PATTERNED ZERO-DIMENSIONAL NANOSTRUCTURES: FABRICATION AND CHARACTERIZATION

BY

NEVILLE LOURENCO ANTONIO DIAS

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor James J. Coleman, Chair
Professor Kent D. Choquette
Professor Elyse Rosenbaum
Assistant Professor Xiuling Li
Professor Catrina Bryce, University of Glasgow
ABSTRACT

Due to the advantages arising from low-dimensional electronic systems, considerable effort has been put into the use of quantum dots and wires as the active media in optoelectronic devices. The realization of quantum dot based devices has been plagued with numerous obstacles. Conventional quantum dots are formed by strain-driven self-assembly. The stochastic nature of the process results in a distribution of dot sizes. If a device is composed of more than one quantum dot, the issue of uniformity becomes critical. Even if the device has only one quantum dot, uniformity is essential to obtain reproducible characteristics across multiple devices. Thus, the geometrical parameters of a quantum dot, such as shape and size as well as the chemical composition, need to be controlled.

In this work, nanoscale selective area metal-organic chemical vapor deposition (MOCVD) has been used to define InAs dot nucleation sites with highly ordered dot-to-dot pitches down to 80 nm corresponding to densities greater than $10^{10}$ cm$^{-2}$, which are among the highest reported for site-defined dots. The fabrication approach avoids modification of the underlying surface, allowing for easier integration into a variety of devices. Patterning of an oxide film by electron beam lithography also allows for creation of arbitrary closely packed arrangements of quantum dots for novel device designs. The resulting quantum dot array has the potential to be used as a template for fabricating multi-stack structures for use in laser and photodetector applications.

Although nano-fabrication methods impose a degree of determinism on the quantum dot size, the lack of coupling between individual dots in an array structure coupled with the size variation is the primary cause for inhomogeneous broadening in quantum dot based devices. In an attempt to address broadening in quantum dots, the nanopore active layer was proposed. The nanopore is in essence an inverse quantum dot structure consisting of a periodically perforated
quantum well that has been filled with a higher bandgap material. In the limit of small pores or large lattice spacing, the nanopore electronic properties approach those of a quantum well. At the other extreme, the nanopore behaves like a quantum dot. Thus the novelty in the nanopore active layer is that it presents an opportunity to design devices covering the continuum between fully three-dimensionally confined quantum dots and one-dimensionally confined quantum wells.

The in-plane periodicity results in miniband formation due to resonant scattering. Theoretical calculations of the intersubband scattering rate in nanopore lattices predict decreased intersubband scattering rates. This is due to the reduced overlap between in-plane components of the initial and final wavefunctions.

We conducted a photoluminescence (PL) study of nanopore lattices as a function of pore diameter while keeping the pitch and material compositions constant. Good agreement is obtained between PL spectra and finite-element calculations of the band structure. We observe increased emission from the higher subbands as the pore diameter is increased, which is a direct experimental verification of theoretical predictions. The decreased carrier cooling rate makes the nanopore useful as a solar cell material in which hot carriers excited by energetic photons can be captured before they decay to lower energy states.
To Mariano, Maria dos Anjos, Roselle, Vanda and Edgar
ACKNOWLEDGMENTS

I would like to thank Jim Coleman for saving me from a sinking ship and giving me the opportunity to work for him. He taught me how to think independently and be less random in my work. I have finally come to truly appreciate his unique brand of tough love. I am grateful to Professors Kent Choquette, Elyse Rosenbaum, Xiuling Li and Catrina Bryce for taking the time to serve on my committee.

I would like to thank Varun Verma for mentoring me in the early stages of my Ph.D. and for measurements performed while he was at the National Institute of Standards and Technology. Uttam Reddy has been a constant friend who spent the last four years trying to teach me the finer things in life. I would also like to thank Akash Garg, Jonathan Young, Nick Choi and Anil Kumar for their help and friendship. My simulation background is thanks largely to Edwin Ramayya, who introduced me to Monte Carlo and scattering rate calculations. Life in our group was more efficient and pleasant thanks to Tonia Siuts. Outside the confines of the Micro and Nanotechnology Laboratory, I was lucky to have great friends (#10, 1105 W. Main), around whom revolved the greater part of my limited social life.

I would like to thank my family for their love and support. Despite being fully convinced of the futility of a Ph.D., my parents never expressed any regret and always encouraged me to strive for greater things and to remember God while doing it. Without Rollie and Ted, God knows where I would have been right now. Whatever I ever do professionally, I owe entirely to them. Vanda and Eddie were always there when I needed anything, and for that I am very grateful.

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences as part of an Energy Frontier Research Center and the National Science Foundation grant (ECCS 08-21979).
# TABLE OF CONTENTS

## CHAPTER 1 – WHY WE DO IT

1.1 Geometrical architecture .................................................................1
1.2 Quantum dots ..............................................................................2
1.3 Nanopores – inverse quantum dots ............................................5

## CHAPTER 2 – HOW IT IS MADE

2.1 Electron beam lithography ............................................................17
2.2 Oxide etching .............................................................................20
2.3 Quantum well etching .................................................................22
2.4 Quantum dot growth .................................................................24

## CHAPTER 3 – OPTICAL AND STRUCTURAL PROPERTIES

3.1 Nanopore photoluminescence .......................................................34
3.2 Time resolved photoluminescence ..............................................38
3.3 Site localized InAs quantum dots ..................................................42
3.4 Dot regrowth optimization ............................................................48
3.5 Stacking of quantum dots ............................................................52
3.6 InAs quantum dots by diblock copolymer lithography ...............57

## CHAPTER 4 – NANOPORE BAND STRUCTURE

4.1 Nanopore geometry .......................................................................64
4.2 Strain theory ................................................................................66
4.3 Nanopore strain distribution .........................................................70
4.4 Effect of strain on band structure ................................................72
4.5 3D vs. 2D band structure simulation .........................................74
4.6 Wavefunction overlaps ...............................................................76
4.7 Nanopore subbands and the free electron model .......................79
4.8 Nanopore stochastic variations .....................................................83

## CHAPTER 5 – CARRIER RELAXATION DYNAMICS

5.1 Acoustic phonon scattering ..........................................................93
5.2 Optical phonon scattering ............................................................96
5.3 Multi-phonon scattering .............................................................100
5.4 Alloy scattering ........................................................................102
5.5 Multiple exciton generation ........................................................104
5.6 Nanopore matrix elements .........................................................109
5.7 Phonon scattering rates ..............................................................115

## CHAPTER 6 – THE FUTURE

References .......................................................................................120
CHAPTER 1 – WHY WE DO IT

1.1 Geometrical architecture

When the dimensions of a semiconductor are reduced in one, two or three dimensions of space below the de Broglie wavelength of a charge carrier, it becomes a nanostructure. Reduction of dimensionality in one, two or three dimensions leads to quantum wells, wires or dots, respectively. The electronic properties of low-dimensional structures depend qualitatively on the dimensionality of the structure and quantitatively on geometric structure (size and shape) (Figure 1.1). The electronic properties in turn control the linear and nonlinear optical and transport properties. Thus “geometrical architecture” opens enormous opportunities for designing novel artificial materials. These possibilities extend above and beyond conventional “chemical architecture,” where properties are modified by varying the chemical composition.

Figure 1.1: Top: Schematic representations of three-, two-, one- and zero-dimensional nanostructures. Middle and bottom: Schematic and functional form of density of states for the given case of dimensionality.

The first practical implementation of the quantum size effect was the quantum well laser [1, 2]. Compared to previous double heterostructure lasers, quantum well lasers possessed lower
threshold currents due to the altered density of states. As shown in Figure 1.1, the density of states in a bulk semiconductor increases as the square root of energy, while in quantum wells, the increase is step-like. A high gain requires population inversion of levels with a high density of states. In a bulk active layer, the lower lying energy levels need to be filled before achieving sufficient gain to overcome the cavity losses. In a quantum well, the peak gain is associated with lowest lying levels, thus leading to a lower laser threshold.

1.2 Quantum dots

Further reduction of the dimensionality leads to quantum wires and quantum dot (QD) structures. The advantages of a QD laser relative to a quantum well laser are due to the δ-like density of states. The oscillator strength $f_{mn}$ is a measure of the intensity of an electronic transition from an initial quantum state $|n\rangle$ to a final quantum state $|m\rangle$. The confinement of electrons and holes in QDs enhances the interaction and leads to giant oscillator strengths. The δ-like spectrum also results in higher maximum material gain and differential gain which can be two to three orders of magnitude larger than analogous values for quantum well lasers. The combination of higher peak density of states and greater oscillator strengths yields low threshold operation in QD lasers. The discrete nature also prohibits thermal broadening of the gain spectrum. Then the peak gain remains fairly constant as a function of temperature at constant injection level, and the threshold current density is temperature independent. QD lasers are also expected to fare better than quantum well lasers in terms of the dynamic figures of merit, the linewidth enhancement factor ($\alpha$), and chirp. The $\alpha$–parameter directly affects the linewidth, since the laser linewidth is $(1 + \alpha^2)$ times the fundamental Schawlow-Townes linewidth. The $\alpha$-parameter represents the coupling of the real and imaginary parts of the complex susceptibility due to a change of injection current. The linewidth enhancement factor is calculated from the
gain spectrum via the Kramers-Kronig relation. For QDs with a Gaussian spectrum, the differential gain \( \frac{dG}{dN} \) where \( G \) is gain and \( N \) is carrier concentration) is symmetric around the peak gain energy. In this case the differential change of refractive index is zero at the peak gain position. Thus \( \alpha = 0 \), and the ideal QD laser operates chirp-free [3].

