing an efficient light source has been problematic. Current silicon laser technologies include a silicon Raman laser [8, 9], which requires a high powered pump laser, and the so-called hybrid laser, where a III-IV laser is wafer bonded to silicon, and is coupled into a silicon waveguide [10, 11]. Currently neither of these methods are both
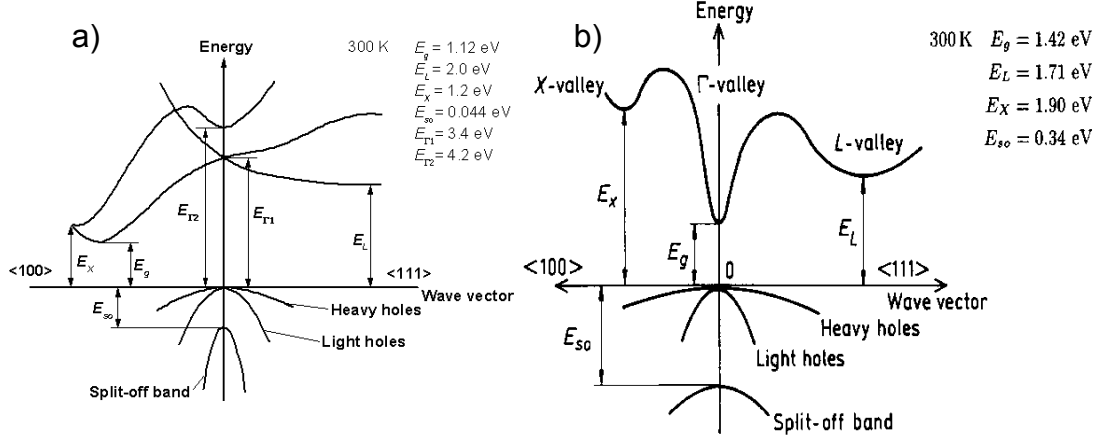
Figure 1.3: The band structure of Si (a) shows that the band gap, E_g, is indirect, meaning radiative emission is mediated by phonons and therefore much less efficient than a direct gap material like GaAs (b).

CMOS compatible and electrically pumped, both of which are necessary to truly achieve chip-scale photonics.

Silicon itself is a very poor light source due to its indirect band gap. Efficient light emitting systems such as GaAs take advantage of their direct band gap (Fig. 1.3b) to radiatively emit when carriers in the conduction band recombine into the valence band. In silicon, however, the indirect gap requires a phonon to make up the momentum mismatch between carriers in the valence and conduction bands (Fig. 1.3a). This three-particle process is extremely inefficient, and as a result carriers in silicon are far more likely to combine non-radiatively.

To this end, two of the most promising research directions towards a CMOS-compatible, electrically pumped laser have been silicon nanostructures and rare earth doping. Silicon nanostructures such as nanosheets [12, 13], nanopillars [14], and nanocrystals [15, 16], all emit in the visible region due to quantum-confined excitons. If excitons can be confined in space, the uncertainty principle makes the excitons more energetic and increases the probability of overcoming the momentum mismatch needed to emit radiatively. If such systems can be fabricated with sufficiently low defect densities, the radiative efficiency increases dramatically [17]. These nanostructures can also be electrically pumped [18, 19] even in SiO_2, which has led to optimism that an electrically pumped laser is possible.

Rare earth doping of silicon materials is another very active field of research. This method involves incorporating rare earth ions into silicon-based materials and using the emission properties of the dopants. Principal among these is erbium, which emits at the silica fiber attenuation minimum at 1.5 μm. Initial studies focused on doping Er into Si directly [20]. While silicon is an excellent host for electrical injection purposes, Er emission is silicon suffers from lack of room-temperature

luminescence, a low solubility limit of Er in Si, and energy back-transfer between the second excited state of Er at 980 nm and the silicon band edge at 1100 nm. More promising were attempts to dope erbium into SiO_2, where the solubility limit of Er increases two orders of magnitude and energy back-transfer becomes negligible. However, electrical injection into such films has generally not been successful.

Lastly, there is the system where erbium and silicon nanocrystals are co-doped into SiO_2. It is this system that will be the subject of the majority of this thesis. By having both erbium and nanocrystals in the matrix, it is possible to both emit at the telecom wavelength of 1.5 μm and have electrical access to the insulating matrix via the nanocrystals. In addition, the nanocrystals also work as sensitizers for erbium, which increases the 1.5 μm luminescence at the expense of the visible nanocrystal luminescence. However, this system must exhibit more robust electrical injection than has previously been exhibited for it to exhibit efficient emission and lasing.

On a very basic level, this goes back to the Si/SiO_2 interface and its excellent insulating properties. The scope of this thesis focuses, ironically, on how to manipulate and move charge across this interface, which has been lauded for its inability to do so. The MOS-field-effect-transistor (MOS-FET) is the basic building block of the microelectronic industry, but the opportunity to use a similar structure as an LED or laser by doping with emitters necessitates we push current through the oxide and therefore violate its insulating properties. But if it can be done efficiently, it would open the door to a fully compatible, electrically pumped light source on silicon, and allow the full integration of optics and electronics on a chip.

1.0.3   Contents of This Thesis

This thesis explores three different systems where the behavior of charge and currents on the nanoscale is of primary interest, each with very different applications. The chapters are organized as follows:

Chapter 2 studies charge transport in silicon nanocrystals on the nanoscale using ultra-high vacuum atomic force microscopy (UHV-AFM). Samples are prepared with a layer of nanocrystals located 10 nm beneath the surface of a SiO_2 film. Such films could replace the floating gate used in traditional flash memory by using nanocrystals as discrete charge storage nodes. To explore this behavior on the nanoscale, a non-contact AFM tip is used to charge the nanocrystals by briefly contacting the film and injecting charge via an applied bias. Subsequent imaging by the same non-contact tip images the charge in the nanocrystal layer, as the electrostatic forces due to the injected charge cause a deflection of the tip. By simulating this interaction, it is possible to determine the magnitude of charge in the film. By analyzing subsequent images in time, it is then possible to 'watch' the charge as it moves in the nanocrystal layer. We find that the injected charge follows a gaussian distribution, and at the point of injection, the charge density in the layer corresponds

to at most one carrier per nanocrystal. Charge retention times for hole retention is on the order of hours in these films, which we attribute to the large valence band offset between Si and SiO_2. That retention in these films is much longer for holes than electrons indicates that a memory device operated by hole programming would offer robust retention characteristics.

Chapter 3 explores a novel approach to hyperspectral imaging by using layered dielectric structures. By fabricating dielectric stacks using materials with dissimilar band offsets, a triangular-shaped energy profile is formed for charges passing through the film. By applying a voltage to this structure, the magnitude of the barrier height is raised or lowered, depending on the bias. In this way the dielectric stack becomes a tunable high-pass filter for electrons. Such a device could be used in a spectral detector where the tunable barrier acts as a filter for photoexcited electrons. The spectrum responsible for the photoexcited carriers could then be extracted with a voltage sweep. The work presented here examines possible structures that would exhibit this barrier-lowering effect. We look at dielectric stacks composed of a variety of candidate films, including $LaAlO_3$, La_2O_3, Sc_2O_3, Al_2O_3, and SiO_2. The film characteristics are determined by current-voltage measurements, internal photoemission measurements, and TEM, and show that the interfacial quality of the films is critical to device operation. Proof of concept is demonstrated by modest barrier lowering in the Sc_2O_3/SiO_2 structure, but an increase in film quality is still needed for practical device operation.

Chapters 4 and 5 theoretically and experimentally explore the erbium-doped silicon nanocluster (Er:Si-nc) system. Chapter 4 presents theoretical work which examines the viability of Er:Si-nc as a gain medium in a silicon slot waveguide laser. The silicon slot waveguide has a significant effect on the radiative rates of the Er and Si-ncs in the slot, which affect the power needed to pump the Er to inversion. Also studied is the effect of electrical pumping, which introduces carriers that excite the Si-ncs. These excited Si-ncs can absorb the Er emission with a characteristic cross section, σ_{ca}, and this effect is studied for the values of σ_{ca} found in the literature. We find that these two effects limit the gain achieved in the system, even as the Er is highly inverted. To mitigate this we present a pulsed excitation scheme. By pulsing excitation, the nanocrystal excitations are allowed to decay while the pulse is off, which eliminates carrier absorption. Meanwhile, the longer-lived Er ions are then free to produce gain in the system. The calculations show that 1 db/cm of gain is possible in the system under pulsed excitation for a 10 nm film. This number is roughly the same as the loss demonstrated in silicon waveguides, which is reason for optimism that a laser can be fabricated in this system.

Chapter 5 presents experimental work concerning electroluminescence in the Er:Si-nc system. We present data arising from two separate sets of films, the first exhibiting luminescent center (LC) sensitization, the second using more traditional Si-nc sensitization. Electroluminescence data is presented for both sets of films, with the LC films achieving higher current densities than the Si-nc films. The Si-nc films were far more robust to biasing, however, which allowed more thorough

characterization. Time-dependent luminescence data was used to extract an electrical excitation cross section of 2.5×10^{-13} cm^2, which is two orders of magnitude larger than that found under optical pumping. A careful inspection of the rise times of these films under electrical excitation shows a delay between the voltage pulse and electroluminescence, which is likely due to charging dynamics in the film after the pulse. A similar effect is not seen in the lifetime, which may mean that the timescale of discharging is much smaller than that for charging. This would be beneficial in a pulsed injection scheme for laser operation, where we wish to keep the carriers in the film to a minimum. We finish by revisiting the modeling of Chapter 4 with the experimental parameters obtained from our experiments. The modeling shows that for the low lifetimes and low current densities observed in our films, Er inversion is not possible. But if the film quality can be improved to allow higher lifetimes and more current, inversion is possible, although it is still likely that a pulsed injection scheme would be needed to observe gain.

Chapter 6 presents a brief summary and outlook of the topics considered in this thesis, with some proposed future directions for the work presented.

Chapter 2

Charge Transport in Nanocrystal Ensembles

The cell phones, digital cameras, and tablet devices that are becoming more common today have increased the demand for large capacity data storage with minimal power consumption. Flash memory is the dominant data storage mechanism on modern portable devices due to its reliability, simplicity, and non-volatility. However, as the individual memory elements become smaller to increase data storage density, one problem they face is increasing leakage current. Smaller devices require thinner dielectric films in their device structure, which increase the tunneling current in the memory elements. This increases the power consumed by the device, and is a problem which must be addressed to increase the storage capacity in flash memory.

In this chapter we explore the viability of silicon nanocrystals as charge storage nodes for use in flash memory devices. We start with an introduction to data storage and flash memory, then describe the merits of a nanocrystal-based flash memory, including why it should decrease leakage current. We then present experimental work on charge retention characteristics of silicon nanocrystal layers using non-contact AFM (nc-AFM) under high-vacuum conditions. By modeling the tip-surface interaction, it is possible to map the charge density in a localized area of a nanocrystal film in a non-destructive way. By repeatedly scanning an area, it is possible to 'watch' the charge as it migrates in the nanocrystal layer. We finish with an exploration of the temperature dependence of charge migration. It should be noted that the work presented in this chapter is an extension of the work presented in Tao Feng's thesis in 2005.

2.1   Flash Memory

Data storage has historically taken place on magnetic media such as tapes or hard drives, where a magnetic media is encoded with a moving tip, which also reads data off the disk. Data access is achieved by moving the magnetic material and/or tip to the place where the data is stored for
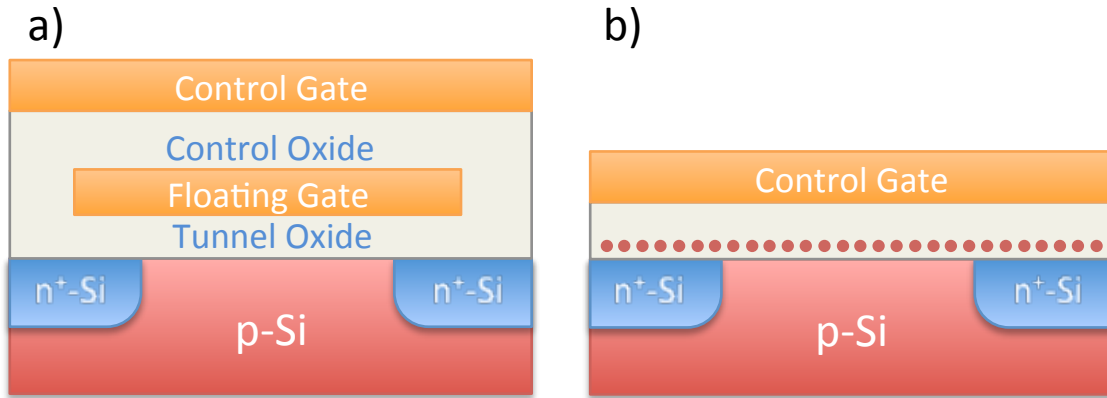
Figure 2.1: a) Traditional flash memory, where charge is held by an electrically isolated floating gate. b) Nanocrystal memory, where charge is held by a layer of nanocrystals (red dots)

readout. These magnetic devices still provide the highest storage capacities, but require energy to move the media by spinning disks and winding tapes. More energy is needed to move the magnetic head in a hard drive to read the data. In 1980, Fujio Masuoka invented the first flash memory, a completely new technology which stored and accessed data electrically, without using the moving parts inherent in magnetic media. Instead of using magnetic media to encode data, trapped charge is used as the encoding element, and read and write operations are performed electronically.

The device architecture for traditional flash memory is shown in Fig. 2.1a. It consists of a traditional transistor structure, but with an electrically isolated floating gate between the channel and gate. To encode data, the floating gate is charged by applying a gate bias, which causes charge to flow across the thin tunnel oxide into the floating gate. The charge contained in the floating gate changes the source-drain threshold voltage, V_T, of the transistor, so the presence of charge can be read by applying a voltage to the gate which is small enough not to program or erase the cell. This shift in threshold voltage is read as the 0 or 1 bit value of the cell. Because the floating gate is electrically isolated, this information will be stored for a very long time with no power applied to the device, which makes it excellent for nonvolatile memory applications. This technology has dominated the hand-held storage market because it uses significantly less power than disk-based methods because there are no moving parts, and advances in miniaturization have made storage capacities large enough for practical use.

In a nanocrystal memory architecture, pictured in Fig. 2.1b, the floating gate is replaced by a layer of silicon nanocrystals in a silicon oxide matrix. The concept of discretizing the floating gate using nanocrystals was first proposed by Tiwari, et al in 1995 [21]. One common failure mechanism for flash memory is a short across the tunnel oxide which would discharge the floating gate. If the charge in the floating gate is instead contained in discretized nanocrystals, it would take many tunnel oxide shorts before the cell would fail completely.

This increased reliability would allow use of a thinner tunnel oxide. The storage capacity of a memory device is given by the areal density of memory elements on the chip. Making each memory element smaller laterally requires that the tunnel oxide be scaled proportionally to keep the capacitance of each element constant. Traditional floating gate memory requires a tunnel oxide of about 7 nm to ensure device reliability. However, if a thinner tunnel oxide could be used, the footprint of each device could be made smaller, increasing storage capacity.

The use of a thinner oxide also opens the possibility of using direct tunneling current to program the floating gate. Current devices use n-type channels that can be programmed using Fowler-Nordheim tunneling across the 3.1 eV conduction band barrier between the Si substrate and the tunnel SiO_2. Direct tunneling would allow the gate to be programmed using holes from a p-type channel, which is not practical under Fowler-Nordheim tunneling because the valence band barrier is 4.7 eV. If this barrier could be traversed using direct tunelling at low voltages, the larger barrier would lead to more robust charge retention. As we will show in this chapter, hole retention in nanocrystals is indeed much longer in nanocrystals than electron retention, making a p-type memory a possible path to more reliable devices.

2.2   Nanocrystal Fabrication

The nanocrystal layers studied in this chapter were fabricated via ion implantation of Si into a thermal SiO_2 layer. This method accelerates silicon ions directly into a substrate to create a silicon-rich region below the surface. Upon annealing, the excess Si phase separates due to Ostwald ripening to form clusters, which crystallize at annealing temperatures around 900 °C. The annealing step also repairs oxide damage incurred during implantation. The nanocrystals can vary in size from a couple nanometers to tens of nanometers, depending on the ion dose and annealing conditions. This method is fully CMOS compatible, and the high temperature steps occur early in the process flow.

Our films were fabricated starting with 100 nm of thermally-grown SiO_2 on a Si substrate. These samples were implanted with 5 keV Si^+ ions at a dose of 1.27×10^{16} cm^{-2}. The Si excess peaks around 10 nm below the surface, according to TRIM [22]. The films were then annealed at 1080 °C in an Ar ambient to precipitate nanocrystals. Etching and AFM studies indicate that this produces nanocrystals on the order of 2 nm in size, with an areal density of 4×10^{12} cm^{-2} [23].

2.3   Imaging Charged Layers Using Non-Contact AFM

The standard non-contact atomic force microscopy (nc-AFM) technique uses an oscillating cantilever to map a sample surface by measuring the change in the cantilever's resonance frequency as the cantilever is brought near the surface. The van der Waals force between the surface and tip alters
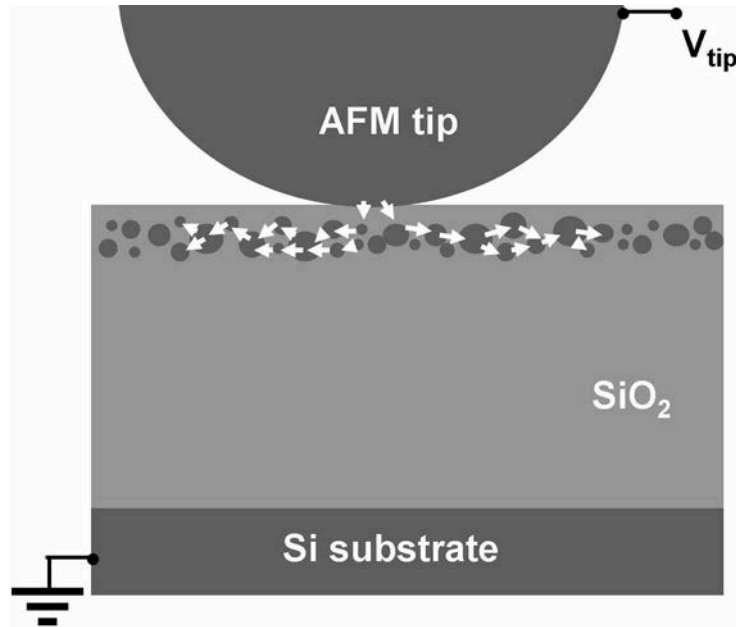
Figure 2.2: Schematic of charge injection into a Si-nanocrystal layer using a nc-AFM tip

the tip resonance, and the tip is moved up or down accordingly to keep the tip force gradient constant, and in this way a topographic image is mapped out without physically touching the surface.

This method can also be used to image charge on the surface. When the surface to be scanned is also charged, there is an extra force between the tip and sample due to the electrostatic force between the sample charge and the induced tip charge. This will cause the tip to be attracted to the surface, and the tip will be retracted to reduce the force gradient. Hence, the resulting nc-AFM image will show a topographical feature, a 'hump', indicative of the charge at the surface [24]. Here the non-contact nature of the scan is preferable because by keeping the tip physically separated from the sample, the tip is unable to transport or remove charge from the surface, and the data taken is due entirely to charge migration in the sample.

To analyze the charge distribution in the nanocrystal layers, we used a n^+ Si nc-AFM tip under ultra-high vacuum (UHV) conditions. The experiments were carried out on an Omicron UHV-AFM system with a base pressure of 10^{-9} torr. To charge the nanocrystal layer, we brought the nc-AFM tip into intimate contact with the sample and biased the sample at either +10 V or -10 V for 10 s to inject holes or electrons, respectively, into the nanocrystal layer. This process is shown schematically in Fig. 2.2 (taken from Feng's thesis). Electron injection in this system has been explored by Feng [23, 25], while hole injection will be discussed here.

On a real nanocrystal sample, the injection process is shown in Fig. 2.3. Part a is an image of the sample before injection. The scale bar goes to 2 nm, indicating the white areas in the image are due to sample roughness. Fig. 2.3b shows the same area after injection at +10 V, and shows
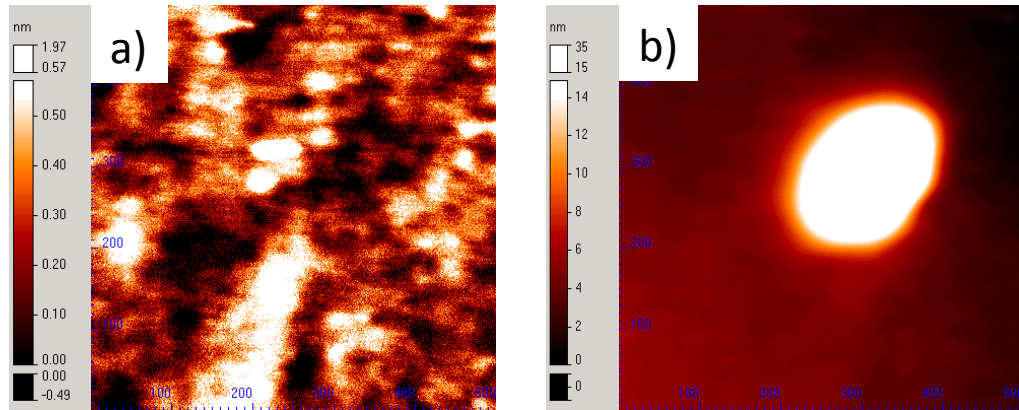
Figure 2.3: a) A typical nc-AFM scan over a 500 nm^2 x 500 nm^2 area before charge injection at +10 V. The sample roughness is about 2 nm. b) The same area after injection. The bright spot is over 30 nm in height, and is due entirely to charge in the nanocrystal layer.

a protrusion corresponding to the charged area. Note now the scale bar goes to 35 nm. The 10 V bias from the tip caused holes to tunnel into the nanocrystal layer. It is important to note that the protrusion shown is not due to any physical topographical feature, but due only to the force on the tip from the charged nanocrystals 10 nm below the surface. This charge is responsible for deflecting the tip over 30 nm across the charged area, which is roughly 200 nm in diameter.

By successively scanning the same area, it is possible to see the charge migrate in real time. Specifically, the protrusion showing in Fig. 2.3b changes shape and decreases in size. This evolution is shown in Fig. 2.4, where the first contour is taken 6 minutes after injection, and representative contours from 2.4 hours, 5.1 hours, and 18.5 hours are shown in succession. In general, the protrusion loses height far more quickly than it broadens. This is shown in more detail in Fig. 2.5 (taken from Feng's thesis), where line scans through the center of the protrusion are plotted. Initially the tip is deflected by more than 30 nm by the charge, which remains for a timescale of hours. The inset plots the peak height alone as a function of time, and fits it to a single exponential. The intrinsic decay constant for the peak to drop by 63% is found to be 2.4 hours for hole injection.

It should be noted that this long timescale is specific to hole injection (positive tip bias). Under electron injection, the timescale is much shorter, resulting in complete charge dissipation in less than one hour [26]. A full study of this phenomenon is provided in Tao Feng's thesis work [25], but this is attributed primarily to the asymmetry in the valence and conduction band offsets for silicon in contact with silicon oxide. The valence band offset, or the energy barrier for holes, is 4.7 eV, while the barrier for electrons is only 3.1 eV [27]. This opens the possibility of using holes as the memory element of the floating gate due to their longer retention characteristics.
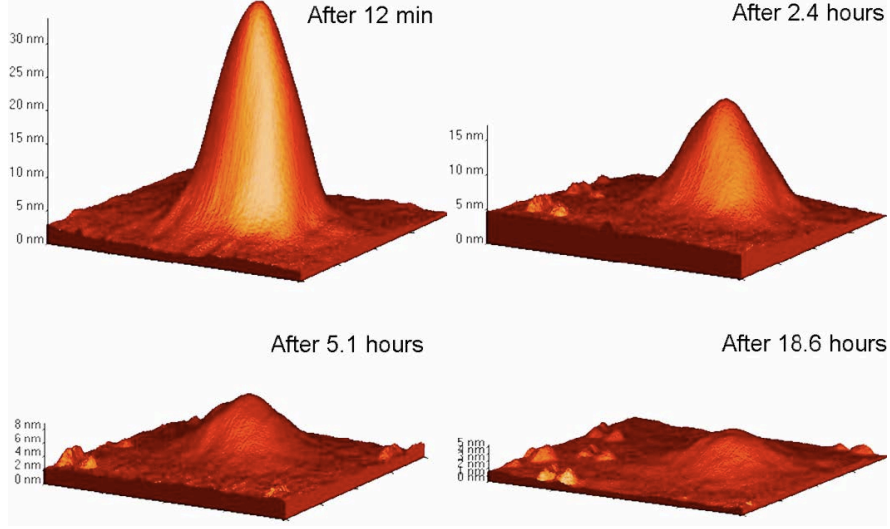
Figure 2.4: Time evolution of a charged layer of silicon nanocrystals imaged under nc-AFM. The four panels show the same area 0.1 hours after charge injection up to 18.5 hours later. The protrusion decreases in height as the charge diffuses away from the charged area.

2.4 Modeling of the Tip-Surface Interaction

Non-contact AFM is an effective and non-destructive method for getting a qualitative look at charged layers and their evolution in real time [24]. However, by simulating the interaction between the tip and charged surface, it is possible to extract a charge density distribution from a single nc-AFM scan [26, 28]. By fitting to many scans spaced in time, we can find the time dependence of the charged layer. Using this model we are able to quantify the charge in the layer, finding that the highest charge density corresponds to roughly one charge per nanocrystal. We can also predict a conductivity for the nanocrystal layer.

During a nc-AFM scan, the electronics of the microscope work to keep the force gradient in the z-direction, dF/dZ, constant by moving the tip vertically. This generates the image during scanning. For our simulation [26], we assume there are only two forces on the tip during scanning, the van der Waals interaction between the tip and sample and the electrostatic force due to the charge at the surface. We assume the tip to be a grounded sphere, the surface to be flat, and the charge distribution after injection to be Gaussian.

Using these forces it is possible to simulate a nc-AFM scan by using the fact that the force gradient during scanning is kept constant. This is stated mathematically by:

$$\frac{d\mathbf{F}}{dz}_{tot} = \frac{d\mathbf{F}}{dz}_{vdW} + \frac{d\mathbf{F}}{dz}_{ES} = contant \tag{2.1}$$

where the *constant* is the set-point force gradient for nc-AFM operation.

The van der Waals force is always present, and for our work is obtained by integrating the
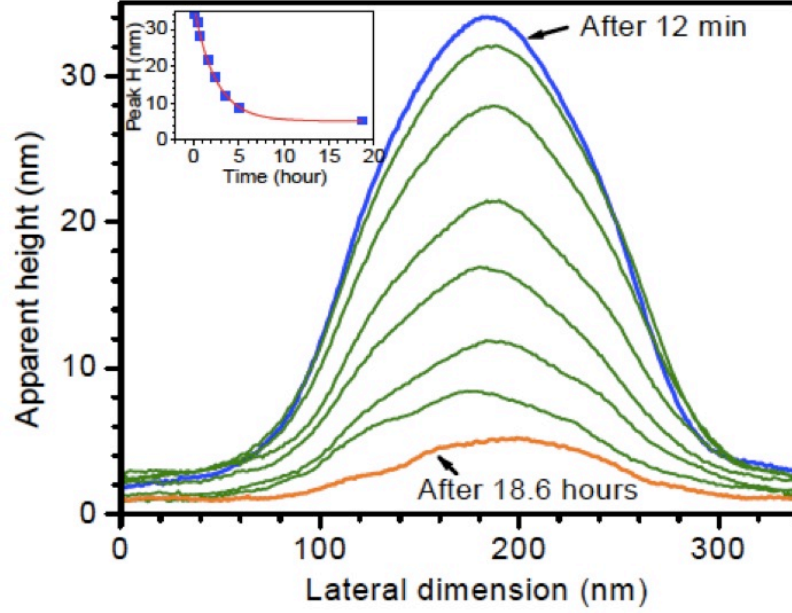
Figure 2.5: A succession of line scans through the center of the protrusion imaged in Fig. 2.4. The height of the protrusion falls on a timescale of hours as charge dissipates from the injection region. The inset shows the evolution of peak height, and fits to a single exponential with a decay constant of 2.4 hours.

Lennard-Jones potential between a sphere and plane where $z >> R$ [29]. The van der Waals force is given by the following equation:

$$\mathbf{F}_{\text{vdW}} = \frac{HR}{6z^2}(-1 + \frac{\sigma^6}{30z^6})\hat{\mathbf{z}} \tag{2.2}$$

where z is the distance between the tip and sample, which is 10 nm during scanning; R is the tip radius, measured to be about 100 nm; H is the Hamaker constant, $H = 3 \times 10^{-19}$ J; and $\sigma = 0.34$ nm is a parameter in the Lennard-Jones potential for argon. See Boer's thesis for a more thorough treatment of this potential and its parameters.

The electrostatic force between the tip and charge stored in the nanocrystal layer is due to the attraction between the charge at the surface and the opposite charged induced at the tip. It is calculated using the method of images, and has the following form:

$$\mathbf{F}_{\text{ES}} = \sum_{i,j} \frac{q_{s_i} q_{t_j}}{4\pi\epsilon_0 D^2_{ij}} \hat{\mathbf{d}}_{\mathbf{ij}} \tag{2.3}$$

where q_s is a charge in the sample, q_t is a charge in the tip, and D_{ij} is the distance between them. Feng has written a comprehensive algorithm for calculating these forces, and it is well explained in his thesis.

By taking the derivatives of equations (2.2) and (2.3) with respect to z, it is possible to simulate
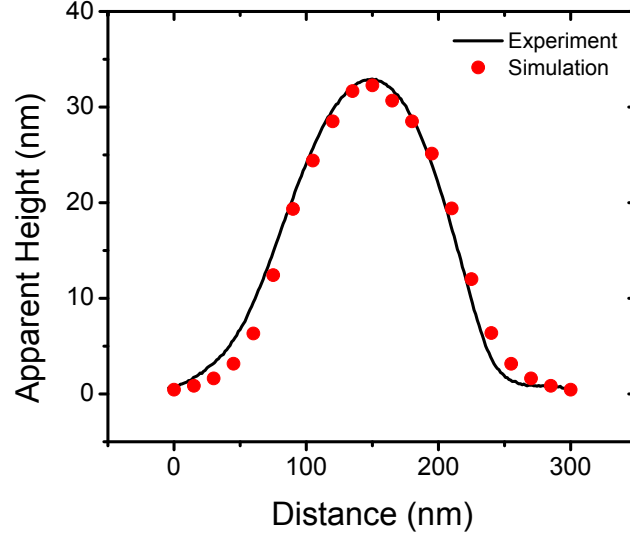
Figure 2.6: Experiment and simulation of nc-AFM line scan. The black line is an experimental line scan through a region charged with nanocrystals. The red dots are nc-AFM simulation fits assuming an underlying Gaussian charge distribution.

a nc-AFM line profile through a charged area using equation (2.1). Our approach to modeling was to simulate a nc-AFM tip scanning an area with an assumed charge density present. We then compared our simlated line profile with a line profile taken from experiment. Then, by changing the assumed charge distribution and iterating, the process is repeated until a sufficient fit is found. Calculations for this model were done in C+ in Feng's thesis and are shown in the appendix there. I have adapted his code for MatLab for my own use, and it is contained in the appendix here. The fits in this chapter were done using this code.

A fit to a line profile from the data set shown in Fig. 2.4 is shown in Fig. 2.6. The solid black line is an experimental nc-AFM line profile obtained 25 minutes after injection, and shows a 30 nm deflection of the tip as it scans the surface. The red dots represent a nc-AFM simulation over a surface with an assumed Gaussian charge distribution. The Gaussian used for this fit has a central peak charge density of 3.7×10^4 holes/μm^2 and a full-width at half-maximum (FWHM) of 127 nm. A gaussian with these input parameters contains 685 holes of total charge, indicating this is the amount of charge present in the nanocrystal layer when this scan was taken. The qualitative fit of the simulation is quite good, indicating that the underlying charge distribution is indeed Gaussian in nature. Studies of the nanocrystal densities in these films [23] show the areal nanocrystal density in the layer is 4×10^4 holes/μm^2, which means that the charge density at the center of the distribution is almost one hole per nanocrystal, which is expected because the energy required to doubly charge a nanocrystal is extremely high [25, 30]. This is significant because it means the amount of charge in a nanocrystal floating gate is directly tied to the nanocrystal density in the film.
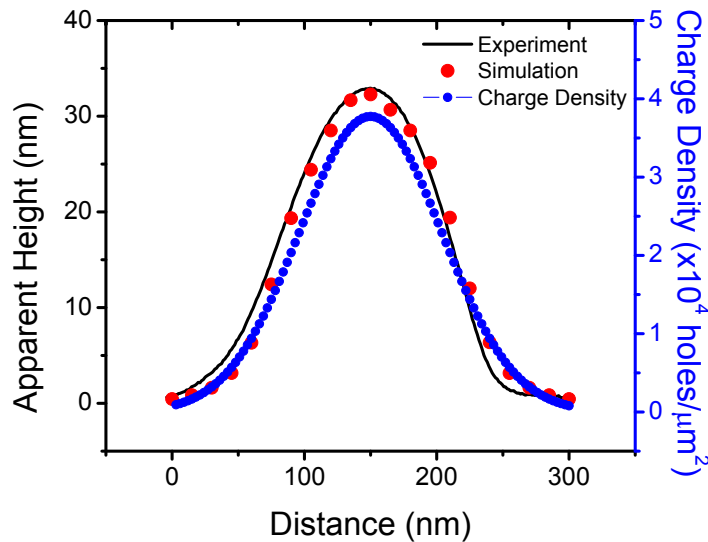
Figure 2.7: Experiment and simulation of nc-AFM line scan with charge density overlaid. The black line is an experimental line scan through a region charged with nanocrystals. The red dots are nc-AFM simulation fits assuming an underlying Gaussian charge distribution shown by blue dots. The charge distribution follows the nc-AFM pretty closely in shape.

The nc-AFM scan shows a protrusion hundreds of nanometers in extent, and the charge density assumed by the simulation has similar dimensions. This is shown in Fig. 2.7, where the assumed charge density is added to the right axis of Fig 2.6. Here we see that the shape of the charge density closely follows that of the nc-AFM line scan, which shows that our experimental technique is fairly sensitive to injected charge. There is a base level of charge that our system cannot detect, but it is encouraging to see that to first order, the deflection of the scanning tip is proportional to the injected charge in the layer directly below the tip.

Using this simulation, we can calculate the charge contained in the layer as a function of time by fitting a charge distribution to a number of scans taken over the same area. This plot is shown in Fig. 2.8. The decrease in charge is actually fairly linear in time, aside from an initial sharper decrease. This is due to charged surface states that dissipate faster than the charge contained in the nanocrystal layer below the surface. This is confirmed by Feng on a sample with ion implantation damage at the surface, but was not annealed to form nanocrystals, and a full explanation can be found in his thesis. Also of note is the long time scale of the linearity, which stays linear for over 18 hours. The last points in this data set were taken after waiting overnight, when we were lucky enough to find the charged area when coming back the next day.

In addition to calculating total charge, we can look at the evolution of the charge distribution. The nc-AFM scans show that while the charge dissipates in time, the charge distribution also flattens out, indicating that the charge is moving laterally in the nanoocrystal layer. The Gaussian
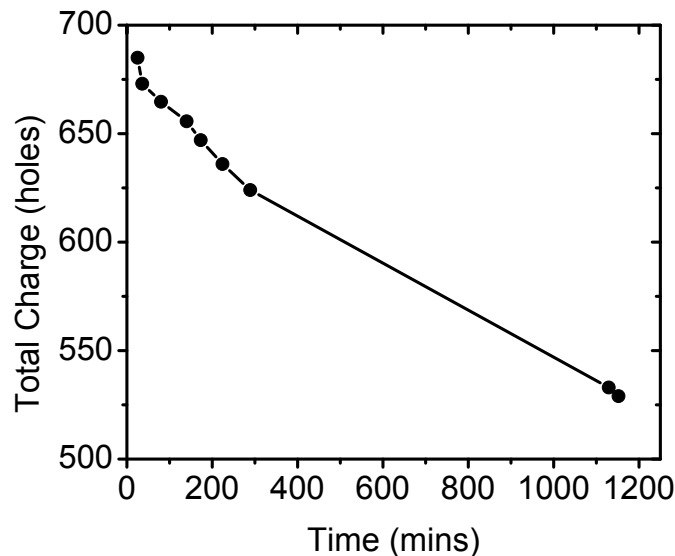
Figure 2.8: Total charge in the nanocrystal layer as a function of time

distribution assumed for this work has the form

$$\rho(r) = \rho_0 \times \exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{2.4}$$

where ρ_0 is the peak charge density and σ is related to the full-width at half-maximum (FWHM) by FWHM $= 2\sigma\sqrt{2\ln 2}$. Figure 2.9 plots these two inputs as a function of time, with ρ_0 on the left axis and σ on the right.

The peak charge density initially decreases sharply before leveling off after about 5 hours, after which ρ_0 levels off to around 1700 holes per square micron, or about one third of its initial value. Conversely, the value of σ increases sharply over the same time frame ρ_0 decreases, showing the distribution flattening out as a function of time before settling at 70 nm, or a FWHM of 165 nm.

This analysis raises a very important question, namely 'where does the charge go?' Presumably, there are three places the charge can go once it enters through the tip. First, it can go to the surface and dissipate laterally through surface states; second, it can tunnel down out of the nanocrystal layer into the substrate; and third, it can diffuse, or 'hop', from nanocrystal to nanocrystal and diffuse laterally below the surface.

The timescale for surface transport is less than an hour or so, as shown in Feng's thesis on samples without nanocrystals. This is likely why the total charge drops sharply in the first few minutes, as shown in Fig. 2.8. But fact that these samples retain charge longer than an hour show that charge within the nanocrystal layer is not diffusing to the surface to dissipate laterally there. Transport into the substrate is ruled out because of the thick oxide layer beneath the nanocrystals.
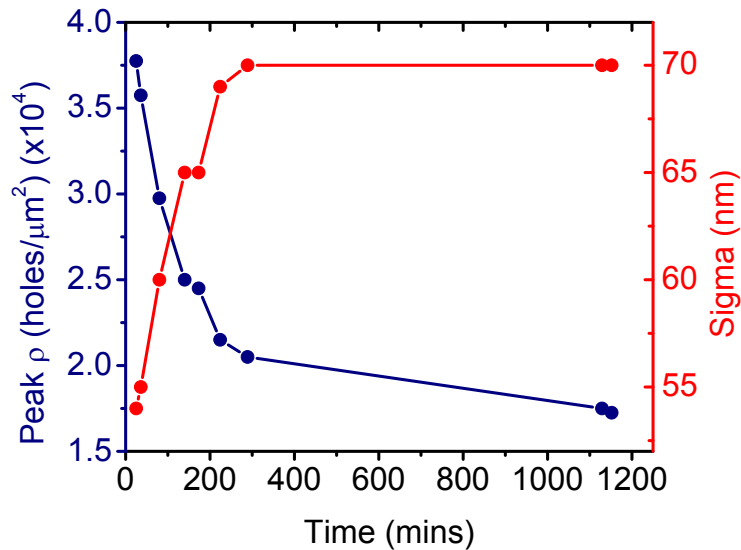
Figure 2.9: Gaussian charge distribution parameters as a function of time. The left axis measures the peak of the distribution in time, which decreases as charge migrates away. The right axis measures the spread of the distribution, σ, which increases in time as the distribution flattens out.