Due to the above advantages, considerable effort has been put into the use of QDs as the active media in optoelectronic devices. The realization of QD-based devices has been plagued with numerous obstacles. Conventional quantum dots are formed by strain-driven self-assembly [4].

Figure 1.2 shows the three modes of heteroepitaxial growth: Volmer-Weber (VW), Stranski-Krastonow (SK) and Frank-van der Merwe (FvdM). They represent island growth (VW, 3D), layer-by-layer plus island growth (SK) and layer-by-layer growth (FvdM, 2D). The particular growth mode for a specific system is a function of the interface energies and lattice mismatch.
The stochastic nature of the nucleation process in SK mode results in a distribution of island (quantum dot) sizes (Figure 1.3 (a)). If a device is composed of more than one quantum dot, the issue of uniformity becomes critical. Even if the device has only one quantum dot, uniformity is essential to obtain reproducible characteristics across multiple devices. Thus, the geometrical parameters of a quantum dot, such as shape and size as well as the chemical composition, need to be controlled. There remain, however, fundamental limitations to the realizable benefits due to fabrication variations [5].

Figure 1.3: Self-assembled (a) versus patterned quantum (b) dots.

Several methods have been proposed to achieve controlled nucleation of quantum dots, including the use of vicinal substrates, etched substrates and patterned buffer layers. Oxide apertures at the micron scale have also been used to achieve tunable self-organized quantum dots without explicit control over the nucleation sites. Selective area epitaxy has previously been applied for growth of InGaAs quantum dots. The degree of mismatch between the dot and substrate resulted in Frank-van der Merwe growth. Moving to higher indium compositions
causes a transition to the Stranski-Krastonow growth regime, which also results in smaller dots and a shifting in the emission wavelength to the relevant 1.3 \( \mu \text{m} \) range.

In this work, we have employed the selective area epitaxy process for site-localized nucleation of InAs quantum dots on GaAs substrates (Figure 1.3 (b)). Compared to other developed techniques, this method does not require modification of the underlying substrate and hence can be used to integrate quantum dots into any base structure. Also, after the growth itself, the substrate remains planar with the exception of the dots themselves, rendering it amenable to further processing for integration into a device structure.

A novel hybrid etch process is developed to fabricate the growth inhibition mask. The ability to nucleate either single or multi-dot clusters inside each aperture is demonstrated. This behavior can further be controlled based on the growth conditions and mask geometry. After the actual dot growth, capping is a very critical process. A technique to minimize dot reflow during capping is developed.

1.3 Nanopores – inverse quantum dots

While quantum dots show very strong 3D quantization effects, they are often sparse. The spatial separation between adjacent dots ensures that, except for a small amount of tunneling, the dots do not share a common carrier pool. The result of this is dots that are essentially uncoupled, and any optical emission observed from the dots is inhomogeneously broadened. In an attempt to address inhomogeneous broadening in quantum dot lasers, it is a logical step to create an inverse quantum dot structure in which a layer of lower energy material – a quantum well – has embedded in it bits of higher bandgap material that are small enough and close enough together to introduce 3D quantization. This inverse quantum dot structure, henceforth called the nanopore, can also be thought of as an electronic analog of a photonic crystal.
To obtain sufficient 3D quantization, the feature sizes should be comparable to the De Broglie wavelength of an electron in a crystal, which is about 10 nm. Figure 1.4 shows the position occupied by the nanopore in the class of periodic structures. The periodicity is an order of magnitude more than that for atomic lattices but smaller than that for photonic crystals. Ideally, the periodicity should be of the order of 10 nm to observe strong quantization effects. However, due to fabrication limitations, the current work is limited to the 80-100 nm period range.

Figure 1.4: Illustration of a nanopore lattice as compared to an atomic lattice and a photonic crystal showing the relevant length scale for the periodicity.

The active medium is a continuous quantum well layer which is periodically perforated with wide bandgap material (see Figure 1.5). The geometry of the nanopore layer, namely the diameter of the pores and pitch of the lattice, determines the degree of 3D quantization as well as the form of the energy band structure. In the limit of small widely spaced pores, there will only be the 1D quantization of the quantum well. In the limit of high pore density such that the pores
touch, the behavior will be that of conventional quantum dots. For most geometries in between these limits, there will be appropriately scaled 3D effects.

The nanopore active layer exhibits a quantized energy structure similar to that of quantum dots. However, the carrier wavefunctions are much less localized as compared to quantum dots, due to the effectively weaker in-plane barrier. This coupling between unit cells results in the formation of a band structure similar to that observed in photonic crystals. Also, the effective density of states for the nanopore structure is larger than that of quantum dots, due essentially to the larger areal fill factor. This structure merits study because of its potential to produce quantum-dot-like intraband states, while providing a higher achievable gain than quantum dots.

Figure 1.5: Schematic of nanopore structure showing quantum well periodically perforated by wide bandgap barrier material to form a matrix of weakly coupled quantum dots.

The nanopore is characterized by its geometrical and material properties (Table 1.1). Nanopore fabrication can be either by a top-down method in which a quantum well is etched periodically or by a bottom-up method in which the nanopore is selectively grown by selective
area epitaxy. The in-plane dimensions of the nanopore are a function of the pore diameter and the pitch. The use of electron beam lithography (EBL) for nanopore patterning imposes a limit on the pitch due to the proximity dosing effect.

Table 1.1: Typical values of nanopore parameters and their limitations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>pitch</td>
<td>80 – 100 nm</td>
<td>EBL</td>
</tr>
<tr>
<td>pore diameter</td>
<td>40 – 70 nm</td>
<td>EBL + etching isotropy</td>
</tr>
<tr>
<td>quantum well</td>
<td>InGaAs</td>
<td>critical thickness limited</td>
</tr>
<tr>
<td>barrier</td>
<td>GaAs/AlGaAs</td>
<td>none</td>
</tr>
<tr>
<td>pore pattern</td>
<td>rhombic or square</td>
<td>none</td>
</tr>
</tbody>
</table>

In the top-down fabrication method, the pore diameter is limited by the minimum feature size attainable by e-beam lithography and etching isotropy. As feature sizes shrink, the pore-to-pore variation increases, which in turn causes inhomogeneous broadening of the emission spectrum. Etching through a quantum well to define the nanopore results in widening the pores due to the isotropy of the low-damage wet etch process. Thus the final pore diameter is a function of the e-beam dose and the wet etch depth. In bottom-up approaches, the pore diameter is limited by the oxide aperture dimensions which depend upon the e-beam dose and etching process used to transfer the resist pattern to the oxide mask. As a rule of thumb, the pore diameters that can be consistently achieved are in the 40-70 nm range for an 80 nm pitch pattern.

The nanopore properties are also strongly dependent upon the material composition of the nanopore active layer and the barrier layers. The band offsets between these layers define the
degree of carrier confinement inside the nanopore and hence the band structure. The quantum well composition and thickness are critical thickness limited. This critical thickness is computed by performing a force balance between the force in a strained film over a substrate and the force associated with a misfit dislocation and finding the point at which the force due to the strain exceeds the misfit force [6].

Assuming an InGaAs nanopore active layer, either GaAs or AlGaAs can be used as the barrier layers. The oxidation of Al-bearing alloys restricts the ability for regrowth of GaAs or InGaAs on AlGaAs layers. In-situ etching in an oxygen-free atmosphere becomes necessary for regrowth on exposed AlGaAs surfaces. In the current work, GaAs was used as the lower and upper barrier layers.

The novelty of the nanopore active layer lies in the number of degrees of freedom available for defining the electronic structure and hence the spectral properties. In addition to the above mentioned parameters, the presence of strain in the InGaAs/GaAs system introduces another knob that can be used to tailor the emission properties. The ability to create designer band structures and to control the overlap between the subband wavefunctions makes the nanopore a good candidate for intersubband devices in which precise control over subband relaxation processes is critical.

Lasers fabricated with the nanopore active layer have been shown to lase preferentially at excited states [7]. Figure 1.6 shows the experimentally obtained spectrum for a nanopore laser as a function of injection current between 0 and 200 mA measured at 77 K. The spectrum can be deconvoluted into three distinct peaks corresponding to the first three subbands in the nanopore.
Lasing initiates on the second and third subbands at around 10 mA. The computed joint density of states reveals a higher maximum gain for the first compared to the second subband. The initial lasing of the second subband suggests that the relaxation rate from the higher subbands could be lower in a nanopore structure. It has been suggested that this effect is similar to the phonon bottleneck postulated for relaxation in quantum dot systems where the intersubband energies do not match the LO phonon energy.

The aim of this work is to gain further insight into the relaxation processes in the nanopore and characterize the effect of nanopore geometry on the emission spectra. A finite element model-based simulation setup is used to compute the nanopore band structure. The strength of the finite element method as compared to finite difference and other analytic methods is its easy applicability to arbitrary geometries. The wavefunctions thus calculated are used to compute the transition rates between wavefunctions.
CHAPTER 2 – HOW IT IS MADE

Two approaches can be followed to create a nanostructure as shown in Figure 2.1. The top-down method is based on a previously grown perfect material which is sculpted to yield the desired nanostructure. As Michaelangelo said, “Every block of stone has a statue inside it and it is the task of the sculptor to discover it.” While sculptors use chisels and hammers, nanotechnologists use electrons, photons and ions as tools of the trade.

![Figure 2.1: Schematic of top-down and bottom-up methods for nanostructure fabrication.](image)

In the top-down method, the resist pattern is transferred to the underlying substrate by etching. In etching techniques, the mask pattern is transferred exactly onto the material surface. The etching process should therefore be anisotropic and ideally proceed only downwards into the material. If the etching is isotropic, a large undercut occurs and the mask pattern is not transferred exactly.

Anisotropic etching with near-vertical side walls requires the use of dry etching for pattern transfer. Dry etching can have a physical basis (ion milling), a chemical basis (plasma etching) or a combination of the two (reactive ion etching). In processes that rely predominantly on the
physical mechanism of sputtering, the strongly directional nature of the incident energetic ions allows substrate material to be removed in a highly anisotropic manner. Unfortunately, etch selectivity is almost non-existent due to the inertness of the process. On the other hand, purely chemical etch processes exhibit very high etch selectivities. Such purely chemical etching mechanisms, however, typically etch in an isotropic fashion. By adding a physical component to a purely chemical mechanism, reactive ion etching achieves the best of both worlds. In addition to etch anisotropy, an important consideration for device fabrication is the degree of damage created by exposure to ions, electrons or UV light during etching [9, 10]. Contamination originating from polymer formation during etching or materials sputtered from the etching chamber can also influence device performance.

Chemical wet etching is generally isotropic. Therefore the pattern transfer fidelity is poor. However, the etching proceeds by an electrochemical process and produces no etch damage. In many cases the etching of crystalline materials is selective, depending upon the material, its composition and/or its crystal orientation. For example in GaAs, a bromine-methanol solution is an anisotropic etchant and produces patterns with (111) A facets [11]. Precise control of critical dimensions is made difficult by the strong dependence of wet etch rates on temperature and solution agitation.

Bottom-up fabrication methods tend to be readily scalable to large batches but typically offer limited control over size, size distribution and particle shape. Selective epitaxy and epitaxially grown nanowires are prime examples of bottom-up techniques. Selective area epitaxy involves growth on dielectric-masked semiconductor wafers. Growth is inhibited in regions of the sample that are covered by the dielectric, allowing the regions where growth occurs to be controlled by the mask. Material conservation causes growth enhancement in the unmasked
regions. Minimization of the surface energy results in faceted growth. This technique has been previously applied to the growth of InGaAs quantum wires, dots and inverse dots on GaAs [7, 12, 13].

Vapor-liquid-solid growth of nanowires is also a widely used technique for bottom-up formation of nanowires. In this process, the semiconductor material is adsorbed by liquid nanodroplets of an appropriate catalytic material such as gold. These nanodroplets, located at the nanowire tips, serve as seeds for nanowire growth. The semiconductor material condenses at the interface between the droplet and the nanowire, thereby extending the length of the nanowire.

In general, bottom-up methods have the disadvantages of growth enhancement as well as compositional enhancement when applied to ternary and quaternary alloys. Dopant incorporation can also be problematic because the presence of the dopant species drastically modifies the growth kinetics. The size variation results in inhomogeneous broadening of the emission spectrum. Thus the advantages gained from patterning are somewhat offset by the compositional and geometrical variation. To this end, a top-down approach based upon wet etching of a grown quantum well was developed for nanopore patterning.

The overall schematics of the fabrication methods for the nanopore and the selectively grown quantum dots are shown in Figures 2.2, 2.3 and 2.4. The nanopore process involves patterning an oxide film with a pore pattern and then transferring this pattern into the underlying quantum well by wet etching. Oxide is chosen as the wet etch mask due to its resistance to attack by acids. Attempts at using poly(methyl methacrylate) (PMMA) as the etch mask have been unsuccessful due to attack of the PMMA by acids and parasitic etching due to porosity in the PMMA film. The thickness of the oxide film was chosen to be 10 nm for the nanopore process. The oxide film should not be too thick or else the quantum well etching may become diffusion
limited. Thicknesses under 10 nm cannot be reliably deposited due to variations in the plasma enhanced chemical vapor deposition (PECVD) process.

Figure 2.2: Processing steps for nanopore and quantum dot fabrication starting from the growth of the base structure up to mask patterning by electron-beam lithography. The underlined text indicates a change in the process specifically for the quantum dot fabrication.
The selective area epitaxy method for QD fabrication uses a dielectric as the growth inhibition mask. Either silicon dioxide (SiO$_2$) or silicon nitride (Si$_x$N$_y$) can be used as the dielectric. The use of SiO$_2$ in this process is due to the consistency in the PECVD deposition process. The choice of thickness for the quantum dot process is to allow for use of a hybrid dry and wet etching process to transfer the resist pattern into the oxide while minimizing the damage to the underlying substrate. Further details on this technique are given in section 2.2.
A glance at the process flow for both the nanopore and quantum dot fabrication reveals a few key process steps that are the most important in determining structural properties. Geometry is controlled by electron-beam lithography and etching, while composition is controlled by the
MOCVD growth process. Thus it becomes necessary to gain some understanding of these process steps to fully harness their potential.

2.1 Electron beam lithography

Use of electrons instead of photons for lithography overcomes the theoretical Rayleigh diffraction limit on the resolution. Electron beam lithography (EBL) systems have three main sub-systems: (1) electron gun, (2) electron focusing optics and (3) a mechanical stage for sample placement. The electron source is either thermionic or field-emission based. The electron focusing optics consist of electrostatic and magnetic elements that accelerate, shape and deflect the electron beam. The mechanical stage is used to accurately position the sample under the electron beam. The position of the stage is critical and is measured to nanometer accuracy using laser interferometric techniques. Alignment to pre-existing structures on the sample is achieved by registration of alignment marks on the sample. High Z (atomic number) elements should be used for the alignment mark to obtain sufficient contrast during the registration process. As a rule of thumb, at least 150 nm of Au is necessary on the alignment mark for efficient mark detection.

The primary limitation on EBL resolution is the scattering of electrons in solid materials. High energy electrons interact with the constituent atoms and electrons in the resist. Two types of scattering can occur: small angle scattering (forward scattering), which tends to broaden the initial beam diameter, and large angle scattering (back scattering), which causes the proximity effect (Figure 2.5). The increase in effective beam diameter due to forward scattering is empirically given by $d_f = 0.9(R_t/V_b)^{1.5}$, where $R_t$ is the resist thickness in nanometers and $V_b$ is the beam voltage in kilovolts. Forward scattering is minimized by using the thinnest possible resist and the highest available accelerating voltage. Back scattered electrons from the substrate may return into the substrate at a significant distance from the incident beam causing additional
resist exposure. The fraction of electrons that are back scattered increases with increasing atomic weight of the substrate. The proximity dosing can be characterized by a double Gaussian given by

\[
 f(r) = \frac{1}{(1+\nu)\pi} \left\{ \frac{1}{\alpha^2} e^{-r^2/\alpha^2} + \frac{\eta}{\beta^2} e^{-r^2/\beta^2} \right\}
\]

(2.1)

where \( r \) is the distance from the center of the beam, \( \eta \) is the ratio of back scattering to forward scattering and \( \alpha \) and \( \beta \), respectively, are the range of forward and back scattering [14].

Figure 2.5: Schematic of electron scattering in substrate during electron beam lithography showing the primary beam and forward scattering and back scattering profiles.

The most widely used e-beam resist is poly-(methyl methacrylate) (PMMA) which is a positive resist. Electron exposure causes scission of the polymer chains, thus decreasing the average molecular weight compared to the unexposed resist, which makes the exposed part more soluble in developer. The resist thickness involves a tradeoff between resolution and etch resistance. As the thickness increases, the achievable resolution decreases while the mask is more resistant to dry etching. The ratio of the minimum feature size to the resist thickness should
be less than 3 to 4 to prevent collapse of delicate resist features. Also, as the resist thickness decreases, the dose required increases because a lesser fraction of electrons scatter in the resist and contribute to resist exposure. Typical resist thickness for the nanopore and quantum dot processes are 50 nm obtained by spinning 2% 495K PMMA dissolved in anisole at 4000 rpm. SEM images of 100 nm pitch patterns are shown in Figure 2.6. While the pitch is controlled by the input pattern file, the pore diameter is regulated by modulating the dose of each exel (electron pixel).

![SEM images of 100 nm pitch nanopore patterns with two doses showing variation in pore diameter.](image)

Figure 2.6: SEM images of 100 nm pitch nanopore patterns with two doses showing variation in pore diameter.

The electron beam lithography system used in this work is a JEOL JBX-6000FS system. The accelerating voltage is 50 kV and the minimum pitch that can be obtained consistently is 80 nm. For finer pitches, the optimal dose window is too small to achieve uniform resist patterns. A minimal change in the exposure dose causes an abrupt transition from under-dosed patterns to over-dosed patterns. Further optimization of the resist composition (PMMA molecular weight), thickness and development procedure is hence necessary for consistent reproduction of sub-80 nm nanopore patterns.
2.2 Oxide etching

After the pore pattern has been formed on the PMMA resist, the next process is to transfer this pattern to silicon dioxide film either to serve as an etch mask for the nanopore etching or as the growth inhibition mask for the selective area growth of InAs dots. As with all types of etching, either wet or dry etching can be used for the pattern transfer. The most commonly used wet etching chemistry is based on hydrofluoric acid (HF). The reaction is \( \text{SiO}_2 + 4 \text{HF} \rightarrow \text{SiF}_4 + 2 \text{H}_2\text{O} \). Since this reaction consumes HF, the reaction rate decreases with time. To avoid this, HF is used in conjunction with a buffering agent such as ammonium fluoride (\( \text{NH}_4\text{F} \)) which maintains a constant concentration of HF. Standard buffered HF is a mixture of HF with \( \text{NH}_4\text{F} \) mixed in a ratio of 1:5. This etches thermal silicon dioxide at about 80 nm per minute and PECVD deposited films at a much faster rate. HF however attacks the PMMA mask and hence cannot be used for transferring the e-beam pattern into the oxide film.