The probability of a hole tunneling through 90 nm of oxide is extremely small, leaving hopping as the most likely mechanism for charge mobility.

However, if charge is primarily hopping from nanocrystal to nanocrystal, then charge should be conserved as it moves through the nanocrystal layer. If this were the case then the total charge would be constant in time and Fig. 2.8 would be flat. However it is not, and we attribute this mainly to the sensitivity of the nc-AFM technique. The RMS roughness of our sample is about 1–2 nm, and once the charge density is low enough to make a tip deflection of that magnitude it becomes indistinguishable from the background. A subsequent fit to the 'hump' region of a nc-AFM line scan would ignore that charge, and give a total charge reading that is less than what is actually there. The density where this happens is around a few hundred holes per square micron, as can be estimated from the plots in Fig. 2.7 by observing the magnitude of charge necessary to deflect the tip 1–2 nanometers.

2.5   Conductivity in Nanocrystal Layers

It is also possible to calculate a conductivity for charge in these films, which is an important parameter for these films if they are to be used for non-volatile memory. It is important to be able to get charge in and out of the films for memory operation, but the conductivity needs to be low enough for long-term information storage. Using the charge densities calculated by the method above, it is possible to extract a scalar conductivity by starting with Ohm's law:

$$\mathbf{J} = \sigma\mathbf{E} \tag{2.5}$$

where σ is the scalar conductivity connecting the current density, $\mathbf{J}$, and the electric field, $\mathbf{E}$. The electric field, $\mathbf{E}$, is related to the charge density, ρ, by Gauss's law:

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon}\rho \tag{2.6}$$

where ϵ is the permittivity of the material. For this work, we assume the permittivity of SiO_2. The current density, $\mathbf{J}$, is also related to ρ by the continuity equation:

$$\nabla \cdot \mathbf{J} = -\frac{\partial\rho}{\partial t}. \tag{2.7}$$

After substituting and simplifying, the following equation relates σ and ρ

$$\sigma\rho = -\epsilon\frac{\partial\rho}{\partial t}. \tag{2.8}$$

This is a scalar equation that relates the conductivity to the time evolution of the charge density. By going go one step further and integrating over area the following is obtained:

$$\int_A \sigma\rho = \int_A -\epsilon\frac{\partial\rho}{\partial t} \tag{2.9}$$

$$\sigma\int_A \rho = -\epsilon\frac{\partial}{\partial t}\int_A \rho \tag{2.10}$$

and if A is chosen to be the area over which the nc-AFM detects charge, then the integral $\int_A \rho$ becomes the total charge calculated in Fig. 2.8. We then approximate the partial derivative in the following manner:

$$\sigma q_f = -\epsilon\frac{q_f - q_i}{t_f - t_i} \tag{2.11}$$

where the subscripts f and i refer to final and initial charge states, corresponding to fits to successive nc-AFM scans over the charged area. In this way a conductivity is obtained for each pair of successive scans.

The result of this calculation is shown in Fig. 2.10. The plot shows a average conductivity $\sigma \sim 2 \times 10^{-16}$ S/m, which is quite low compared to the values for either silicon ($\sigma \sim 640$ S/m) or even thermal SiO_2 ($\sigma \sim 10^{-12}$). This is because the charge is trapped in the nanocrystals, and must exceed the barrier height for holes of 4.7 eV to escape into the surrounding oxide's conduction band. This low value for the conductivity is encouraging for non-volatile memory applications, where the
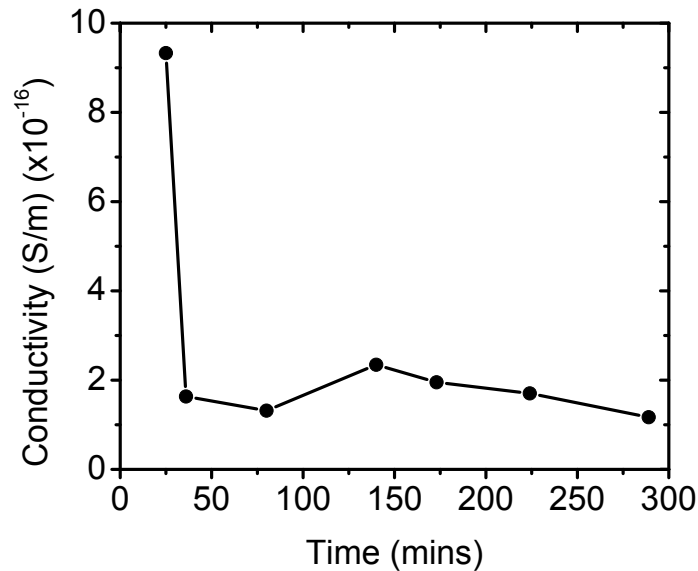
Figure 2.10: Conductivity at room temperature of a nanocrystal layer injected with holes. The sharp decrease at short times is due to more mobile charge at the oxide surface. The low magnitude shows that hole retention is very good at room temperature, which is necessary for non-volatile memory applications.

industry standard retention time for injected charge is 10–30 years. The steep drop in the figure at short times is due to the surface charge present after injection, which dissipates at a much faster rate than the trapped nanocrystal charge.

2.6 High-Temperature Measurements

The timescale for hole retention in these nanocrystal layers is quite long, which is beneficial for an actualized memory device. However, for a more thorough understanding of charge transport in these films, we attempted to take data at elevated temperatures. From a scientific standpoint, high temperature measurements allow an insight into the transport phenomenon between isolated nanocrystals by adding a thermally-assisted charge transport mechanism. From a practical standpoint, any device built using these layers must withstand the possibly high temperatures present in modern electronics, and it is important to characterize the effect of elevated temperatures on charge retention.

The Omicron VT-AFM system used for this work is equipped with a resistive heating element on the sample mounting bracket. This orientation puts the heating element very close to the sample, but also introduces uncertainty into the measurement because the electrical connections to the heater are reset each time the sample stage is moved or changed. As a result there is significant uncertainty in the temperature when the heater is used, which we did not realize until testing the temperature
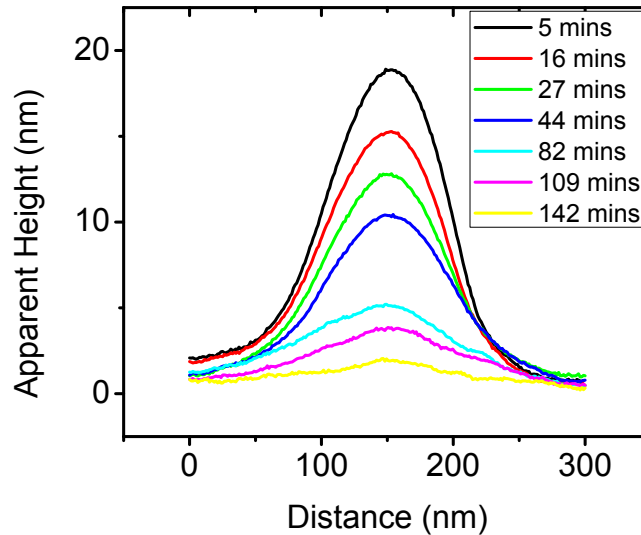
Figure 2.11: Line scans through a charged nanocrystal layer at elevated temperature. The timescale for charge dissipation is shorter than at room temperature due to thermally assisted charge hopping.

calibration with a thermocouple sample mounted to the heating sample mounting bracket.

Unfortunately, due to technical difficulties with the system these experiments could not be reproduced after identifying this variability. As a result, the data presented here is illustrative of the general dynamics of nanocrystal films at an elevated temperature, even though the exact temperature is unknown.

An elevated temperature study was performed on the same sample described in the previous section. The input parameter to the heater is the current across the heating element, which in this case was 319 mA. The Omicron calibration table puts the sample at 177 °C, however our thermocouple studies indicated this temperature varied between 67 °C and 131 °C. We attribute this variability to the contact made between the heater and its electronics, because each time the heater was turned on, it stabilized at a constant temperature, but that temperature varied between 67 °C and 131 °C.

The line scans at elevated temperature are shown in Fig. 2.11. The scans at this temperature show a much faster decrease in apparent height, showing complete dissipation in less than three hours, compared to significant retention at over 18 hours for room temperature samples. This is due to thermally assisted transport, as the increased thermal energy increases the probability of a confined hole jumping the 4.7 eV barrier into the oxide conduction band.

We can perform the same analysis on this set of data as for the data set at room temperature to obtain the charge and conductivity of the layer at elevated temperature. Figure 2.12a shows the calculated total charge present after injection as a function of time, and confirms the shorter
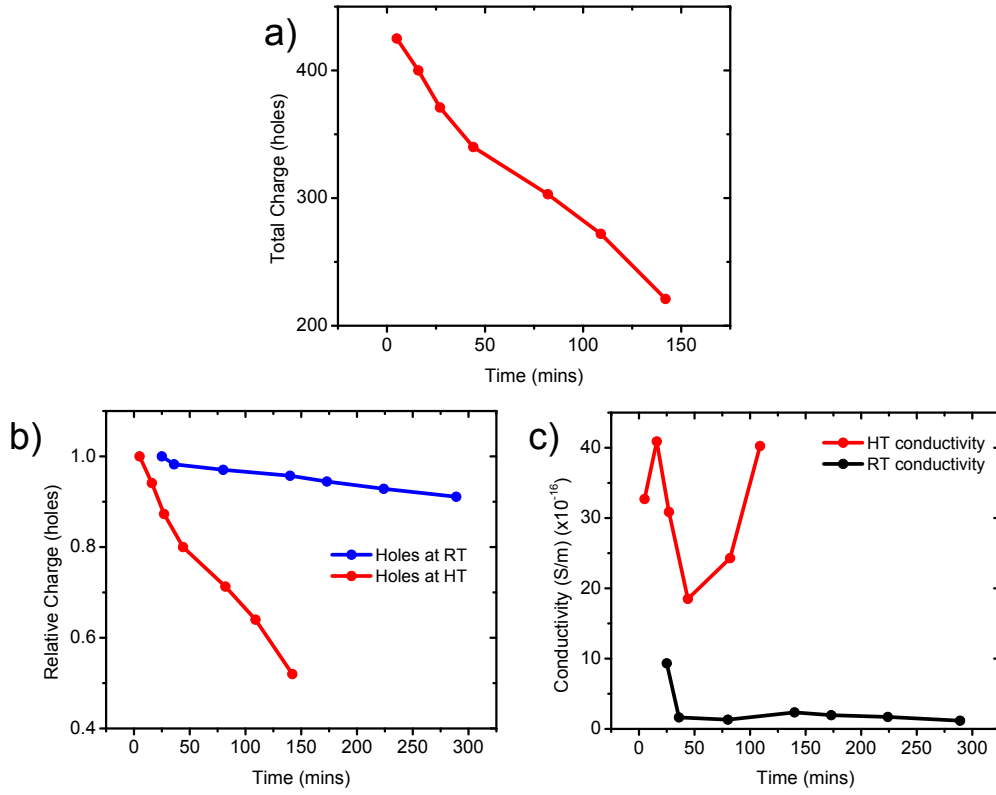
Figure 2.12: a) Total charge contained in a nanocrystal layer at high temperature as a function of time. b) The normalized charge at both room temperature and high temperature, showing the faster dissipation in the nanocrystal layer. c) The calculated conductivities of the film at both low and high temperature. The increased temperature raised σ by over an order of magnitude.

timescale for charge dissipation. The magnitude of charge injected is about half that observed at room temperature, and this is likely due to the variability of the contact between the tip and sample during injection. Other data sets show variability in the size of the observed 'hump' in the line scans at all temperatures. In general, however, we expect the injected charge to be slightly smaller simply because there is a delay of 5–7 minutes between injection and the time the tip scans over the area the first time, during which more charge diffuses away due to the higher temperature.

For a comparison to the room-temperature data, the relative charge is plotted in Fig. 2.12b, with each set normalized to 1. Here we clearly see the effect of high temperature on retention, as each data set decreases fairly linearly, but with a greater slope at high temperature. We also compare the electrical conductivities in Fig. 2.12c. The high temperature conductivity is noisier than at room temperature, but does increase by an order of magnitude to $\sim 3 \times 10^{-15}$ S/m. But even with this increase, the conductivity is still well below that for thermal oxide.

2.7 Conclusion

We have analyzed charge injection into layers of silicon nanocrystals using noncontact AFM under ultra-high vacuum conditions. This method provides special insight to the transport dynamics in these layers by giving submicron scale visualization of the charge in the layer. We are able to directly image the charge with the nc-AFM tip, and quantify it by simulating the tip-surface interaction. This microscale picture is something that cannot be obtained with other, more common methods such as I-V or C-V measurements, which only measure the macroscale charge dynamics.

What we learn from this method is that upon biasing the layer at a single point — the AFM tip — the nanocrystal layer assumes a Gaussian charge distribution whose peak density corresponds to a single charge per nanocrystal on average. This information is very important for scaling concerns, as now the amount of charge contained in a nanocrystal floating gate is directly controlled by the nanocrystal density.

By watching the charge distribution in time, we show that the charge does not dissipate laterally at a large rate, although this rate can be increased by increasing the temperature. Hole injection in these films lead to very robust retention times, due largely to the high barrier for holes between Si and SiO_2. This behavior opens the door to a memory device that uses hole injection instead of electron injection. However, the long retention times provided by this high barrier height also increase the voltage needed to write and erase these layers. However, that could be mitigated by using a carefully engineered tunnel oxide, which will be explored in the next chapter.

Chapter 3

Layered Tunnel Barriers For Hyperspectral Imaging

3.1 Introduction

In this chapter we present work towards a voltage-tunable hyperspectral detector by using the novel approach of layered tunnel barriers. Such a device would be able to take an image using a traditional pixel array, and allow spectral data to be obtained with a voltage sweep of the pixel array. This project was funded by NASA and was a collaboration between Caltech, the Darrell S. Schlom group at Penn State University, and Doug L. Bell at the Jet Propulsion Laboratory. We start with a description of hyperspectral arrays and describe how using a layered tunnel barrier would provide voltage-dependent wavelength selectivity. We then describe our work towards a realized structure using lanthanum oxide films, scandium oxide films, and silicon oxide stack structures. We show a voltage-dependent barrier height in the Sc_2O_3/SiO_2 system before ending with an outlook for this technology going forward.

3.2 Hyperspectral Imaging

Hyperspectral imaging, or imaging spectroscopy, is the process of obtaining the spectral content of an image. This is especially useful for deep space observation, where we wish to take pictures of deep space, and also want the spectrum at each point in our image. That spectral information gives the chemical composition of distant objects via elemental absorption lines, as well as their temperature via the blackbody radiation spectrum.

However, this information represents a 3D 'data cube' — two dimensions for the spatial image and another for the spectrum. Recording on a traditional two-dimensional pixel array has traditionally been done in two ways: the first is to put a grating in front of the array which will put the spectrum on one axis of the array, as shown in Fig. 3.1a. The other axis is then used for a single spatial
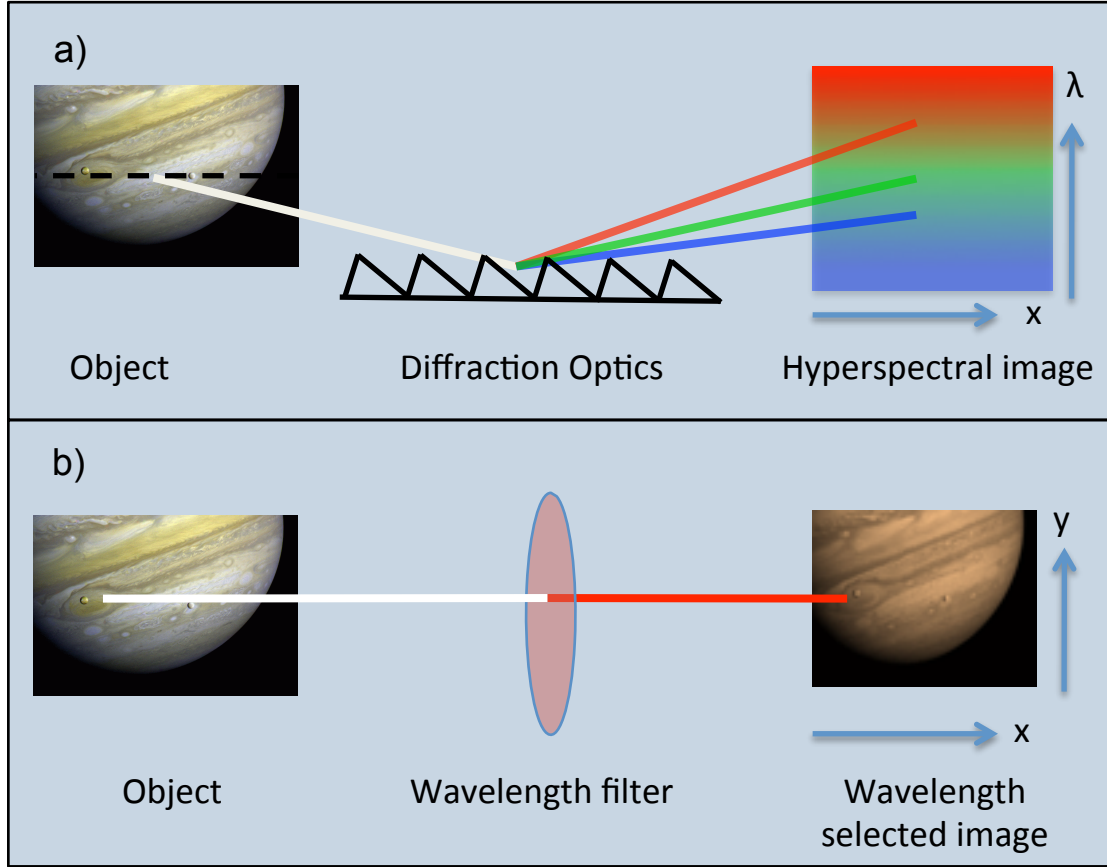
Figure 3.1: Traditional hyperspectral imaging. a) Hyperspectral imaging using a grating to separate out the spectral content of the source. The resulting image has wavelength along the y-axis, and maps the dotted line to the x-axis. The full data set is compiled by sweeping the dotted line vertically. b) Hyperspectral imaging using wavelength filtering. Here the filter gives a two-dimensional image, but only for a small wavelength range, and the full data set is obtained by compiling different wavelength ranges. Our proposed device eliminates the need for either a grating or filter.

dimension. In this way a single 'image' on the pixel array shows the spectral content of a single line in space, shown by the dotted line in Fig. 3.1a. Then this line is moved around using the optics on the setup to sweep the second dimension. This is the so-called 'pushbroom' or 'whiskbroom' method. The second method is to use the array for a two-dimensional image, and use a set of filters to get the spectral data, shown schematically in Fig. 3.1b. In this way the spectrum is swept out in time, and the 'true' image is obtained by compiling all the images from different wavelength bins.

For this project, we propose a memory element with a voltage-tunable barrier height. Such a barrier, described in more detail below, would filter photoexcited carriers by energy by changing the barrier height for photoemission with a voltage sweep. The spectral content is obtained by comparing the current due to photoexcited carriers for different voltages, because the barrier height corresponds to the energy of the photons that excited the carriers. This device moves the burden of wavelength selectivity from an external filter or grating to a barrier integrated with the memory

pixel. This greatly reduces the complexity of the device by removing the external filters and gratings, as well as reducing the weight and power required by the device, both of which are at a premium for extraterrestrial missions.

3.3   Layered Tunnel Barriers

To achieve a barrier with a voltage-tunable barrier height, we use a dielectric stack using materials of different barrier heights. This concept comes first from Likharev [31], who proposed a triangular barrier for voltage tunability. This concept is explored further by Casperson, et al. [32], and the schematic for the barrier is reproduced from that paper in Fig. 3.2.

The energy diagram for a single-layer barrier is shown in Fig. 3.2a (solid line), with silicon to the left, a metal contact to the right, and a dielectric of barrier height of ϕ_B in between. Applying a voltage, V, to the stack pulls the right edge down (dotted line), and a current density, J_e, flows through the triangular portion of the barrier via Fowler-Nordheim tunneling [33]. The barrier height stays the same height, it simply changes shape, allowing more current to pass through the thinner area near the peak of the barrier. Fig. 3.2b shows a theoretical triangularly shaped barrier of the same height ϕ_B. When a voltage V is applied, the total height of the barrier decreases, allowing electrons with smaller energies to traverse it. Practically, it is difficult to have a constantly graded barrier, but by using a stack with differing barrier heights, we can approximate a triangular barrier by putting the largest barrier between two materials with smaller barriers. This energy diagram is pictured schematically in Fig. 3.2c, with its real-space counterpart in Fig. 3.2d. Here we see that by using the dielectric stack with graded barrier heights, the total barrier height lowers under applied voltage.

This type of structure is ideal for a hyperspectral detector because the applied voltage turns the device into a high-pass filter for photoexcited electrons. An electron will only have the energy to traverse the barrier if it absorbs a photon of energy equal to or greater than the barrier height. The differential current obtained by changing the barrier height is due to photons only at the energy corresponding to the barrier height. In this way the photocurrent measures the wavelength of incoming radiation, and the cell takes spectroscopic data with a simple voltage sweep without gratings or filters. Furthermore, data storage can be directly integrated into this structure by making this layer stack the tunnel barrier in a flash memory cell, described in the previous chapter.

3.4   Material Selection and Band Offsets

The wavelength range accessible by this method is dependent entirely on the barrier heights involved and how much barrier lowering is possible for a given dielectric stack. For the tri-layer stack proposed
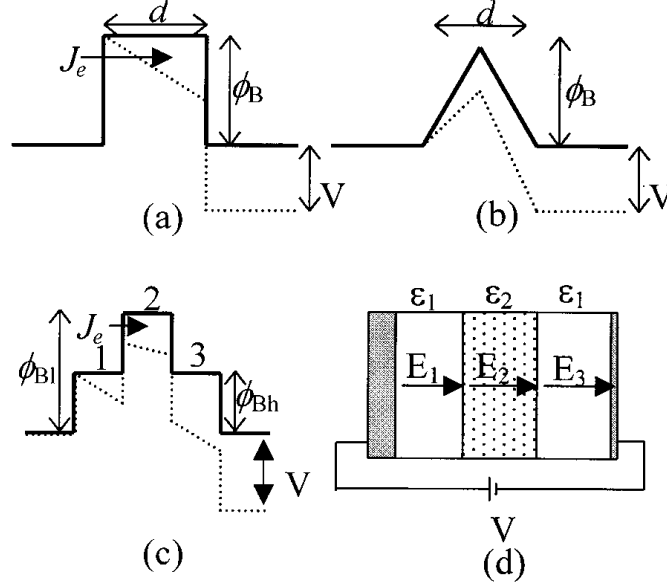
Figure 3.2: Voltage-tunable barriers based on triangularly graded energy profiles

in Fig. 3.2c, the middle layer (labeled '2') must have a larger band gap than the outside layers, and its band offsets must also be larger to create the triangularly shaped potential barrier.

Band offsets are most commonly measured using internal photoemission (IPE) [34], where a tunable laser or lamp illuminates a dielectric stack with electrical contacts. Current is generated in the structure by photoexcited carriers in the stack, and the onset of current at a given excitation energy corresponds to the energy barrier of the stack. This is shown schematically in Fig. 3.3. Figure 3.3a shows the flatband condition for a metal-dielectric-semiconductor (MDS) capacitor without illumination. The semiconductor has a band gap E_{SC} and the dielectric has a larger band gap, E_D. The conduction band offset, ΔE_c, is the energy barrier between the semiconductor conduction band and the dielectric conduction band, and the valence band offset, ΔE_V, is the barrier between the semiconductor and dielectric valence bands.

During an IPE experiment, the sample is illuminated with light of a specific energy, $h\nu$, shown in Fig. 3.3b and c by the arrow. This energy, when absorbed in the semiconductor, can excite carriers over the barrier in the dielectric where they can be collected by the metal. When a hole absorbs a photon of sufficient energy, it can escape from the conduction band and over the barrier as shown in Fig. 3.3b. Its total barrier is labeled Φ_H and is equal to the sum of ΔE_V and E_{SC}.

$$\Phi_H = \Delta E_C + E_{SC} \tag{3.1}$$

Likewise if an electron in the valence band absorbs the incident light, it can traverse the barrier as shown in Fig. 3.3c. The total barrier for electron conduction is Φ_E, where
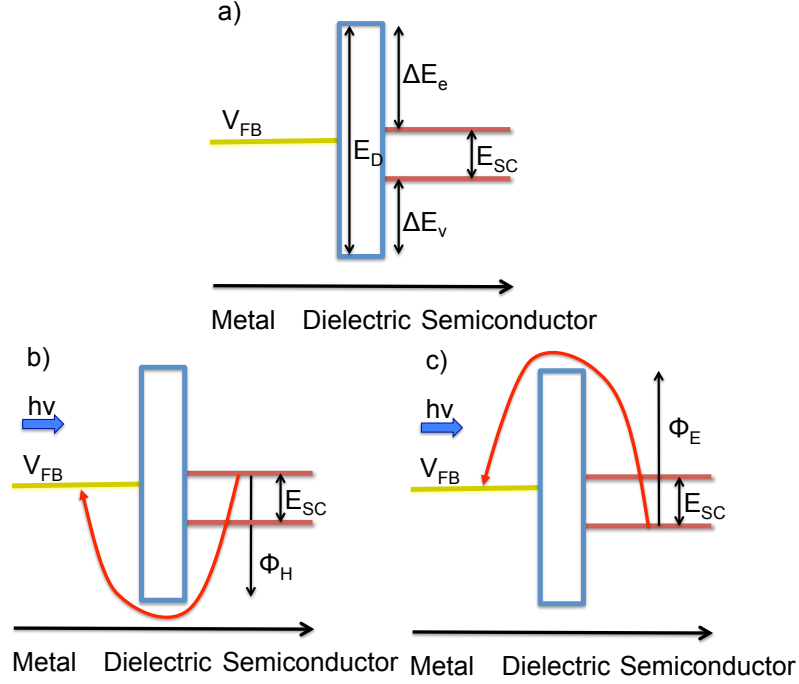
Figure 3.3: Schematic for internal photoemission (IPE). a) Flatband diagram showing band offsets, ΔE_V and ΔE_C, between a dielectric and semiconductor. b) Illumination at wavelength ν excites a hole from the conduction band of the semiconductor over the barrier and into the metal. c) Electrons traversing the barrier from the valence band into the metal

$$\Phi_E = \Delta E_V + E_{SC}. \tag{3.2}$$

For these types of experiments, the top metal contact is generally kept optically thin. This allows the light in to be absorbed by the semiconductor, and keeps electrons and holes from being excited in the metal, where they could traverse the barrier in the opposite direction. Avoiding these processes greatly simplifies the data analysis.

Data of this kind has been taken for many different materials, as this data is of great interest to researchers looking for high dielectric constant ('high-κ') materials to replace silicon oxide in CMOS devices. Please see Robertson [35] for a more complete discussion of this topic. For our purposes we seek a pair of materials with dissimilar band gaps and band offsets to create a triangularly shaped barrier. Figure 3.4 is a table taken from Afanas'ev [27] summarizing IPE work on various dielectric materials on silicon. Most of the candidate materials listed here have similar band gaps around 5.6 eV, with conduction band offsets around 2 eV and valence band offsets around 2.5 eV. However, Al_2O_3 and SiO_2 are both promising candidates for a middle layer due to their larger band gaps.

Oxide	E_g	Φ_e	ΔE_C	Φ_h	ΔE_V
SiO_2	8.9	4.25	3.15	5.8	4.7
Al_2O_3	6.2	3.25	2.1		
ZrO_2	5.5	3.1	2.0	3.6	2.5
HfO_2	5.6	3.1	2.0	3.6	2.5
Sc_2O_3	5.6	3.1	2.0	3.7	2.6
$LaAlO_3$	5.7	3.1	2.0	3.7	2.6
$LaScO_3$	5.7	3.1	2.0	3.6	2.5
$GdScO_3$	5.6	3.1	2.0	3.6	2.5
$DyScO_3$	5.7	3.1	2.0	3.6	2.5
$La_2Hf_2O_7$	5.6	3.2	2.1	3.5	2.4
	5.9	3.15	2.0	3.7	2.6
Lu_2O_3	5.8	3.2	2.1	3.7	2.6
	5.4	3.1	2.0	3.3	2.2
$LaLuO_3$	5.5	3.1	2.0	3.4	2.3
Gd_2O_3	5.8	3.2	2.1	3.9	2.8

Figure 3.4: Band gaps and band offsets for various candidate dielectrics, from Afanas'ev [27]. The offsets ΔE_V and ΔE_C are measured with respect to silicon, and the total barriers Φ_e and Φ_h are obtained by adding the Si band gap.

3.5 Lanthanum Oxide Films

For our first set of experiments we chose to use $LaAlO_3$, La_2O_3, and Al_2O_3 films grown on n$^+$-Si, with the La films being candidates for the lower offset barriers and the Al_2O_3 a candidate for the higher offset film. Our films were grown by molecular beam epitaxy (MBE) at Penn State in the group of Darrell Schlom, who was a collaborator on this project. His group has extensive experience growing these oxides [36] and can grow them on silicon without oxidizing the Si at the interface [37].

We first performed current-voltage measurements on single-layer films by evaporating a 15 nm gold layer on top of a 40 nm $LaAlO_3$ film and a 20 nm La_2O_3 film. The thin gold film functions both as a contact for I-V and a conductive transparent window for IPE measurements. The I-V results are shown in Fig. 3.5, and show that the leakage current in the La_2O_3 film is over four orders of magnitude lower than the $LaAlO_3$ film at 0.5V, even though the La_2O_3 film is only half as thick. Practically, this makes the $LaAlO_3$ layer unsuitable for a tunable barrier because the leakage through the film would not only make the photoemission data noisy, but would leak power from the device. The reason for the disparity in leakage currents is likely due to the amorphous nature of the $LaAlO_3$, which increases the scattering of hot electrons present in the film. The La_2O_3 film, however, is epitaxially matched to the Si, which reduces hot carrier scattering.
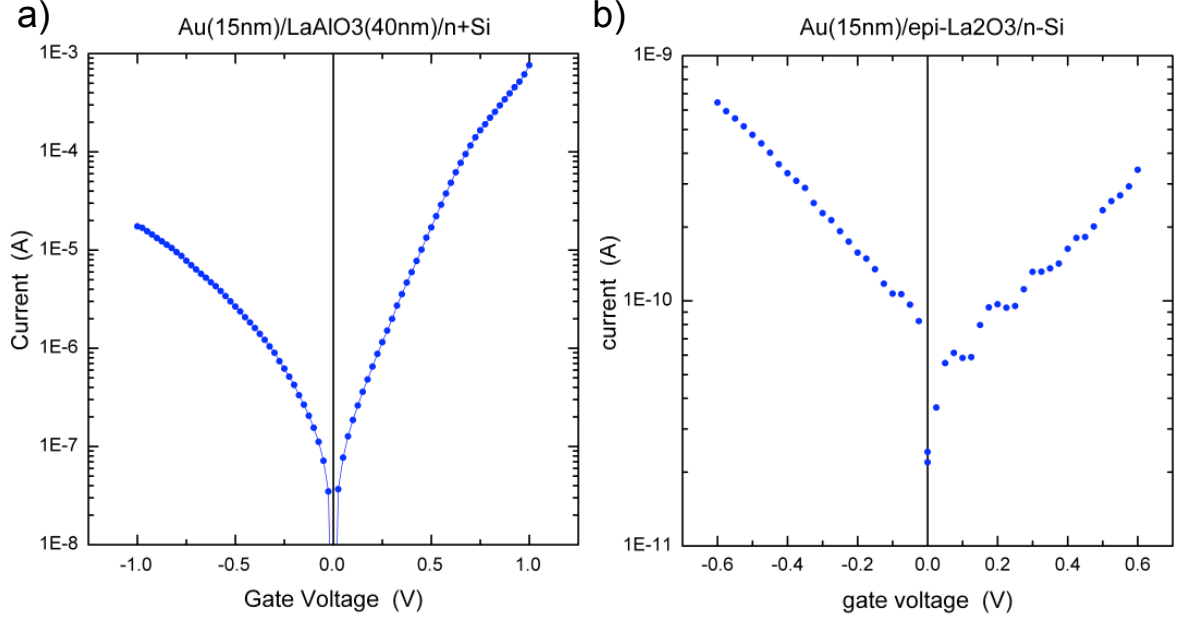
Figure 3.5: IV curves for LaAlO$_3$ and La$_2$O$_3$ on Si. The LaAlO$_3$ is too leaky for practical use, while the La$_2$O$_3$ shows much smaller leakage.

We also performed IPE measurements on these films to measure band offsets and band gaps. All the IPE measurements for this work were performed by Doug Bell at JPL. Due to the high leakage, no useful IPE data could be obtained for the LaAlO$_3$ film. We were able to obtain IPE data for the La$_2$O$_3$ film, and it is shown in in Fig. 3.6 at zero applied bias. The data shows the square root of the yield as a function of photon energy, from 2.5 eV to about 6 eV. The yield has units of A/W, or current per power incident on the device. IPE thresholds generally take the form of a power law of the form:

$$Y(h\nu) = A(h\nu - \Phi)^p \tag{3.3}$$

where Y is the yield, ν is the frequency of the excitation beam, Φ is the threshold for current conduction, and the exponent p is dependent on the functional form of the excited carrier energy distribution at the emitter surface — in our case the emitting surface is silicon [34]. Generally the data is plotted assuming a quadratic ($p = 2$) or cubic ($p = 3$) dependence, and when plotted as such will show a straight line in the plot of yield. Figure 3.6 shows a straight line using a quadratic form, with the a x-intercept at 3.58 eV. Due to the sign of the current, this is the threshold for emission of holes over the valence band barrier. Subtracting the Si band gap we obtain a ΔE_V of 2.5 eV. There is also a higher energy threshold around 5–6 eV, which corresponds to band gap excitation in the La$_2$O$_3$.

Even though leakage current through this layer is small, the efficiency of hot carrier transport
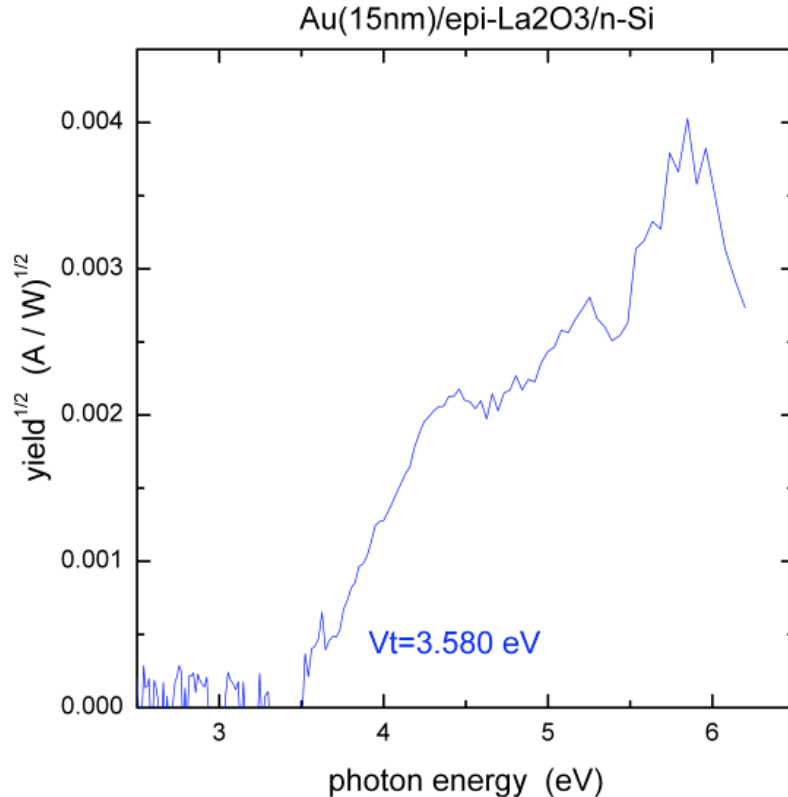
Figure 3.6: Internal photoemission of La_2O_3 on n-Si

under illumination is quite small, as shown by small magnitude of the yield in Fig. 3.6. We think this is likely due to the reactivity of lanthanum films with atmospheric water [38]. To remedy this we also fabricated a structure with 20 nm of La_2O_3 capped by 20 nm of Al_2O_3. The Al_2O_3 works as a diffusion barrier which prevents the La_2O_3 from exposure to the ambient, and also allows us to test a bilayer structure that may exhibit barrier lowering.

To measure barrier lowering, we perform IPE at a range of applied biases to the structure. If the two materials have offsets that are dissimilar enough, the conduction threshold, Φ, will change as a function of applied bias, indicating that the voltage has pulled the barrier down (or pushed it up). For these experiments the leakage current can be low enough so that current flow in the device is primarily due to photoexcited carriers.