A side effect of dry etching is etch damage. The presence of damage manifests itself in reduced photoluminescence. The energy of the ions used in reactive ion etch is usually under 1 keV, and so is not sufficient to render the material amorphous. Interpolating from the dependence of ion penetration depth on ion energy, it is found that these relatively low energy ions stop within a nanometer from the surface. This analysis however does not account for the crystallinity of the sample. Incoming ions incident along crystalline directions penetrate much more deeply – a phenomenon known as channeling. This channeling can create damage deep inside the surface. Thus the use of a single dry etch step for oxide patterning renders the sample optically inactive.

To limit the damage to the underlying substrate, a hybrid process based on a combination of dry and wet etching was developed for the selective area process. A schematic of the hybrid
etching process is shown in Figure 2.7. After electron beam lithography, a CHF$_3$ reactive ion etch is used to transfer the PMMA pattern into the oxide. The etch depth is chosen such that about 10 nm of oxide remains at the base. Care should be taken to calibrate the etching times because the etch rate for small features can be significantly less than that for larger features. The PMMA mask is then removed by an acetone/methanol/isopropanol (IPA) rinse followed by a low power oxygen descum. A dilute buffered HF (1HF:50NH$_4$F) etch is then used to completely open the apertures. The dilute BHF etch rate is about 17 nm/min. An SEM micrograph of the oxide mask after etching is shown in Figure 2.8. It should be noted that damage is not a particularly serious concern for the nanopore fabrication because the damaged region is eventually etched away during the nanopore wet etching step. Hence, this hybrid etch process is not used in the nanopore fabrication process.

![Figure 2.7: Schematic of hybrid etching process for selective area epitaxy mask fabrication.](image-url)
2.3 Quantum well etching

Nanopore properties are strongly dependent upon the material composition of the nanopore active layer. The quantum well composition and thickness are critical thickness limited as described by the Matthews-Blakeslee (MB) model. This model derives a critical thickness by performing a force balance between the force in a strained film over a substrate and the force associated with a misfit dislocation [6]. The force due to a misfit stress is given by

\[ F_H \approx G \left[ \frac{1 + \nu}{1 - \nu} \right] b h e \]  \hspace{1cm} (2.2)

where \( G \) is the shear modulus, \( b \) is the slip distance and \( h \) is the thickness of the film. The force in a dislocation line is given by

\[ F_D \approx \frac{G b^2}{4\pi(1-\nu)} \left[ \ln \left( \frac{h}{b} \right) + 1 \right] \]  \hspace{1cm} (2.3)

If \( F_H > F_D \), dislocations move and hence the critical thickness is determined by the equality of \( F_H \) and \( F_D \)

\[ h_c \approx \left( \frac{b}{e} \right) \left[ \frac{1}{4\pi(1+\nu)} \right] \left[ \ln \left( \frac{h}{b} \right) + 1 \right] \]  \hspace{1cm} (2.4)
The lateral dimensions of the nanopore are defined by etching the quantum well with either resist or oxide as a mask. The key parameters to be considered in selecting an etch process are rate, uniformity, selectivity, critical dimension control, feature profile and etch damage. In general, slower etches are preferred. Patience is a virtue in etching, as in most of life, and a few more minutes spent etching are preferable to overetched and ruined devices.

The excessive damage created by reactive ion etching prevents its use for optically active materials. Numerous wet etching chemistries have been developed for the GaAs-AlGaAs system. The underlying chemical mechanism of most wet etching of GaAs is the oxidation of the surface to form Ga and As oxides, followed by the dissolution of these oxides by chemical attack with acids or bases [15]. Many GaAs etches call for the addition of hydrogen peroxide (H₂O₂) to serve as an oxidizing agent to promote the formation of the surface oxides. Other chemicals that have been used as the oxidizing agent for etching GaAs are HNO₃, K₂Cr₂O₇, CrO₃, I₂ and Br₂.

The multi-step nanopore etch leverages the cleaning nature of the dilute NH₄OH etch and the uniformity of the phosphoric-based etch to yield reproducible nanopore patterns. The pre-cleaning step is a 30 s rinse in 1:10 NH₄OH:DI followed by two rinses of 10 and 20 s in DI water. The 1:4:495 H₃PO₄:H₂O₂:DI etch is calibrated by etching a stripe pattern in a GaAs wafer for 5 min. The etch depth is measured in a profilometer and the etch rate measured. The typical etch rate is around 0.4-0.5 nm/s. The nanopore etch is then performed for the calibrated time. An oxidizing step in H₂O₂ is then performed followed by etching with HCl:DI to remove residues left at the base of the etched pore. Figure 2.9 shows a scanning electron microscope image of the etched nanopore pattern using both a 1:10 NH₄OH:DI pre-cleaning step and a 1:1 HCl:DI precleaning step.
Figure 2.9: Etched nanopore pattern after (a) 1:10 NH₄OH:DI pre-cleaning and (b) 1:1 HCl:H₂O precleaning.

2.4 Quantum dot growth

Strain plays an important role in heteroepitaxial growth in determining the structural properties of overgrown layers. In the InAs/GaAs(100) system, the 7.1% lattice mismatch results in Stranski-Krastonow growth behavior. The transition from 2D to 3D is not abrupt and occurs over a delivery range of 1.45 – 1.74 monolayers, characterized by coexisting 2D clusters, quasi-3D clusters and 3D islands, and involves a varying mass transfer between the 2D and 3D features as the 3D islands evolve and stabilize [16]. The choice of growth conditions controls whether the process is limited by either mass transport, surface kinetics, or thermodynamics. The most significant growth parameters are the growth temperature, the III-V ratio, the growth rate, indium coverage and the substrate orientation. In addition to the parameters involved in the actual growth of the dots, subsequent steps such as the growth interruption time and choice of the cap layer also affect dot properties.
2.4.1 Growth Energetics

In lattice matched systems, the growth mode is solely governed by the interface and surface energies. If the sum of the epilayer surface energy $\gamma_2$ and the interface energy $\gamma_{12}$ is lower than the energy of the substrate surface, $\gamma_2 + \gamma_{12} < \gamma_1$, i.e. if the deposited material wets the substrate, the FvdM growth mode occurs. A change in $\gamma_2 + \gamma_{12}$ alone may drive a transition from the FvdM to the VW growth mode. For a strained epilayer with small interface energy, initial growth may occur layer by layer, but a thicker layer has a large strain energy and can lower its energy by forming isolated islands in which the strain is relaxed. Thus the SK growth mode occurs. The relaxation of the elastic energy due to the formation of coherently strained islands is related to the Asaro-Tiller-Grinfeld instability of a strained layer against a long-wavelength corrugation of a surface [17, 18].

In addition to elastic relaxation, strain energy can also be reduced by the formation of dislocations. The formation of an island from a uniform film is accompanied by a relaxation of the elastic energy $\Delta E_{\text{elastic}}^V > 0$, and by a change of the surface area, $\Delta A > 0$. The size of the corresponding change of the surface energy depends on the formation of side facets of the islands and on the disappearance of certain areas of the planar surface. It has been shown by Vanderbilt and Wickham that the morphology of the mismatched system is determined by the relation between $\Delta E_{\text{surf}}$ and the energy of the dislocated interface $E_{\text{interf}}^{\text{disl}}$ [19]. The ratio of these two energies, denoted by $\Gamma = E_{\text{interf}}^{\text{disl}} / E_{\text{surf}}$, is the control parameter that governs the morphological phase diagram (Figure 2.10).

If $\Delta E_{\text{surf}}$ is positive and large, or the energy of the dislocated interface is relatively small, the corresponding value $\Gamma$ on the phase diagram is smaller than $\Gamma_0$. Then formation of coherently
strained islands is not favorable. With the increase of the amount of the deposited material, a transition occurs from a uniform film to dislocated islands and coherently strained islands do not form.

If $\Delta E_{\text{surf}}$ is positive and small, or the energy of the dislocated interface is relatively large, the corresponding value $\Gamma$ on the phase diagram is larger than $\Gamma_0$. With increasing material deposition, a transition from a uniform film to coherent islands occurs.

![Phase diagram showing morphology versus amount of deposited material $Q$ and $\Gamma$ parameter showing the regions of uniform film growth and formation of either coherent or dislocated islands.](image)

Figure 2.10: Phase diagram showing morphology versus amount of deposited material $Q$ and $\Gamma$ parameter showing the regions of uniform film growth and formation of either coherent or dislocated islands [19].

If the change in the surface energy due to island formation is positive and small, coherent islands can be formed and the total energy of the island can be written as

$$E_{\text{island}} = E_{\text{edges}} + \Delta E_{\text{surf}} - E_{\text{relax}}^V,$$

where $E_{\text{edges}}$ denotes the short-range energy of the island edges, $\Delta E_{\text{surf}}$ denotes the excess surface energy of the island, and $-E_{\text{relax}}^V$ denotes the energy of elastic relaxation of the volume strain caused by lattice mismatch with the substrate. The scaling behavior is $E_{\text{edges}} = AL$, $E_{\text{surf}} = BL^2$, and $E_{\text{relax}}^V = CL^3$, where $L$ is a characteristic dimension of
the island. If the surface energy dominates the edge energy term, island formation is energetically favorable only for cases where the island size exceeds a critical value of 

\[ L_{\text{crit}} = \frac{2B}{3C} \].

Figure 2.11: Sequence of strain relief mechanisms and associated structures in strained heteroepitaxy.

The last stage of island growth occurs via Ostwald ripening in which small islands coalesce into larger islands. This is a spontaneous process that occurs because larger islands are more energetically favorable than smaller islands. While the formation of many small islands is kinetically favored (i.e. they nucleate more easily), large islands are thermodynamically favored (i.e. they have a lower surface energy). In general, the ripening process increases the average volume of the islands and decreases the island density. Once a given island reaches a certain critical volume, it transitions from a coherent to a disclocated state. The complete process is shown graphically in Figure 2.11.
2.4.2 InAs Coverage

The three stages of dot growth with increasing InAs coverage are shown in Figure 2.12. The coverage is defined in terms of monolayers (ML). One monolayer is defined a layer whose thickness equals half the lattice constant of the material.

![Graph showing the transition from 2D islands to incoherent dots with changing InAs coverage for atmospheric pressure MOCVD growth of InAs dots on GaAs (100). The 2D-3D transition occurs at around 1.63 ML.](image)

The first stage occurs for coverage under 1.5 ML. Submonolayer depositions result in the formation of 1 ML islands. Islands oriented along the [0-11] direction having a uniform width of ~4 nm have been measured using STM. MBE studies have shown that the wetting layer structure depends upon the growth temperature and consists of an InGaAs alloy component along with uncovered sections of the substrate and by InAs at high coverages. During the intermediate stage
between 1 and 1.5 ML, the surface consists of a combination of a monolayer of InAs and multiple monolayer islands which change in shape and size as the InAs coverage increases.

When the coverage is increased beyond 1.5 ML, an abrupt transition occurs from 2D islands to 3D coherent dots. The saturation number density is temperature dependent, but is typically in the range of $5 \times 10^9$ cm$^{-2}$ to $2 \times 10^{11}$ cm$^{-2}$. Reducing separation between dots results in dots coalescing into larger 3D islands to further reduce the strain energy. These large 3D islands tend to be incoherent (dislocated) and result in degraded luminescence efficiency.

### 2.4.3 Growth Temperature and Rate

Typical growth temperatures are between 400 °C and 550 °C. In this temperature regime, pyrolysis of the precursors can be the limiting step in the growth process. Figure 2.13 shows the temperature dependence of pyrolysis of tri-methyl-gallium (TMGa), tri-methyl-indium (TMIn) and arsine (AsH$_3$). At 450 °C, all the TMIn and TMGa is decomposed while AsH$_3$ pyrolysis is around 50%. The growth conditions are generally arsenic rich with V-III ratios of around 30, and hence no arsenic deficiency problems exist. However, the partial pyrolysis of AsH$_3$ means that the actual V-III ratio at the growth surface is different from that calculated from the respective flows.

Size and density of freely grown InAs/GaAs dots depend strongly on the growth rate and temperature. By increasing the growth temperature and decreasing the growth rate of InAs, the density of islands becomes low and the size of islands becomes large. On the other hand, decreasing the growth temperature and increasing the growth rate of InAs, the island density becomes higher and the size smaller (Figure 2.14).
Higher desorption rates and longer migration lengths at higher temperatures cause the larger size and lower densities of dots [22]. The growth rate dependence of InAs island formation can be explained by considering adsorption and desorption of InAs. Higher growth rate means higher adsorption rate. The density of In adatoms becomes higher at higher adsorption rate. This leads to large nucleation probability. Then, even a small island can stably remain on the surface at higher growth rate.

Figure 2.13: Decomposition of TMGa, AsH$_3$ and TMIn [20, 21].

Figure 2.14: Pictorial representation of the dependence of quantum dot size and density on growth temperature and rate [22].
2.4.4 V-III Ratio

The V-III ratio on the GaAs surface is key to determining the size and density of InAs dots [23]. Figure 2.15 shows 1 μm × 1 μm atomic force microscope scans of InAs dots grown with arsine flow rates of 15, 5, 3 and 0 sccm corresponding to V-III ratios of 33, 11, 6.7 and 0.

In addition to the reduction in dot size, the dot density also increases with decreasing arsine flow. Under a high V-III ratio, many arsenic atoms are available to bind to indium atoms. Because the arsenic atoms serve as intermediaries between indium atoms in the formation of InAs, polycrystalline InAs islands are formed in an environment of abundant arsenic atoms. Because some of the indium and arsenic atoms attached to the islands are so far above the surface that the atomic spacing no longer fits the crystal lattice of the GaAs substrate, the newly formed InAs islands are polycrystalline and cause defects when the GaAs cap layer is deposited.

![AFM images of InAs islands deposited at 520 °C on a GaAs surface with arsine flow rates of (a) 15, (b) 5, (c) 3, and (d) 0 sccm with the corresponding V-III ratios of 33, 11, 6.7, and 0 [23].](image)

When the V-III ratio is low, few arsenic atoms are available except the dangling arsenic bonds on the GaAs surface. As a result, the arriving indium atoms tend to bond with the GaAs surface instead of with one another. With indium atoms spreading across the surface for attachment instead of congregating together, the InAs dot density increases and the average dot
size decreases. If the V-III ratio is too low, arsenic desorption from the surface reduces the number of sites for indium atoms to bond, which then reduces the dot coverage. Even though sufficient indium atoms are available at the GaAs surface, the process is limited by the lack of arsenic atoms and dangling bonds at the surface. Without enough arsenic atoms to bind with, the number of dots is reduced and the size becomes even smaller.

2.4.5 Growth Interruption Time

A growth interrupt (GRI) involves temporarily stopping the growth for a short period of time by removing all precursors. Due to the high mobility of In adatoms and the high strain involved, material continues to redistribute between the QDs and the wetting layer after the sources are switched off. Typically, the deposition is stopped and a GRI introduced just prior to QD formation (i.e., just before the 2D to 3D transition). If the GRI is too short, the InAs wetting layer fails to reorganize into islands and remains as a highly strained thin film.

Once the wetting layer is consumed to form dots, further island growth occurs by coalescing of smaller islands into large dislocated islands by a process called Ostwald ripening. Atoms detach more readily from smaller particles, which have a higher vapor pressure, and condense more readily on larger particles. Thus, larger particles grow at the expense of smaller particles. The size distribution of the particles broadens with time, and the area density of particles decreases as the smallest particles dissolve. When applied to dot growth, Ostwald ripening results in reduced dot densities and red-shifting of the peak emission. Further ripening causes dots to exceed the critical size for dislocation formation.

It is also worth mentioning that the presence of arsine during a GRI leads to strong island ripening and a dramatic decrease in photoluminescence intensity. It has been suggested that higher arsine flows (and/or its associated atomic hydrogen) destabilize the wetting layer, leading
to the formation of large defective islands [24]. Also, higher free arsenic concentrations in gas phase may aid in incorporation of indium into the islands [23]. The stability of islands during growth interrupts without arsine may be due to quick removal of any residual arsine from the system. Islands form rapidly by consuming any previously absorbed arsine at the surface. Continued island ripening is then hindered due to the lack of further arsine supply, essentially freezing the island morphology.
3.1 Nanopore photoluminescence

The continuous wave photoluminescence is obtained using an Ar ion laser emitting at 488 nm. The beam is focused by a microscope objective onto the sample placed in a liquid N$_2$ cryostat. The luminescence signal is collected by the same microscope objective, dispersed by a monochromator and detected by a liquid N$_2$-cooled Ge detector (Figure 3.1).

![Figure 3.1: Schematic of in-house photoluminescence setup.](image)
Nanopore patterns with a constant pitch of 80 nm and diameters ranging from 55 to 70 nm were fabricated and measured. Figure 3.2 (a) shows the peak wavelength and corresponding photon energy as a function of pore diameter. As the pore diameter increases, the nanopore transitions from a quantum-well-like structure to a quantum dot system. The effective area of the triangular region formed by three surrounding nanopores decreases with increasing pore diameter. This increasing confinement results in a quadratic increase in ground state energy, which is in agreement with established theory (Figure 3.2 (b)) [25].

![Graph showing peak emission wavelength and emission peak energy shift versus pore diameter.](image)

Figure 3.2: (a) Peak emission wavelength and emission peak energy shift versus pore diameter. (b) Schematic showing effect of pore diameter on nanopore wavefunction confinement.

To study the effect of the nanopore on excited state emission, the quantum well was designed to have two interband transitions around 1008 nm and 928 nm at 77 K. Figure 3.3 shows the comparison between the normalized photoluminescence from a quantum well and a nanopore pattern with pitch of 80 nm and a pore diameter of 60 nm. Most significantly, the ratio of the emission from the band at 928 nm to the emission at 1008 nm is greater in the nanopore compared to the reference quantum well.
The nanopore wavefunctions can be expressed as

\[ \Psi(x, y, z) = \Phi(x, y) \times Z(z) \]  

(3.1)

where \( \Phi_{xy} \) refers to the in-plane variation of the wavefunction and \( Z \) refers to the out-of-plane variation of the wavefunction. Although the nanopore is not strictly uniform in the \( z \) direction like a quantum well, the variable separation assumption yields accurate results for pore diameters in the 80 nm range. This has been validated by finite element analysis that yields identical band structure for a 2D simulation disregarding the \( z \)-dimension and a full 3D calculation.