The results of this experiment are shown in Fig. 3.7. In Fig. 3.7a, which plots the yield linearly, we see that the collection efficiency is two orders of magnitude larger than for the sample with La_2O_3 alone, which we attribute to the Al_2O_3 capping layer keeping the La_2O_3 from reacting with the ambient. We also see a large feature around 5–6 eV, which we attribute to band edge absorption in either or both layers. The onset of conduction happens at a much smaller value of around 3 eV, and this data is shown in more detail in Fig. 3.7b. As the applied voltage increases from 1.5 V (top
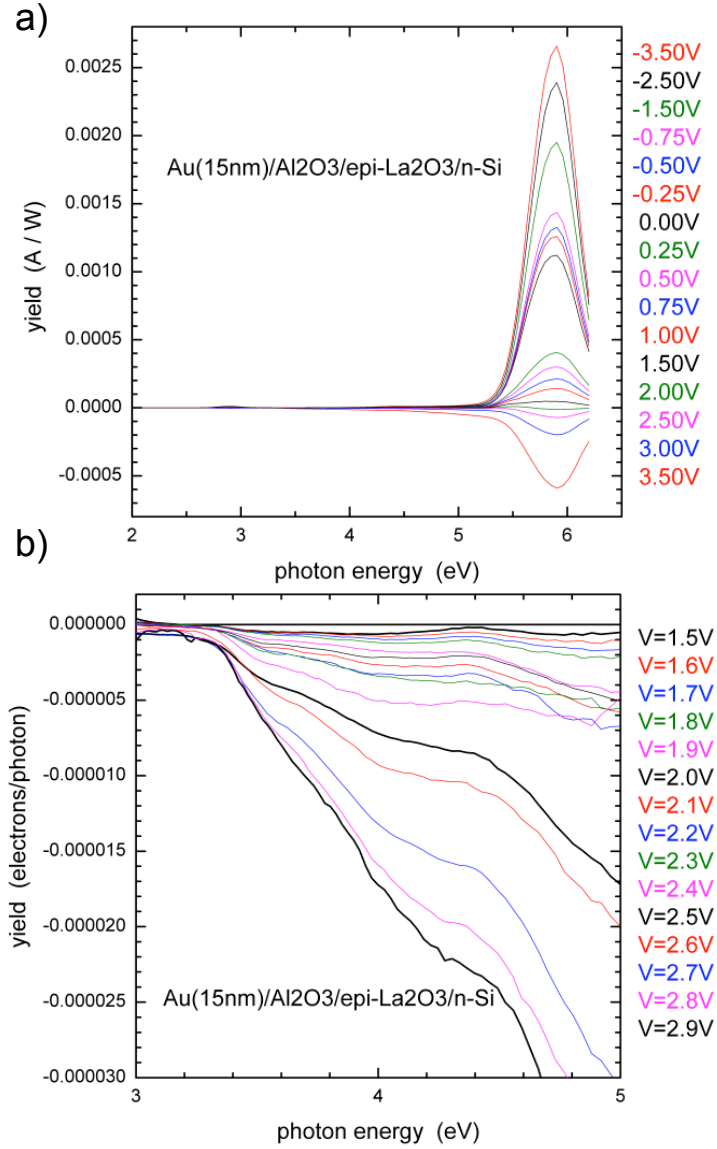
Figure 3.7: Internal photoemission of a La_2O_3/Al_2O_3 bilayer. a) The feature around 6 eV is due to band edge absorption. The collection efficiency is much higher than for the single La_2O_3 sample, indicating the Al_2O_3 has protected the La_2O_3 film. b) The onset of conduction around 3 eV shows no barrier-lowering effect, meaning the band offsets of the two materials are very similar.
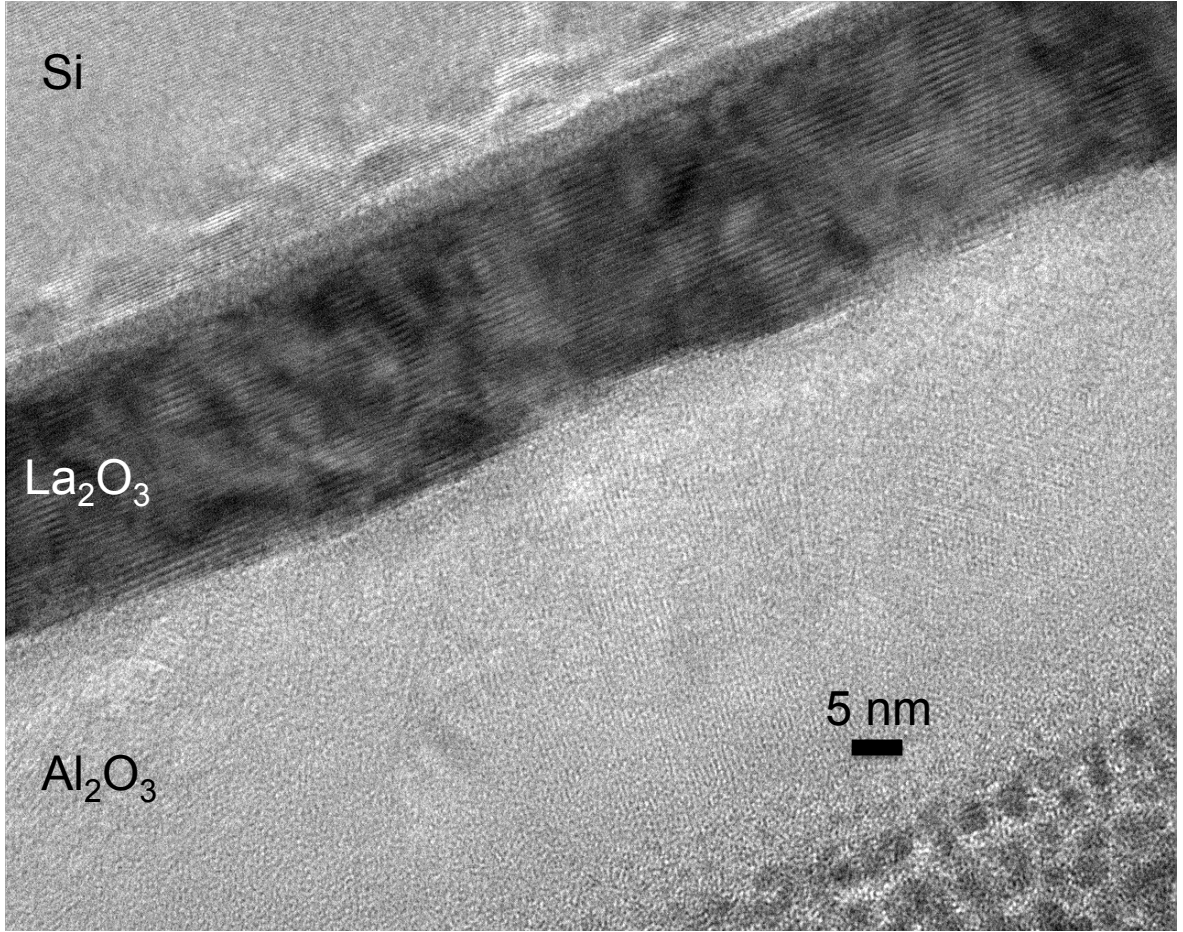
Figure 3.8: Transmission electron micrograph of $Si/La_2O_3/Al_2O_3$ structure showing epitaxial regions of La_2O_3 and Al_2O_3, as well as amorphized interfacial regions at each interface

black line) to 2.9 V (lowest black line) we see that the yield, now plotted as a quantum efficiency of electrons per photon, increases linearly in the negative direction. Because the current is negative, we are seeing electron transport from the silicon substrate to the metal contact, and the threshold at 3.3 eV corresponds to the total energy barrier for electrons. This indicates that the total conduction band offset is around 2.2 eV for this structure.

However, this experiment shows that there is no barrier lowering because at all voltages the threshold energy is 3.3 eV. This likely means that the conduction bands of the two materials line up very closely, and the increase in conduction as a function of voltage is due to Fowler Nordheim tunneling alone.

To get a better understanding of these films we analyzed the La_2O_3/Al_2O_3 bilayer sample using transmission electron microscopy (TEM). A TEM image of the $Al_2O_3/La_2O_3/Si$ structure is shown in Fig. 3.8. The image shows lattice fringes in the bulk areas of the Si, La_2O_3, and Al_2O_3, but at each interface is an amorphized layer, possibly due to the growth kinetics of the dielectrics during
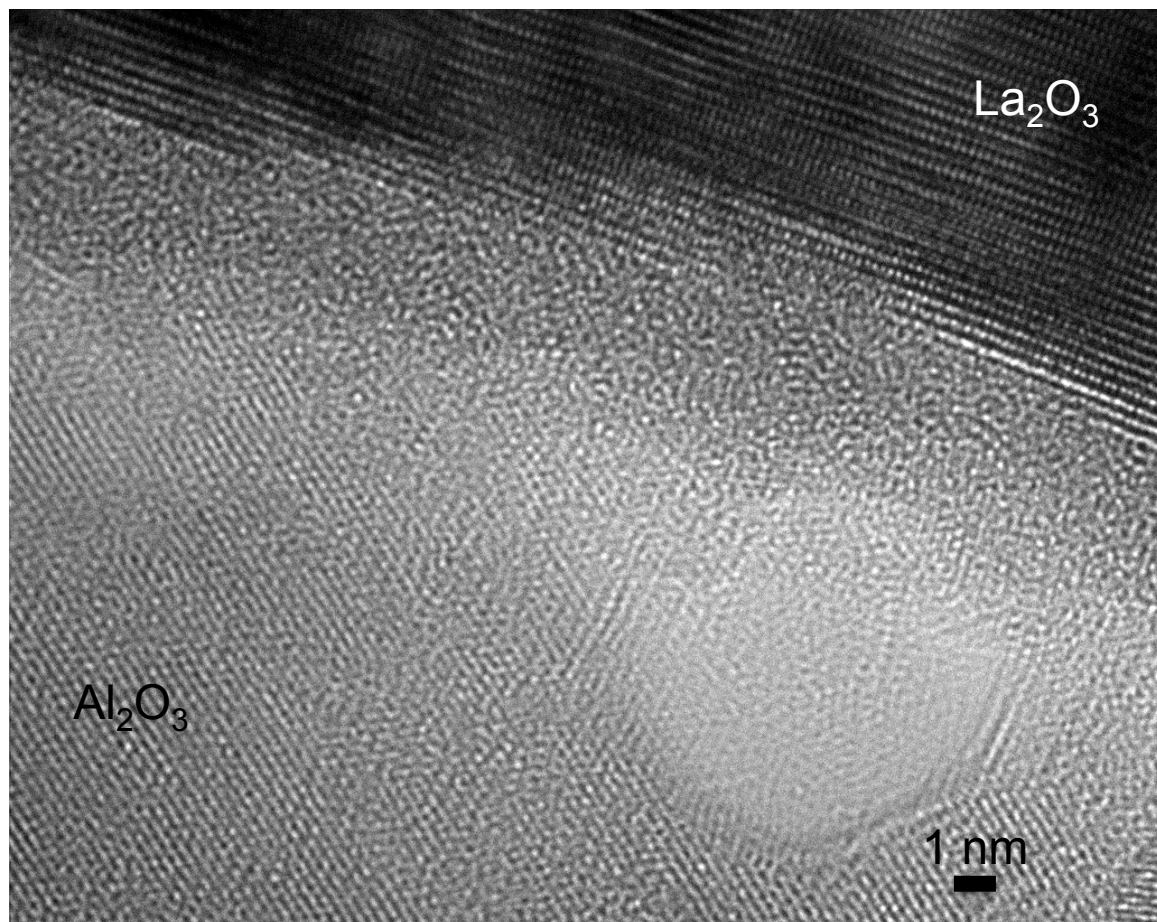
Figure 3.9: Transmission electron micrograph of the La$_2$O$_3$-Al$_2$O$_3$ interface showing voids in the Al$_2$O$_3$, as well as an amorphized region at the interface

deposition.

At the Si/La_2O_3 interface, it is possible that exposure to atmosphere during TEM sample preparation is partly responsible for the interfacial layer there. As mentioned before [38], La_2O_3 is reactive to oxygen and water in atmosphere causing film degradation and roughness. SEM images taken on our single-layer La_2O_3 on Si sample show clear bubbling and delamination, possibly due to atmospheric contamination. We also see bubbling in the bilayer sample of La_2O_3 with an Al_2O_3 cap. This could be due to stresses induced between the films during deposition, degradation of the La_2O_3 film, or possibly poor adhesion of the La_2O_3 layer to the Si substrate.

The interface between the La_2O_3 and Al_2O_3 is shown in a magnified TEM image in Fig. 3.9. Here the lattice fringes showing the crystallinity of the two materials are more apparent, as is the 2–4 nm amorphized layer between the two. Also present at this interface are voids about 5 nm in diameter. These voids may also be nucleation sites for delamination of the films.

Even though we expected that the offsets of these two materials would enable barrier lowering simply by looking at the offsets shown in Fig. 3.4, we can speculate as to why this did not happen. One may be the comparative lack of data on band offsets between various dielectrics. The values for the offsets given in Fig. 3.4 are all taken relative to silicon. There is also extensive data growing these dielectrics on GaAs. This is mainly because the primary research interest in these films are as high-κ replacements for SiO_2 in the CMOS platform [35], which has caused great interest in the interfaces of these films with semiconductors. However the band alignment between the dielectrics themselves has been largely unexplored, and the interface kinetics between the dielectrics is as important as the kinetics between the dielectric and semiconductor, because in both cases we are interested in tightly controlling the current flow between the layers. The TEM results on this structure showing amorphous regions at the interfaces is likely altering current conduction between the films by trapping charge. The amorphous layers also change the energy profile of the stack. The barrier-lowering effect is based on putting layers of different band offsets together, and while we know something about the band alignment of bulk La_2O_3 and Al_2O_3 layers, the band alignment of these amorphous layers is likely different and unknown at this time. The presence of voids and delamination in the films could also be a contributing factor. It is also possible that stresses in the bilayer structure could be changing the band structures.

3.6   Barrier Lowering in Sc_2O_3 Films

As a result of the measurements on La_2O_3 which showed no barrier lowering, we decided to try another epitaxial dielectric, Sc_2O_3, for voltage tunability purposes. Sc_2O_3 does not degrade as much in air as La_2O_3, and is well grown and characterized by our collaborators in Prof. Schlom's group at Penn State [39]. Furthermore, the MBD epitaxy method used deposits films directly in Si,
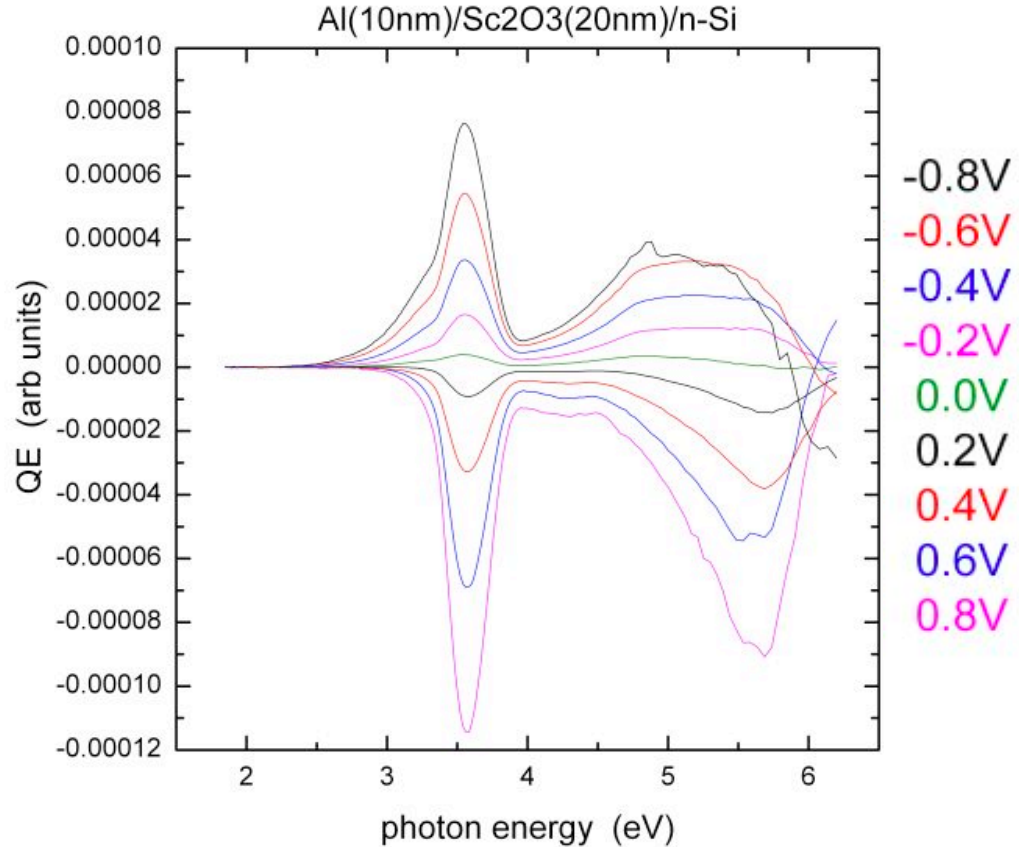
Figure 3.10: IPE of Sc_2O_3. The high energy peak corresponds to a band gap energy of 3.9 eV, while the lower energy peaks correspond to excitation over the valence and conduction band barriers. The values extracted for ϕ_{VBO} and ϕ_{CBO} are 1.14 eV and 1.63 eV, respectively.

without an interfacial layer [37].

The Sc_2O_3 layers were grown by the Schlom group at Penn state and are 20 nm thick on n^+ Si. IPE characterization was carried out by Doug Bell at JPL and is shown in Fig. 3.10. The broader peak at 4–6 eV corresponds to band-edge absorption and carrier generation in the Sc_2O_3 layer. The threshold for this behavior is the band gap and has a value of 3.9 eV for the films used here. Note that this number differs significantly from a value of 5.6 eV in the literature [27] and is actually closer to the value for amorphous Sc_2O_3 [40], even though these layers are epitaxial [39].

The peaks in Fig 3.10 between 3 and 4 eV correspond to carrier generation in the substrate and subsequent conduction over the Sc_2O_3 barriers. Negative gate voltages result in positive induced current, and represent holes being photoexcited from the silicon conduction band over the valence band barrier. A cubic fit yields a threshold barrier energy of 2.26 eV. Subtracting the silicon band gap energy of 1.12 eV gives a valence band offset of 1.14 eV. Positive gate voltages give a negative induced current, as electrons jump from the silicon valence band to the metal, and the threshold for this action is 2.75 eV. Subtracting the silicon band gap gives a conduction band offset of 1.63 eV.
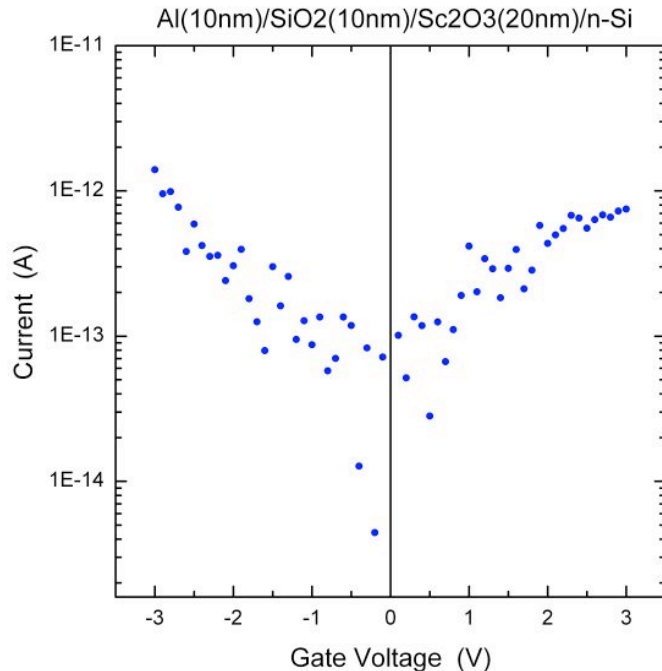
Figure 3.11: IV curve of Sc_2O_3 with SiO_2 capping layer. The current is much lower than Al_2O_3-capped samples due to the larger band gap of SiO_2.

Adding these offsets to the silicon band gap should give the total band gap, and indeed they do add up to the band gap energy of 3.9 eV found by IPE in the same data set.

The data taken from this film is already much cleaner than a similar experiment performed on La_2O_3 (Fig. 3.6). We are able to bias these these films without excessive charging or leakage, which had been problems for La-based dielectrics. These biasing measurements allow us to calculate both band offsets and a band gap that are all consistent with each other, which we also could not do with the La films. Furthermore, the magnitude of the Sc_2O_3 offsets are ideally suited for a tunable barrier — both offsets are greater than 1 eV, which is necessary to prevent leakage during biasing, while still being low enough that we expect to see lowering by using a material with higher offsets, such as Al_2O_3 or SiO_2.

In our first attempt to observe barrier lowering, the Schlom group at Penn State fabricated structures with 20 nm Sc_2O_3 capped with 20 nm of Al_2O_3, similar to the La_2O_3 structure analyzed in the previous section. Unfortunately, IPE studies on those samples also showed no barrier lowering, so we decided to try using SiO_2 as the high-barrier layer. SiO_2 is a viable candidate due to its large band gap of 8.9 eV, and its large band offsets with silicon [41]. However, SiO_2 is an unknown quantity with respect to most of the candidate layers proposed for this work, because most studies of these films are done for the express purpose of replacing SiO_2 due to its low dielectric constant. But for barrier-lowering purposes, the dielectric constant is of secondary concern, as we wish to exhibit only a voltage tunable barrier.
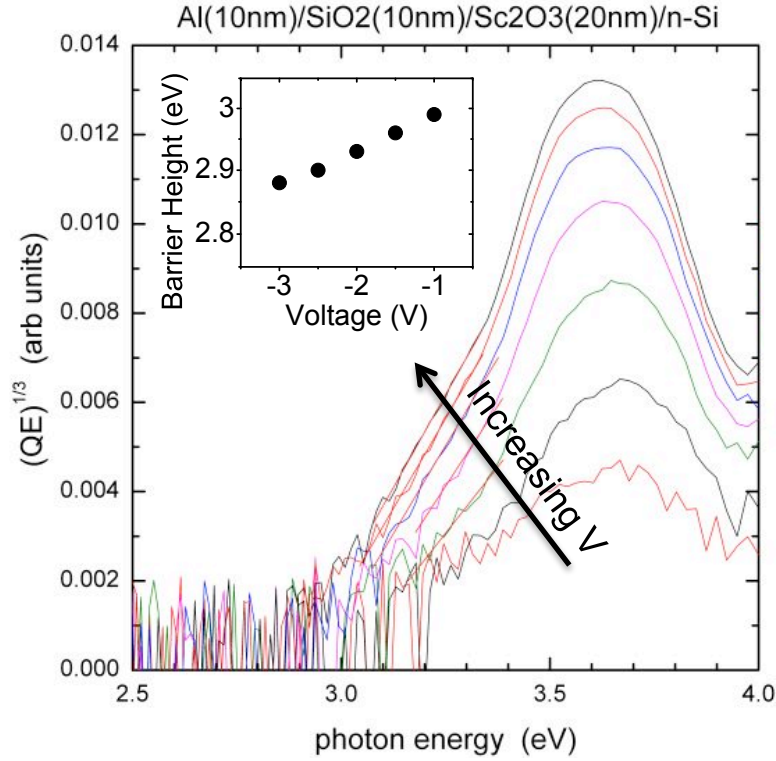
Figure 3.12: Voltage dependent IPE of Sc_2O_3 with SiO_2 capping layer. The barrier height decreases with increasing negative voltage as the barrier is lowered. The inset shows the dependence of the barrier height on voltage.

Deposition of the SiO_2 layers was done at Caltech using radio-frequency (RF) magnetron sputtering on Sc_2O_3 films grown by the Schlom group by MBE. The films were 10 nm thick and deposited at room temperature. A current-voltage scan is shown in Fig. 3.11 which shows a very low current out to 3 V. This is due to the larger band gap of SiO_2.

The internal photoemission results on this sample are shown in Fig. 3.12. The peak shown is positive photocurrent for a range of negative voltages, which is the result of holes excited in the silicon traversing the valence band barrier into the metal. The threshold for this is around 3 eV, or the valence band offset is 1.9 eV. The presence of the SiO_2 layer has increased the valence band offset by over 0.7 eV from its value of 1.14 eV for a single Sc_2O_3 layer. The IPE traces at various negative voltages are shown ing Fig. 3.12 on a cubic set of axes. The extracted barrier heights are shown in the inset, and show a linear decrease in the barrier height as the voltage is swept from -1.0 V to -3.0 V. This is the barrier-lowering effect expected from a bilayer structure with differing offsets. The valence band barrier has changed by 0.11 eV between the applied voltages of -1.0 V and -3.0 V.

While this structure shows a larger valence band offset and barrier lowering at negative voltages, at positive voltages we see the same energy barrier and hence the same conduction band offset as

that of a single Sc_2O_3 layer. This indicates that the conduction bands of the two materials are very closely aligned, and as a result we see no barrier lowering in the conduction bands.

This paints an energy band picture that is very different from what is expected for SiO_2. Our experiments show the valence band offset for the bilayer Sc_2O_3/SiO_2 stack to be 1.9 eV while the conduction band offset is 1.6 eV. Adding these to the silicon substrate band gap of 1.1 eV gives a total band gap of 4.6 eV. While this gap is larger than the 3.9 eV band gap of Sc_2O_3, it is significantly less than the common 8.9 eV band gap of SiO_2. We attribute this discrepancy mainly to the oxide quality of the sputtered SiO_2. During biased measurements we observed significant charging in the bilayer structure that was not present in the single layer Sc_2O_3 layer, which is likely due to a high defect density in the deposited SiO_2 film.

We attempted sputtering the oxide layer at a higher temperature of 400C without seeing a significant improvement in the layer transparency or barrier energy, suggesting that high temperature deposition doesn't reduce the defects in the layer enough to improve the device quality. It is likely that at even higher deposition temperatures a higher quality film could be grown because high quality oxides are usually produced by oxidizing silicon at 900–1200 °C, but at such high temperatures the Sc_2O_3 film breaks down. As a result we must conclude that these candidate films as grown are unsuitable for a voltage tunable detector.

3.7 Conclusion

In this chapter we explored a series of candidate films for use in a voltage-tunable hyperspectral detector using layered tunnel barriers. By creating a crested energy barrier, it is possible to change the barrier height by applying a voltage. This voltage-tunable barrier can be used as a high-pass filter for photoexcited electrons, making useful as a pixel element in a hyperspectral detector. A device based on these layered barriers would eliminate the need for external gratings or filters typically used to resolve spectral data. Our proposed device would isolate the spectral data in time using a voltage sweep, greatly reducing the complexity of the device.

We tested $LaAlO_3$, La_2O_3, and Sc_2O_3 for use as small band gap, small offset films in conjunction with higher offset Al_2O_3 and SiO_2 films to create prototype crested barrier structures. The $LaAlO_3$ layers tested exhibited high leakage currents that made IPE measurements difficult and make these films unsuitable for barrier-lowering purposes. We also tested La_2O_3 films and found that the leakage was acceptable in these films for IPE measurements. However, our experiments using Al_2O_3 at a top layer did not result in any barrier lowering. Our TEM results suggest this is due to amorphous layers at both the Si/La_2O_3 and La_2O_3/Al_2O_3 interfaces.

Experiments using Sc_2O_3 yielded much better results, as these films exhibited low leakage currents and were robust to a wide range of biasing during IPE measurements, and we were able to

well characterize their band alignment. We measured a band gap of 3.9 eV with a conduction band offset of 1.14 eV and valence band offset of 1.63 eV relative to silicon. We demonstrated barrier lowering in the Sc_2O_3/SiO_2 structure, although the magnitude of lowering is less than expected or required for a practical device.

We attribute the lack of suitable barrier lowering in our films mainly due to the growth quality of the films. That we see barrier lowering in these films is evidence that with a proper choice of materials a voltage tunable barrier could be fabricated, but the amorphous layers between La_2O_3/Al_2O_3 shown by TEM and charging effects in Sc_2O_3/SiO_2 bilayers suggest that the lack of performance in these films are due to defects at the interfaces of these films which cause hot carrier scattering. As such we believe that the performance of these films could be greatly improved by using higher quality materials with more tightly controlled interfaces. It is also possible that other candidate films would yield better results, but there is also the issue that, simply put, most of these materials have very similar band gaps and offsets. This greatly reduces the space of materials we can combine to create a crested barrier structure.

One alternate path forward would be to use material systems where the band gap can be varied by adjusting the film composition, such as the AlInGaN systems, where the band gap can be varied from 0.7 for InN to 6.2 for AlN [42, 43]. In fact, the AlGaN system has shown very controlled barrier-lowering behavior by constructing a crested barrier with varied Ga levels [44]. Such systems can be grown epitaxially, which reduces hot carrier scattering at the interfaces; by using materials in the same system, the band offsets are better understood and more tightly controlled.

Chapter 4

Erbium-Based Silicon Laser Design

4.1 Introduction

The development of high-speed, high-bandwidth fiber optic cabling allows us to transfer large amounts of data at light speed all around the world. Transmitting information using light instead of electricity is faster, more energy efficient, and has ushered in the age of instant global communication. Furthermore, the desire to connect these optical components to the chip scale electronics that run cell phones and computers has inspired research toward a chip-scale, silicon-based light source [45, 46]. A silicon light source would allow easy integration between microprocessors and optically based long-haul components. It would also open the door to chip-scale optical computing, which potentially could mean faster processors that use less energy than current electrical devices.

In this chapter we present theoretical modeling of a prototype CMOS-compatible, Er-based, electrically pumped laser on the silicon platform. We start with an introduction to silicon photonics before describing the basic elements of our proposed structure: erbium-based gain materials and silicon slot waveguides. We then present a rate equation approach to erbium inversion in a silicon waveguide. We show that lasing is possible in this system after accounting for carrier absorption and local density of optical states (LDOS) effects only by using a pulsed injection scheme. The single biggest technological hurdle for an actual device is efficient electrical excitation of erbium, and experimental work towards electrical injection of Er is presented in Chapter 5.

4.1.1 Light From Silicon

Semiconductor lasers were first developed using materials like GaAs [47, 48, 49] and GaAsP [50]. These devices work under electrical pumping because the radiative recombination rate of electron-hole pairs, or excitons, is very high in the bulk. This is due to the direct band gap structure in these materials, which allow excitons to combine directly and emit photons. Silicon, however, has an indirect band gap, which means that photon emission via exciton annihilation requires a lattice

phonon to conserve crystal momentum. This three-particle mechanism makes the radiative efficiency of silicon very low, and lasing very difficult.

A laser has been demonstrated on the silicon platform in two ways already. The first is via the Raman scattering mechanism [8, 9], where a pump laser is efficiently scattered by Si to a longer wavelength. The output wavelength is tied to the pump wavelength by the magnitude of the Stokes shift in silicon. This is an intrinsically optical process which requires a high powered pump laser.

The second method involves wafer bonding a III-V based laser diode to a silicon waveguide [10, 11]. This method evanescently couples the output of a AlGaInAs laser into a silicon waveguide. This hybrid laser has the advantage of being fully electrically pumped, but currently suffers from a relatively low maximum operating temperature of 40 °C.

There are also two main areas of research towards a fully CMOS-compatible silicon light source. The first uses silicon nanostructures. By confining the exciton in space, the momentum of the particle is decentralized by the uncertainty principle, and it is then easier to bridge the momentum mismatch needed for radiative emission from silicon's indirect band gap. The first example of this is the observation of visible photoluminescence (PL) from porous silicon [51]. Since then the field has moved towards using silicon nanocrystals embedded in SiO_2 [15], where the emission wavelength is controlled by the size of the nanocrystal, varying from about 600–900 nm. Studies on these films have shown gain [52] and stimulated emission [53] under optical pumping, and can be electrically pumped via impact excitation [54, 55] and bipolar injection [18, 19, 56]. However, an electrically pumped silicon laser based on nanocrystals is still lacking.

A second avenue to a silicon light source uses rare earth doping of silicon materials, and this method is the focus of this chapter. Of particular interest is erbium, which emits around 1.5 μm. This is the standard wavelength for the telecommunications industry because the loss in silica fiber is lowest at 1.5 μm. A laser using Er as the gain medium would be instantly compatible with current telecommunications technology. Silicon is also transparent at 1.5 μm, which enables the use of silicon as a waveguiding element for integrated optical components.

4.1.2 Erbium-Based Gain Materials

Erbium is the standard gain material for long-haul fiber optic communication due to its emission at 1.5 μm. The 1.5 μm emission in Er is due to the ground state atomic transition between the $^4I_{13/2}$ and $^4I_{13/2}$ levels. The energy diagram for a free atomic Er is shown in Fig. 4.1. One of the curious properties of erbium is that, similar to all rare earths, its lowest atomic transitions occur in the 4f manifold, but this manifold is shielded by the $5s^25p^6$ electrons. This gives rare earths an atomic-like optical behavior in solids, because the 4f electrons involved in luminescence are partly shielded from bonding to the matrix. In fact, most rare earths bond in a triply ionized state by losing their $6s^2$ electrons and one of their 5d or 4f electrons. So in most host materials, erbium is
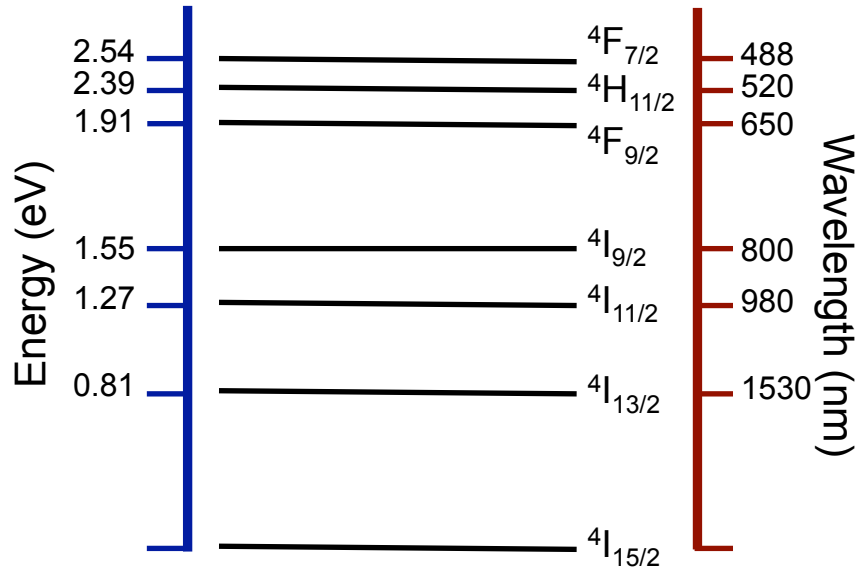
Figure 4.1: Energy levels of erbium

found in the Er^{3+} state. This atomic-like behavior also means that Er must be resonantly pumped to emit light efficiently, and generally this is done by pumping the $^4I_{13/2}$ or $^4I_{11/2}$ manifolds with expensive lasers operating at either 1480 nm or 980 nm, respectively.

Another important property of Er is that, to first order, the intra-4f transitions are parity forbidden. This restriction is partially lifted by the host matrix, which allows the transitions to occur, but also results in very long luminescent lifetimes for erbium. Values in the few millisecond range are common, with values as high as 15 ms in glass [57]. These long lifetimes make it easier to invert erbium for lasing and amplification applications.

However, the reason a chip scale Er-based laser has not been developed is because the gain achievable by Er in thin films is actually quite small. This is mainly due to two effects: first, the excitation cross section, σ, is very small for Er, generally on the order of 10^{-21} cm^2 [58]. The second effect is that Er clusters at concentrations of more than about 2% [57], which leads to proximity effects like concentration quenching [59] and cooperative upconversion [60], which decrease the radiative lifetime of Er. These effects are mitigated in fiber amplifiers by using very low Er concentrations and long fibers. The low concentration limits Er proximity effects, and meter-scale fiber increases the interaction length so the Er can be inverted despite its low cross-section.

The prospect of a thin film Er gain material was expanded greatly when Fujii et al. observed Er sensitization by silicon nanocrystals in 1997 [61]. Silicon nanocrystals show optical properties in the visible spectrum between 600–900 nm [15], and Fujii showed that when these films were co-doped with erbium, a decrease in the visible nanocrystal luminescence was accompanied by an increase in the infrared Er emission. The plot from that paper is shown in Fig. 4.2. The Er emission was also up to two orders of magnitude larger than that seen in films that had no excess silicon. Infrared emission
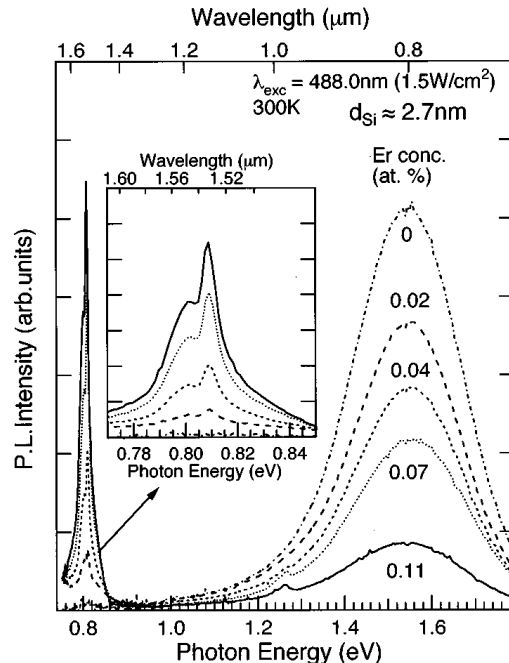
Figure 4.2: Erbium sensitization by silicon nanocrystals. From Fujii et al. [61]. Films with only nanocrystals show visible luminescence, which decreases as Er is added. A corresponding increase in the infrared Er luminescence is evidence of sensitization by the nanocrystals.

was also seen for pump wavelengths that do not correspond to Er energy levels. This all suggests broadband absorption in the nanocrystals, followed by energy transfer to nearby Er ions. Subsequent work by Franzo et al. showed that the the excess silicon need not be crystallized for sensitization to take place. This allowed further optimization of the coupling by fabricating Er-doped silicon nanocluster layers, where it was found that generally smaller, more numerous amorphous silicon clusters were the best way to sensitize Er in the film [62].

This method of indirectly pumping erbium generated a new area of research. It was found that the effective cross section, σ_{eff}, of this process was very close to the nanocrystal cross section, $\sigma \sim 10^{-16}$ [63]; five orders of magnitude higher than that for Er in SiO_2. This also proved to be a way to electrically access the Er ions, as the conductive silicon inclusions increased the electrical conductivity of the films [64]. There is also evidence of optical gain in a rib waveguide [65], although this result has not been repeated. Papers were also devoted to measuring and increasing the optically active fraction of Er ions that are sensitized by silicon nanoclusters [66, 67, 68]. Even though the fraction of sensitized Er is still below the 50% required for lasing, that number is increasing, and these results have increased optimism that a device could be built that takes advantage of this sensitization to electrically invert erbium and produce gain in a thin film structure.

a)

Si

Er:Si-nc

Si

SiO$_2$

b)

Dominant field orientation:

TE-like Mode

58 nm

800 nm

SiO$_2$

Si

125 nm

25 nm

SiO$_2$

Si

149 nm

SiO$_2$

$n_{\text{eff}} = 2.84$

TM-like Mode

SiO$_2$

Si

SiO$_2$

Si

SiO$_2$

$n_{\text{eff}} = 2.06$
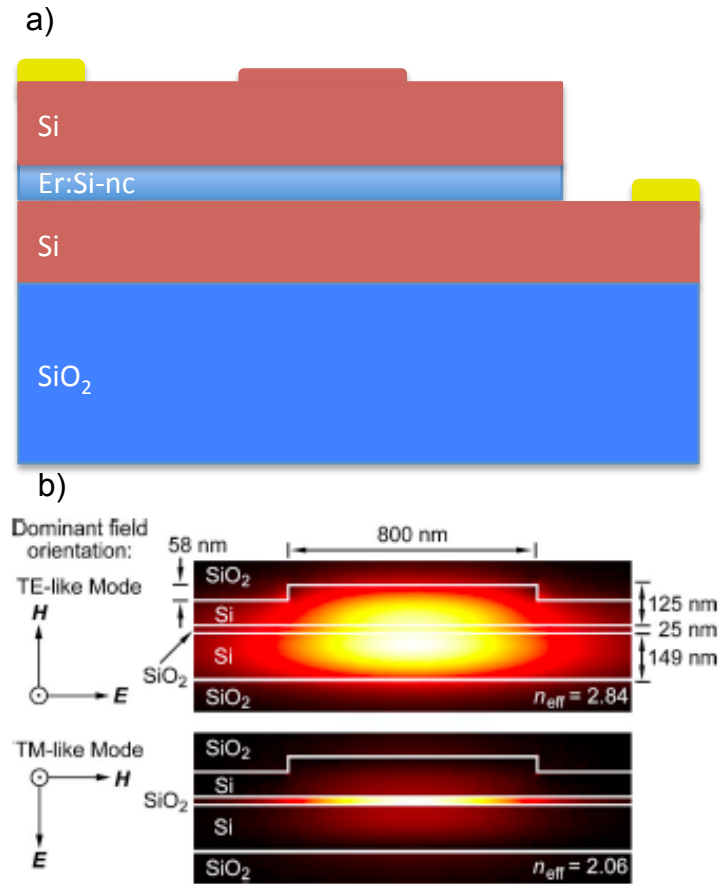
Figure 4.3: Schematic of laser structure using a slot waveguide. a) The waveguide consists of an Er:Si-nc slot layer between two Si layers. Light is confined vertically to the two silicon layers, and laterally by the small ridge on top. Electrical injection is provided by the metal contacts, spaced away from the lasing mode. b) From Briggs [69], showing TE- and TM-like modes in a slot waveguide, showing tight confinement in the slot layer for the TM-like mode

4.1.3 Silicon Slot Waveguides

We propose the use of a silicon slot waveguide for use in our silicon-based laser. A schematic of the device is given in Fig. 4.3a. The slot waveguide design puts an active gain medium inside a silicon waveguide. We use Er-doped silicon nanoclusters (Er:Si-nc) for the gain medium. The silicon waveguide is transparent to the Er emission, and the two slabs of Si act as contacts for current injection [70]. Lateral confinement of the lasing mode is provided by the small ridge on top of the device, and a cavity can be formed by patterning the ridge into a ring on top of the device.