The band at 1008 nm arises from the in-plane subbands formed for the first excited \( Z \) wavefunction (henceforth denoted as \( \Phi_1 \)) while the band at 928 nm corresponds to the in-plane subbands formed for the second excited \( Z \)-wavefunction (denoted as \( \Phi_2 \)).

Electrons in the \( \Phi_2 \) band can decay in two possible ways: either by relaxation into the \( \Phi_1 \) band or by direct electron-hole recombination with the corresponding valence subbands. Generally, in quantum wells the interband relaxation processes tend to be very efficient due to phonon scattering, thus limiting the efficiency of emission from the second excited states.
Figure 3.4: Photoluminescence spectrum at 77 K for hexagonal 80 nm pitch nanopore lattices with pore diameter from 50 to 70 nm.

In the nanopore, the in-plane periodicity modifies the $\Phi_{xy}$ wavefunctions. This in-plane variation reduces the overlap between the wavefunctions, thus reducing the probability of transitions from one subband to the next. This has been verified previously in nanopore lasers which exhibit emission from the third $\Phi_{xy}$ subband [7].

As the pattern pitch is varied from 55 nm to 70 nm keeping the pitch constant at 80 nm, the emission from the $\Phi_2$ band relative to the $\Phi_1$ band increases as shown in Figure 3.4. This follows from the tendency of the nanopore to approach quantum dot type behavior as the pore diameter becomes equal to the pitch.

The above measurements for all the nanopore samples were done for a constant pump intensity of 5 kW/cm$^2$. It should also be noted that the process of performing the nanopore etch results in an effective decrease in the available active gain volume (quantum well) (Figure 3.5). For a 55 nm pore diameter pattern, the remaining InGaAs material is 57% of the original quantum well. This decreases to 30% for a 70 nm pore diameter pattern. Thus the observation of enhanced emission from the $\Phi_2$ band as the pore diameter increases could be an artifact of state
filling. However, leakage of carriers into the GaAs barrier for larger pore sizes causes more emission from the barrier layers and counteracts some of the state filling concerns.

![Figure 3.5: Area fill factor of 100 nm pitch nanopore pattern versus pore diameter.](image)

### 3.2 Time resolved photoluminescence

Optical spectroscopy of semiconductors provides invaluable information on electronic band structures, phonons, single-particle spectra and defects. But information on the details of interaction processes among the elementary excitations is often much more difficult to obtain. Here, ultrafast optical spectroscopy can do much more. Indeed, an optical excitation has the ability to generate nonequilibrium carrier and exciton distributions, and time-resolved spectroscopy provides the best means for determining the temporal evolution of such distribution functions. Figure 3.6 shows a typical time-resolved photoluminescence setup.

A high-repetition-rate pulsed laser source (in this case a Ti:sapphire laser operating at 795 nm with an 82 MHz repetition rate) sends photons onto the sample. One part of the laser sample is sent to a fast photodetector coupled to a constant-fraction discriminator. The generated electronic signal, called the “start” signal, is sent to a time-amplitude converter to trigger the data collection. Photons emitted by the sample are scattered by a monochromator and detected by an avalanche photo diode (APD). The intensity is recorded as a function of time. Alternatively, the
decay rate can be determined by performing a statistical average of the arrival times of photons. The excitation level is adjusted in such a way that a single event is seen by the APD, generating a “stop” signal. The time interval between the absorption of a photon and the re-emission is computed. By making this measurement a great number of times and storing each result, the decay profile of the excited sample is constructed.

Figure 3.6: Schematic of time-resolved photoluminescence setup.

Figure 3.7 (a) shows the photoluminescence emission observed from a reference quantum well at a pump power of 1 µW and a nanopore sample at a pump power of 50 µW. Three peaks are observed at 845 nm, 875 nm and 915 nm along with a distinct shoulder centered around 975 nm. The peak at 845 nm appears at first to be from the GaAs substrate. The temperature dependence of the band gap is given by the Varshni equation

\[ E_g(T) = E_g(0) - \frac{\alpha T^2}{\beta + T} \]  

(3.2)

where \( E_g(T) \) is the energy gap at temperature \( T \) and \( \alpha \) and \( \beta \) are constants. For GaAs \( \alpha = (5.405 \pm 0.2) \times 10^{-4} \) eV/K, \( \beta = (204 \pm 45) \) K and \( E_g(0) = (1519.2 \pm 0.2) \) meV [26].
Figure 3.7: Photoluminescence emission at 4 K from (a) QW at pump power of 1 μW and (b) nanopore at pump power of 50 μW.

Assuming a stage temperature of 4 K, the GaAs $n=1$ exciton peak is expected at 827 nm. The peak at 845 nm could be a weakly confined quantum well state. This state is unlikely to be visible at higher temperatures due to the high probability of remaining carriers leaking into the GaAs barrier instead of residing in the weakly bound state.

The peak at 875 nm could be attributed to free-electron to neutral-impurity transitions occurring in the GaAs barrier or substrate, which become evident in very low temperature luminescence studies. The peaks at 915 nm and 980 nm can be identified as the $e_2$-$h_2$ and $e_1$-$h_1$ transitions. The fact that the $e_2$-$h_2$ transition is stronger than the $e_1$-$h_1$ transition suggests that the pump intensity may be causing bleaching of the $e_1$-$h_1$ transition. Similar peaks are observed in the nanopore emission (Figure 3.7 (b)) with the peak at 980 nm being more prominent.

Figure 3.8 shows the instrument response that is obtained by tuning the monochromator to the pump laser wavelength. This characteristic response curve considers both the width of the pump pulse and the response time of the APD. The instrument response function must be deconvolved from all time decay data to obtain the actual material lifetime.
Figure 3.8: APD response when monochromator is tuned to pump laser wavelength.

The emission intensities as a function of time for the nanopore are shown in Figure 3.9. The limit of the x-axis is defined by the repetition rate of the laser that was set to 12 ns. The decay curves for wavelengths below 980 nm (the lowest energy peak in the spectra) show a significant number of background counts. This is a sign that the lifetime is significantly longer than the Ti:sapphire repetition period (~12 ns) and the carriers are not recombing completely between pulses. Similar results have been observed for time resolved photoluminescence of etched QDs and may indicate the presence of radiative defect states with very long lifetimes on the order of a microsecond. The emission at 980 nm is most likely due to the nanopore and has a biexponential decay with a fast component of ~0.4 ps and a slow component of 2 ns.
Figure 3.9: Photoluminescence transients induced at (a) 845 nm, (b) 875 nm, (c) 915 nm and (d) 980 nm by a 795 nm pump pulse in a nanopore sample.

3.3 Site localized InAs quantum dots

Epitaxial self-assembled quantum dots (QD) are attractive due to their ease of fabrication and precise positioning within optical cavities for use in quantum computing devices [27]. The stochastic nature of the QD nucleation process in the Stranski-Krastonow growth regime, however, complicates the deterministic positioning of a single QD within a cavity [28]. If a device is composed of more than one quantum dot, the issue of uniformity becomes critical. Even if the device has only one quantum dot, uniformity is essential to obtain reproducible characteristics across multiple devices. Thus, the geometrical parameters of a quantum dot such
as shape and size as well as the chemical composition need to be controlled. Numerous techniques have been demonstrated for site-controlled nucleation of QDs including lithographic patterning, natural or artificial nano-structuring of the surface or exploiting facet-dependent migration [29-38].

Figure 3.10: Scanning electron micrograph of 80 nm patterned quantum dot array.

The selective area epitaxy method has been adapted to achieve directed self-assembly of InAs QDs on patterned GaAs substrates by metal-organic chemical vapor deposition (MOCVD). Electron beam lithography was used to define apertures of the order of 50 nm in an oxide growth inhibition mask with pitches down to 80 nm. This bottom-up approach provides site and spectral control without the need for modification of the underlying epitaxial structure [12]. Also, the planar nature of the sample after growth is compatible with regrowth steps required for final device integration. We have refined the process to reduce etching damage induced by the mask patterning process. Figure 3.10 shows an SEM micrograph of an 80 nm pitch quantum dot pattern which corresponds to a dot density of $1.8 \times 10^{10}$ cm$^{-2}$.

Figure 3.11 shows an atomic force microscope (AFM) image of a QD array with a pitch of 150 nm and a mask diameter of 70 nm. The height of the QD is about 10 nm considering a 2 nm
base GaAs post arising from the regrown 1 nm GaAs layer. The dots are 60-65 nm in diameter at the base and have a distinct pyramidal structure.

![Atomic force micrographs of 150 nm pitch nanopore patterns showing transition from (a) mostly single dots to (b) multi-dot clusters.](image)

Figure 3.11: Atomic force micrographs of 150 nm pitch nanopore patterns showing transition from (a) mostly single dots to (b) multi-dot clusters.

A statistical analysis of large areas of regrown QDs reveals that about two-thirds of the apertures contain only a single dot with the remaining ones showing early signs of multiple dot nucleation (Figure 3.11 (a)). The number of apertures with single dots tends to 100% as the aperture diameter is further decreased. Increasing the mask diameter to 90 nm by modulation of the e-beam dose results in growth of QD clusters instead of predominantly single QDs (Figure 3.11 (b)). Each cluster contains between two and five quantum dots with the number increasing as the aperture widens. Dot nucleation proceeds due to minimization of the strain energy from a 2D film to a 3D island. The nucleation is governed from a competition between the strain and surface energies. Theoretical simulations of the effect of the mask dimensions on SK growth have confirmed regions in which nucleation of multi-dot clusters may be energetically more favorable [39]. The nucleation dynamics are a function of the mask dimensions and the total amount of deposited material. For a fixed mask dimension, as the deposited material is increased,
the nucleation evolves from single dots to multi-dot clusters and then back to large islands due to relaxation of multiple dots into a single island.

Figure 3.12: Photoluminescence emission from patterned quantum dots at pump power of 100 kW/cm².

Photoluminescence (PL) data was collected at room temperature with an Ar-ion pump laser at 488 nm using a nitrogen-cooled Ge-detector. The PL spectrum from a 150 nm pitch, 90 nm diameter patterned dot region is shown in Figure 3.12. Performing a Gaussian fit yields two peaks at 1 eV and 1.17 eV with linewidths of 186 meV and 109 meV respectively. This bimodal distribution has been observed in regrown InAs quantum dots and has been attributed to indium migration effects during the regrowth process [40].

The above measurement could not be performed at low pump powers due to signal-to-noise ratio limitations. To scale down to lower pump powers, measurements were performed using a Ti:sapphire laser (795 nm wavelength) and detected using a single photon detector. The pump beam was focused down to a spot size of approximately 4 µm with a 0.6 numerical aperture objective. Light emitted from the QD was collected by the same objective and focused onto the input slit of a 0.75 m monochromator. PL spectra for the patterned QDs at a pump intensity of
10 µW (power density 20 W/cm²) are shown in Figure 3.12. Peaks can be observed at 1270 nm (0.977 eV), 1170 nm (1.06 eV), 1068 nm (1.162 eV) and 975 nm (1.27 eV). The peak at 975 nm is due to emission from the wetting layer.

![Comparison between emission from patterned and self-assembled quantum dots showing the excitonic and wetting layer peaks.](image)

Figure 3.13: Comparison between emission from patterned and self-assembled quantum dots showing the excitonic and wetting layer peaks.

Solid-phase diffusion dramatically alters the shape of the QDs during regrowth. Research is ongoing in performing transmission electron microscopy (TEM) analysis on the capped dots to determine their exact dimensions to develop a model to explain the spectral features. It is possible that the peaks at 1270 nm, 1170 nm and 1068 nm correspond to the first, second and third excited states of the dots.

Figure 3.13 shows a comparison between the emission from a 150 nm pitch patterned dot region and a region of self-assembled dots at a pump power of 10 µW. Although the density of dots in the self-assembled region is of the order of $10^{10}$ cm², the randomness in size and shape results in a broadening of the spectrum. This broadening cannot be accurately determined due to the overlap of the emission from the excited state emission that smears out the Gaussian line shape. In comparison, the emission from the patterned region shows clear peaks at 1270 nm, 1170 nm and 1068 nm. The emission intensity from the patterned region is expected to be greater
than that from the self-assembled region due to narrowing of the Gaussian distribution. However, the damage introduced by the processing steps preceding the patterned growth diminishes emission intensity. Despite the damage, similar intensities were observed from both self-assembled and patterned dot regions. Further optimization of the fabrication process is necessary to reduce processing-induced damage and leverage the full potential of our approach.

Figure 3.14: Photoluminescence from patterned quantum dots at 10 µW and 50 µW pump powers.

Figure 3.14 shows the comparison between a patterned QD sample pumped at 10 µW and 50 µW. The emission is increased at the longer wavelengths at the expense of the emission at 1170 nm. The discreteness of the QD spectra results in state-filling effects as the pump intensity is increased. At low intensity only ground-state levels are observed because of the fast intersubband relaxation processes. As the intensity is increased, a progressive saturation of the lower energy transitions is combined with the emergence of the emission peaks originating from the excited-state interband radiative transitions. These are observed as intersubband carrier relaxation towards the lower level is slowed due to the reduced number of available final states.
3.4 Dot regrowth optimization

An often overlooked aspect of epitaxially grown quantum dots is that the capping process can result in drastic changes in the dot geometry. The deposition of the capping layer is itself a lattice mismatched heteroepitaxial process and is therefore associated with strain release, segregation, faceting, intermixing and strain enhanced diffusion. These phenomena take place at the dot surface and strongly modify the dot morphology and composition. Since the optical and electronic properties of the dots strongly depend upon their size, shape and stoichiometry, an understanding of the overgrowth process is essential.

The intermixing and shape change in QDs can be understood considering the strain energies of the whole system. In a QD, strain relaxation occurs mostly at the apex of the dot, while the base remains lattice matched to the GaAs lattice due to the pseudomorphic interface. The lattice spacing at the apex is close to the value of bulk InAs, decreasing slightly towards its base edges and corners. During capping, the amount of relaxed strain energy in principle has to be reinvested. In the directly surrounding GaAs matrix, the local lattice parameter becomes slightly lower than the lattice constant of bulk GaAs. Only far from the QDs does the GaAs capping material adopt its bulk lattice constant.

Growth of GaAs on the dot apex is energetically very unfavorable. Hence, it diffuses towards the sides of the QD where it finds more preferential sticking positions, dragging indium from the apex along with it (Figure 3.15 (b) and (c)). Thus, the dot apex gets removed; gallium and indium intermix and adhere to the sides of the dot, increasing the slope of the side facets [41].

On the one hand, this process leads to a global strain energy reduction compared to a capped pyramid. On the other hand, the local lattice constant on the dot apex slightly decreases since the
distance to the base becomes shorter. During further capping the GaAs layer is higher above the wetting layer than above the dot, and the dot itself gets covered first at its side and finally at its center. This is a result of the competition between the strain energy and the step energies of a rough growth surface. Finally, the dot gets completely covered and the GaAs growth surface flattens again.

Figure 3.15: Schematic of QD capping process [41].

In addition to the transition from a pyramid to a lens shape, the aspect ratio of the dots is also observed to change during the capping process [42]. Even at the early stages of capping a free-standing dot (one monolayer of GaAs), strong modifications take place in island morphology. After depositing three monolayers of GaAs, the dot elongates in the $[1\bar{1}0]$ direction of the substrate. Further deposition causes complete disappearance of the faceted regions and a steady increase in the length of the island elongation.
The results obtained for self-assembled dots are not strictly applicable to capping of selective epitaxy dots grown in the absence of the growth inhibition template. In the conventional method for selective dot epitaxy, the sample is removed from the reactor after growth of the QD layer, the oxide mask is etched away and the sample loaded back into the reactor for regrowth of the capping layer. This is in stark contrast to self-assembly in which capping is initiated immediately after the growth interruption time without any temperature ramps in between.

![AFM image and typical cross section](image)

Figure 3.16: Atomic force microscope image (a) and typical cross section (b) of as-grown 200 nm pitch uncapped InAs quantum dots.

Before the capping initiates, the sample temperature is ramped up to 450 °C in an arsine atmosphere. The arsine overpressure is required to prevent desorption of the volatile arsenic atoms from the surface at elevated temperatures [43]. Thermal stability of the quantum dot structure during the temperature ramp is crucial if they are to be used in device structures. An AFM image of a 200 nm pitch dot array after removal of the oxide template is shown in Figure 3.16 (a). Figure 3.16 (b) shows the typical cross section of the dots grown in one aperture. To
study the effect of the temperature ramp on the dot morphology, the temperature was ramped up to 450 °C in 100 seconds and immediately ramped down to room temperature. Figure 3.17 (a) and (b) show an AFM image and the typical profile of a single aperture after the thermal treatment. Diffusion of the InAs in the $[1\bar{1}0]$ direction results in redistribution of the material into double dot-like structures. Annealing at higher temperatures or for longer times results in complete disappearance of the selective area dots due to diffusion. For denser patterns, adjoining dots merge and no distinct pattern can then be observed. Limiting this morphological transformation during the regrowth temperature ramp is essential to utilizing these QDs for device applications.

![AFM image and cross section](image)

Figure 3.17: Atomic force microscope image (a) and typical cross section (b) of 200 nm pitch uncapped InAs quantum dots after ramping temperature to 450 °C.

Selective area epitaxy occurs since adatoms preferentially avoid nucleation on the mask surface and grow instead on the exposed GaAs surface. After removal of the growth inhibition mask, the island configuration is not stable with respect to the uncovered portion of the sample. This instability is related to both the strain and compositional gradients occurring at the interface.
between the island and the surroundings. When the temperature is raised, the system has a tendency to achieve its equilibrium configuration by strain and composition driven mass transport.

The heat of formation for InAs and GaAs is 114 kcal/mol and 128 kcal/mol respectively [44]. The heat of formation is an indication of the group III element (gallium or indium) to arsenic bond strength. As indicated, the In-As bond is weaker than the Ga-As bond. A weaker bond suggests increased surface mobility during growth. Hence, it is safe to assume that uncapped InAs has a higher tendency to diffuse than GaAs or an InAs-GaAs intermixed alloy.

To limit this surface diffusion, the dots were capped with 2 nm of GaAs after the growth interruption step. Shown in Figure 3.18 (a) and (b) are an AFM image of a GaAs capped InAs QD array and a profile of a typical oxide aperture. The GaAs capping increases the height and simultaneously rounds of the sharp peaks observed in the uncapped samples. After annealing at 450 °C using the same procedure as discussed above, a very small change was observed in the dot shape (Figure 3.19 (a) and (b)). Thus, the surface diffusion was effectively eliminated by capping the InAs QDs with a small GaAs cap immediately after dot growth.