By choosing the dimensions of the waveguide we can ensure single-mode operation. Specifically, the TM_0 mode is very tightly confined to the active layer as shown by the electromagnetic simulation in Fig. 4.3b (from Briggs [69]). This ensures higher modal gain due to the high field in the active layer [71]. Most silicon optical devices use TE modes instead of TM modes because TM modes are more sensitive to bending loss and lateral leakage in ring resonators. This effect can be mitigated by a clever design of the waveguide by fixing the ridge width to the so-called 'Magic Width' [72] which causes destructive interference of the lateral leakage, leading to TE-like loss characteristics of TM modes.

For this laser to work, the gain provided by the Er:Si-nc must be greater than the loss inherent in this system. Low-loss waveguides have been fabricated in single-crystal silicon [73], amorphous silicon [74], and wafer-bonded silicon [69]. The last two methods are of interest here, as the fabrication of the slot structure can be achieved by either depositing amorphous silicon onto the active layer, or by wafer bonding a crystalline silicon slab to the active layer. Proximity effects limit the maximum Er concentration to values of order of 10^{20} cm^{-3}, and as a result the maximum gain we can expect from such a device is, to first order, less than 10 dB/cm. This is much less than the gain achievable in III-V structures, and puts strong requirements on the device loss.

The slot waveguide structure also dramatically changes the local density of optical states (LDOS) of any emitters placed in the slot. The LDOS at the emitter location is determined by the dielectric environment nearby, which leads to a change in the radiative rate of spontaneous emission of an emitter in the slot compared with vacuum [75, 76]. This effect has been described theoretically in silicon slot waveguides by Jun and Briggs [77], and demonstrated experimentally by our group and others [78, 79]. This effect has been accounted for in the design and modeling of our proposed device, as described below.

4.1.4 Carrier Absorption

Also of great importance is the less understood phenomenon of carrier absorption in Er:Si-nc layers. This is the analog to free carrier absorption in semiconductors, where free carriers absorb photons in the bulk [80]. In an electrically pumped Er:Si-nc layer, charged carriers will occupy the nanoclusters,

and these carriers can absorb the Er emission. What is not completely known is how the confined nature of the nanoclusters affects this absorption. This effect is intrinsically different from other sources of loss because it can't be engineered away like, for example, proximity effects in Er or waveguide scattering losses. It also can't be mitigated by increasing the pump, as the carrier density would also increase.

This process is characterized by a carrier absorption cross section, σ_{CA}, and its value at 1.5 μm has only been estimated to this point. A careful study of carrier absorption in silicon nanocrystals was carried out by Kekatpure et al., using microdisk resonators with silicon nanocrystals [81]. This method measured σ_{CA} by optically pumping the disks to generate carriers in the nanocrystals, then measuring the quality factor, Q, of the disks, and attributing the loss to carrier absorption. However, by measuring Q via nanocrystal emission, this study only directly measured σ_{CA} in the visible. An extrapolation of the data to 1.5 μm gives an estimate of $\sigma_{\mathrm{CA},1.5\mu m} \sim 8 \times 10^{-17}$ cm^2. There is also an estimate in the literature from Navarro-Urrios et al. [82], where they measure pump-induced loss in a Er:Si-nc waveguide and estimate a cross section of $\sigma_{\mathrm{CA},1.5\mu m} \sim 4 \times 10^{-19}$ cm^2. These values straddle the value at which a CW laser is expected to operate, as calculated by Pacifici et al. in 2003 [83]. The work in this chapter expands this work by accounting for the radiative rate enhancement due to the slot waveguide structure, and uses a pulsed excitation scheme to mitigate the effects of carrier absorption.

4.2 Rate Equation Modeling of Er:Si-nc Layers

We develop here a rate equation model for the coupled Er:Si-nc system, first as a bulk medium, and then as the slot layer in a silicon slot waveguide. The energy level scheme for the interaction is shown in Fig. 4.4, and the relevant parameters used in the calculation are listed in Table 4.1.

The Si-nc is depicted as a two-level system with energies E_{b} and E_{a} having populations n_{b} and n_{a}, respectively. The Si-nc cross section is σ_{nc}, and the Si-nc lifetime, τ_{nc}, is dependent on the dielectric environment via the LDOS.

The Er^{3+} ions are depicted as a three-level system with energies E_3, E_2 and E_1 having populations N_3, N_2, and N_1, respectively, where the highest energy level, E_3, is sensitized by the Si-ncs. This is a simplification of the Er spectrum shown in Fig. 4.1, which is justified because the lifetimes of the upper levels in the Er manifolds are much shorter than the lifetimes of these first levels; any excitations into the upper levels decay to E_3 and E_2 much faster than E_3 and E_2 decay to the ground state E_1, leaving these three levels as the only levels with significant populations. The lifetimes of these levels are τ_{21}, τ_{31}, and τ_{32}, with values given in Table 4.1. The values of τ_{31} and τ_{21} were measured experimentally by Pacifici [83] and are used here as the bulk values. The value of 3 ms for τ_{21} is an estimate consistent with that measured by other groups [68]. Also, because the magnitude
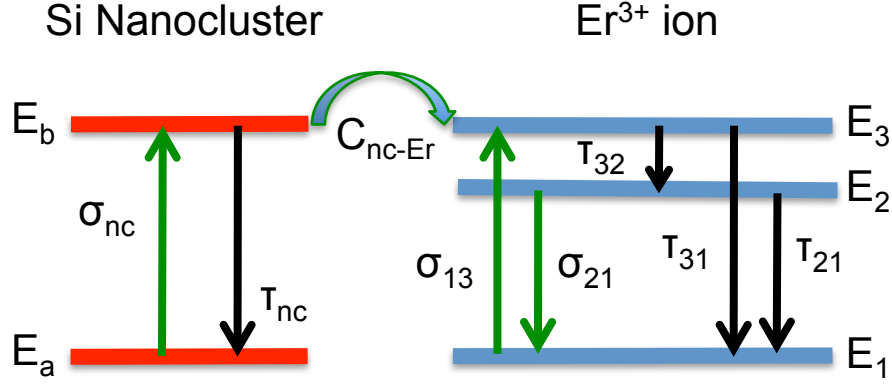
Figure 4.4: Energy levels for the Er:Si-nc interaction used for modeling. Erbium is shown schematically as a three-level system, the nanocluster as a two-level system, and they are coupled by C_{nc-Er}.

Table 4.1: Rate Equation Parameters

	Bulk	10 nm Slot: Optical Pumping	10 nm Slot: Electrical Pumping
τ_{nc} (μs)	50	7.1	7.1
τ_{21} (ms)	3	0.291	0.291
C_{nc-Er} (cm^{3}s^{-1})	3×10^{-15}	3×10^{-15}	3×10^{-15}
σ_{nc} (cm^2)	10^{-16}	10^{-16}	10^{-14}
σ_{13} (cm^2)	10^{-19}	10^{-19}	$-$
σ_{21} (cm^2)	8×10^{-21}	8×10^{-21}	8×10^{-21}
τ_{32} (μs)	2.38	2.38	2.38
τ_{31} (μs)	714	714	714

of τ_{31} is larger than τ_{32}, excitations into E_3 will preferentially decay to E_2 instead of E_1. Because E_2 decays to E_1 by emitting a photon at 1.5 μm, E_2 is the level we wish to invert for stimulated emission.

The decay rates of Si-nc films have been measured for bulk media [17]. Here we use a bulk Si-nc lifetime, τ_{nc}, of 50 μs. The nanocluster concentration is assumed to be 10^{19} cm^{-3}, which corresponds roughly to 2 nm nanoclusters [23], and the assumed Er concentration is 2×10^{20} cm^{-3}, chosen to be just below the onset of concentration quenching [84]. The coupling coefficient between the excited nanocluster level and the highest Er^{3+} level, C_{nc-Er}, has been also been measured by Pacifici and is assumed to be independent of device geometry.

Two pumping mechanisms are considered: optical pumping at 488 nm with a photon flux of 10^{20} cm^{-2}s^{-1}, which excites both Si-ncs and Er, and electrical pumping, which is assumed to only excite Si-ncs. Under optical pumping, excitons are generated in the nanoclusters, and transfer their energy to Er. Under electrical excitation, two mechanisms may explain the excitation: either the current generates excitons as they pass through the film and charge the nanoclusters, or a bipolar injection

causes excitons to form in the nanoclusters. Which mechanism occurs in this system is unclear at this point, but we assume that under both optical and electrical excitation the nanoclusters act as sensitizers for Er, and in the process are left with carriers of populations n_a and n_b for the ground state and excited state, respectively. Under optical pumping at 488 nm, the Er^{3+} absorption cross section is σ_{13}, whose value of 10^{-19} cm^2 has also been determined experimentally [83]. The parameter $\phi(t)$ is the time-dependent pump flux of either photons or electrons. The time dependence of the energy level populations are given by the following equations:

$$
\begin{aligned}
\frac{dn_{\mathrm{b}}}{dt} &= -\frac{n_{\mathrm{b}}}{\tau_{\mathrm{nc}}} + \sigma_{\mathrm{nc}}\phi(t)n_{\mathrm{a}} - C_{\mathrm{nc\text{-}Er}}n_{\mathrm{b}}N_1, \\
\frac{dn_{\mathrm{a}}}{dt} &= \frac{n_{\mathrm{b}}}{\tau_{\mathrm{nc}}} - \sigma_{\mathrm{nc}}\phi(t)n_{\mathrm{a}} + C_{\mathrm{nc\text{-}Er}}n_{\mathrm{b}}N_1, \\
\frac{dN_3}{dt} &= -\frac{N_3}{\tau_{32}} - \frac{N_3}{\tau_{31}} + C_{\mathrm{nc\text{-}Er}}n_{\mathrm{b}}N_1 + \sigma_{13}\phi(t)N_1, \\
\frac{dN_2}{dt} &= \frac{N_3}{\tau_{32}} - \frac{N_2}{\tau_{21}}, \\
\frac{dN_1}{dt} &= \frac{N_2}{\tau_{21}} + \frac{N_3}{\tau_{31}} - C_{\mathrm{nc\text{-}Er}}n_{\mathrm{b}}N_1 - \sigma_{13}\phi(t)N_1.
\end{aligned}
\tag{4.1}
$$

The population of each level is increased by decay from higher levels, and decreased by decay into lower levels. These are the n/τ terms. The pumping terms with $\phi(t)$ drive carriers from the ground states to the highest levels, and the coupling term $C_{\mathrm{nc\text{-}Er}}$ causes both a decay of the nanoclusters and an excitation of the Er levels. MatLab was used to solve the coupled differential equations, and the code is given in the appendix.

4.3 Results

4.3.1 Excitation in Bulk Media

Using the present rate equation model for bulk media under a flux of 10^{20} cm^{-2}-s^{-1}, it is possible to achieve over 90% inversion of the Er^{3+} in continuous wave (CW) optical pumping, as shown in Fig. 4.5. The excited Si-nc fraction reaches 13% in steady state. The Er curve saturates at about 3 ms, which is consistent with the assumed radiative lifetime. The Si-nc curve initially doesn't rise as rapidly as the Er curve, and this is due to the energy transfer between the two species. Once the Er^{3+} reaches a saturation point, the Si-nc level also saturates. This model implicitly assumes that all the Er^{3+} ions are both active and coupled to Si-ncs. Maximizing the coupling between nanoclusters and erbium is a significant engineering hurdle and is an active area of experimental research [85, 66, 68].

Even though the Er^{3+} is efficiently inverted in this model, net gain in the system is only possible
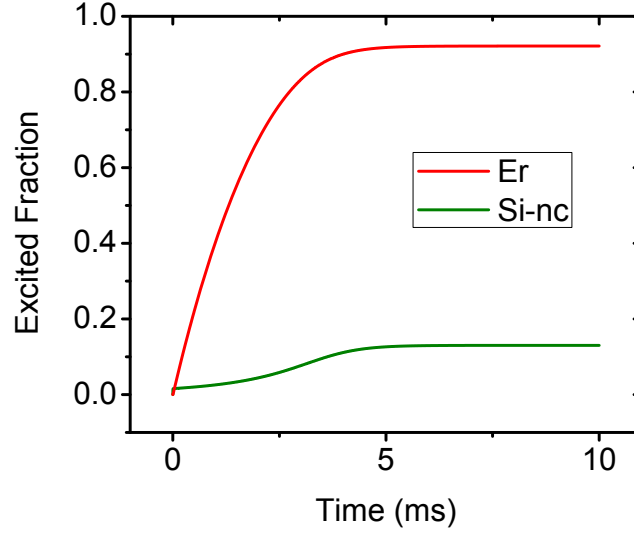
Figure 4.5: Excited fractions of Si-ncs and Er^{3+} under optical pumping of bulk layers. Under a flux of 10^{20} cm^{-2}-s^{-1} the Er^{3+} shows over 90% inversion, with 13% of the Si-ncs being excited.

if gain from inverted Er^{3+} is greater than loss due to carrier absorption. Here we define a figure of merit, χ, which is the ratio of gain due to inverted Er^{3+}, g, to absorption by excited Si-ncs, α_{ca}:

$$\chi = \frac{g}{\alpha_{ca}} = \frac{\sigma_{21}(N_2 - N_1)}{\sigma_{ca}n_b}, \tag{4.2}$$

where σ_{21} is the emission cross section of Er^{3+} at 1.5 μm, $N_2 - N_1$ is the inverted population of Er^{3+} ions, σ_{ca} is the absorption cross section of nanoclusters at 1.5 μm, and n_b is the population of excited Si-ncs. The Er^{3+} emission cross section at 1.5 μm, $\sigma_{21} = 8 \times 10^{-21}$ cm^2, comes from Mertens et al. [86], and is assumed to be the same as the absorption cross section at 1.5 μm, σ_{12}. This number is more than an order of magnitude lower than that used by Pacifici [83], reflecting a newer experimental result.

If $\chi < 0$, then $N_2 < N_1$ and the Er^{3+} is not inverted. If $0 < \chi < 1$, then the Er^{3+} is inverted, but carrier absorption overwhelms gain from Er^{3+}. But if $\chi > 1$, then gain is possible in the system because the Er^{3+} gain is greater than carrier absorption loss. Figure 4.6 shows the χ parameter under a 488 nm optical flux of 10^{20} cm^{-2}s^{-1} as a function of time for two values of σ_{CA}: a 'high' value found by Kekatpure of $\sigma_{ca} = 8 \times 10^{-17}$ cm^2 [81], and a 'low' value found by Navarro-Urrios of $\sigma_{ca} = 4 \times 10^{-19}$ cm^2 [82]. For the high value of σ_{ca}, χ peaks at about 2 ms, where the Er excitation fraction is high and the Si-nc fraction is low. As the Si-nc fraction increases and the Er fraction levels off, χ decreases again before reaching a steady-state value. At no point in time does χ rise above 1, meaning gain is not possible in this system if σ_{ca} has a value this high. However, if σ_{ca} takes the lower value, χ peaks at 3.6 and achieves a steady-state value of 2.6, meaning the excited
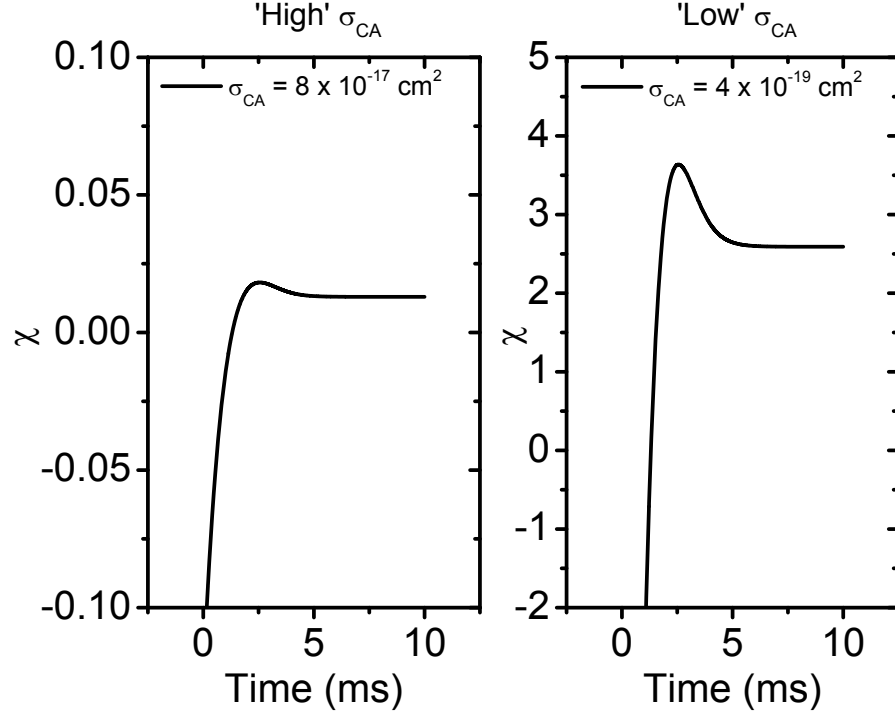
Figure 4.6: Plot of χ as a function of time for differing values of σ_{CA}. For lasing to occur χ must be greater than one, meaning gain is possible in bulk media with the lower value, but not the higher value.

nanoclusters aren't absorbing all the Er^{3+} emission. These results are consistent with modeling done by Pacifici in 2003 [83].

Plots 4.5 and 4.6 assumed a single pump flux of 10^{20} cm^{-2}s^{-1}. By varying the pump flux it may be possible to maximize χ for a given value of σ_{ca}. Figure 4.7 shows that the steady-state Er^{3+} inversion (red line) reaches its maximum of over 90% under a flux of about 10^{20} cm^{-2}s^{-1}, but doesn't increase much for higher pump powers. All the extra incident radiation is absorbed by the Si-ncs (green line), increasing their excited fraction to 78% at a flux of 10^{21} cm^{-2}s^{-1}. These carriers lead to increased loss in the system under higher pumping. This behavior is different from traditional laser systems, where generally a stronger pump makes for higher gain — this is not the case for the Er:Si-nc system described here.

This is also shown by a plot of χ versus pump power, given in black on the right axis in Fig. 4.7. The plot for the low value of σ_{ca} (solid line) peaks sharply at lower pump powers, where the Er^{3+} is inverted without a large fraction of excited Si-ncs. However, at higher powers the Si-nc fraction increases without a significant increase in the Er^{3+} inversion. The increased absorption by Si-ncs causes χ to decrease with pump power, eventually dropping below 1 at a flux of 2.3×10^{20} cm^{-2}s^{-1}, meaning that even though Er^{3+} inversion is over 90%, lasing is not possible due to carrier absorption. Furthermore, we find that for the higher value of σ_{ca} (dotted line), χ doesn't reach 1 at
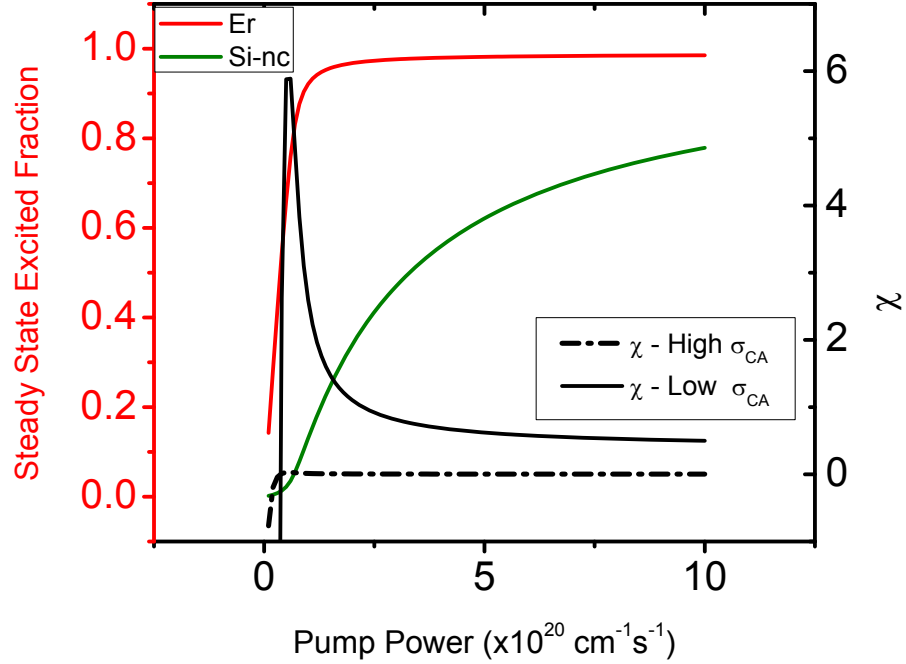
Figure 4.7: In color: Bulk excited fractions of Er^{3+} and Si-ncs as a function of incident pump flux. In black is χ (right axis). For high σ_{ca}, χ never reaches 1, while for low σ_{ca}, χ peaks strongly at low pump powers, and drops below 1 at 2.3×10^{20} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

any pump power, peaking at 0.03 at a flux of 0.6×10^{20} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

4.3.2 Excitation in Slot Waveguides

The picture changes significantly when we wish to excite Er^{3+} in a silicon slot waveguide structure. First, the thin slot layer allows for electrical injection, which is not practical in a bulk medium. Second, the slot waveguide geometry significantly modifies the local density of optical states (LDOS) of an emitter in the slot, which changes the emitter's radiative rate [78, 77].

Recent work has explored using the Er:Si-nc system as a gain medium in silicon slot waveguides, exploring electrical injection [70], TM mode confinement and emission [87], and radiative rate enhancement [78, 79]. We model a slot waveguide with a 10 nm Er:Si-nc layer of index $n = 1.57$ between two 140 nm Si layers on SiO_2, as shown on the left in Fig. 4.8. The two Si layers form a single-mode TM waveguide, and use an active layer thin enough to pump electrically via tunneling current. To calculate the effect the LDOS has on the emitters, we start with the equation

$$\Gamma_{\mathrm{tot}} = \rho \times \Gamma_{\mathrm{rad}} + \Gamma_{\mathrm{nrad}}, \tag{4.3}$$

where Γ_{tot} is the inverse of the total lifetime, Γ_{rad} is the radiative rate in bulk media, Γ_{nrad} is the
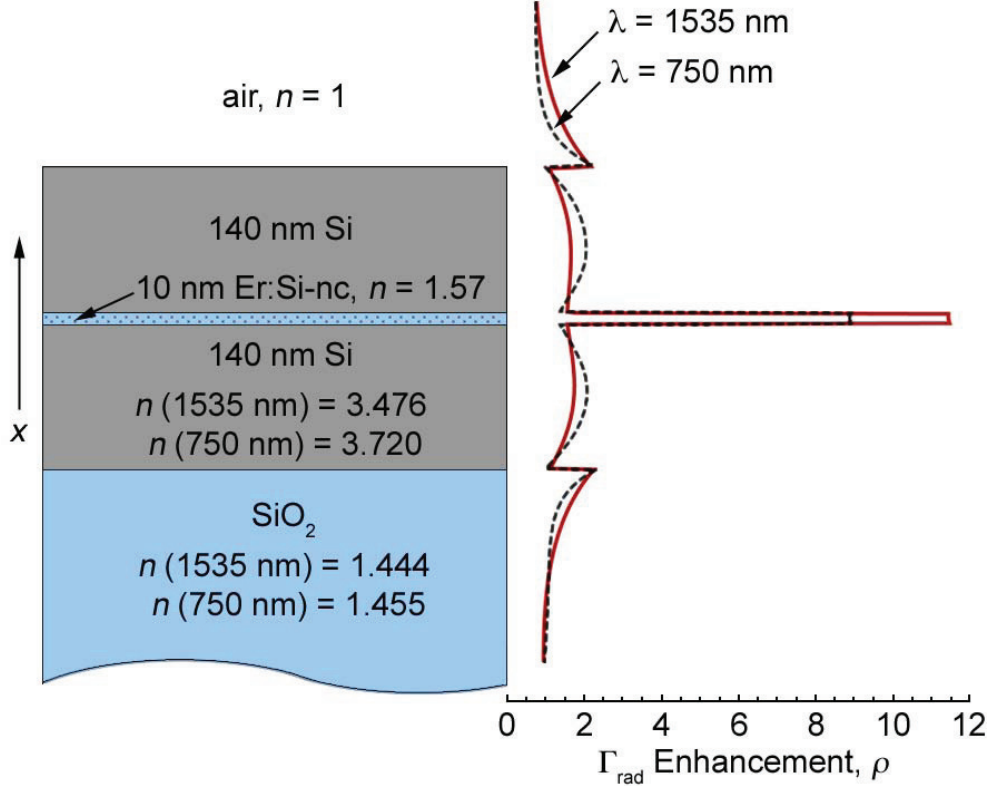
Figure 4.8: On the left is a diagram of the slot waveguide structure simulated here, with the LDOS enhancement plotted on the right. In the slot the radiative rate of the Er^{3+} is increased by 11 times, and the radiative rate of the Si-ncs by 7.

decay rate due to non-radiative processes, and ρ is the calculated radiative rate enhancement factor due to the dielectric environment. The enhancement factor ρ was calculated in the slot structure for both the Er^{3+} and Si-ncs, and plotted on the right in Fig. 4.8. Details for this type of calculation can be found elsewhere [77, 88]. At the Er^{3+} emission wavelength of 1535 nm the LDOS enhancement relative to bulk material with $n = 1.57$ is 11.43, which means that if there were no non-radiative effects, the Er^{3+} would be expected to emit over 11 times faster in the slot than in bulk. At the nanocrystal emission wavelength of 750 nm, the radiative rate enhancement is 8.92.

The total decay rate for Er^{3+} in Si-rich oxide is found by using equation (4.3) with the ρ value of 11.43 calculated for the slot, the bulk radiative rate of 59 Hz found by de Dood et al. [89] (corrected for the higher index of the layer), and an assumed bulk non-radiative rate. For our calculations we assume a bulk non-radiative rate of 274 Hz, which gives a total bulk ($\rho = 1$) decay rate of 333 s^{-1}, or our assumed lifetime of 3 ms. We also assumed the non-radiative rate in the slot increases tenfold due to the presence of the silicon layers, consistent with experimental results found by this group and others [78, 79]. The mechanism for this increase in decay rate is unknown at this time, and is not explained by LDOS effects alone. Putting these numbers into eq. (4.3) gives a total decay rate,

Figure 4.9: The bottom axis (blue) shows optical pumping of the Er:Si-nc slot, where χ never reaches 1 at any pump power. A similar result is found under electrical pumping (red), with χ reaching a maximum of 0.6 in both cases. Note this plot uses the smaller value of σ_{ca}.
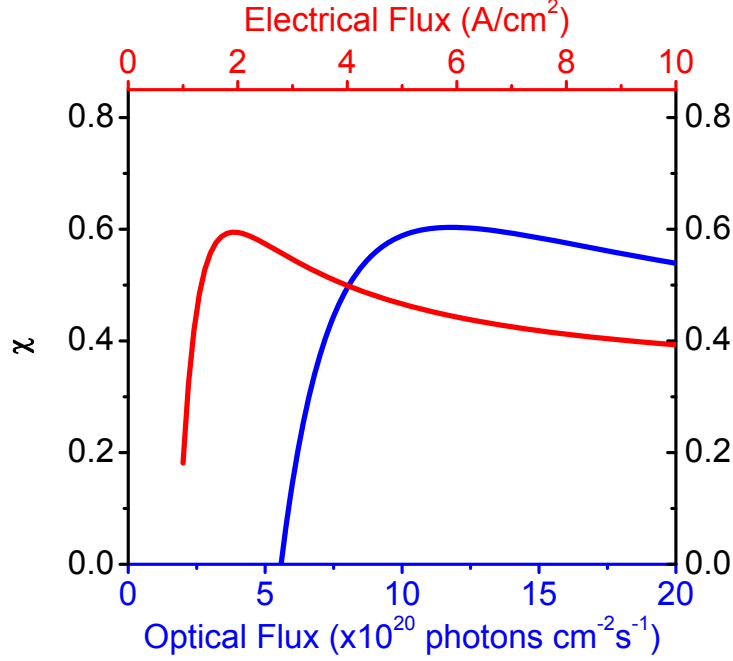
$\Gamma_{tot\text{-}Er}$, of 3431 Hz for Er^{3+} in a 10 nm slot. This corresponds to a lifetime, τ_{21}, of 0.29 ms.

A similar calculation was performed for the Si-ncs, using 9.4 kHz as both the bulk radiative rate and non-radiative rates in vacuum [17]. At the peak emission wavelength of the nanoclusters, assumed to be 750 nm, $\rho = 8.92$ (Fig. 4.8), and scaling the radiative rate by the index of the layer, $n = 1.57$, a total rate of 141 kHz is obtained for the Si-ncs. This is a lifetime of 7.1 μs, which is still longer than the transfer time between the Si-ncs and Er, which is around 1 μs [83]. This is very important because in the coupled Er:Si-nc system, the Si-ncs must transfer their energy to Er before they emit photons themselves. Also, because the energy transfer is assumed to be a Forster-type dipole interaction, it is not affected by the dielectric environment, so the LDOS enhancement is not expected to increase this transfer rate [90].

Performing the rate equation analysis with these new rates shows that under optical pumping, inversion ($\chi = 0$ in Fig 4.9) is only achievable for photon fluxes above 5.6×10^{20} $cm^{-2}s^{-1}$. However, in a slot waveguide geometry, electrical pumping is possible. Under electrical excitation, the Si-nc cross section increases by two orders of magnitude to $\sigma_{nc} = 10^{-14}$ cm^2 [64], which increases the Er^{3+} effective cross section by the same factor. These calculations show 85% Er^{3+} inversion at 10 A/cm^2 of current density, while the excited nanocluster fraction increases to 74%. But due to the increased excited Si-nc fraction, $\chi < 1$ for all current densities. In fact, using the smaller value for σ_{ca}, Fig. 4.9 shows χ reaches a maximum of 0.6 at 2 A/cm^2 current density (red curve, top axis), implying that

CW gain is impossible using electrical pumping. For currents greater than 2 A/cm^2, the increase in the excited Si-nc population is greater than the increase of the excited Er^{3+} population, decreasing the value of χ. Figure 4.9 shows a similar result for optical pumping (blue line, bottom axis), with χ reaching a maximum of 0.6 at 1.2×10^{21} cm^{-2}s^{-1}, again implying gain is not achievable under CW excitation.

4.4   Pulsed Excitation in Slot Waveguides

Here we propose a pulsed excitation scheme for inverting Er^{3+}. By pulsing the excitation, it is possible to invert the Er^{3+} during the excitation pulses, and Er^{3+} will remain inverted for some time after the pulse is turned off. Since $\tau_{21} >> \tau_{nc}$, carriers in the Si-ncs will decay quickly and there will be a period of each excitation cycle where carrier absorption is thus very small.

Figure 4.10 shows the time dynamics of the Er:Si-nc system under pulsed 10 A/cm^2 pumping. The pulse rate is chosen to be 4 kHz, and the pulse length is 0.1 ms, hence the duty cycle is 40%. Pulsed electrical excitation is considered here, however the concept also applies to optical pumping if sufficiently high pump powers are available. Figure 4.10a shows the Er^{3+} and Si-nc excited fractions, plotted as a function of time with the pulse train. The Er^{3+} reaches inversion after very few excitation pulses, and remains inverted in the steady state. However, the Si-nc excited fraction peaks around 0.7 and decays to near zero quickly once the pulse is turned off. Figure 4.10b is the figure of merit, χ, for this pulse train, plotted on a logarithmic scale as a function of time. Note this figure uses $\sigma_{ca} = 8 \times 10^{-17}$ cm^2 [81], which is the larger value, and results in increased loss. The figure of merit, χ, grows rapidly as the Si-nc population decreases, and stays above 1 for over 40% of the pulse period. During the fraction of the pulse cycle where $\chi > 1$, gain is possible because of the very low loss due to carrier absorption. This pulsing scheme is also relatively insensitive to the value of σ_{ca}, because changing σ_{ca} only scales the value of χ, which is much greater than 1.

4.4.1   Modal Gain

For the pulsing scheme described above, it is possible to calculate the modal gain for the fundamental TM mode of the structure described in Fig. 4.8. For light propagating in the z-direction, the modal gain coefficient, α_m, is given by the following expression

$$\alpha_m = (\sigma_{21}(N_2 - N_1) - \sigma_{ca}n_b) \times \frac{\int_{slot} 2c\varepsilon_0 n(x)\mathbf{E} \cdot \mathbf{E}^\star dx}{\int_{-\infty}^{\infty} (\mathbf{E} \times \mathbf{H}^\star + \mathbf{E}^\star \times \mathbf{H}) \cdot \hat{\mathbf{z}}dx} \quad . \tag{4.4}$$

The first term in eq. (4.4) represents the sources of gain and loss in this system, assumed to be only gain from excited Er^{3+} and absorption from ground state Er^{3+} and excited carriers in Si-ncs. The
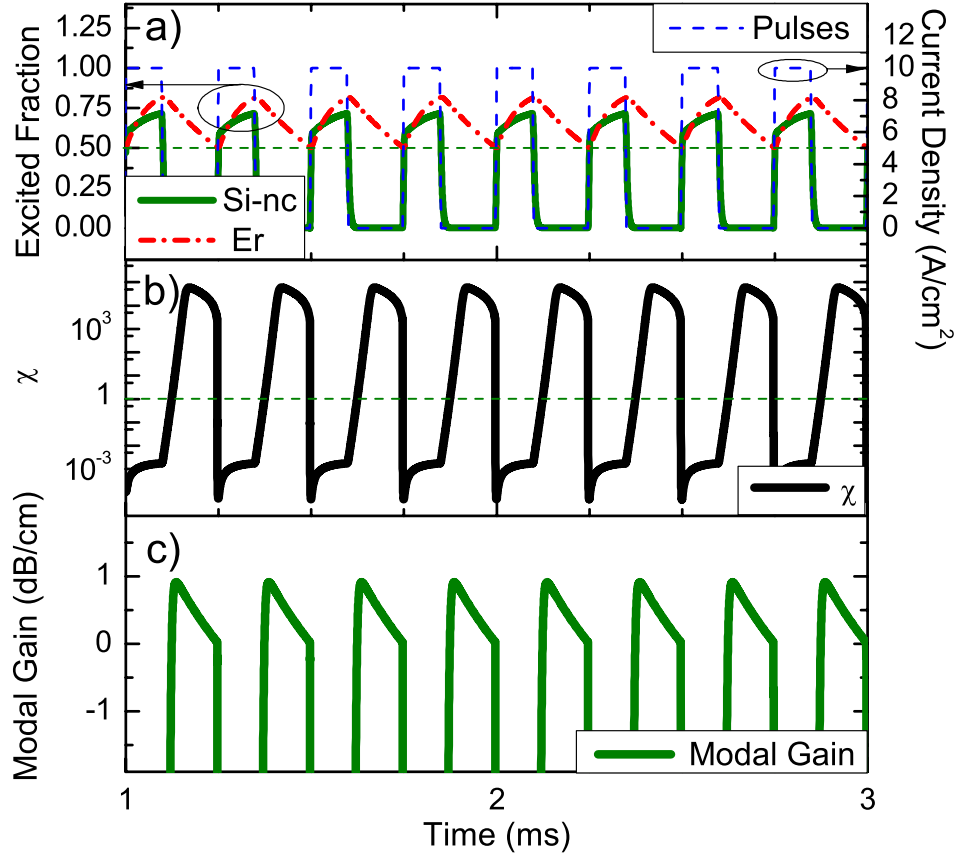
Figure 4.10: Pulsed electrical excitation of an Er:Si-nc slot. (a) Pulsed excitation (blue curve) allows steady state inversion of the Er^{3+} (red curve), while after each pulse the nanocluster excited fraction drops to zero (green curve). (b) $\chi > 1$ for a large part of each pulse. Gain from Er^{3+} overcomes carrier absorption loss as excited carriers in the nanoclusters recombine. (c) Modal gain in the slot waveguide reaches 0.9 dB/cm after each pulse.

second term is the modal confinement factor for the given geometry [73], where c is the speed of light, ε_0 is the permittivity of free space, $n(x)$ is the refractive index profile of the slot structure, **E** is the electric field and **H** is the magnetic field. This term is integrated over the Er:Si-nc slot, and for a 10 nm slot has a value of 0.285.

The result of this calculation for pulsed electrical excitation is given in Fig. 4.10c. During each pulse the gain reaches a maximum of just over 0.9 dB/cm before decaying. Note that for this slot waveguide structure, gain is not achievable under CW pumping due to carrier absorption alone. This plot uses the larger value for σ_{ca}, but when using the smaller value, the two order of magnitude decrease in σ_{ca} only increases this peak value of gain to 1.1 dB/cm (Fig. 4.11). Because the Si-ncs decay faster than Er, carrier absorption is negligible shortly after the excitation pulse is turned off, which makes the modal gain relatively insensitive to the magnitude of σ_{ca}.

Figure 4.11: Maximum gain for Er:Si-nc in a slot waveguide. As the slot thickness increases, more of the mode overlaps the active region, resulting in increased modal gain.

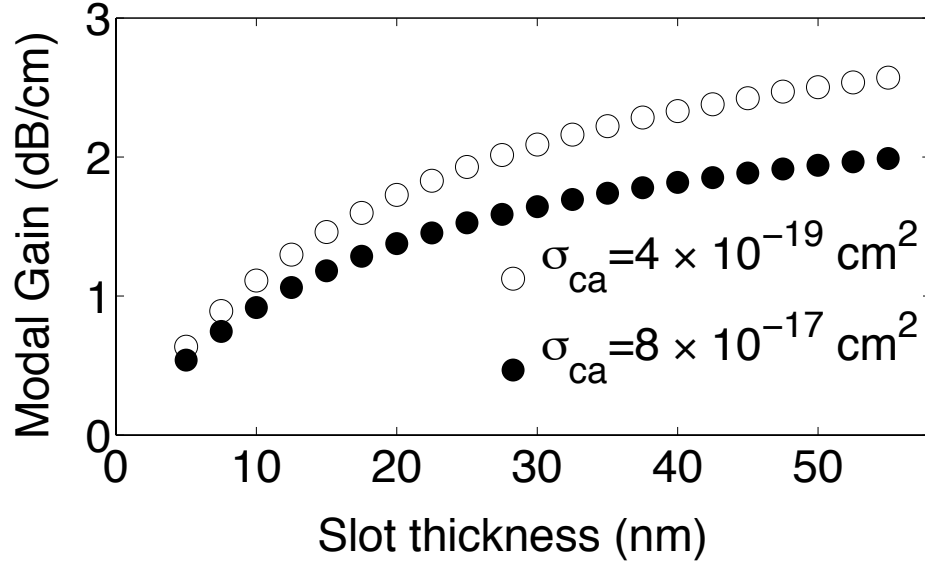Achieving net gain in a waveguide with lateral confinement puts stringent requirements on the intrinsic waveguide loss. Passive Si waveguides have been fabricated with losses on the order of 0.4 dB/cm for TE modes and 1 dB/cm for TM modes [73]. To achieve net gain and lasing, the modal gain in the TM mode must be at least 1 dB/cm to overcome waveguide losses alone. Practically, the modal gain must be even larger than 1 dB/cm to overcome other sources of loss, such as loss due to integration of electrical contacts or cooperative upconversion in the Er:Si-nc layer. The simplest way to increase gain would be to increase the Er concentration, which was not done in this work to avoid the parasitic effect of concentration quenching. However, it is possible that improved thin film fabrication can lead to higher Er incorporation [68].

Another way to increase gain is to make the active layer thicker. This allows the active layer to overlap a larger fraction of the confined TM mode, increasing the gain via the integral factor in eq. (4.4), which increases from 0.285 for a 10 nm slot to 0.603 for a 50 nm slot. Figure 4.11 plots the maximum gain as a function of slot thickness for each value of σ_{ca} used in this work. The increase in gain is most pronounced at smaller thicknesses and saturates at larger thicknesses. For a 50 nm slot, the peak gain reaches 2 dB/cm for the higher value of σ_{ca} and 2.5 dB/cm for the lower value, which is a significant fraction of the theoretical maximum assuming complete inversion and zero carrier absorption, $\sigma_{21} \times N_{Er} = 6.95$ dB/cm. Despite the large difference in magnitude for the values of σ_{ca}, the difference of only 0.5 dB/cm again shows the insensitivity of the modal gain to the value of σ_{ca} in a pulsing scheme. A 10 nm thickness was chosen for this work with the intention of using tunnel injection as the pumping mechanism, however this need not be the case. It is possible that

bipolar injection could be used to pump thicker layers [18, 91].

For this calculation, contributions due to electrically excited carriers in the silicon have been neglected. For a background doping level of 10^{15} cm^{-3}, this introduces a negligible change in the index for our purposes. During the accumulation phase of the pulse, carriers in the silicon can reach very high levels; however, during these times carrier absorption in the Si-ncs is significant, so net gain is not achievable during this fraction of the excitation pulse.

4.4.2   Other Pulsing Schemes

In the previous section, a single pumping scheme — 100 μs pulses at 4 kHz — was used to show that gain on the order of 1–2 db/cm can be achieved in an Er:Si-nc slot [92]. This frequency and pulse length was chosen to illustrate the advantage of pulsed excitation. In Fig. 4.10a, the pulses are chosen to maximize the time during which the Er remains inverted while the pump is off. As a result we clearly see that the elimination of carrier absorption by this method leads to modal gain around 1 db/cm.

However, this pulsed injection laser need not be limited to this single frequency. So long as the Er can be efficiently inverted and the pump turned off to eliminate carrier absorption loss, the laser can be engineered to operate at a range of pumping frequencies. To first explore this, we examine the behavior of the figure of merit, χ, as a function of frequency and pulse length. Specifically, we look at the length of time, or 'window', where $\chi > 1$ during each period, and gain is possible. Graphically, this is the width of the spike in Fig 4.10b, measured at $\chi = 1$. All plots in this section assume the same 10 A/cm^2 pump used in Fig 4.10, and the larger value of σ_{ca}. Using the higher σ_{ca} gives a conservative estimate of the gain achievable. However, as shown by the gain plot in Fig. 4.11, the gain achievable under pulsed excitation is not very sensitive to the value of σ_{ca}.

This plot gives insight to the competing mechanisms of this pulsed excitation scheme: having the pump on long enough to invert Er while having it off long enough for lasing to occur. At low frequencies, the curves are fairly flat — this is because at low frequencies both the Er and Si-ncs are allowed to decay completely between pulses. The window where $\chi > 1$ is determined entirely by how long the Er stays inverted after each pulse. For a pulse length of 20 μs, the pulse is short enough so that the Er isn't inverted by a single pulse, and hence the window is zero. As the pulse length increases to 40 μs, the Er is only briefly inverted, so the window is very small. As the pulse length is increased, it eventually reaches a saturation value of 0.13 ms, which is the time needed for Er to decay from almost complete inversion to an excited fraction of 50%, or the onset of inversion. Note that this value of 0.13 ms is consistent with the Er lifetime in the slot of 0.29 ms.

As the frequency increases, the time between pulses decreases, so for short pulses it becomes possible to increase the length of Er inversion by pumping before the Er has decayed completely. This is shown by the increase in the window with frequency for small pulse lengths.
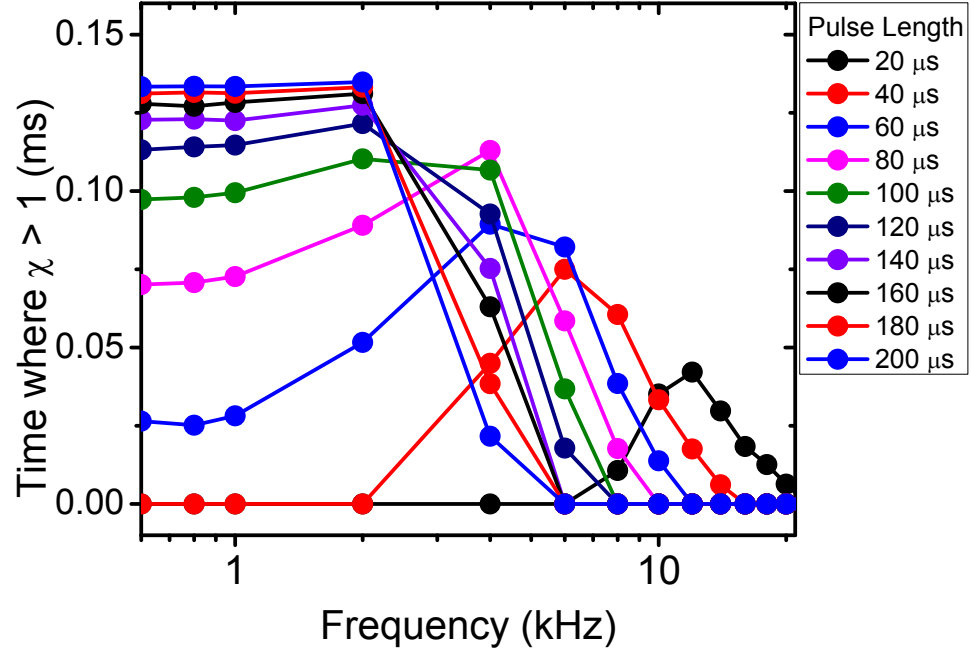
Figure 4.12: Dependence of χ on pulse width and frequency. Plotted is the time during each period where $\chi > 1$ and lasing is possible.

However, for longer pulse lengths, the Er has already reached maximum inversion, so the increase in frequency eventually introduces pulses that occur before the Er loses inversion, until at very high frequencies the pulse length exceeds the period. At this point the excitation is no longer pulsed, it is operating in CW. For this reason the window goes to zero because, as shown before, gain is impossible in CW.

In essence, what Fig 4.12 shows is that for each pulse length, there is a frequency which is optimal for keeping Er inverted. Depending on which frequency is required for operation, a specific pulse length can be chosen, with smaller pulse lengths required for higher frequencies. For example, the shortest pulse length given, 20 μs, is not long enough to invert the Er with a single pulse, but if used at a high enough frequency, around 10 kHz, lasing can be achieved.

Also of interest is the modal gain as a function of pulse length and frequency. Plotted in Fig. 4.13 is the peak modal gain during each pulse cycle for 10 nm and 50 nm active layers. It follows trends similar to Fig. 4.12, with the gain increasing with pulse length at low frequencies up to a maximum around 1 db/cm for the 10 nm film and 2 db/cm for the 50 nm film. At low pulse lengths the peak gain increases with frequency before reaching a peak and decreasing again. The maximum gain achievable for a given pulse length decreases as the pulse shortens, but this maximum occurs at higher frequencies.
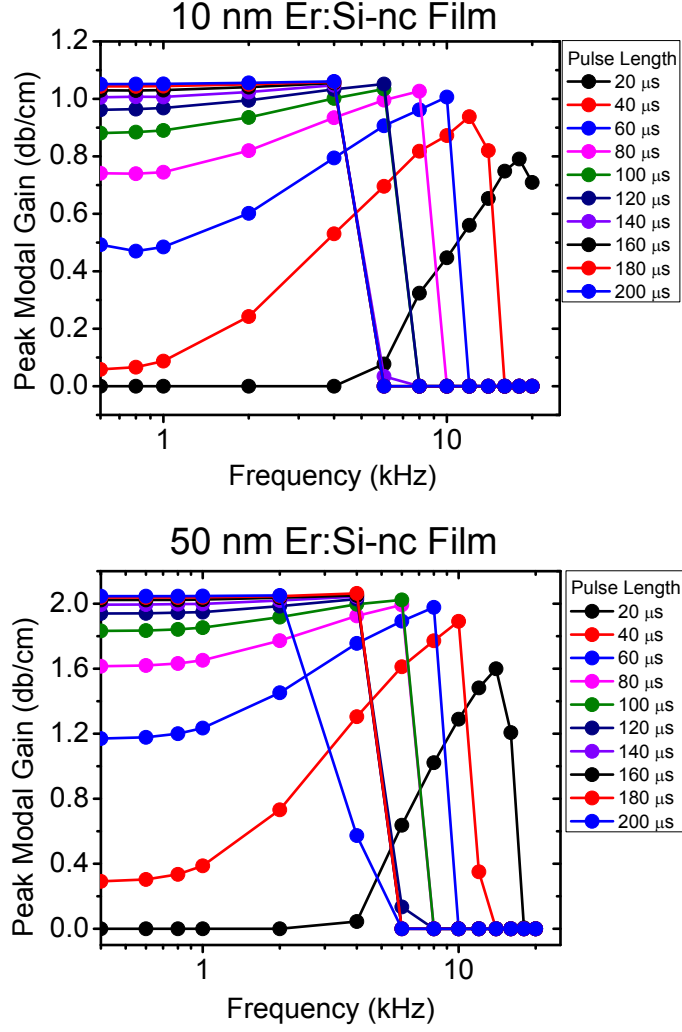
Figure 4.13: Peak modal gain as a function of frequency and pulse length for 10 nm and 50 nm Er:Si-nc films

The curves in Figs. 4.12 and 4.13 describe the two extreme regimes of pulsed excitation. At one end is the long-pulse, low-frequency regime, where the pulse length is long enough to maximally invert the Er. The limiting behavior is the maximum Er inversion, which is why the windows in Fig. 4.12 show saturation behavior. In this case pulsed excitation is necessary to allow the nanoclusters to decay enough for laser operation. At the other end is the short-pulse, high-frequency regime, where the frequency is high enough to use short pulses to invert the Er. In this case the limiting behavior is the time needed for the nanoclusters to decay enough for laser operation. The curves in Fig. 4.13 have maxima for low pulse lengths because at high enough frequencies another pulse is introduced before the carrier absorption has decayed enough for laser operation.

4.5 Conclusion

We have applied rate equation modeling to silicon slot waveguides using Er:Si-nc as the active media, and have determined that gain is likely not achievable under CW excitation due to carrier absorption by excited nanoclusters. However, using a pulsed injection scheme, it is possible to take advantage of the comparatively short excited state lifetime of Si-ncs and achieve gain in this system. With pulsed excitation, the modal gain in the Er:Si-nc slot waveguide can be in the dB/cm range, depending on slot thickness, and furthermore is comparatively insensitive to the value of σ_{ca}. We have also explored the frequency dependence of this pulsed behavior, and found that device can be operated at any frequency below about 10 kHz, which is the frequency at which there is insufficient time for the nanoclusters to depopulate enough to minimize carrier absorption.

The LDOS effect in the slot waveguide has been characterized and incorporated into the model. For laser applications, the increased radiative rate of Er^{3+} raises the threshold for lasing, and carrier absorption results in net material loss in CW operation. While CW gain may be possible in bulk layers if σ_{ca} is sufficiently small, layers in slot waveguides require pulsed excitation to achieve gain.

Aside from the significant problem of coupling a high fraction of Er to Si-ncs, all the other parameters in this paper have been chosen conservatively. There is no fundamental barrier to the fabrication of films with lower non-radiative rates and higher active concentrations than those assumed here, which would only increase the gain in the Er:Si-nc slot waveguide system. The modal gain predicted here using pulsed excitation is slightly greater than passive Si waveguide losses that have been previously demonstrated, which is reason for optimism that net gain in a silicon-based laser can be realized with this gain medium. The other major obstacle to the proposed device is efficient electrical pumping of Er:Si-nc layers, and is the subject of the experimental work in the next chapter.

Chapter 5

Luminescence of Erbium-Doped Films

5.1   Introduction

In this chapter we explore the experimental fabrication and characterization of $Er:SiO_2$ and Er:Si-nc films, with an eye towards optimizing the optical and electrical properties of the films. A significant amount of work has been devoted to the study of photoluminescence of these films, but comparatively little work has been done exploring electrical excitation in these films. We first present results obtained from our co-sputtered $Er:SiO_2$ and Er:Si-nc films by describing their photoluminescent properties. We then focus on erbium electroluminescence in two different sets of Er:Si-nc films. The first set of films exhibit 'luminescent-center-(LC)-mediated sensitization', which was first defined by Peter Kik's group [93]. The second set are more common nanocluster-sensitized films grown using reactive oxygen sputtering. We present results which, to our knowledge, show the first electroluminescence in LC-sensitized films.

For comparison we then present electroluminescence from reactive oxygen sputtered films and compare the two sets. We find that reactively sputtered films are far more robust than LC-sensitized films, but allow smaller current densities before breakdown. From the reactively sputtered films we were able to extract an electrical excitation cross section that is two orders of magnitude larger than the optical cross section.

We then study the time-dependent luminescence of reactive oxygen sputtered films under electrical pumping as a way to explore the electrical excitation mechanisms present in these films. We find that electrical pumping has no obvious quenching effect on the luminescent lifetime, but do observe charging effects during the rise time. We finish by revisiting the rate equation modeling presented in Chapter 4 by using the parameters measured here. We find that to invert erbium in these films the EL luminescent lifetime must be increased by an order of magnitude, and the maximum current density must be increased.

5.2   Erbium-Doped SiO$_2$

5.2.1   Fabrication and Characterization

The films described in this chapter were fabricated using ion-beam co-sputtering. Ion beam sputtering works by ionizing an inert gas, and accelerating the gas toward a target, which sputters ions and atoms from the target and onto a sample some distance away. A neutralizer was used to deionize the accelerated gas before it sputters the target. Our system uses an Ion Tech 3 cm ion source, and the base pressure of our system is 10^{-7} torr. Film deposition occurs at room temperature with a gas pressure of 2×10^{-4} torr.

To make Er:SiO$_2$ we used accelerated argon to sputter films off a four inch SiO$_2$ target, to which we attached pieces of Er. The relative Er content was determined by the placement and size on the target. We also added pieces of Corning 2947 soda lime glass (SLG) because we found they increased the luminescent intensity and lifetime of our deposited films. Results suggesting the benefits of using SLG were found by Snoeks [59] and Miniscalco [94].

The Er concentration of these films was verified by Rutherford backscattering spectroscopy to be 0.4 atomic percent, or 2.6×10^{-20} cm^{-3}, while soda lime contents like Na or Ca were found to be below 1 percent. Film thicknesses were determined by ellipsometry.

5.2.2   Photoluminescence of Er:SiO$_2$

The first films developed in this system were Er:SiO$_2$. These films are easily characterized, and they serve to describe the luminescent properties of erbium in its simplest form, before studying the coupled Er:Si-nc system.

Photoluminescence measurements were carried out on these films by exciting Er films using the 488 nm line of an Innova 70C Ar ion laser. The incoming laser was passed through a TeO$_2$ acousto-optic modulator (AOM), which split the beam into a central unmodulated beam and modulated side lobes. The first diffracted lobe was used as the modulated excitation source, and standard lock-in techniques were used to detect the signal with a SR830 lock-in amplifier. The modulation frequency is 9 Hz unless otherwise noted. Spectral data was taken by focusing the emission into an Oriel MS257 monochromator, and infrared detection was done using a nitrogen cooled Hamamatsu photomultiplier tube (PMT), which was kept at -80 °C. Time-dependent data was taken using a SRS 430 multi-channel scaler.

Figure 5.1 shows the typical erbium luminescence in the infrared band around 1530 nm. The sample shown was deposited as described above and is 43 nm thick. To activate the Er, the film was annealed in a tube furnace at 800 °C for one hour in a nitrogen ambient. This reduces the number of defects in the oxide and allows the erbium to bond in the Er^{3+} state, which luminesces. The 488 nm laser used to pump the sample is resonant with the $^4I_{7/2}$ level of the erbium ion, which quickly
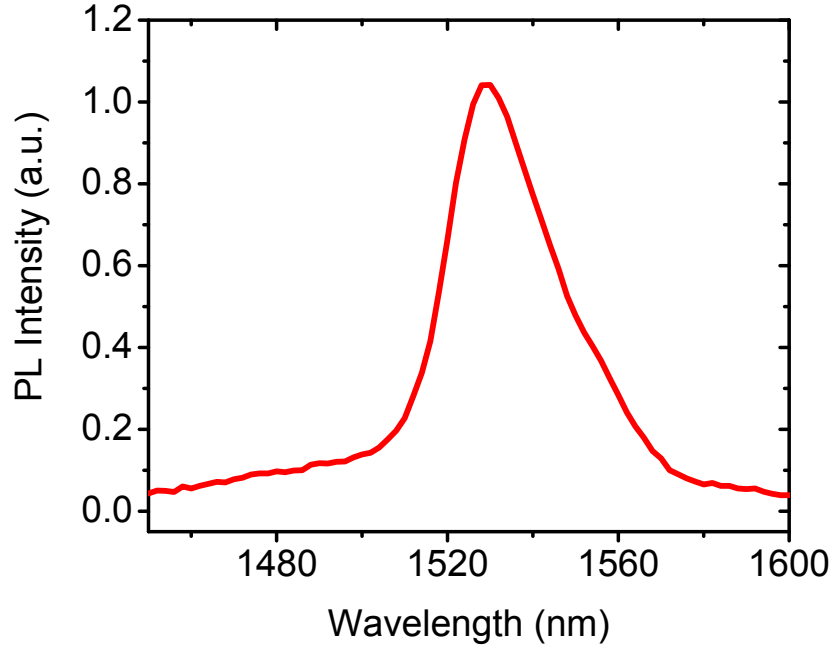
Figure 5.1: Photoluminescence of an Er:SiO$_2$ film, deposited on Si and annealed for one hour in N$_2$ at 800 °C

decays to the metastable $^4I_{13/2}$ level. Transitions between this level and the $^4I_{15/2}$ ground state give rise to the emission around 1.5 μm shown in Fig. 5.1.

Transitions between the 4f levels of the erbium atom are parity forbidden, and this selection rule is weakly violated when erbium atoms are placed in a matrix that ionizes them into the Er^{3+} state. As a result the $^4I_{13/2} - ^4I_{15/2}$ transition becomes weakly allowed, and the lifetime of the metastable $^4I_{13/2}$ is quite long, on the order of milliseconds. The time-dependent characteristics of Er:SiO$_2$ after annealing at 1100 °C are shown in Fig. 5.2. The plot shows the luminescence as a function of time after the pump is turned off at t = 0. The luminescence decay follows a single exponential curve, with a 5.9 millisecond lifetime. The measured lifetime is a reflection of the oxide film quality, as the erbium is quenched by non-radiative defects in the film, which are partly eliminated by annealing. It is also a measure of the quantum efficiency of the erbium in the film, as the oxide defects introduce non-radiative defects that directly compete with the radiative rate of the erbium. In general, a high measured lifetime indicates that the erbium in the film is well coordinated and luminescing efficiently.

One way to observe this phenomenon on the macroscale is to vary the annealing temperature of the sputtered Er:SiO$_2$ films. A study of this kind is given in Fig. 5.3. The Er luminescence at 1530 nm is given in red, and peaks for an annealing temperature, T$_A$, of 800 °C before decreasing for
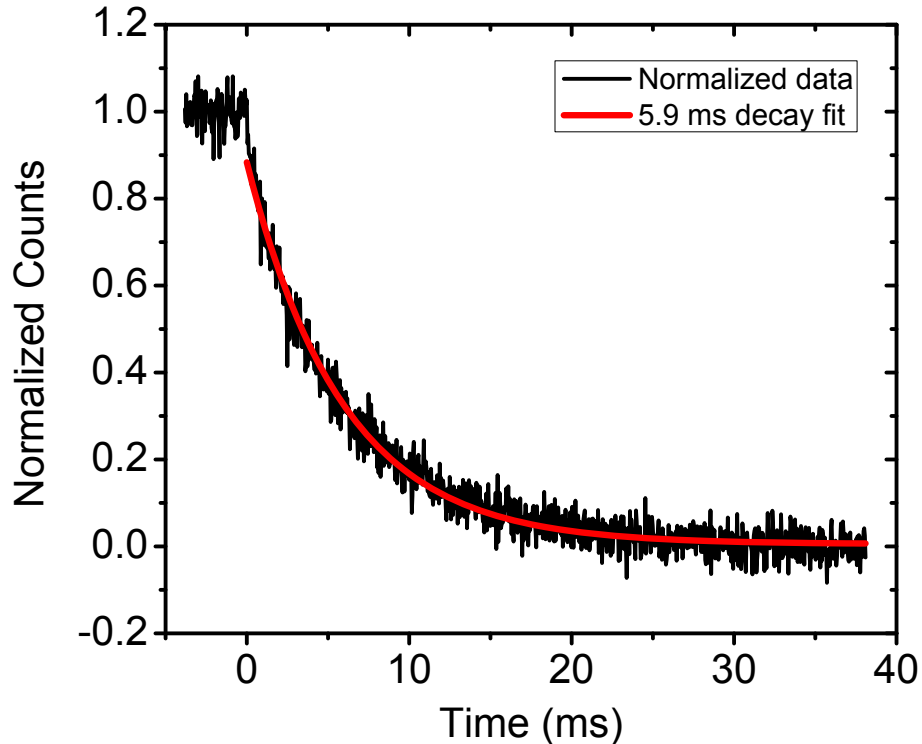
Figure 5.2: Lifetime of Er:SiO$_2$ film after annealing at 1100 °C. The decay is single exponential with a lifetime of 5.9 ms.

higher temperatures. This is due to erbium clustering at higher temperatures, and is fairly common in other Er:SiO$_2$ films [57]. The lifetimes of the films, however, increase monotonically with T$_A$, up to 5.8 ms at 1100 °C. This is due to an increase in the oxide quality for higher T$_A$. So while the luminescence decreases, partly due to erbium clustering [95], the erbium that is left unclustered is better coordinated, leading to higher lifetimes.

The results presented here are very much in line with current understanding of Er:SiO$_2$ as a gain material. However, there are some results found while studying these films that do not exactly fit in with this paradigm. We experimented with annealing these films in oxygen, and found that lifetimes were significantly larger under oxygen annealing than nitrogen annealing. We suspect that the reason for this is that Er^{3+} is thought to be formed by bonding with oxygen in the matrix [96], so introducing oxygen at high temperatures would create an environment where erbium would be more likely to bond in its ideal state for luminescence.

We also find that films deposited on SiO$_2$ showed significantly higher lifetimes than those deposited directly on Si. We can only speculate at this time why exactly this is, but one reason may be that the dielectric environment of the substrate is affecting the radiative rate of the erbium in the deposited film. A more thorough treatment of this phenomenon is explored in Ryan Briggs' thesis using these films [97].
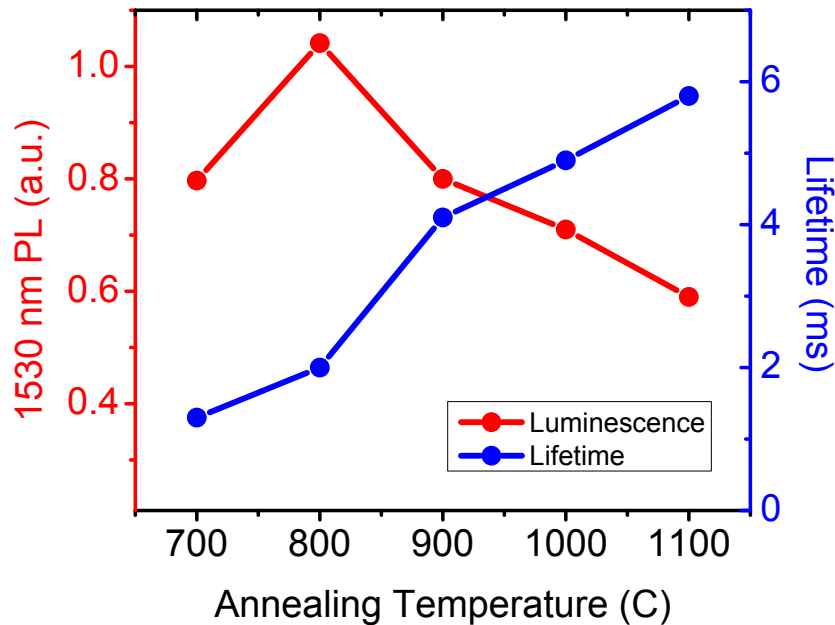
Figure 5.3: The Er luminescence (red) peaks for a T_A of 800 °C, while the luminescent lifetimes (blue) increase monotonically with T_A due to increased oxide quality

Also of interest is the fact for thin films deposited on silicon, high annealing temperatures in oxygen will significantly oxidize the silicon substrate, growing a layer of undoped SiO_2 beneath the $Er:SiO_2$ film. However, these films still did not show lifetimes as high as samples deposited on SiO_2 layers before annealing. In all cases the photoluminescence intensities of these films were similar but less than their nitrogen counterparts.

From here we move to nanocluster-sensitized films as a way to electrically access erbium. Our attempts to characterize these $Er:SiO_2$ films electrically produced no electroluminescence, but because silicon nanoclusters can be electrically accessed and exhibit visible electroluminescence [18, 19, 91], we show here work towards using electrically pumped Er:Si-nc as an electroluminescent gain medium.

5.3 Er:Si-nc Film Fabrication and Characterization

5.3.1 History of Er:Si-nc Films

As mentioned previously, Er:Si-nc emerged as a gain medium for silicon photonics in 1997, when Fujii et al. first published a paper showing the sensitizing action of silicon nanocrystals on erbium [61]. This sensitizing action resulted in an increase of the Er luminescence by over an order of magnitude. Furthermore, the material system was found to respond to broadband sensitization, as Er luminescence was observed for non-resonant pump wavelengths. This was further evidence

of sensitization, as nanocrystals also exhibit broadband absorption. Further work established that Er need not be sensitized by nanocrystals, and that amorphous nanoclusters also act as efficient sensitizers [62]. These films were also discovered to have the same absorption cross section of their component nanoclusters, $\sigma_{\text{eff}} \sim 10^{-16}$ cm^{-2}, which is four orders of magnitude larger than the absorption cross section of Er:SiO$_2$, where $\sigma_{\text{Er}} \sim 10^{-20}$ cm^{-2} [94].

Films containing erbium and silicon nanoclusters are most commonly grown via three methods. The first is ion implantation, where Si and Er ions are implanted directly into a thermal SiO$_2$ film. This method was used frequently for studying gain media using silicon nanocrystals [98]. Si crystallizes in SiO$_2$ at temperatures above 1000 °C, while Er is generally best activated at around 800–900 °C, and begins to cluster as well at temperatures above 1000 °C [57], so separate ion implantations of both Si and Er allow each to be activated at their preferred temperature. Secondly, films have also been prepared using chemical vapor deposition (CVD) [98], but suffer from limited Si mobility in the matrix [99], and an excess of Er-quenching OH$^-$ groups [100]. However, the only paper to our knowledge showing gain in this system used electron-cyclotron resonance enhanced CVD [101].

The third, and most common, method of fabricating Er:Si-nc is cosputtering [61, 68, 87]. Because the Si inclusions need not be crystalline to efficiently sensitize Er [62], the field has moved toward co-sputtering SiO$_2$ films with Si and Er. By annealing at temperatures lower than that needed to crystallize silicon, the density of the sensitizing Si inclusions increased, while their size decreased. This method grows homogenous films, which is more difficult to engineer in ion-implanted films. It is also far easier to control and vary the relative concentrations of Si and Er, and avoids hydrogen precursors that lead to OH$^-$ inclusions in CVD-grown films.

5.3.2   Photoluminescence and Luminescent-Center-Mediated Sensitization

To fabricate Er:Si-nc films, pieces of Si were added to the previous Er:SiO$_2$ target to create silicon-rich oxide. These films were then annealed to form Er:Si-nc. We also added more erbium to the target to increase luminescence, and RBS shows the new Er content is 1.1 atomic percent. The first set of films we describe here were sputtered in a pure Ar environment on silicon substrates. XPS shows they have a Si excess of 0.8%, and ellipsometry indicates a refractive index of 1.65, which is larger than that of pure SiO$_2$ (1.44) due to the silicon excess. Annealing is done in an AccuThermo AW 610 rapid thermal annealing (RTA) furnace for 5 minutes at 500 °C under nitrogen ambient.

One of the characteristics of sensitization is luminescence arising from pumping at wavelengths that are not resonant with erbium atomic transitions. Figure 5.4 shows the infrared emission under both 488 nm (green) and 476 nm (blue) excitation. Pumping at 488 nm is resonant with the Er atomic levels, and shows the characteristic Er peak. Pumping at 476 nm however is not resonant with Er, and that the emission is almost identical to that under 488 nm is evidence of energy transfer
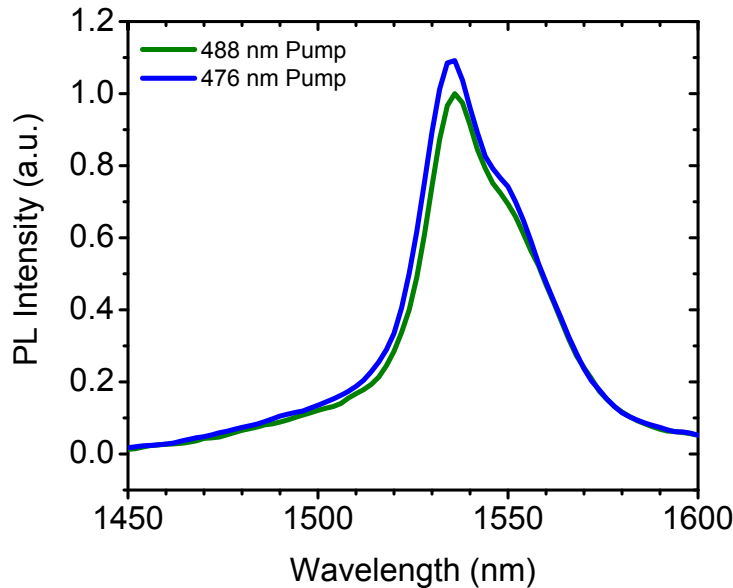
Figure 5.4: Erbium photoluminescence under resonant 488 nm pumping (green) and non-resonant 476 nm (blue) pumping. That the two curves are almost identical shows the sensitizing action of Si-ncs present in the film.

from the Si-nc in the film to Er. Pumping our Er:SiO$_2$ films at 476 nm shows almost no luminescence. It should also be noted that at 488 nm, there are actually two pathways to Er luminescence in Er:Si-nc: absorption by the Er directly by the resonant 488 nm laser, and sensitization by the Si-ncs, which also absorb at 488 nm. The fact that the two traces at 488 nm and 476 nm are very similar in intensity means that absorption and sensitization by Si-ncs is the dominant mechanism for Er luminescence in these films. This is because the effective cross section for the Si-ncs is much larger than that for Er alone in SiO$_2$, as will be discussed below.

The annealing temperature used for this sample, 500 °C, is much lower than that seen in Er:SiO$_2$ films (Fig. 5.3), which generally show a peak in their luminescence for annealing temperatures between 800- 900 °C. This is due partly to the excess silicon added to the films, but we believe these films are exhibiting luminescent-center (LC) mediated sensitization. This is a term first used by Kik and Savchyn [93] to describe films that are characterized by low peak annealing temperatures, short luminescent lifetimes, and blueshifted visible emission from the silicon nanostructures. Figure 5.5 shows the dependence of the peak luminescence on annealing temperature, showing a clear peak at 500 °C, consistent with the results found by Savchyn.

Further evidence of LC sensitization comes from studying the visible luminescence in these films. While erbium has sharp luminescent bands corresponding to its atomic energy levels, Si-ncs have a broad emission band in the visible, which is partly quenched by the sensitizing action involving Er. Luminescence measurements were carried out, and spectra in the visible wavelength range were taken with a Spec 10 silicon CCD system. Figure 5.6a shows the visible luminescence due to LCs in the
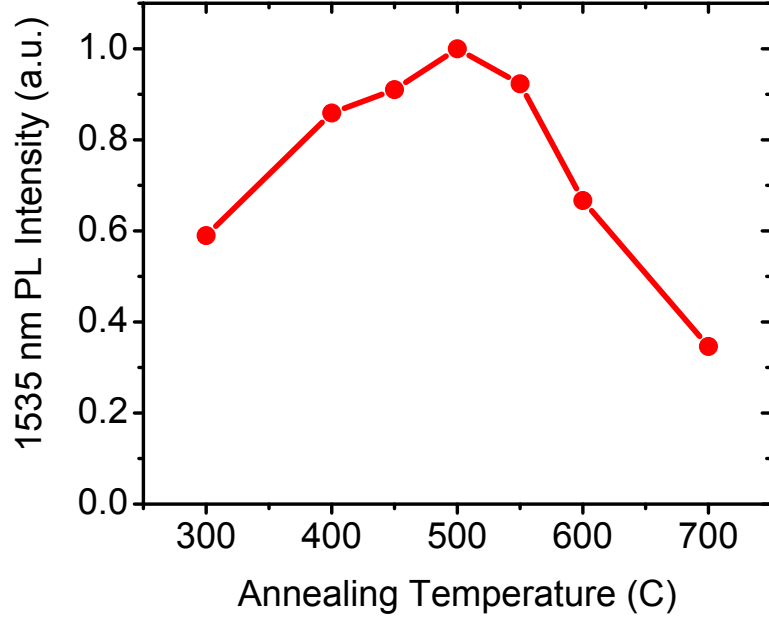
Figure 5.5: Peak erbium photoluminescence under 488 nm pumping, as a function of annealing temperature, T_A. The peak at low T_A is consistent with luminescent center mediated sensitization.

wavelength range between 550–800 nm. The luminescence peak is centered around 680 nm, which is blueshifted from the common Si-nc luminescence band and is a characteristic of LC-sensitized films. The peak also decreases with T_A, which is a trend opposite that seen in nanocrystals. This may be because LCs are more closely related to silicon defect complexes than fully formed clusters, and would be annihilated by high temperature annealing. The cutoff at 550 nm is due to a wavelength filter for the 488 nm excitation source. There is also a sharper peak centered at 980 nm due to the second excited state of Er, or the $^4I_{11/2} -^4 I_{15/2}$ transition. This peak shows a maximum for a T_A of 500 °C, which is consistent with the behavior in the infrared region shown in Fig. 5.5.

The last hallmark of nanocluster sensitization is an effective cross section on the order of that for silicon nanocrystals, $\sigma_{\text{eff}} \sim 10^{-16}$ cm^{-2}. To measure σ in our films, we measure the luminescent rise time as a function of pump power. In general, the photoluminescence intensity is proportional to the number of excited Er^{3+} ions, N^*, and inversely proportional to the luminescent lifetime, τ.

$$I_0 \propto \frac{N^*}{\tau} \tag{5.1}$$

Using a rate equation formalism, the proportion of excited Er^{3+} ions is increased by pumping at a flux ϕ, and decreased by emission with a lifetime τ

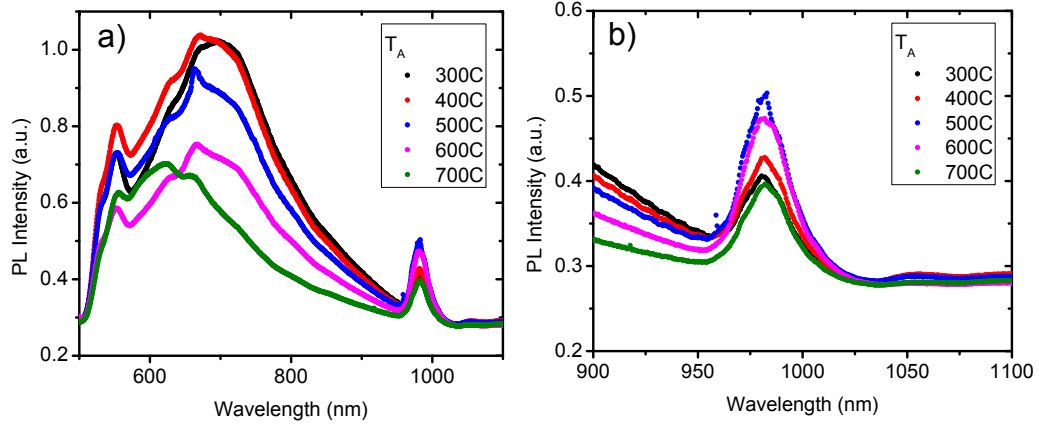$$\frac{dN^*}{dt} = \sigma\phi(N - N^*) - \frac{N^*}{\tau} \tag{5.2}$$

Figure 5.6: a) Er:Si-nc films show luminescent bands in the visible range between 500–1000 nm, with the large band between 600–800 nm due to Si-nc luminescence, and the small peak at 980 due to Er luminescence. The Si-nc peak decreases with T_A, while the Er-related 980 nm peak (b) follows the same trends as the Er infrared emission shown in Fig. 5.5.

where σ is the effective cross section, or the probability that an incoming photon will excite an Er^{3+} ion, and N is the total Er concentration. Solving the equation gives the time dependence of luminescence, and it is given by

$$I(t) = I_0(1 - e^{-t/\tau_{on}})$$
(5.3)

where I_0 is the steady-state intensity, and the measured rise time, τ_{on}, is given by

$$\frac{1}{\tau_{on}} = \sigma\phi + \frac{1}{\tau}.$$
(5.4)

This equation allows the cross section to be extracted from the linear dependence between the rise rate, $1/\tau_{on}$, and the flux, ϕ. An experiment of this type was performed on the Er:Si-nc films described above, and is shown in Fig. 5.7a). At time t = 0 the 488 nm pump laser is turned on at pump powers varying between 25 mW and 132 mW. The laser is focused to a spot size of 0.6 mm^2, and pulsed at 45 Hz for lock-in purposes. The resulting rise times increase with pump power, as expected from equation (5.4). The rise times are all single exponential, and the corresponding rise rates are plotted as a function of flux in Fig. 5.7b. The rates follow a linear relation, and using eq. (5.4) to find the effective cross section (plotted in dark blue) gives a value of $\sigma_{eff} = 1.8 \times 10^{-16}$ cm^2. This is much larger than the values for Er:SiO$_2$, which are of order 10^{-20} [94]. The linear fit has two parameters, the first is the cross section, and the second is the intercept at zero flux, which should correspond to the measured lifetime. The $\phi = 0$ intercept of the fit in Fig. 5.7b corresponds to a lifetime of 47 μs. The measured lifetime in this film is 50.6 μs, in agreement with the fit value.
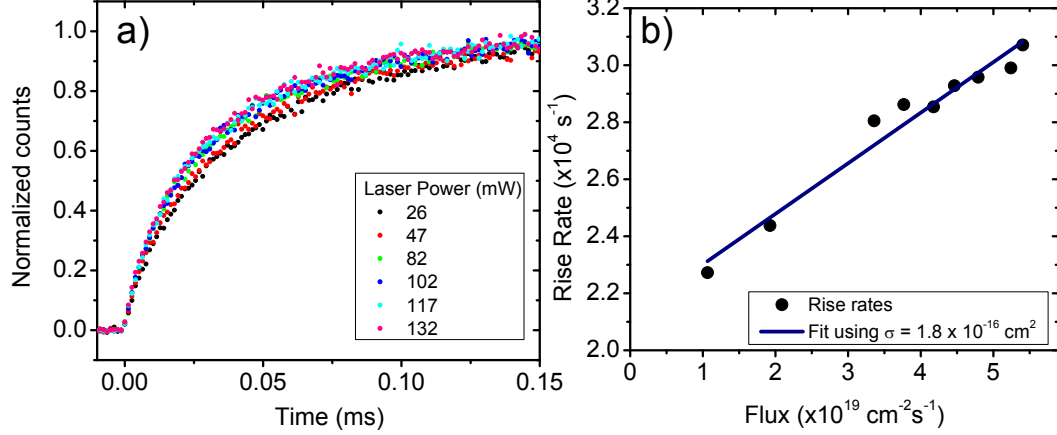
Figure 5.7: a) Rise times for Er:Si-nc film at various pump powers. b) The excitation cross section of Er:Si-nc, σ_{eff}, is extracted from a linear fit of the rise rates to their corresponding pump fluxes. The high cross section is common for Er:Si-nc films, and the intercept at $\phi = 0$ corresponds to a lifetime of 47 μs, in agreement with the observed lifetime of 50.6 μs.

The effective cross section can also be extracted by observing the photoluminecence intensity as a function of pump power [102]. The steady-state solution of eq. (5.2) gives a steady-state intensity of

$$ I = I_0 \frac{\sigma_{\text{eff}} \tau \phi}{\sigma_{\text{eff}} \tau \phi + 1} \tag{5.5} $$

where I_0 is the saturation intensity corresponding to excitation of all the active Er^{3+}. A study of this kind is given in Fig. 5.8. The extracted σ_{eff} has a value of 1.2×10^{-16} cm^2, which is in good agreement with the value found by fitting rise times in Fig. 5.7. The normalization parameter, I_0, from this fit is found to be 3.12, meaning that a pump flux 3.12 times higher would saturate the PL, exciting all of the optically active Er^{3+}. This was not done here due to the power limits of our laser source. It should be noted that this does not mean the pump flux is a factor of 3 away from inverting the entire erbium population, because it is unlikely that all the erbium is coupled to nanoclusters. This number represents the pump flux needed to excite all the active, coupled Er^{3+} ions in the film.

One characteristic of these films is that their lifetime is very short, in the tens of microseconds. This is significantly lower than the lifetimes found in Er:SiO$_2$ (Fig. 5.3), or even in traditional nanocrystal-sensitized films [84]. This is likely due to the low annealing temperatures used in LC-sensitized films, and the films grown in the Kik group show similar lifetimes [93]. The lifetimes of these films also exhibit a stretched exponential behavior, meaning the decay time is described by

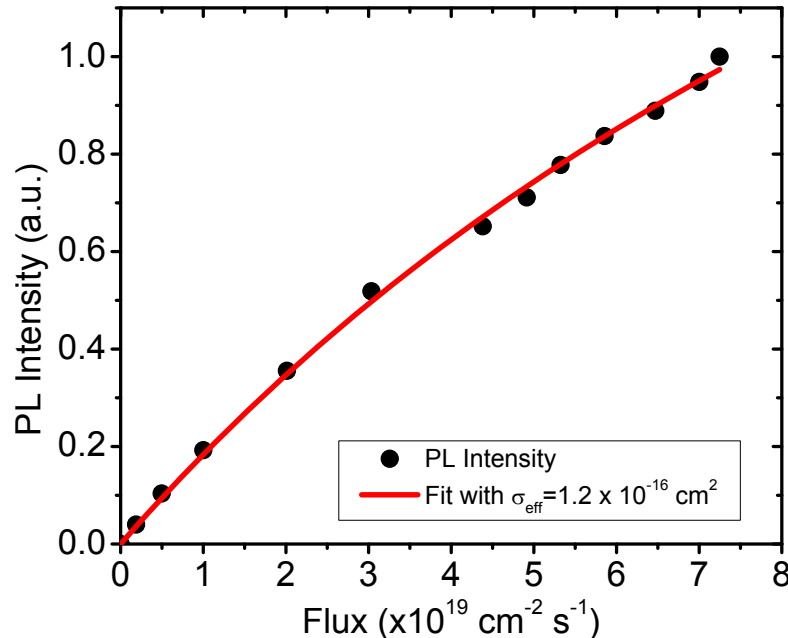$$ I(t) = I_0 \times \exp(t/\tau)^\beta \tag{5.6} $$

Figure 5.8: The excitation cross section of Er:Si-nc is extracted by fitting the PL intensity to the pump flux. The cross section is in good agreement with that found by measuring rise times.

where τ is the lifetime, and β is a dimensionless parameter between 0 and 1. This form is used to describe luminescence from silicon nanocrystals, where luminescence arises from nanocrystals that vary in size, and β phenomenologically describes the spread in lifetimes that results from this. For nanocrystal films its value varies between 0.7 and 0.9 [103, 17]. In our Er:Si-nc film $\beta = 0.8$, which is more evidence that infrared erbium luminescence is sensitized by nanoclusters. For comparison, our Er:SiO$_2$ films all exhibited single exponential behavior, meaning that for Er:SiO$_2$, $\beta = 1$.

5.4   Electroluminescence in LC-Sensitized Films

Here we present work on electroluminescence (EL) of these LC-sensitized films. Our attempts to observe EL in Er:SiO$_2$ films failed, and we believe that the excess silicon inclusions will allow for a greater current density to be passed through these films. Silicon nanocrystal films have exhibited electroluminescence [18, 19], and Er:Si-nc films have also exhibited EL [104, 64, 105, 106]. There is also evidence that the electrical cross section of Er:Si-nc is significantly larger than the optical cross section [64]. But to date we are not aware of any erbium electroluminescence arising from LC-sensitized films.
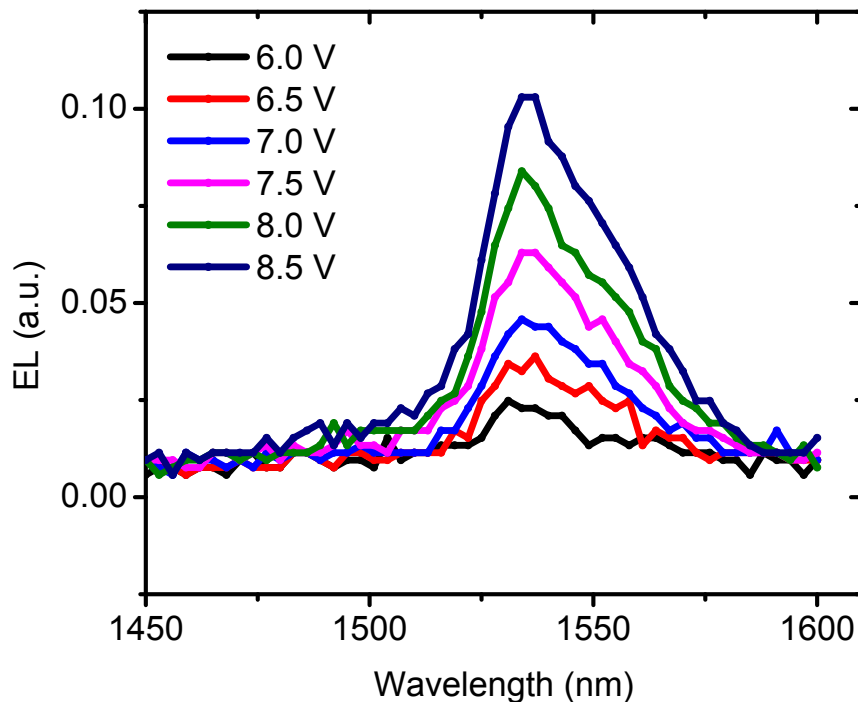
Figure 5.9: Electroluminescence in a 14 nm Er:Si-nc Film. The EL is due to Er^{3+} emission and increases monotonically with voltage up to breakdown.

5.4.1 Experiment Setup

To explore the electrical properties of our Er:Si-nc films, we fabricated N-I-P diodes by depositing Er:Si-nc films on degenerately doped p^+ silicon substrates that were prepared using traditional RCA cleaning. After deposition, the erbium was activated using the same thermal treatment used for films in the previous section — RTA annealing at 500 °C in nitrogen for 5 minutes. Photolithography was used to define top contacts, and 100 nm of indium tin oxide (ITO) was then sputtered as the top contact in a AJA RF sputtering system. ITO was chosen as the top contact because ellipsometry shows it is 91% transparent at the Er^{3+} emission wavelength of 1.5 μm, and 94% transparent at the pump laser wavelength of 488 nm. ITO can also be heavily doped n-type, and hall measurements on these films showed a carrier concentration of 6.5×10^{19} cm^{-3}, with a mobility of 48 cm^2/V-s. Finally, a thick aluminum contact was evaporated on the back side of the sample.

Electroluminescence experiments were carried out on the same stage as the previous PL experiments, which allowed the existing setup to be used for alignment purposes. Contact to the back contact of the device was made by clipping to the electrically isolated vacuum stage. Contact to the ITO top pads was achieved by a probe tip mounted to the sample stage. Electrical characterization was performed by a Kiethley 236 source measurement unit (SMU).
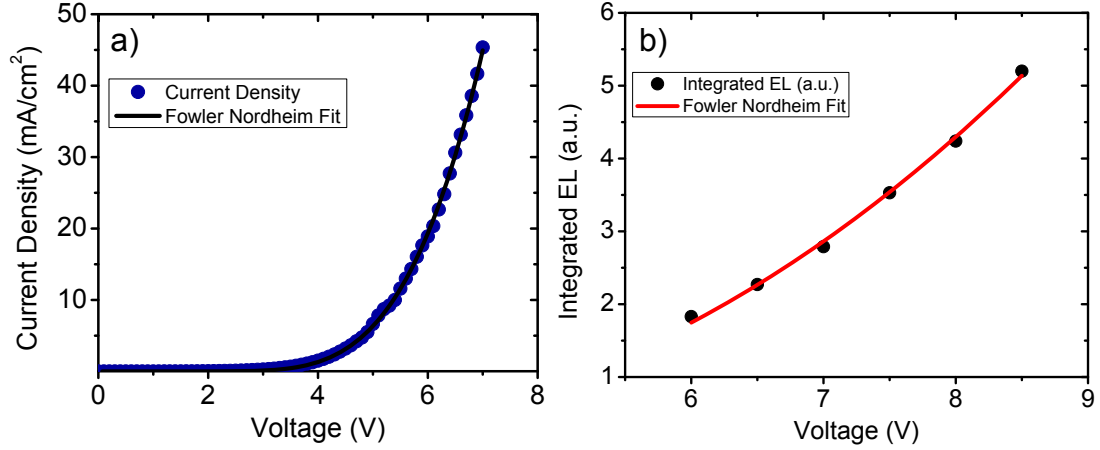
Figure 5.10: a) Current voltage characteristic of 14 nm Er:Si-nc film (dots), with a Fowler-Nordheim tunneling fit (black line) indicating FN tunneling is the current conduction mechanism. b) Integrated EL as a function of voltage (dots) with a Fowler-Nordheim fit (red line), indicating that EL also follows FN-like behavior.

5.4.2 Electroluminescence and Fowler-Nordheim Conduction

To take EL data, a DC bias was placed across the device, and a chopper at 45 Hz was used at the monochromator input to use the lock-in technique. Figure 5.9 shows the erbium EL output for a 14 nm Er:Si-nc film, for voltages between 6 V and 8.5 V. The emission has the same spectral shape as Er^{3+}, and increases monotonically with voltage up to 8.5 V. This device exhibited breakdown shortly after the 8.5 V spectrum was taken, meaning that these devices do not emit until close to the breakdown voltage of the oxide. Taking the breakdown voltage to be 8.5 V, the breakdown field across the 14 nm film, E_{BD} is $E_{BD} \sim 5.9$ MV/cm. For comparison purposes, a high quality thermally grown SiO_2 film breaks down between 8–14 MV/cm [107]. The lower breakdown field is likely indicative of the lower quality film grown by sputtering, as well as the higher defects still in the film due to the lower annealing temperature.

For some insight to the electrical properties, current-voltage (IV) scans were also taken on these films. The IV scan for the 14 nm Er:Si-nc film is given in Fig. 5.10a. The top contact size for these films is 5.16 mm^2, which gives a current density of up to 45 mA/cm^2 at 7 V. The current density is diode-like, with a current onset around 4 V for this film.

In the literature, there have been multiple mechanisms that have given rise to conduction in Er:Si-nc films, such as Poole-Frenkel emission [106], space-charge limited conduction [108], and impact excitation [64]. We find that our films fit best to Fowler-Nordheim (FN) conduction [33], which is described by
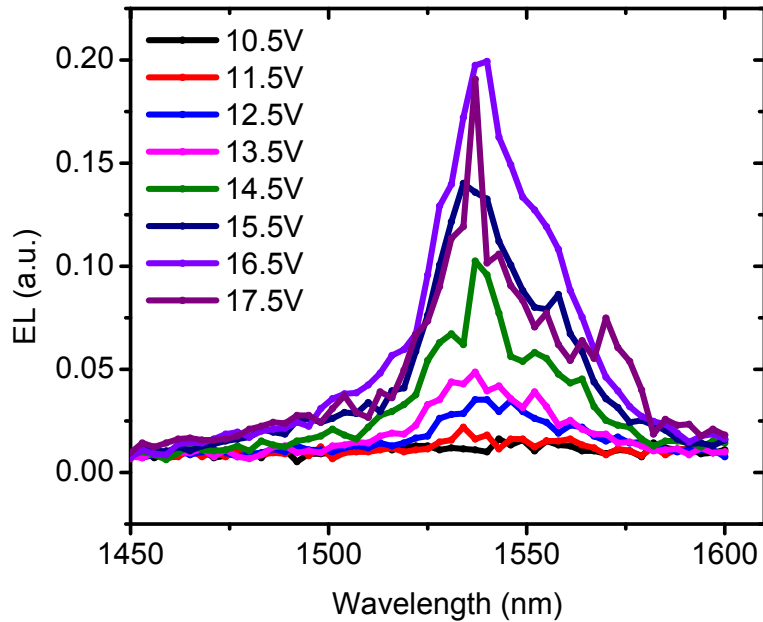
Figure 5.11: EL spectra for 35 nm Er:Si-nc Film. The spectra occur at higher V due to the thicker film, but also monotonically rise with V and exhibit breakdown at a similar field to the 14 nm film.

$$J \propto V^2 \times \exp(B/V) \tag{5.7}$$

where B is a fitting parameter. A fit of this type is shown by the black line in Fig. 5.10a, and agrees well with the data. It is also informative to plot the integrated EL as a function of voltage. The EL was integrated over the emission band between 1450 nm and 1600 nm, and plotted in Fig. 5.10b for voltages between 6 V and 8.5 V. The FN fit well approximates the exponential rise in the EL, which supports our theory the FN tunneling is responsible for both current conduction and EL in these films.

Erbium EL was also observed when using a 35 nm Er:Si-nc film, prepared the same way as the 14 nm film, only with a 0.81 mm^2 top ITO contact. The EL spectra for this thicker film is given in Fig. 5.11. The onset of EL now occurs at a higher voltage, 11.5 V, due to the increased film thickness. Like the 14 nm film, breakdown occured after taking the spectrum at 17.5 V, which means the breakdown field is 5.1 MV/cm, similar to the value found for the thinner film. The current density in this film is also well described by FN conduction, as shown in Fig. 5.12a. The integrated EL is also fit well by a FN model as a function of voltage, as shown in Fig. 5.12b.

The spectra in Fig. 5.11 are noticeably noisier than the spectra given for the 14 nm film in Fig. 5.9. This is not due to noise in our detection system, as the absolute magnitude of the emission is slightly higher in the thick film than the thinner film. We believe that the emission is, to first order,
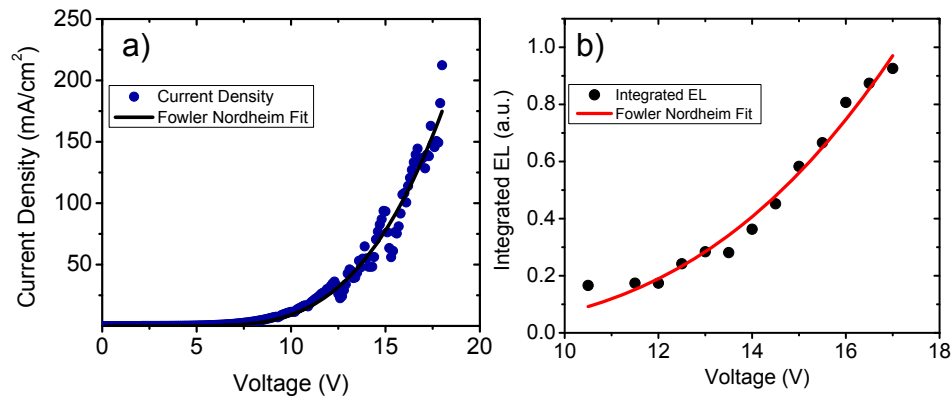
Figure 5.12: IV characteristic (a) and EL characteristics (b) of a 35 nm Er:Si-nc film, shown with FN fits to both curves (solid lines)

directly proportional to the electron flux passing through the film. This means the noise in each spectrum is most likely due to variability in the conduction at a fixed voltage. This is seen in the IV curve in Fig. 5.12a, and this variability in current is manifested in the EL spectra. Each spectrum takes 8 minutes to obtain, and we attribute the noise in each spectrum directly to the time variation of the current.

5.4.3 Annealing Temperature Dependence of EL

The films in the previous section were annealed at 500 °C because that was the temperature that maximized the photoluminescence intensity. However, there is reason to believe that the optimal T_A for PL would not necessarily be the same for EL. Photoluminescence in these films is thought to arise from light interacting with well-coordinated Si-ncs which are coupled to Er^3. Under electrical excitation, charge must pass through a semi-insulating film to excite Si-ncs and Er^{3+}. This fundamentally requires that there be electrically active pathways in the films. These pathways may be defects that would decrease luminescence under photo-excitation, and they may evolve with T_A in a way that is different from that found under PL.

To explore this question, we fabricated diodes in the same way as the previous section, but varied the annealing temperature between 400 °C and 600 °C. All anneals were done in the RTA system under nitrogen for 5 minutes. We also fabricated a sample with no annealing for comparison purposes. All of these films are 23 nm thick.

The results are plotted in Fig. 5.13, with Fig. 5.13a showing the integrated EL for the various samples as a function of voltage. It is interesting that the voltage onset of EL actually decreases as a function of T_A. Figure 5.13b shows the measured I-V curves for each sample. In each case, the onset of current directly corresponds to the onset of EL. This is shown more explicitly in Fig. 5.13c,
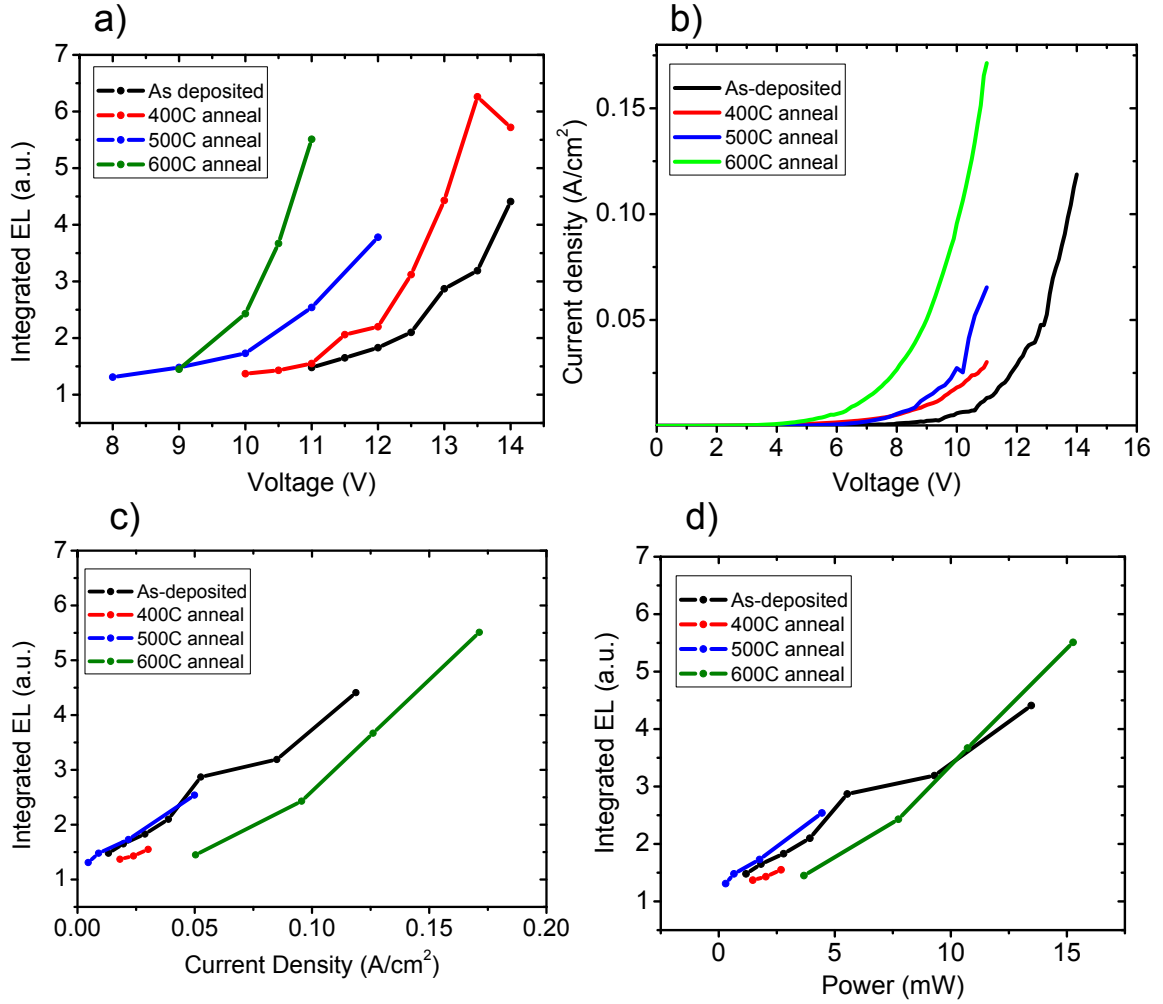
Figure 5.13: a) Integrated EL as a function of voltage for films annealed at 400 °C, 500 °C, 600 °C and as-deposited. b) I-V traces for all samples. The onset of current corresponds to the onset of EL. c) The integrated EL versus current density is linear for all samples. d) Input power versus integrated EL, showing little variation due to T_A

which plots the integrated EL as a function of current density. In this case all the curves are linear, indicating a direct proportionality between the electron flux and EL. The last curve, in Fig. 5.13d, plots the integrated EL as a function of input power, which is computed by multiplying the current and voltage for each point.

The EL versus voltage curves and the I-V curves show a very clear trend with annealing temperature, namely that current and EL appear at lower voltages with increasing T_A. Interestingly, this suggests that the films are made more conductive by annealing up to 600 °C. This may be due to more complete Si-nc formation, which could provide a more efficient conduction pathway than oxide defects. However, this trend is not as clear when plotting EL versus current. The linear curves in Fig. 5.13c roughly show the same trend, but it is not conclusive. The picture is less clear when plotting EL versus input power. In this case the curves merge closer together, which means that from a power efficiency standpoint it does not seem to matter which T_A is used in these films. This is obviously very different than that seen in PL, where there is a very clear maximum in luminescent efficiency with T_A. Fixing the current density and looking at the EL output shows no conclusive data, and this is mainly due to film breakdown, which prevented a more thorough characterization of these films. This is why some of the current data for the 400 °C film is missing in Fig. 5.13.

5.4.4   Discussion

The noise in the spectra in Fig. 5.11 speak to a principle problem incurred while studying these films, and that is that they are presently unreliable. In general, the results produced here were not completely repeatable, with many samples showing no EL, or showing brief EL before shorting. None of the films studied survived biasing long enough to obtain a time-dependent trace, which did not allow us to study lifetimes or rise times, or extract an effective electrical excitation cross section for these films. However, we believe this to be the first exhibition of EL in the LC-sensitized Er:Si-nc film system.

The main impediment towards realizing more efficient devices is oxide breakdown. This is the reason the curves in Fig. 5.13 all stop where they do. Presumably, if all the Er^{3+} could be accessed electrically these curves would saturate at some current density as the Er^{3+} becomes inverted. That this is not observed is evidence that we are not close to inversion. And practically, the magnitude of the EL intensities in these films is much less than what we can observe in PL. This is due to the much higher photon flux that can be applied to the films.

These devices must be operated at close to breakdown to see luminescence, and even then the current densities are much less than an A/cm^2. For reference, in Chapter 4, our rate equation modeling showed that 10 A/cm^2 was the electron flux needed for gain under pulsed operation. Interestingly, other films in the literature show higher electron fluxes, with densities as high as 2 A/cm^2 in sputtered films [108], and 20 A/cm^2 for PECVD films [64]. If this system is to exhibit

efficient, bright electroluminescence then the electron flux through these films must be increased.

5.5   Reactive Oxygen Sputtering

To put the films described here in a better context, we fabricated Er:Si-nc films that behave in a more traditional manner to compare the EL characteristics. These films were grown using reactive oxygen sputtering, which is the process of introducing oxygen into the sputtering chamber during deposition. This method was explored by Priolo's group in 2010 [109], and in their paper they show a dramatic increase in the Er^{3+} EL using this method over their previous samples grown using PECVD [64]. They also anecdotally suggest these films are more robust under bias and show greater lifetimes. This work served as the inspiration for the films grown in this section.

5.5.1   Fabrication

Our previous films were sputtered in a pure Ar atmosphere, with the oxygen content in the film coming from the SiO_2 on the sputter target. To fabricate the films in this section, a mixture of Ar and O_2 was plumbed into the chamber during deposition. There are two gas sources going into the system, one of pure argon, and a second of 10% O_2 in Ar. The oxygen content in the chamber during deposition was controlled by varying the ratios of these two input sources. Both input gases were plumbed directly into the ion source, so that the O_2 would be 'reactive' and actively sputter the source target.

The target used for these films was a silicon wafer with pieces of Er and soda lime glass on it. The target change was necessitated because the addition of oxygen during deposition oxidized the excess silicon in the previous target.

5.5.2   Photoluminescence of Reactive Oxygen Sputtered Films

The first films studied using reactive oxygen sputtering consisted of Er:Si-nc films grown using 10% oxygen in an argon ambient. XPS was used on these films to determine a silicon excess of 1.6 atomic percent. The films were annealed at temperatures between 500 °C and 1000 °C, in both nitrogen and oxygen ambients. Data was also taken for three different excitation wavelengths: 476 nm, which is nonresonant with erbium, but excites Si-ncs; 488 nm, which is resonant with erbium and excites Si-ncs; and 980 nm, which is resonant with erbium but is outside the Si-nc absorption band, so should primarily excite Er^{3+} alone. The curves were normalized for input pump power, which was nominally 50 mW. The photoluminescence spectra were integrated over the range 1450–1600 nm and are plotted in Fig. 5.14.

The first plot in Fig. 5.14a shows PL results for samples annealed in nitrogen. The PL is near zero for $T_A = 500$ °C, and peaks strongly for T_A of 800 °C and 900 °C. This is more common
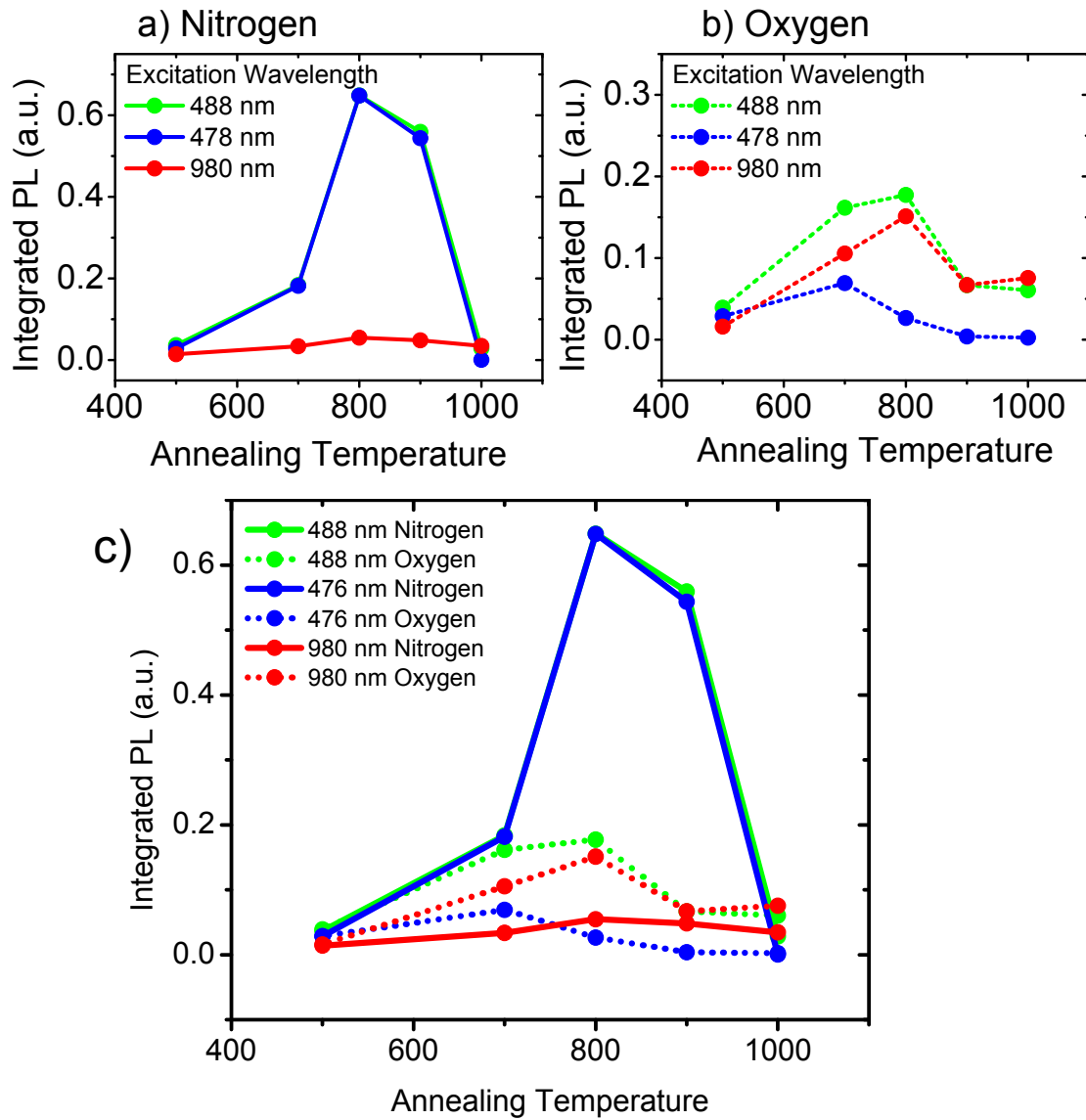
Figure 5.14: a) PL as a function of annealing temperature in nitrogen under 488 nm, 476 nm, and 980 nm excitation, showing sensitization. b) PL for samples annealed in oxygen, which are not sensitized at high T_A. c) PL for both sets of samples, showing that the sensitizing action of Si-ncs greatly increases the PL at 800 °C.

for Er:Si-nc films and immediately distinguishes them from our previous LC-sensitized films. The PL intensity under non-resonant excitation at 476 nm (blue) is quite high, indicating sensitization by the Si-ncs in the film. The 476 nm curve is very close to the curve at 488 nm (green), which is resonant with Er^{3+}. That these two curves overlap is evidence that luminescence in this film is primarily due to sensitization, not to direct absorption by Er^{3+}. PL under excitation at 980 nm (red) is much lower than the other curves, and we attribute this to the lack of sensitizing action at 980 nm, as the Si-nc absorption band generally cuts off around 900 nm. It should be noted that no visible luminescence was observed for these films, which is also evidence that the Si-ncs are preferentially sensitizing Er^{3+} rather than emitting in the visible.

The second plot in Fig. 5.14b is the PL obtained for films annealed in oxygen. In these films the oxygen annealing results in two competing processes. The first is the formation of sensitizing Si-ncs, which we expect for silicon rich samples under high temperature. The second is oxidation of the Si-ncs due to the oxygen present during annealing. At $T_A = 700$ °C there is nonzero PL under all three excitation wavelengths, which first suggests that 700 °C is a sufficient temperature to activate some of the Er^{3+} in these films. The PL under 476 nm excitation indicates that there is sensitizing action going on in these films, meaning some Si-ncs have formed and have not been completely oxidized in the film. PL under 980 nm excitation is also present, indicating the oxygen anneal has allowed some of the Er^{3+} to be excited resonantly. The 488 nm PL is more intense than either the other two, which we expect because it excites Er^{3+} via both mechanisms. As T_A increases, more of the Si-ncs are oxidized away, and the 476 nm PL goes to zero, while the 980 nm PL approaches the 488 nm value, consistent with the film becoming less sensitized and the emission due more to direct Er^{3+} absorption.

Figure 5.14c plots both data sets on the same axis. Here it is clear that the sensitizing action present under nitrogen annealing results in a PL increase of over 3 times at $T_A = 800$ °C. In fact, under 476 nm and 488 nm excitation, all of the films emit more brightly by annealing in nitrogen. But the PL under 980 nm excitation is increased by annealing in oxygen, and we attribute this to a better coordination of Er^{3+} under oxygen annealing, which would lead to brighter luminescence due to direct excitation of Er^{3+}. This effect is seen only under the 980 nm pump due to the lack of excitation of the sensitizing Si-ncs.

Other evidence to support this theory comes from from a study of the lifetimes, which are shown in Fig. 5.15. Parts b and c show the behavior of the lifetime and stretch parameter beta (defined in eq. (5.6)) as a function of annealing temperature under nitrogen and oxygen, respectively. For samples annealed in nitrogen, the lifetimes are around 1 ms, and increase with T_A, as does the beta parameter. The data point for T_A=1000 °C is slightly anomalous, due to the very low PL observed for this sample (Fig. 5.14a). The data for oxygen-annealed samples shows a much clearer picture. For these samples the lifetime increases smoothly with temperature to 8 ms at 1000 °C. The beta
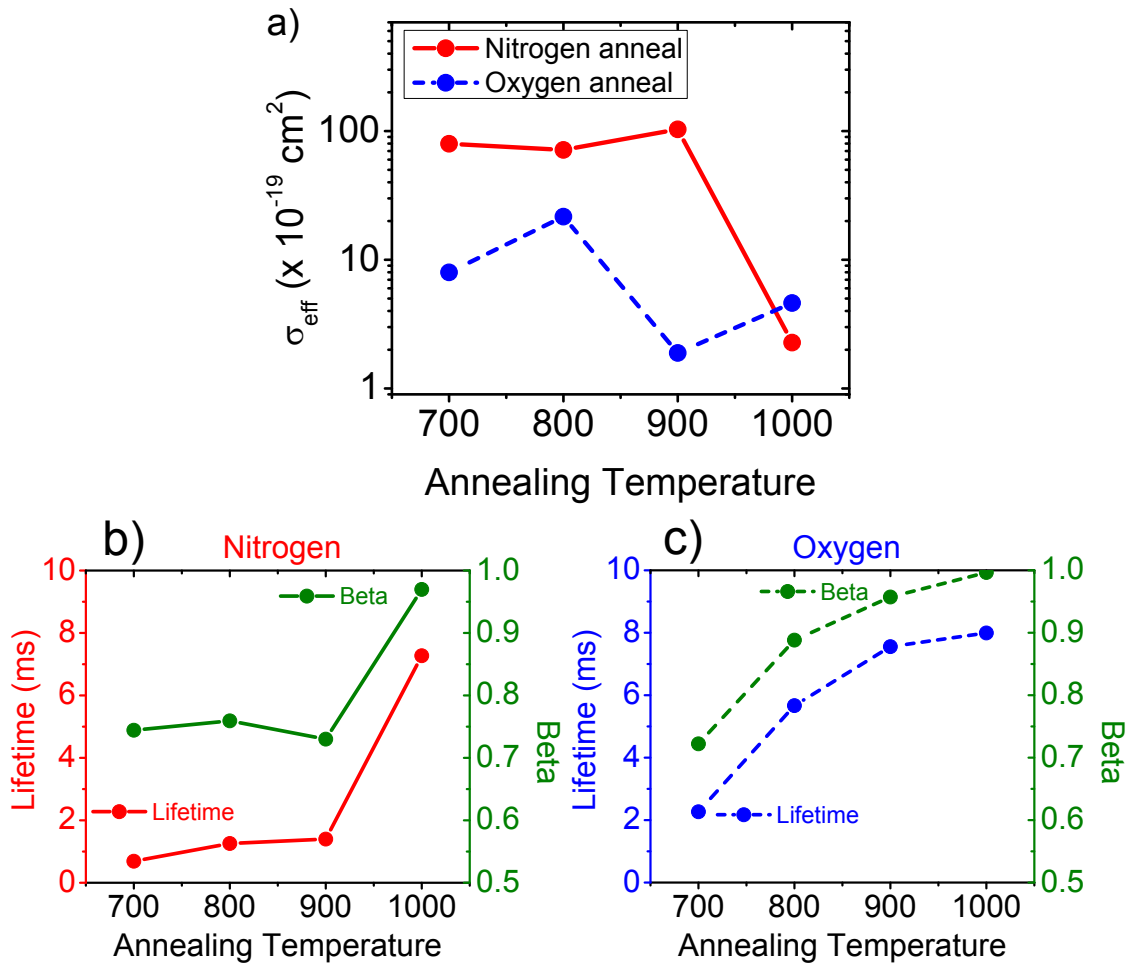
Figure 5.15: a) Effective cross sections under nitrogen (red) and oxygen (blue) annealing. b) and c) describe the lifetime dependence on T_A under nitrogen and oxygen, respectively.

parameter also increases monotonically, which we expect for higher temperatures. In oxygen, the Si-ncs will become oxidized, and in the limit all the excess silicon is oxidized, the film becomes Er:SiO$_2$, and beta should become 1, which is what is observed at 1000 °C. Looking at the data between 700 °C and 900 °C, where there is significant PL for all samples, the samples annealed in oxygen have significantly higher lifetimes. We attribute this again to the extra oxygen in the film leading to better-coordinated Er^{3+}.

Cross-section measurements were also carried out for these films, using the same rise time technique described in Section 5.3.2. The results for each of the samples are shown in Fig. 5.14a. The samples annealed in nitrogen (red) have an effective cross section, σ_{eff}, around an order of magnitude larger than those annealed in oxygen (blue), which is expected by the sensitizing Si-ncs. The magnitude of this effect is less than that generally seen in the literature, where the increase is 3–5 orders of magnitude. This is because at low T$_A$, films annealed in oxygen still contain Si-ncs, and are thus sensitized to some extent. This is why the cross sections measured in the oxygen films at 700 °C and 800 °C, $\sigma_{\text{eff}} \sim 10^{-18}$, are larger than that found for Er:SiO$_2$ films, which are around 10^{-19}–10^{-20} cm^2 [94]. However for the samples annealed at 900 °C, σ_{eff} drops to 10^{-19} under oxygen annealing. At this temperature the film is no longer sensitized, as shown by the lack of emission at 476 nm (Fig. 5.15b).

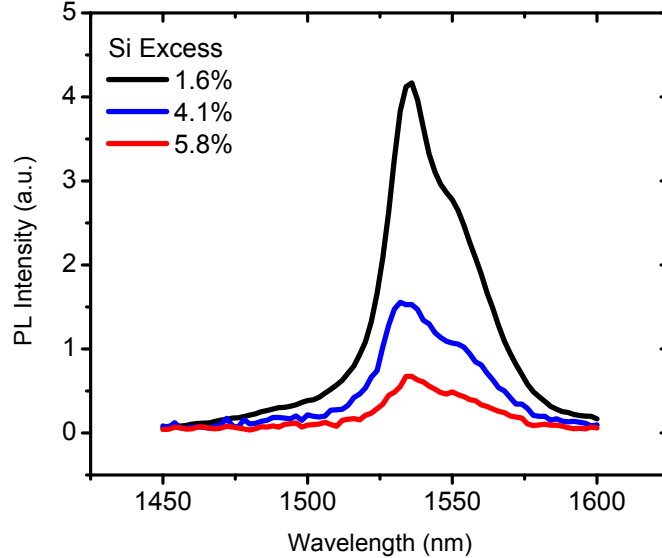### 5.5.3 Silicon Content and Electroluminescence

Figure 5.16: PL intensities of thin films of varying Si excess. The PL intensity drops as the Si excess increases, with the 5.8% excess sample being slightly above the detection limit of our collection optics.

Table 5.1: Properties of Er:Si-nc by Si Excess

Si excess (at. %)	PL (a.u.)	Lifetime (μs)	EL?
1.6	16.1	328	No
4.1	7	68	No
5.8	3	25	Yes

While the films described in the previous section are very efficient Er:Si-nc films for photoluminescence, our attempts at making diodes and observing electroluminescence failed. The oxide in these films only passed a very small amount of current before breakdown, which did not result in any measurable EL.

The films described in the previous section had a silicon excess of 1.6%, as derived from XPS measurements. To drive more current through these films, the silicon content in the films was increased by reducing the oxygen present during deposition. For comparison, EL studies by other groups observed erbium EL from silicon excesses of 6% [109], 13% [64], 22% [106], and 27% [108].

Films were deposited with varied oxygen content during deposition, and the resulting films were characterized using XPS to determine the excess silicon content. This method measures the energy of electrons ejected by X-ray radiation to determine the elemental composition of thin films. Its absolute resolution is on the order of a few atomic percent, which makes it unsuitable for identifying the Er or soda-lime content in the films. However, by looking at the ratio of signal due to oxygen and silicon, it is possible to estimate the ratio of these elements in the film.

Three silicon contents were studied, where each film was 30 nm thick, prepared for EL and annealed at 800 °C for 5 minutes in nitrogen. The PL intensities of these films are given in Fig. 5.16 and show that the added Si content is decreasing the photoluminescence efficiency of these films. At 5.8% excess, the PL intensity is only slightly above the noise level of our detection optics. The luminescent lifetimes of these films also decrease precipitously as the silicon content is added. The lifetimes are given in Table 5.1 and show a decrease from 328 μs for the 1.6% film to 25 μs at 5.8% excess.

Both the PL and lifetime data suggest that adding silicon to these films in this manner degrades the luminescence, but at 5.8% excess EL is observed for these films. This behavior describes the inherent trade-off between PL and EL, as it would appear that the best films for PL are unable to conduct enough to observe EL, while adding silicon to increase conduction degrades the PL properties. The sample using 5.8% excess produced bright, stable PL, and will be further explored in the next section.
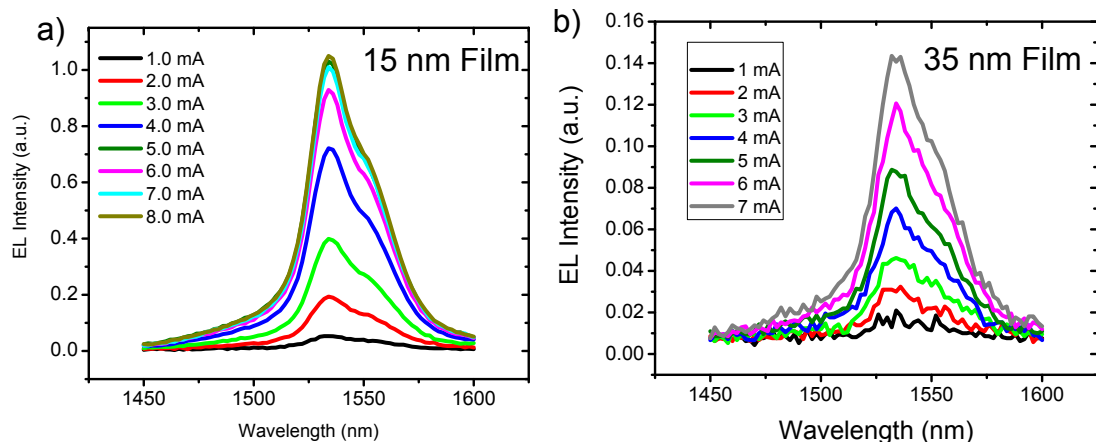
Figure 5.17: EL obtained from films containing 5.8% Si excess. The 15 nm film (a) has brighter EL than the 35 nm (b) film.

5.5.4 Electroluminescence in Reactive Oxygen Sputtered Films

Samples containing 5.8% silicon excess exhibited bright, stable erbium electroluminescence. Reactive-oxygen sputtered Er:Si-nc films were deposited on degenerately doped p^+ silicon substrates that were RCA cleaned and then briefly dipped in hydrofluoric acid to remove the native oxide. After deposition, the films were annealed at 800 °C in nitrogen for 5 minutes to activate the Er^{3+}. ITO pads 0.32 cm^2 in size were sputtered through a physical mask onto the sample, and aluminum was evaporated as a back contact.

Two film thicknesses were studied, 15 nm and 35 nm, which are the two thicknesses where EL was observed in our previous LC-sensitized films. The EL results from the two films are given in Fig. 5.17. Electroluminescence spectra from the 15 nm film (a) are shown for currents through the device between 1 mA and 8 mA. The spectra were much cleaner than any of the previous EL spectra. This is due partly to the increased reliability in these films and partly due to the large contact pad size which allowed a greater absolute signal to be collected. The spectra increased in intensity up to 5.0 mA before saturating. The I-V data also followed a Fowler-Nordheim-like conduction, where the sample was biased up to 16 V before breakdown, indicating a breakdown field $E_{BD} \sim 10$ MV/cm. This was almost double that found for the LC films, and is evidence of the increased oxide quality obtained by reactive sputtering and higher annealing temperatures.

The 35 nm film (b) exhibited similar characteristics, supporting currents up to 7.0 mA as the EL increased monotonically. No EL saturation was observed in this sample before breakdown, and the intensity was actually less than that seen in the thinner film by a factor of 7. This is seen more clearly from a plot of the integrated EL intensity versus current density given in Fig. 5.18. This plot shows a super-linear dependence of EL on current density, and was likely due to multiple excitations
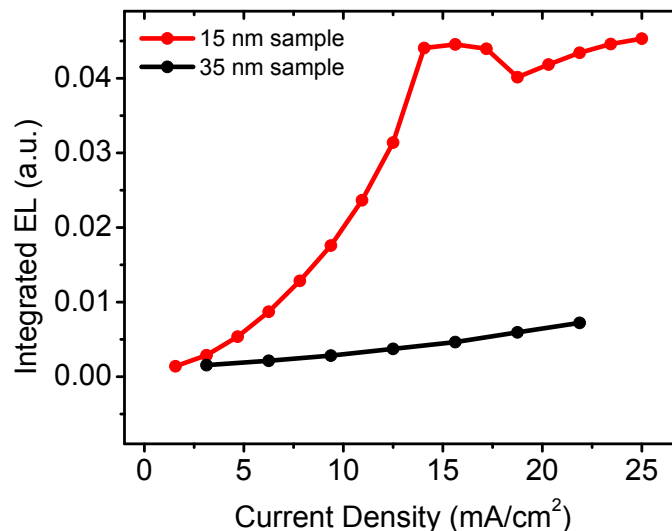
Figure 5.18: Integrated EL of the films shows much brighter EL coming from the 15 nm film (red) than the 35 nm film (black). Both films show a super-linear dependence on current density, with the 15 nm film showing saturation at 14 mA/cm^2.

per electron passing through the film. At high voltages each electron has a high energy, and because electrons aren't absorbed like photons are, each electron is free to interact with multiple Si-ncs as it passes through the film.

These films were much more reliable than the LC-sensitized films described earlier, and as a result time dynamics were obtained for the EL in the 15 nm film. The function generator used to pulse the excitation for time-dependent studies had a maximum output of 14 V, which is the voltage at which the 14 nm film luminesced brightly, but had not reached saturation. The rise time and lifetime at 14 V are given in Fig. 5.19. The EL rise time (a) shows a delay around 15 μs between when the pulse began at t = 0 and when the luminescence started to rise. After this offset, the EL rose with a characteristic time of 30 μs at 14 V. Fig. 5.19b shows the lifetime after the pulse is turned off, and has a characteristic decay time of 14 μs.

Data of this type was obtained for pump voltages of 12.5 V, 13 V, 13.5 V, and 14 V. These data sets allow the calculation of an effective electrical excitation cross section, σ_{EL}, by analyzing the rise time data for the associated current densities. This data is plotted in Fig. 5.19c, and Fig. 5.19 d shows the calculated rise rates (dots) plotted as a function of electron flux. The electron fluxes were calculated from the current densities at the various pump voltages. The red line is a linear fit of the fluxes to the rise rates, using equation (5.4), and shows that for our films $\sigma_{\mathrm{EL}} = 2.5 \times 10^{-13}$ cm^2. The PL cross section for this film was calculated to be 9×10^{-16} cm^2, which means the cross section is two orders of magnitude larger under electrical excitation than under optical excitation. This number agrees with the cross section estimated by Iacona [64], which was found to be two orders of

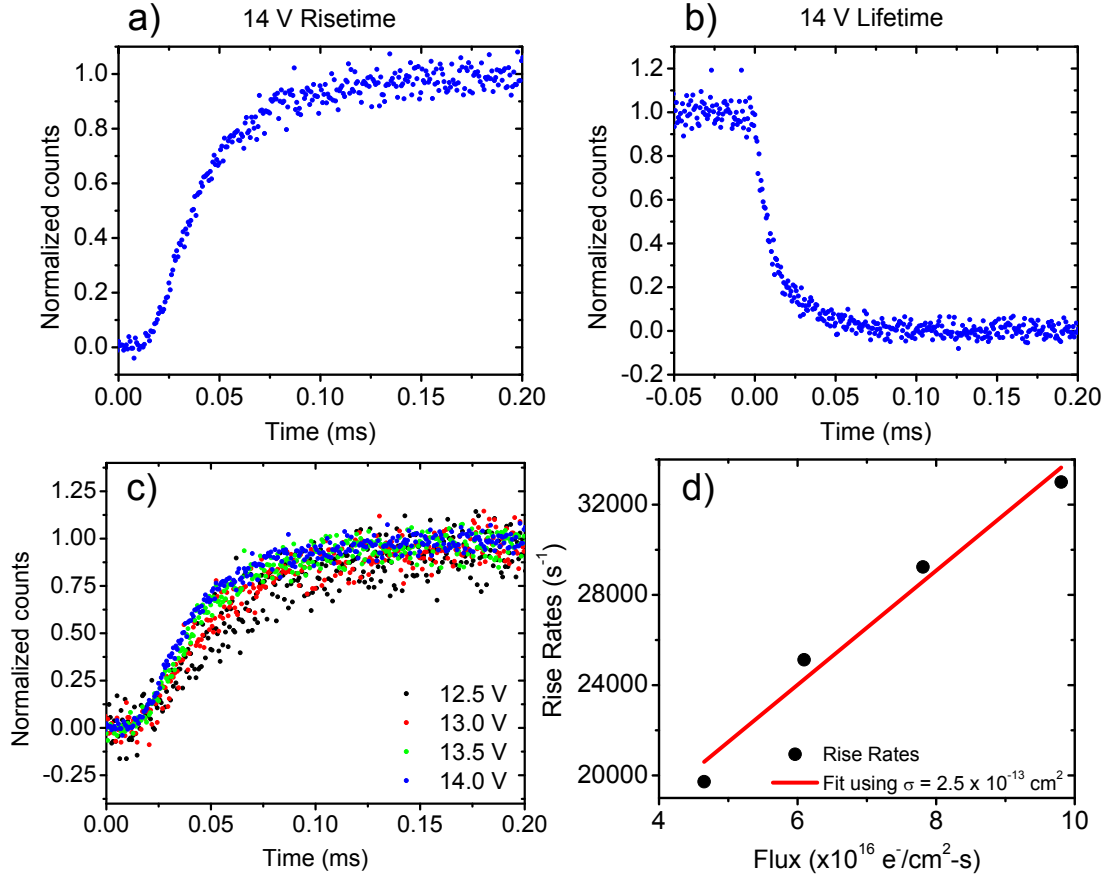Figure 5.19: EL time dynamics of the 15 nm film. a) There is a slight delay between when the pulse starts at t = 0 and when luminescence starts to rise. This is partly due to the RC charging dynamics of the diode. b) The measured EL lifetime at 14 V is 14 μs. c) Rise time data for voltages between 12.5 V and 14 V. d) Inverse rise times plotted to extract the electrical excitation cross section, σ_{EL}
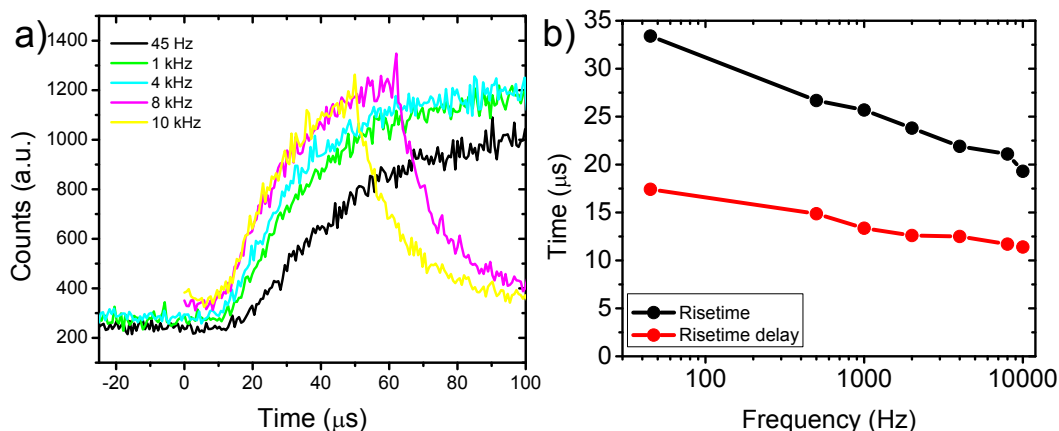
Figure 5.20: a) Frequency dependence of EL output for frequencies between 45 Hz and 10 kHz. b) The frequency dependence of rise times are given in black. Shown in red is the delay between the voltage pulse and EL. Both curves decrease with frequency, suggesting the films become easier to pump at high frequency.

magnitude larger than the PL cross section of their films.

The fit used in Fig. 5.19 to extract σ_{EL} intercepts with the y-axis at $\phi = 0$ at a rise rate of 8833 s^{-1}, which corresponds to a lifetime of 113 μs. For PL cross sections extracted via optical excitation this intercept agrees well with the measured lifetime, but the measured EL lifetime is only 14 μs, which is an order of magnitude lower than the value found by fitting EL rise time measurements. The reason for this is unknown at this time, but may be due to film degradation during sputtering of the top contacts; the PL lifetime for this film away from the ITO contacts is 80 μs. The PL lifetime of the finished device was not obtained because the PL intensity was too low. The lifetime is also affected by the higher dielectric constant of the ITO contact. And lastly, the charging dynamics may have affected the luminescent lifetime. Evidence of charging phenomena is shown by the delay in the EL rise time (Fig. 5.19), and the literature suggests that Auger processes may lead to luminescence quenching by injected carriers [110].

5.5.5 Frequency Dependence of Electroluminescence

In Chapter 4 an argument was made that carriers in the film significantly limit the output of the film when put in a slot waveguide due to carrier absorption. To explore the charging time dynamics of our experimental system, frequency-dependent EL was taken for the 15 nm sample. By pumping in a time domain similar to the rise time of the system, we could observe the effects of pulsed excitation and charging in these films.

The previous EL time dependent data was taken at a pulse rate of 45 Hz, or a 11 ms pulse followed by 11 ms off. This was chosen so that each pulse would be isolated, and the time-integrated

data would represent transitions between the steady-state on and off positions. For the 15 nm EL sample described in the previous section, the sample was pumped at 14 V with frequencies between 45 Hz and 10 kHz, and data was taken using the multi-channel scalar. In each case the duty cycle is 50%. The results are plotted in Fig. 5.20a, and show the rise times under the various pumping frequencies. The rise times increase with frequency, with each trace showing a delay between the pulse turn-on at t = 0 and the onset of luminescence. This is similar to the data found in Fig. 5.19. The background noise level of each trace also increased with frequency, being most pronounced at 8 kHz and 10 kHz. This is due to the fact that at these frequencies, the luminescence does not have time to completely decay before the next pulse turns on.

The delay time between the voltage turn-on and the onset of EL also decreased with frequency. This behavior is plotted with the measured rise times in Fig. 5.20b, on a semilog plot. The decrease in both the delay and rise time are linear in this plot, and both show that the Er^{3+} pumped more quickly at higher frequencies.

While the rise times indicate that the films are more easily pumped at higher frequencies, at some point the high frequency will result in a pulse length that is simply too short to see significant Er^{3+} emission. To look at this more closely we plot the EL intensity versus frequency in Fig. 5.21a. These points were calculated by fitting the lifetime corresponding to each frequency, and measuring the number of counts between the maximum value (when the pump is turned off) and the noise level extracted from each lifetime fit. In this way the EL intensity plotted reflects the actual EL output over the system noise level. A simple peak-to-valley measurement was insufficient because at high frequency the Er^{3+} has not decayed completely to the system noise level before the next pulse. This is also why the lock-in technique is not applicable here — without saturation at the 'on' and 'off' states, the lock-in amplifier filters out the EL signal.

Using this method, the EL reached a maximum at 8 kHz, which corresponds to a frequency where the pulse length is long enough to saturate the EL output. This is more explicitly shown in Fig. 5.21b, where the EL output is plotted as a function of pulse length. The 8 kHz data corresponds to a 62.5 μs pulse length, which is longer than the exponential rise time of 20 μs. At this point the EL was fully saturated and the curve reached a maximum, which stays constant for higher frequencies. To show this, the raw data at 4 kHz is shown in Fig. 5.21c, where the curve saturated in both the on and off states. Figure 5.21d shows the raw data for 8 kHz, which reached the same output level as the 4 kHz data, but the pulse is cut off just as it reached saturation.

To further study the charging dynamics in these films, an oscilloscope was put in line with the function generator to measure the voltage across the device as a function of time during the frequency measurements. The oscilloscope used was a Tektronix TDS 3054B. It was used to verify that the time resolution of the function generator is 40 ns during 14 V output, which is much shorter than the timescales observed here. The results of this measurement for films pumped at 45 Hz are
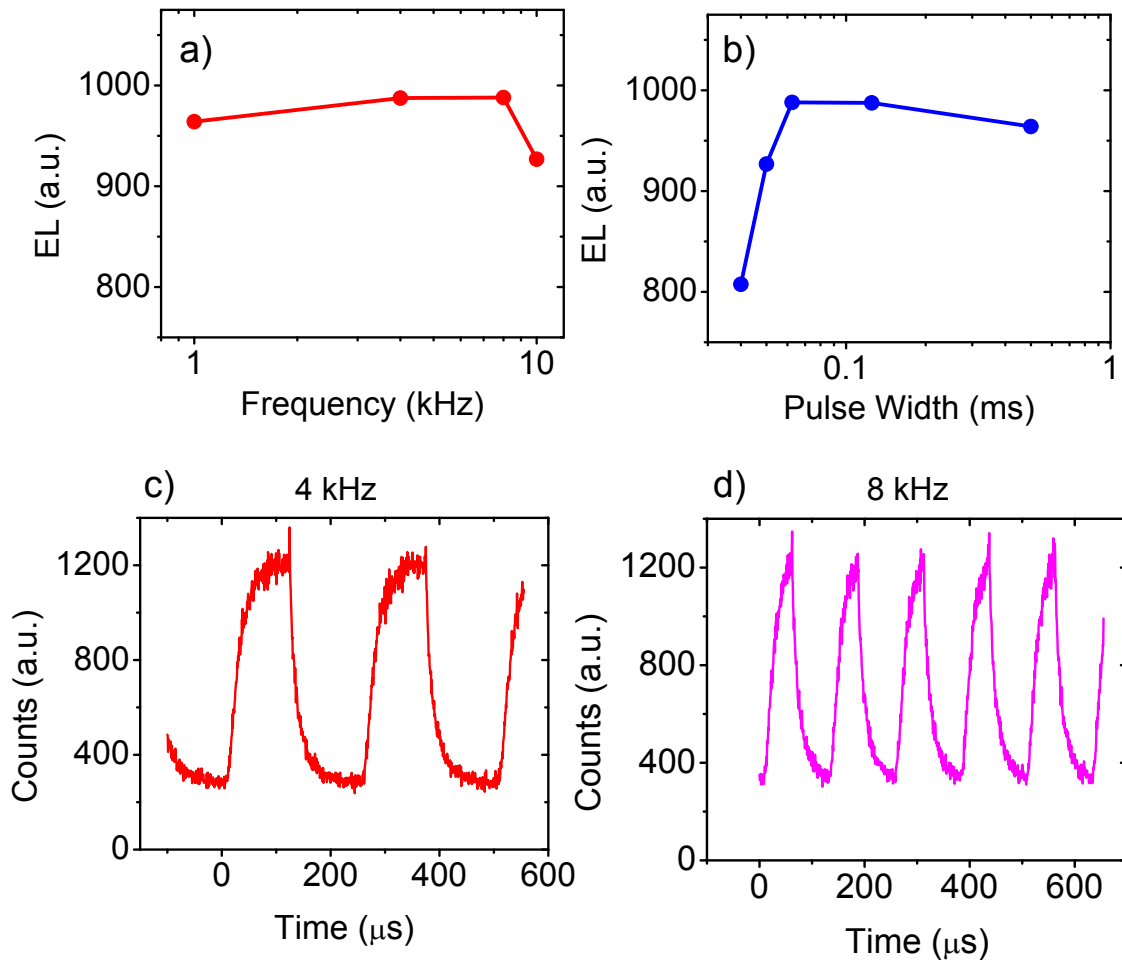
Figure 5.21: a) Maximum EL output as a function of frequency, with a maximum at 4 kHz. b) Maximum EL output as a function of pulse length. The maximum at 0.25 ms corresponds to 4 kHz and is the point at which the pump is sufficiently long to reach EL saturation. The EL outputs at 4 kHz (c) and 8 kHz (d) show this directly.

shown in Fig. 5.22. Part a is the rise time under a pulse starting at t = 0. Shown in black is the voltage measured by the oscilloscope in the circuit, which rises with a characteristic rise time of 5 μs. Meanwhile, the blue dots are the resulting Er^{3+} luminescence, which reach saturation long after the voltage pulse has reached saturation. The rise time of the voltage is 5μs, which can be attributed primarily to RC charging of the devices. If we assume the device to be a RC circuit powered by the 14 V pulse, the measured 5 μs rise time corresponds to a capacitance of a few nF, which is of the same order as the measured capacitance of similar devices. However, this 5 μs delay does not entirely explain the delay in the onset of EL, which is 17 μs. At this point this is unexplained, but may be due to some charging or trapping phenomenon occurring in the film before the onset of EL.
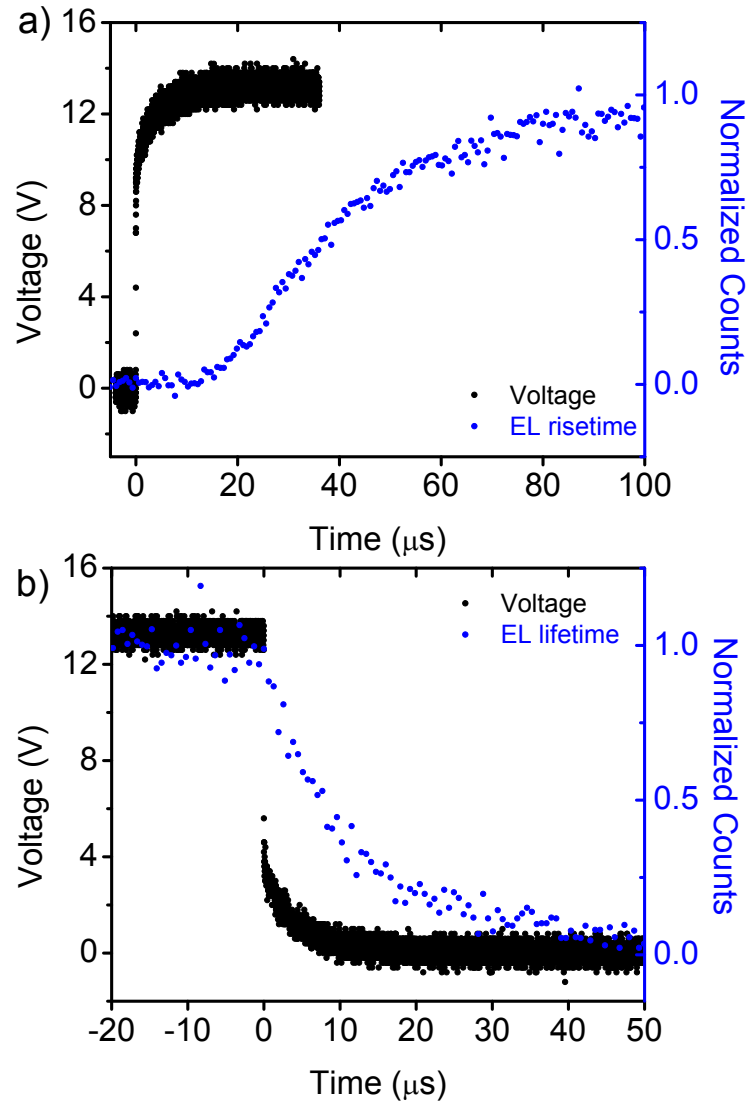

Figure 5.22: Time dynamics of excitation at 45 Hz. Part a) shows the voltage output (black) and the resulting EL (blue dots) during the rise time. Part b) shows the lifetime. The time dynamics of the voltage are attributed to RC charging of the devices.

The lifetime data and voltage are plotted in Fig. 5.22 b. The voltage shows the same 5 μs decay as the rise time, further evidence of RC charging. The EL plot is overlaid, and has a lifetime of 14 μs. However, in this case the lifetime has no obvious correlation with the decaying device voltage. The lifetime has the same value during the first 10 μs while the voltage decays as it does when the voltage has dropped to zero.

The Priolo group has suggested that Auger processes between excited Er^{3+} and injected carriers may quench Er^{3+} luminescence [110]. The Auger process presented for their films resulted in two distinct lifetimes: a short one corresponding to Auger quenching, followed by a longer one when the carriers are swept away, which corresponds to the intrinsic Er^{3+} decay time. If this happened in these films, the Auger component occurred on a timescale that completely encompasses the lifetime decay, meaning the 14 μs decay would be due entirely to Auger processes. This may indeed be the case because the short measured lifetime is of the same order as the Auger process. However, the voltage decays to zero before the luminescence does, meaning that the carriers causing all the quenching would have had to be few enough to not register as a measured voltage. There is also the fact that the PL lifetime is similarly low, where presumably the optical excitation would leave far fewer residual carriers than electrical excitation. Last, this paper shows explicitly that an applied current partly quenched PL in their samples. This did not happen in these films, and in fact, the EL and PL intensities added to one another. Unfortunately, these samples all shorted before a thorough exploration of this phenomenon could be pursued.

The dynamics observed at 45 Hz in Fig. 5.22 don't change significantly at higher frequency. The rise time (a) and lifetime (b) for the data obtained at 8 kHz is shown in Fig. 5.23. The voltage rise and fall times were 5 μs, just like at 45 Hz, but the EL data was not normalized because of the lack of saturation in the on and off states. The rise time behavior was very similar, with a 12 μs delay between the pump pulse and the EL onset. However, the EL level for $t < 10$ μs was due in part to noise, and in part to Er^{3+} emission from the previous pulse. The lifetime behavior is almost identical to that found at 45 Hz, with the measured lifetime of 12 μs. However, the delay in the rise time at 8 kHz was 12 μs, which is slightly less than the 17 μs found at 45 Hz. These magnitudes are not conclusive, but do represent some evidence that if some charging effect occurs during the rise time, then whatever converse effect occurs during decay is not as efficient at quenching luminescence, since the lifetime decrease is less than the rise time delay decrease, and is very close to the noise limit of these measurements.

That the behavior was so similar between the high and low frequency pulsing is evidence that either there are no significant charge quenching effects happening here, or that the charge effects are the dominant mechanism for the Er^{3+} decay in these films. The only explicit behavior related to possible charging dynamics was the decrease in the rise time with frequency, and the decrease in the delay between the voltage pulse and the onset of EL. That a similar or complimentary process
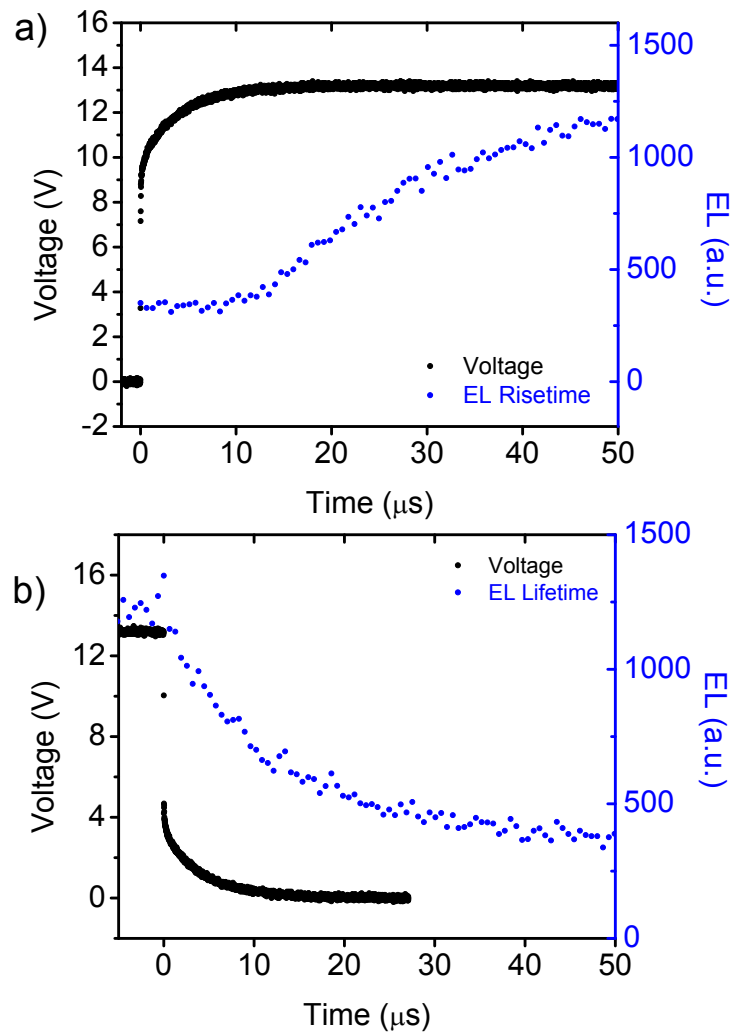
Figure 5.23: The time dynamics of EL at 8 kHz are very similar to that at 45 Hz. The EL is not normalized due to the lack of a saturated signal during the on and off pulses. Again the rise and fall times of the voltage are 5 μs, and have no apparent effect on the EL lifetime.
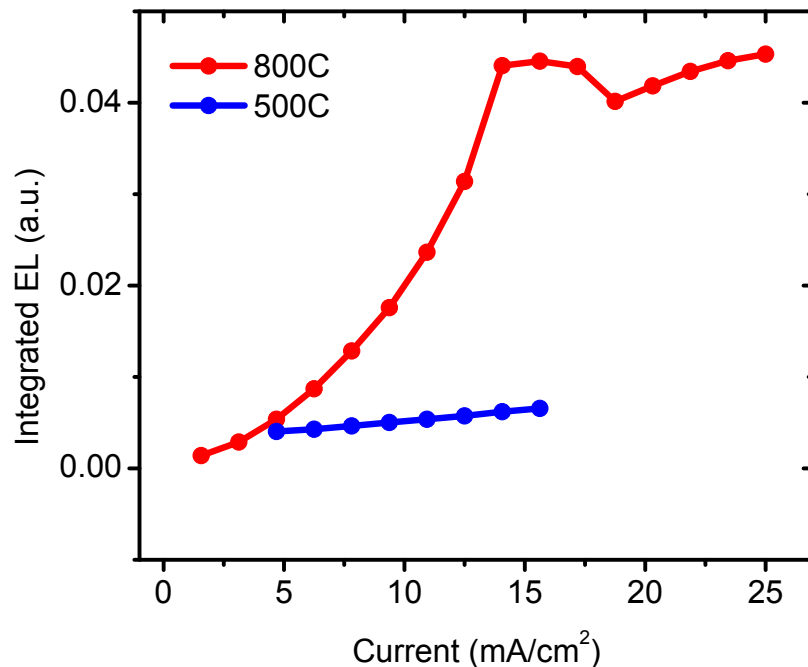
Figure 5.24: The EL intensity of the film processed at 800 °C is greater than an identical film annealed at 500 °C. This strong dependence on T_A was not seen in LC-sensitized films.

did not obviously appear after the pulse is turned off would suggest that any charge effects are either very short lived ($< 5 \mu$s) or very long lived ($> 100 \mu$s). Unfortunately, the intrinsically short lifetimes of these films limit study of any long-lived components that may have been present. To that end, I would argue that the main aspects limiting luminescence in these films were the high number of defects introduced by adding silicon to the matrix, as well as some degradation due to sputtering of the top ITO contact.

5.6 Comparison of LC- and Si-nc- Sensitized Films

In this chapter EL was observed from films exhibiting LC sensitization, and from films exhibiting Si-nc sensitization. The main characteristics of the LC-sensitized films were an insensitivity to annealing temperature, and a roughly linear dependence of the EL on the current density. The Si-nc sensitized films were far more robust than the LC films, which allowed lifetimes, rise times, and an electrical excitation cross section to be obtained. These films also were fabricated using a higher annealing temperature.

Both films were grown at two thicknesses, 15 nm and 35 nm, and in each case the luminescence from the 15 nm film was more efficient than that from the 35 nm film. Both films also exhibited very small luminescent lifetimes in PL. But EL from the Si-nc films exhibited a super-linear dependence
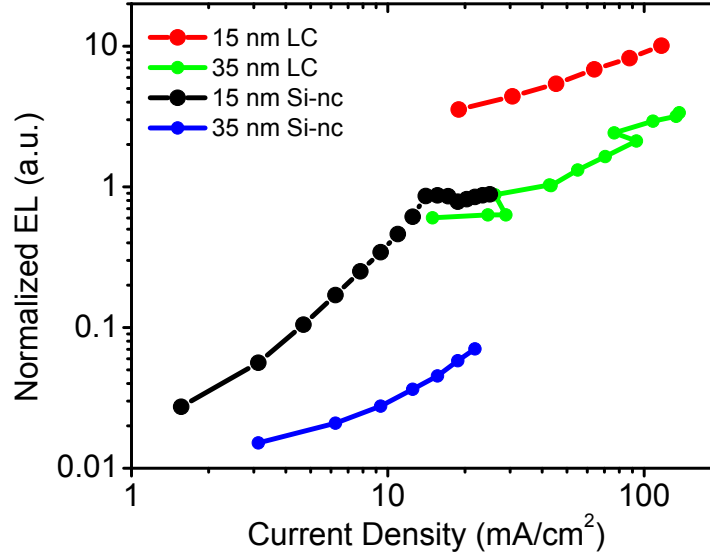
Figure 5.25: The EL for LC and Si-nc sensitized is plotted, normalized for film thickness. The LC films allow higher current densities, but the Si-nc films are more robust.

on current density, which was not seen in the LC films. The other hallmark of the LC films is an insensitivity to the annealing temperature. To further distinguish the films, a Si-nc-sensitized film was fabricated with an annealing temperature of 500 °C. EL was obtained from this sample, and the integrated EL is plotted as a function of current density in Fig. 5.24, along with the curve for the sample annealed at 800 °C (red). The sample annealed at 800 °C was much brighter than the 500 C sample, and this dependence on annealing temperature is something that was not seen in the LC films.

The last distinction to make between the sets of films is the absolute current density. The Si-nc films used larger contacts to obtain the EL data, and had lower current densities for observed EL than the LC films. Fig. 5.25 is an attempt to normalize out all the factors affecting the data by plotting the integrated EL for both LC- and Si-nc-sensitized films, normalizing the EL by the film thickness. This plot shows that while the the Si-nc films were more robust, the current densities achievable in these films was almost an order of magnitude lower than the LC films. Figure 5.25 also shows clearly how the thinner films were more efficient at EL by normalizing out the film thickness. This is encouraging for work towards building a slot waveguide structure, where a thin film will need to be accessed and electrically pumped. This plot indicates that going to thinner films is better for the EL properties, which is not the case under PL, where Er^{3+} intensity scales with the thickness. At this point one can only speculate as to why this would be, but it is possible that for thin films the carriers are more able to tunnel into the entire thickness of the film, where they could form excitons or perform some other sensitizing action. It is also possible that for thick films, biasing
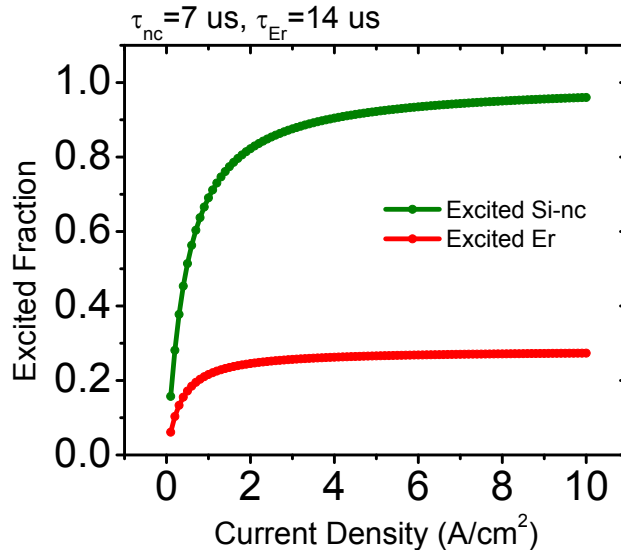
τ_{nc}=7 us, τ_{Er}=14 us

Figure 5.26: Rate equation modeling using the values obtained experimentally indicates the excited Er^{3+} fraction is very low for our observed pump flux of 0.015 A/cm^2, and would only reach a maximum of 27% if higher current densities were possible.

results in more of a filamentary-style conduction, which would be more likely to create shorts in the film without accessing all the Er^{3+}. A third option is that perhaps the thin films allow more of the defects and traps in the film to be filled with charge, allowing a more even distribution of the current passing through the film.

5.7   Rate-Equation Modeling of Experimental Data

In Chapter 4, a rate-equation model was presented that analyzed the effects of putting Er:Si-nc in a silicon slot waveguide and trying to invert it electrically for laser operation. The numbers used for that work were taken from literature and were chosen to represent properties observed in fabricated films. The main assumptions made were that all the Er^{3+} was coupled to Si-ncs, and that an arbitrary amount of current could pass through the films. Here we revisit those calculations, using the experimental values found in this chapter. Specifically, an electrical cross section was measured, $\sigma_{EL} = 2.5 \times 10^{-13}$ cm^2, an Er^{3+} lifetime of 14 μs was observed for EL, and maximum current densities of 0.1 A/cm^2 and 0.015 A/cm^2 were found for for LC-sensitized and Si-nc-sensitized films, respectively.

To revisit the modeling, the erbium concentration was set at 0.4 atomic percent, or 2.5×10^{20} cm^{-3}, and the number of Si-ncs was kept at 10^{19} cm^{-3}. The decay rate of the Si-ncs was assumed to be 7 μs, which is smaller than that seen for well-formed crystalline Si-ncs, but is reasonable for nanoclusters processed at lower temperatures, such as the 800 °C annealing treatment used in the
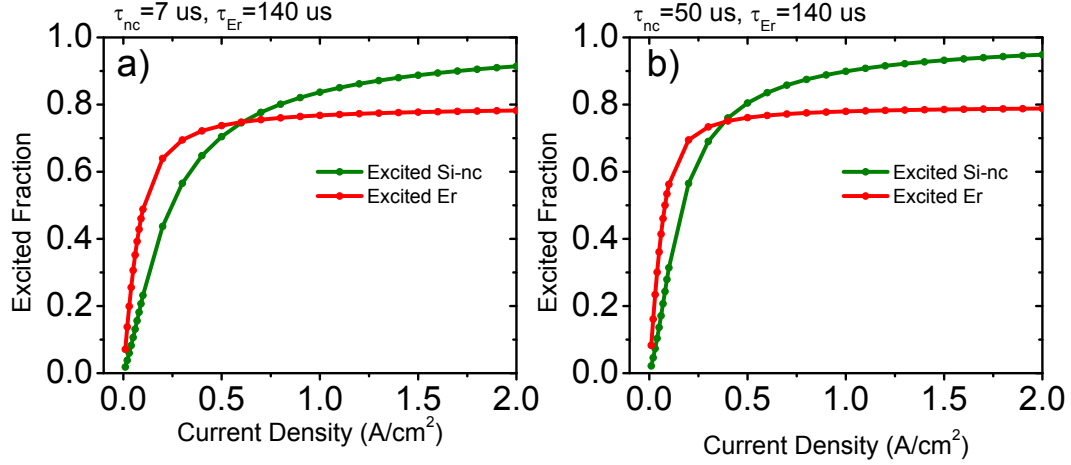
Figure 5.27: a) The excited fractions of Er (red) and Si-nc (green) as a function of pump power, calculated from a rate equation model using parameters obtained experimentally. Plot b) increases τ_{Er} an order of magnitude to see Er^{3+} inversion, and plot c) then increases τ_{nc} an order of magnitude.

previous section. Full coupling between Si-nc and Er^{3+} is still assumed, and the measured values for τ_{Er} and σ_{EL} were used.

One argument made during discussion of EL in the fabricated films is that the EL output was limited by the current density that could be achieved in these films before breakdown. To examine this, the current density was varied in the rate-equation model, using our experimentally obtained numbers under CW pumping. The output is given in Fig. 5.26. The Er^{3+} excited fraction is predicted to maximize at 27%, even if pumping at 10 A/cm^2 were possible. For reference, the maximum current density observed experimentally was 0.02 A/cm^2 for the Si-nc-sensitized film, which in this plot corresponds to an Er^{3+} excited fraction of 10%. The reason for the low excited Er^{3+} fraction is largely due to the small lifetime of 14 μs observed in these films.

EL lifetimes in the few 100 μs range have been measured [64, 109], so it is reasonable to run the simulation with a larger lifetime. Figure 5.27a plots the excited fractions if the Er^{3+} lifetime was increased an order of magnitude, from 14 μs to 140 μs. The Er^{3+} fraction reached inversion at 0.2 A/cm^2, and saturated at at 79% excited fraction. The inversion threshold of 0.2 A/cm^2 has been reached in films of this lifetime in the literature, and the lack of complete coupling is what kept this inversion from being observed experimentally. Fig. 5.27b plots the results arising from increasing the Si-nc lifetime to 50 μs. This plot is very similar to the plot that assumes $\tau_{nc} = 7\mu$s, which suggests that the value of τ_{nc} has little effect on the Er^{3+} inversion, so long as the Si-ncs sensitize Er^{3+} before emitting or non-radiatively decaying.

To further explore the lifetime dependence in these films, Fig. 5.28 plots the excited Er^{3+} and Si-nc fractions as a function of Er^{3+} lifetime under 0.015 A/cm^2 (a) and 0.15 A/cm^2 (b) pumping. Using 0.015 A/cm^2 pumping represents the point at which the EL became saturated in the Si-nc
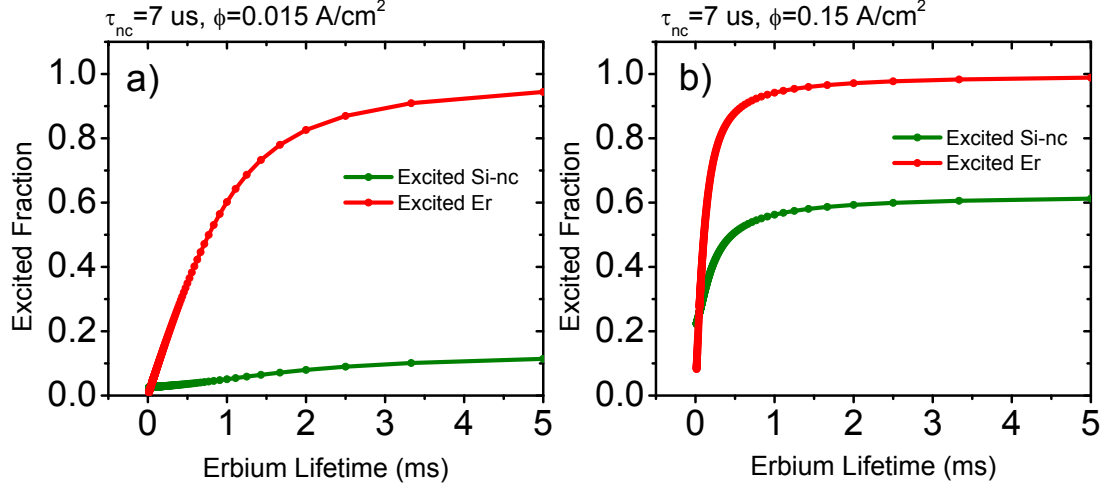
Figure 5.28: a) Excited fraction under 0.015 A/cm^2, as a function of τ_{Er}. b) By increasing the pump to 0.15 A/cm^2 the onset of Er^{3+} can be achieved at lower τ_{Er}.

film (Fig. 5.18), and using 0.15 A/cm^2 represents an order of magnitude increase in the current density, and has been observed in other films. Under 0.015 A/cm^2 pumping, Er^{3+} becomes inverted if $\tau_{\mathrm{Er}} = 0.8$ ms, and can reach almost complete inversion for lifetimes of a few ms. While 0.8 ms is a number that has been achieved experimentally, we are not aware of any measured EL lifetimes in the ms range, though such lifetimes have been observed in PL [68]. If the pumping current density can be increased to 0.15 A/cm^2 (5.28b), then Er^{3+} can be inverted at a lower τ_{Er} value of 109 μs, and reach an excited fraction of 94% if the lifetime is 1 ms.

This behavior gets to the major trade-off observed between lifetime and current in electrically pumping these films, which was hinted at in Table 5.1. Erbium is desirable as a lasing material because of its long lifetimes in oxide, but pumping the oxide electrically via Si-ncs inherently lowers this lifetime. It is hoped that by improving film fabrication this trade-off can be mitigated enough to achieve gain in these films.

While achieving inversion in these films is desirable and necessary for laser operation, electrically pumping these films introduces carriers and excites nanoclusters in a way that could limit gain. This phenomenon was explored in Chapter 4 theoretically by examining a figure of merit, χ, which is the gain due to Er^{3+} divided by absorption due to excited nanoclusters:

$$\chi = \frac{\sigma_{\mathrm{Er}}(N_{Er}^* - N_{Er})}{\sigma_{\mathrm{ca}} n_{\mathrm{nc}}^*}, \tag{5.8}$$

where σ_{Er} is the emission cross section at 1.5 μm, N_{Er} is the number of Er^{3+} ions in the ground state, and N_{Er}^* and n_{nc}^* are the excited fractions of Er^{3+} and Si-ncs, respectively. The carrier absorption cross section is σ_{ca} and is assumed here to be the lower value presented in Chapter 4 of

$\sigma_{\mathrm{ca}} = 4 \times 10^{-19}$ cm^2.

For the calculations presented here, $\chi < 1$ for almost all the plots presented here, meaning that even though Er^{3+} can be inverted, the emission is absorbed in a cavity by the excited nanoclusters. In this case we showed that pulsing the excitation is a way to achieve gain because the Er^{3+} stays inverted after the Si-nc excitations decay away [92].

The only situation where $\chi > 1$ was shown in 5.28a, where the excited Si-nc popuation remained low due to a small pump flux of electrons, while Er^{3+} was inverted. This suggests that if a film could be inverted with such a small current density, it would lase under CW pumping. However, there are a few subtleties to consider here. This model assumes that as electrons pass through the film, they excite the nanoclusters, and this excitation fraction is what is plotted. What is not explicitly considered here is trap charging due to electrical excitation. Presumably the film itself becomes charged under electrical excitation as the defects and traps in the film fill up, and these charges would not 'decay' in the same sense as a luminescent-excited Si-nc would. Under a pulsed injection mechanism, the voltages could be engineered to sweep these extra trapped carriers away, but in CW these charges would be present, and could contribute to absorption in a way very similar to the carrier absorption described in Chapter 4.

So then the question becomes whether the films were being charged, and whether that charging had a significant effect on Er^{3+}, and this was explored by our frequency dependent measurements on EL samples. The measured voltage across the samples after the pump was turned off (Fig. 5.22 b) indicated any residual voltage went away on a timescale smaller than the RC charging time of our devices. If charge remained in the film, it was at a level too low to to detect with our oscilloscope. And if charge was present in the film, it did not seem to have an effect on the luminescence, as shown by the decay lifetime of EL. This hypothesis that the injected charge has no effect on luminescence is supported in the literature, where films grown via PECVD did show active luminescence quenching by injected carriers [110], but films grown by reactive oxygen sputtering did not [109]. This supports the contention that reactive oxygen sputtering grows films with less defects, and that is supported by the work presented here on Er:SiO$_2$, where reactive oxygen sputtering grew bright films with long lifetimes. It is also supported by the EL results, where the oxides supported higher fields before exhibiting breakdown.

5.8   Conclusion

In this chapter we presented work on erbium electroluminescence arising from erbium-doped silicon-rich films fabricated two different ways: sputtered in argon and reactively sputtered with argon and oxygen. The pure argon films resulted in films that exhibited luminescent center sensitization, characterized by short lifetimes and bright erbium luminescence for annealing temperatures around

500 °C. The reactively sputtered films exhibited bright luminescence for annealing temperatures around 800 °C, and had lifetimes in the millisecond range, although these properties were severely degraded by adding enough silicon to observe EL.

Each set of films behaved as sensitized films do, by luminescing for pump wavelengths that do not correspond to erbium energy levels, and with a cross section up to two orders of magnitude larger than in unsensitized, $Er:SiO_2$ films.

Electroluminescence was observed for both sets of films prepared at their optimal annealing temperatures. The reactively sputtered films exhibited EL at lower current densities, and their EL was far more robust than the LC films, such that rise time and lifetime data was able to be taken before device breakdown. These results were used to extract an electrical excitation cross section of 2.5×10^{-13} cm^2, which is two orders of magnitude larger than that found for these films under optical excitation.

Frequency-dependent data was taken to analyze the time dynamics of these films, and to explore what effects may be present due to lingering charge in the films. The luminescent rise times in these films were delayed from the voltage pulse by approximately 15 μs. This effect was only partly due to RC charging of the devices, which was 5 μs, and may have been due to other charging effects in the film that take place prior to the onset of luminescence. The fact that this delay was decreased as the frequency was increased speaks to this as well, as there may have been charge present in the films shortly after the pulse turned off. However, voltage data taken from these samples at high frequencies did not show a significant voltage after the RC discharging time, meaning any residual charge did not manifest as a measurable voltage in our setup. Lifetime studies of this data did not indicate that the processes occurring during the rise time have similar effect once the pump is removed, as the lifetimes showed no significant behavior that would correspond to lingering charge.

Finally, the experimental data presented here was used in the rate-equation model introduced in Chapter 4, and the model was used to explore how the films must be improved to see inversion of Er^{3+} under electrical pumping. If we allow the assumption that all Er^{3+} ions are coupled to Si-ncs, then Er^{3+} inversion in these devices can be achieved by fabricating films with an EL lifetime of 0.8 ms, while very strong inversion requires EL lifetimes in the 2–5 ms range, which may be impractical for these films. However, if the current density achievable in these films can be increased an order of magnitude as well, much smaller lifetimes can lead to strong inversion, with a 1 ms lifetime leading to an excitation fraction of 94%. But even with the high inversion fraction, it is likely that absorption due to excited Si-ncs will limit gain in a lasing system, and therefore a pulsed injection scheme similar to that described in Chapter 4 would be needed.

Chapter 6

Summary and Outlook

The topics presented in this thesis cover a relatively wide array of topics, with charge transport in novel systems being the similar thread between them. Because every thesis raises as many questions as it answers, no thesis is ever truly complete. In that vein, here are some final thoughts for future directions on the topics presented here.

6.1 Nanocrystal Charging

In Chapter 2, we presented a novel way of characterizing charge on the nanoscale through the use of UHV-AFM. Through simulation and analysis it was possible to quantify the charge in the nanocrystal layer and observe it in real time. The method presented and the model used describe the system quite well. The observation that nanocrystals are singly charged when biased was a significant discovery, as it directly relates the charge that can be contained in a nanocrystal layer to the density of nanocrystals. A possible future direction of this work would be to more thoroughly explore a range of nanocrystal densities. The work presented in this thesis was limited to a single nanocrystal density. For our experiments, two densities were fabricated by ion implantation, but only one yielded complete experimental results — the method used to inject charge resulted in a great deal of broken tips, and a series of inconsistent charging results.

One possible way of exploring a range of nanocrystal densities would be to use sputter deposition. The films in Chapter 2 were fabricated using ion implantation. This method introduces damage to the oxide surface and results in a range of nanocrystal sizes due to the implantation profile in the oxide. Sputter deposition would eliminate both of these issues because the oxide surface would not be perturbed in the same way, and the film itself would be more homogenous.

The other aspect not fully explained here is the data taken at high temperatures. Qualitatively these films behaved as one would expect, with the charge dissipation rates increasing with temperature. However, an incomplete temperature calibration prevented the extraction of meaningful data about the temperature-dependent charge transport mechanisms in these films. A more complete

characterization of these films at known high temperatures would be very beneficial for both scientific and practical reasons, as these memory devices must be expected to operate at some elevated temperature.

6.2   Hyperspectral Imaging

The data presented in Chapter 3 represented an idea that actually has applications in two realms: hyperspectral imaging and flash memory. This work was funded by a NASA grant exploring hyperspectral imaging, and thus we focused on this application. But a tunable barrier would be ideal for a flash memory device; programming and erasing could be done over a small barrier, while retention and readout could be done using the full barrier height. This would allow the program and erase times to be greatly reduced and reduce oxide wear, while still retaining long retention.

We found that the interfacial quality in these films was of utmost importance, as charging and leakage effects were common in the films. The La films in particular suffered from degradation and delamination from exposure to the atmosphere. Refined deposition and processing techniques could alleviate some of these issues, as would exploring materials that are more robust to begin with.

But the most difficult part of this project was finding materials that exhibited the proper band offset properties to study. Most of the films in the field have remarkably similar band gaps and offsets, with only Al_2O_3 and SiO_2 being viable candidates for the large-offset films. For this reason we suggest moving to some epitaxial ternary system, where the film composition can be varied to change the band gap [44]. Nitride systems, for example, can vary their band gap between 0.9 eV and 6.2 eV by using In, Ga, and Al. Such films could also be grown epitaxially, which would improve the electronic properties.

6.3   Erbium-Doped Nanoclusters

The results presented in Chapter 5 showed that this material system has a promising future for silicon photonics applications, but that films of higher quality are still required to achieve longer-lived Er in electrically pumped films. While films fabricated in other groups are more efficient than those presented here, the community as a whole still faces significant challenges.

Principal among these is the fraction of optically active Er coupled to Si-ncs. Studies performed by other groups put this number at less than 50% [67, 66, 111], which is the minimum level needed for lasing. While this is strictly an engineering problem at this point, it is still a significant one. The sensitizing agents in these films have been steadily decreasing in size — from nm-scale silicon nanocrystals, to amorphous nanoclusters, to defect-scale luminescent centers — but the fact remains that whatever the sensitizing complexes are, they are sensitizing single Er atoms. This inherent

difference in the scales is of direct interest, as an open question in the literature is how many Er atoms can be sensitized by a single nanocluster.

As the research has moved towards smaller sensitizers, it is also becoming apparent that high temperature deposition of these films has beneficial effects [68, 112]. This was not done in this thesis because our ion beam sputterer does not have a heated sample stage. Our group does have an RF sputtering system with heated substrate capabilities, but efforts to fabricate erbium-doped films in this system yielded films of far lower quality than the ones using ion beam sputtering. A significant amount of time during the course of this thesis was devoted to working with RF sputtering, and I would be remiss to leave that time completely out of this thesis. The RF sputterer in this group is a shared user facility, and we suspect part of the reason that the Er films grown in this system were of lower quality is due to contamination from other materials present in the chamber (the ion beam system is devoted solely to Er-containing films). Another problem encountered with this system was reliably and repeatably fabricating films with erbium contents around 1 atomic percent. Were this work to be revisited, developing a custom Er:SiO$_2$ target for this system would be ideal, as would extra measures to limit cross contamination.

Due to the nanoscale phenomena involved here, I wonder if a deposition technique such as atomic layer deposition (ALD) could be used to engineer the Er-Si-nc interaction at the nanoscale. ALD has been used to explore the Er:Y$_2$O$_3$ system [113], and to dope various glasses [114], but to my knowledge, an attempt at sensitizing Er with silicon in this system is lacking. ALD could possibly allow the direct control on the nanoscale of the two factors most responsible for Er luminescence — oxygen bonding to Er and proximity to Si-related sensitizers.

Bibliography

[1] G. E. Moore *Electronics*, vol. 38, no. 8, 1965.

[2] W. Arden *ITRS*, 2010.

[3] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita *Electronics Letters*, vol. 15, no. 4, pp. 106–108, 1979.

[4] R. Soref and J. Larenzo *IEEE Journal of Quantum Electronics*, vol. 22, no. 6, pp. 873–879, 1986.

[5] V. R. Almeida, C. A. Barrios, R. R. Panepucci, and M. Lipson *Nature*, vol. 431, pp. 1081–1084, Oct. 2004.

[6] W. M. J. Green, M. J. Rooks, L. Sekaric, and Y. A. Vlasov *Optics Express*, vol. 15, no. 25, p. 17106, 2007.

[7] R. Kekatpure, M. Brongersma, and R. S. Shenoy *Optics Letters*, vol. 30, no. 16, p. 2149, 2005.

[8] O. Boyraz and B. Jalali *Optics Express*, vol. 12, pp. 5269–5273, Oct 2004.

[9] H. Rong, R. Jones, A. Liu, O. Cohen, D. Hak, A. Fang, and M. Paniccia *Nature*, vol. 433, pp. 725–728, Feb. 2005.

[10] A. W. Fang, H. Park, O. Cohen, R. Jones, M. J. Paniccia, and J. E. Bowers *Optics Express*, vol. 14, no. 20, p. 9203, 2006.

[11] X. Sun, A. Zadok, M. J. Shearn, K. A. Diest, A. Ghaffari, H. A. Atwater, A. Scherer, and A. Yariv *Optics Letters*, vol. 34, pp. 1345–1347, May 2009.

[12] Z. H. Lu, D. J. Lockwood, and J.-M. Baribeau *Nature*, vol. 378, p. 258, 1995.

[13] S. Saito, N. Sakuma, Y. Suwa, H. Arimoto, D. Hisamoto, H. Uchiyama, J. Yamamoto, T. Sakamizu, T. Mine, S. Kimura, T. Sugawara, M. Aoki, and T. Onai in *IEEE International Electron Devices Meeting*, pp. 1–4, 2008.

[14] L. Cao, P. Fan, E. S. Barnard, A. M. Brown, and M. L. Brongersma *Nano Letters*, vol. 10, pp. 2649–2654, 2010.

[15] T. Shimizu-Iwayama, K. Fujita, S. Nakao, K. Saitoh, T. Fujita, and N. Itoh *Journal of Applied Physics*, vol. 75, no. 12, pp. 7779–7783, 1994.

[16] Q. Zhang, S. C. Bayliss, and D. A. Hutt *Applied Physics Letters*, vol. 66, no. 15, pp. 1977–1979, 1995.

[17] R. Walters, J. Kalkman, A. Polman, H. Atwater, and M. de Dood *Physical Review B*, vol. 73, no. 13, 2006.

[18] R. J. Walters, G. I. Bourianoff, and H. A. Atwater *Nature materials*, vol. 4, no. 2, pp. 143–6, 2005.

[19] M. Peralvarez, C. Garcia, M. Lopez, B. Garrido, J. Barreto, C. Dominguez, and J. A. Rodriguez *Applied Physics Letters*, vol. 89, no. 5, p. 051112, 2006.

[20] A. J. Kenyon *Semiconductor Science and Technology*, vol. 20, no. 12, pp. R65–R84, 2005.

[21] S. Tiwari, F. Rana, H. Hanafi, A. Hartstein, E. F. Crabbe, and K. Chan *Applied Physics Letters*, vol. 68, no. 10, p. 1377, 1995.

[22] www.srim.org.

[23] T. Feng, H. Yu, M. Dicken, J. R. Heath, and H. A. Atwater *Applied Physics Letters*, vol. 86, no. 3, p. 033103, 2005.

[24] E. A. Boer, M. L. Brongersma, H. A. Atwater, R. C. Flagan, and L. D. Bell *Applied Physics Letters*, vol. 79, no. 6, p. 791, 2001.

[25] T. Feng. PhD thesis, California Institute of Technology, 2005.

[26] T. Feng, G. Miller, and H. A. Atwater *Journal of Applied Physics*, vol. 102, no. 3, p. 034305, 2007.

[27] V. V. Afanas'ev and A. Stesmans *Journal of Applied Physics*, vol. 102, no. 8, p. 081301, 2007.

[28] M. Gordon and T. Baron *Physical Review B*, vol. 72, no. 16, 2005.

[29] E. A. Boer. PhD thesis, California Institute of Technology, 2001.

[30] I. Kim, S. Han, K. Han, J. Lee, and H. Shin *IEEE Electron Device Letters*, vol. 20, no. 12, pp. 630–631, 1999.

[31] K. K. Likharev *Applied Physics Letters*, vol. 73, no. 15, p. 2137, 1998.

[32] J. D. Casperson, L. D. Bell, and H. A. Atwater *Journal of Applied Physics*, vol. 92, no. 1, p. 261, 2002.

[33] R. H. Fowler and L. Nordheim *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 119, no. 781, pp. 173–181, 1928.

[34] R. J. Powell *Journal of Applied Physics*, vol. 41, no. 6, p. 2424, 1970.

[35] J. Robertson *Reports on Progress in Physics*, vol. 69, no. 2, pp. 327–396, 2006.

[36] L. F. Edge, D. G. Schlom, S. A. Chambers, E. Cicerrella, J. L. Freeouf, B. Hollander, and J. Schubert *Applied Physics Letters*, vol. 84, no. 5, p. 726, 2004.

[37] L. F. Edge, D. G. Schlom, R. T. Brewer, Y. J. Chabal, J. R. Williams, S. A. Chambers, C. Hinkle, G. Lucovsky, Y. Yang, S. Stemmer, M. Copel, B. Hollander, and J. Schubert *Applied Physics Letters*, vol. 84, no. 23, p. 4629, 2004.

[38] Y. Zhao, M. Toyama, K. Kita, K. Kyuno, and A. Toriumi *Applied Physics Letters*, vol. 88, no. 7, p. 072904, 2006.

[39] W. Cai, S. E. Stone, J. P. Pelz, L. F. Edge, and D. G. Schlom *Applied Physics Letters*, vol. 91, no. 4, p. 042901, 2007.

[40] Z. Xu, A. Daga, and H. Chen *Applied Physics Letters*, vol. 79, no. 23, pp. 3782–3784, 2001.

[41] A. M. Goodman *Phys. Rev.*, vol. 152, pp. 785–787, 1966.

[42] S. Strite and H. Morkoc *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 10, no. 4, pp. 1237–1266, 1992.

[43] V. Davydov, A. Klochikhin, V. Emtsev, D. Kurdyukov, S. Ivanov, V. Vekshin, F. Bechstedt, J. Furthmller, J. Aderhold, J. Graul, A. Mudryi, H. Harima, A. Hashimoto, A. Yamamoto, and E. Haller *Physica Status Solidi (B)*, vol. 234, no. 3, pp. 787–795, 2002.

[44] N. Tripathi, J. R. Grandusky, V. Jindal, F. Shahedipour-Sandvik, and L. D. Bell *Applied Physics Letters*, vol. 90, no. 23, p. 231103, 2007.

[45] L. Pavesi *Materials Today*, vol. 8, no. 1, pp. 18–25, 2005.

[46] B. Jalali *Nature Photonics*, vol. 1, pp. 193–195, Apr. 2007.

[47] M. I. Nathan, W. P. Dumke, G. Burns, J. Frederick H. Dill, and G. Lasher *Applied Physics Letters*, vol. 1, no. 3, pp. 62–64, 1962.

[48] R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, and R. O. Carlson *Phys. Rev. Lett.*, vol. 9, pp. 366–368, Nov 1962.

[49] T. M. Quist, R. H. Rediker, R. J. Keyes, W. E. Krag, B. Lax, A. L. McWhorter, and H. J. Zeigler *Applied Physics Letters*, vol. 1, no. 4, pp. 91–92, 1962.

[50] J. Nick Holonyak and S. F. Bevacqua *Applied Physics Letters*, vol. 1, no. 4, pp. 82–83, 1962.

[51] L. T. Canham *Applied Physics Letters*, vol. 57, no. 10, pp. 1046–1048, 1990.

[52] L. Pavesi, L. Dal Negro, C. Mazzoleni, G. Franzo, and F. Priolo *Nature*, vol. 408, p. 440, 2000.

[53] V. I. Klimov *Science*, vol. 290, no. 5490, pp. 314–317, 2000.

[54] J. Valenta, N. Lalic, and J. Linnros *Applied Physics Letters*, vol. 84, no. 9, p. 1459, 2004.

[55] G. Franzo, A. Irrera, E. C. Moreira, M. Miritello, F. Iacona, D. Sanfilippo, G. Di Stefano, P. G. Fallica, and F. Priolo *Applied Physics A: Materials Science & Processing*, vol. 74, no. 1, pp. 1–5, 2002.

[56] A. Anopchenko, A. Marconi, E. Moser, S. Prezioso, M. Wang, L. Pavesi, G. Pucker, and P. Bellutti *Journal of Applied Physics*, vol. 106, no. 3, p. 033104, 2009.

[57] A. Polman, D. C. Jacobson, D. J. Eaglesham, R. C. Kistler, and J. M. Poate *Journal of Applied Physics*, vol. 70, no. 7, p. 3778, 1991.

[58] W. J. Miniscalco and R. S. Quimby *Optics Letters*, vol. 16, no. 4, p. 258, 1991.

[59] E. Snoeks, G. N. v. d. Hoven, and A. Polman *Journal of Applied Physics*, vol. 73, no. 12, p. 8179, 1993.

[60] E. Snoeks, G. N. v. d. Hoven, A. Polman, B. Hendrikseny, and M. B. J. Diemeery *Journal of the Optical Society of America B*, vol. 12, no. 8, p. 1468, 1995.

[61] M. Fujii, M. Yoshida, Y. Kanzawa, S. Hayashi, and K. Yamamoto *Applied Physics Letters*, vol. 71, no. 9, p. 1198, 1997.

[62] G. Franzo, S. Boninelli, D. Pacifici, F. Priolo, F. Iacona, and C. Bongiorno *Applied Physics Letters*, vol. 82, no. 22, pp. 3871–3873, 2003.

[63] F. Priolo, G. Franzo, D. Pacifici, V. Vinciguerra, F. Iacona, and A. Irrera *Journal of Applied Physics*, vol. 89, no. 1, p. 264, 2001.

[64] F. Iacona, D. Pacifici, A. Irrera, M. Miritello, G. Franzo, F. Priolo, D. Sanfilippo, G. Di Stefano, and P. G. Fallica *Applied Physics Letters*, vol. 81, no. 17, p. 3242, 2002.

[65] K. Han, I. Kim, and H. Shin *Journal of Semiconductor Technology and Science*, vol. 1, no. 1, p. 40, 2001.

[66] B. Garrido, C. Garca, S. Y. Seo, P. Pellegrino, D. Navarro-Urrios, N. Daldosso, L. Pavesi, F. Gourbilleau, and R. Rizk *Physical Review B*, vol. 76, no. 24, 2007.

[67] C. J. Oton, W. H. Loh, and A. J. Kenyon *Applied Physics Letters*, vol. 89, no. 3, p. 031116, 2006.

[68] K. Hijazi, R. Rizk, J. Cardin, L. Khomenkova, and F. Gourbilleau *Journal of Applied Physics*, vol. 106, no. 2, p. 024311, 2009.

[69] R. M. Briggs, M. Shearn, A. Scherer, and H. A. Atwater *Applied Physics Letters*, vol. 94, no. 2, p. 021106, 2009.

[70] C. A. Barrios and M. Lipson *Optics Express*, vol. 13, no. 25, p. 10092, 2005.

[71] J. T. Robinson, K. Preston, O. Painter, and M. Lipson *Optics Express*, vol. 16, no. 21, p. 16659, 2008.

[72] M. A. Webster, R. M. Pafchek, and A. Mitchell *IEEE Photonics Technology Letters*, vol. 19, no. 6, p. 429, 2007.

[73] R. Pafchek, R. Tummidi, J. Li, M. A. Webster, E. Chen, and T. L. Koch *Applied Optics*, vol. 48, no. 5, p. 958, 2009.

[74] A. Harke, M. Krause, and J. Mueller *Electronics Letters*, vol. 41, no. 25, pp. 1377–1379, 2005.

[75] C. Creatore and L. Andreani *Physical Review A*, vol. 78, no. 6, 2008.

[76] E. Snoeks, A. Lagendijk, and A. Polman *Physical Review Letters*, vol. 74, no. 13, p. 2459, 1995.

[77] Y. C. Jun, R. M. Briggs, H. A. Atwater, and M. L. Brongersma *Optics Express*, vol. 17, no. 9, p. 7479, 2009.

[78] R. M. Briggs, G. M. Miller, and H. A. Atwater in *Proceedings of the 6th IEEE International Conference on Group IV Photonics*, pp. 223–225, 2009.

[79] C. Creatore, L. C. Andreani, M. Miritello, R. Lo Savio, and F. Priolo *Applied Physics Letters*, vol. 94, no. 10, p. 103112, 2009.

[80] D. K. Schroder, R. N. Thomas, and J. C. Swartz *IEEE Journal of Solid-State Circuits*, vol. SC-13, no. 1, p. 180, 1978.

[81] R. D. Kekatpure and M. L. Brongersma *Nano Letters*, vol. 8, no. 11, pp. 3787–3793, 2008.

[82] D. Navarro-Urrios, A. Pitanti, N. Daldosso, F. Gourbilleau, R. Rizk, G. Pucker, and L. Pavesi *Applied Physics Letters*, vol. 92, no. 5, p. 051101, 2008.

[83] D. Pacifici, G. Franz, F. Priolo, F. Iacona, and L. Dal Negro *Physical Review B*, vol. 67, no. 24, 2003.

[84] F. Priolo, G. Franzo, F. Iacona, D. Pacifici, and V. Vinciguerra *Materials Science and Engineering: B*, vol. 81, p. 9, 2001.

[85] P. Pellegrino, B. Garrido, J. Arbiol, C. Garcia, Y. Lebour, and J. R. Morante *Applied Physics Letters*, vol. 88, no. 12, p. 121915, 2006.

[86] H. Mertens, A. Polman, I. M. P. Aarts, W. M. M. Kessels, and M. C. M. van de Sanden *Applied Physics Letters*, vol. 86, no. 24, p. 241109, 2005.

[87] M. Galli, D. Gerace, A. Politi, M. Liscidini, M. Patrini, L. C. Andreani, A. Canino, M. Miritello, R. L. Savio, A. Irrera, and F. Priolo *Applied Physics Letters*, vol. 89, no. 24, p. 241114, 2006.

[88] H. P. Urbach and G. L. J. A. Rikken *Physical Review A*, vol. 57, pp. 3913–3930, May 1998.

[89] M. de Dood, L. Slooff, A. Polman, A. Moroz, and A. van Blaaderen *Physical Review A*, vol. 64, no. 3, 2001.

[90] M. de Dood, J. Knoester, A. Tip, and A. Polman *Physical Review B*, vol. 71, no. 11, 2005.

[91] A. Marconi, A. Anopchenko, M. Wang, G. Pucker, P. Bellutti, and L. Pavesi *Applied Physics Letters*, vol. 94, no. 22, p. 221110, 2009.

[92] G. M. Miller, R. M. Briggs, and H. A. Atwater *Journal of Applied Physics*, vol. 108, no. 6, p. 063109, 2010.

[93] O. Savchyn, F. Ruhge, P. Kik, R. Todi, K. Coffey, H. Nukala, and H. Heinrich *Physical Review B*, vol. 76, no. 19, 2007.

[94] W. J. Miniscalco *Journal of Lightwave Technology*, vol. 9, no. 2, p. 234, 1991.

[95] M. W. Sckerl, S. Guldberg-Kjaer, M. R. Poulsen, P. Shi, and J. Chevallier *Physical Review B*, vol. 59, no. 21, p. 13494, 1999.

[96] L.-F. Bian, C. G. Zhang, W. D. Chen, C. C. Hsu, and T. Shi *Applied Physics Letters*, vol. 89, no. 23, p. 231927, 2006.

[97] R. M. Briggs. PhD thesis, California Institute of Technology, 2011.

[98] G. Franzo, D. Pacifici, V. Vinciguerra, and F. Priolo *Applied Physics Letters*, vol. 76, no. 16, p. 2167, 2000.

[99] G. Franzo, M. Miritello, S. Boninelli, R. Lo Savio, M. G. Grimaldi, F. Priolo, F. Iacona, G. Nicotra, C. Spinella, and S. Coffa *Journal of Applied Physics*, vol. 104, no. 9, p. 094306, 2008.

[100] V. P. Gapontsev, S. M. Matitsin, A. A. Isineev, and V. B. Kravchenko *Optics & Laser Technology*, vol. 14, pp. 189–196, Aug. 1982.

[101] H.-S. Han, S.-Y. Seo, and J. H. Shin *Applied Physics Letters*, vol. 79, no. 27, p. 4568, 2001.

[102] O. Gusev, M. Bresler, P. Pak, I. Yassievich, M. Forcales, N. Vinh, and T. Gregorkiewicz *Physical Review B*, vol. 64, no. 7, 2001.

[103] J. Linnros, N. Lalic, A. Galeckas, and V. Grivickas *Journal of Applied Physics*, vol. 86, no. 11, p. 6128, 1999.

[104] G. Z. Ran, Y. Chen, W. C. Qin, J. S. Fu, Z. C. Ma, W. H. Zong, H. Lu, J. Qin, and G. G. Qin *Journal of Applied Physics*, vol. 90, no. 11, p. 5835, 2001.

[105] A. Nazarov, J. M. Sun, W. Skorupa, R. A. Yankov, I. N. Osiyuk, I. P. Tjagulskii, V. S. Lysenko, and T. Gebel *Applied Physics Letters*, vol. 86, no. 15, p. 151914, 2005.

[106] O. Jambois, Y. Berencen, K. Hijazi, M. Wojdak, A. J. Kenyon, F. Gourbilleau, R. Rizk, and B. Garrido *Journal of Applied Physics*, vol. 106, no. 6, p. 063526, 2009.

[107] Y.-B. Park and D. Schroder *IEEE Transactions on Electron Devices*, vol. 45, no. 6, pp. 1361–1368, 1998.

[108] K. Sun, W. J. Xu, B. Zhang, L. P. You, G. Z. Ran, and G. G. Qin *Nanotechnology*, vol. 19, no. 10, p. 105708, 2008.

[109] A. Irrera, F. Iacona, G. Franzo, M. Miritello, R. Lo Savio, M. E. Castagna, S. Coffa, and F. Priolo *Journal of Applied Physics*, vol. 107, no. 5, p. 054302, 2010.

[110] F. Priolo, C. Presti, G. Franz, A. Irrera, I. Crupi, F. Iacona, G. Di Stefano, A. Piana, D. Sanfilippo, and P. Fallica *Physical Review B*, vol. 73, no. 11, 2006.

[111] D. Navarro-Urrios, Y. Lebour, O. Jambois, B. Garrido, A. Pitanti, N. Daldosso, L. Pavesi, J. Cardin, K. Hijazi, L. Khomenkova, F. Gourbilleau, and R. Rizk *Journal of Applied Physics*, vol. 106, no. 9, p. 093107, 2009.

[112] S. Cueff, C. Labbe, J. Cardin, J.-L. Doualan, L. Khomenkova, K. Hijazi, O. Jambois, B. Garrido, and R. Rizk *Journal of Applied Physics*, vol. 108, no. 6, p. 064302, 2010.

[113] T. Hoang, P. LeMinh, J. Holleman, and J. Schmitz *IEEE Electron Device Letters*, vol. 28, no. 5, p. 383, 2007.

[114] G. Dingemans, A. Clark, J. A. van Delft, M. C. M. van de Sanden, and W. M. M. Kessels
Journal of Applied Physics, vol. 109, no. 11, p. 113107, 2011.