3.5 Stacking of quantum dots

Three-dimensional stacking of quantum dots increases the total volume of the active material and allows tuning of the electronic wave functions due to the quantum mechanical coupling across the spacer layers. In this way, quantum dot molecules can be obtained, which are of interest for quantum computing [45]. Interlayer correlations in multilayer structures have been found to yield significant improvements in the size uniformity of the dots [46].
Figure 3.18: AFM image (a) and typical cross section (b) of as-grown 200 nm pitch InAs quantum dots capped with GaAs before oxide mask removal.

Figure 3.19: AFM image (a) and typical cross section of (b) 200 nm pitch InAs quantum dots capped with GaAs before oxide mask removal and subjected to 450 °C annealing.

Interlayer dot correlations occur due to three kinds of mechanisms: (1) chemical processes such as surface segregation, (2) nonplanarized corrugated surface morphologies, or (3) long-range elastic interactions due to the strain fields emerging from the buried dots (the last two of which are illustrated in Figure 3.20). Each of these processes causes a significant nonuniformity in properties such as strain, surface diffusion and surface chemical potential across the spacer.
layer. If sufficiently large, this will affect the nucleation of the subsequent dots and, under favorable conditions, vertical and lateral correlations in multilayers are formed.

![Figure 3.20: Interlayer correlation due to (a) elastic strain fields and (b) surface corrugations of spacer layers.](image)

In the InAs/GaAs system, strain fields are caused by the strong elastic lattice deformation around the dots due to the large lattice mismatch between the dot material and the surrounding material, and they extend up to the epilayer surface. During subsequent dot layer growth, these strain fields impose a bias on the diffusion current of deposited adatoms due to the concomitant gradients in the surface chemical potential. At the strain minima, therefore, a local enhancement of growth rate, and thus preferential island nucleation, will occur. Additionally, preferential dot nucleation may also be enforced by the local reduction of the island nucleation barrier at these minima.

The surface strain $\epsilon_\parallel$ at a radial distance $r$ due to an island buried at a depth $d$ is given by

$$
\epsilon_\parallel = -\frac{P}{d^3} \left[ \frac{2 - r^2/d^2}{\left(x^2 + d^2\right)^{3/2}} \right]
$$

(3.3)

$P$ is given by $P = \epsilon_0 V_0 (1 + \nu) / \pi$, where $\epsilon_0$ is the lattice mismatch between the dot material and the substrate, $\nu$ is the Poisson ratio and $V_0$ is the volume of the dot [47]. Deposited adatoms
diffuse on the surface along the gradients in the surface chemical potential determined by the surface strain fields until the nearest strain energy minimum is reached. Figure 3.21 (a) shows the dependence of the indium atom binding energy and diffusion barrier on strain [48]. These two factors impact the diffusion kinetics of indium for 2D and 3D arrays of strained islands.

When a 2D sheet of QDs is completed and capped by a spacer GaAs layer, the GaAs lattice is expanded in the regions above the buried InAs QDs. As can be seen from Figure 3.21, the binding strength increases with increasing strain. Thus, the local concentration of adatoms, and hence the nucleation probability for a new island, is increased in the region above a buried island where the capping layer surface is under tensile strain. The tendency to nucleate preferentially in tensile strained regions provides a microscopic explanation for the observation that stacked InAs quantum dots tend to grow vertically aligned at regions of the capping surface where the GaAs lattice constant is widened up and thus more closely matches the InAs lattice constant.

![Figure 3.21: (a) Binding energy and (b) diffusion barrier for an indium adatom on a GaAs (100) surface as a function of strain [48].](image)

Figure 3.22 shows a scanning electron microscope image of the second layer of quantum dots overgrown on a 200 nm pitch patterned InAs array. The nominal thickness of the InAs layer
for the stacked dots was 0.89 nm. During the regrowth, the temperature is ramped to 450 °C before starting the GaAs spacer growth. The stacked dots align at the exact position of the first layer of dots in accordance with theory. Higher temperatures could destroy the initial ordering of the dots due to indium diffusion on the surface. No annealing of the sample at elevated temperatures is used for the same reason.

Figure 3.22: Scanning electron micrograph of the second layer of aligned quantum dots with a spacer thickness of 5 nm between the two layers.

Another important consideration in stacked quantum dots is the spacer thickness required to achieve vertically aligned dots while keeping the dots separate. Studies have shown that for GaAs spacer thicknesses of up to 60 monolayers (17 nm), the dots remained electronically coupled [49]. It should be noted that the coupling refers only to electronic coupling between dot states and not to the elastic coupling between dot layers. As has been shown previously, capping of InAs dots results in planarization of the surface by redistribution of indium from the top of the pyramidal dot and selective incorporation of gallium away from the dot.

After the growth surface has become planar, GaAs growth proceeds in a layer-by-layer manner. Due to the strain field created by the first layer of dots, the overlying dots tend to be larger, resulting in a red shift of the emission spectrum. The GaAs spacer layer was chosen to be
5 nm. Since the average height of the first layer of dots is 8 nm, at least 5 nm of GaAs was found to be necessary to ensure complete coverage of the first dot layer with GaAs before indium deposition for the stacked dots. Narrowing the size distribution in the higher layers of stacked dots makes this a viable approach for attaining highly uniform arrays of quantum dots. Further experiments to study the effect of growth conditions, such as growth rate and temperature and indium coverage, on stacking behavior are the next logical step in using stacked dots for device applications.

3.6 InAs quantum dots by diblock copolymer lithography

While electron beam lithography allows for precise control over the location and geometry, the low throughput due to excessive write times makes it sub-optimal for high volume, large-scale device applications. Nanopatterning techniques for achieving enhanced resolution over large areas in a cost-effective manner involve use of nanoimprint lithography or diblock copolymer lithography.

Diblock copolymer (DBC) lithography consists of linking two polymers, each barely miscible in the other. Microphase separation between two species occurs when the magnitude of the enthalpic contribution (usually positive and decreasing with temperature) to the free energy of mixing the two components is greater than the magnitude of the entropic contribution (always negative and ideally temperature independent); phase separation thus minimizes unfavorable interactions between dissimilar species. This mutual immiscibility leads to microphase separation and formation of ordered structures (Figure 3.23).

Confining the block copolymer to a thin film impacts the ordering and morphology via interactions with the confining surfaces. Each surface (air and substrate) usually has a preferential affinity for one block in the copolymer, so typically that block will wet that
particular surface; for morphologies having anisometric microdomains, such as cylinders and lamellae, these preferential surface interactions thereby induce the microdomains to lie parallel to the substrate. Control over the final morphology hence necessitates a chemical pretreatment or monolayer-based coating or “brush” layer of the patterned surface. While on a local scale, the nanodomains formed are extraordinarily regular, it is difficult to control the order on a more global scale.

Figure 3.23: Scanning electron microscope scans of patterned diblock copolymer film.

The diblock copolymer nanofabrication process consists of two series of polymer application, baking and chemical treatments as shown in Figure 3.24. The first spin-on process is to form a neutral surface with regard to the preferential attraction or absorption of one component of the diblock copolymer. This initial polymer layer or “brush” consists of a random copolymer which is attached over the surface presenting both polymer components to the subsequent polymer layer. The brush layer attaches to the surface through a covalent chemical linkage, such as a hydroxyl-termination, or it can be cross-linked at the surface. The self-organizing diblock copolymer is then coated onto the prepared surface in a thin film, typically 20–40 nm thick. The microphase separation is promoted by thermal annealing at relatively low temperatures above the glass transition temperature of the polymer for extended periods of time.
To utilize this diblock copolymer morphology as a mask for nanopatterning the underlying semiconductor, it is necessary to selectively remove one of the blocks of the diblock copolymer, developing the pattern into a more conventional polymer stencil used in later processing. For example, in the case of a PS-b-PMMA polymer, this would consist of exposing the polymer thin film using ultraviolet (UV) irradiation and then selectively removing the PMMA through chemical etching, i.e. acetic acid. The pattern, which can now be transferred to an underlying dielectric or semiconductor film, is highly uniform over large surface areas even though no long-range order is present.

![Diagram](image-url)

Figure 3.24: Process steps for selective area epitaxy (SAE) of quantum dots on DBC patterned substrates.
Transfer of the DBC pattern to the underlying oxide is performed by a FREON-23 (CHF$_3$) etch process. The DBC mask is then removed by sequential sonication in acetone, methanol and isopropanol. A precautionary O$_2$ ashing step at low power (100 W) is also performed to ensure complete removal of all organic residues. An SEM image of the oxide mask after polymer removal is shown in Figure 3.25. The sample is then etched in 1:1 HCl:DI to remove native III-V oxides and transferred into the MOCVD reactor for regrowth of quantum dots in the oxide apertures. Care should be taken to prepattern the substrate with a sparse oxide pattern (squares or stripes) prior to DBC patterning to avoid excessive growth enhancement effects observed in samples with high oxide coverage.

Figure 3.25: Scanning electron microscope scans of oxide mask after removal of DBC polymer.

The thermal processing of the SiO$_2$ mask prior to growth is required to provide stability during the subsequent processing and growth steps [50]. Samples annealed at the lower temperatures suffer from peeling-off or cracks after the growth of quantum dots. The optimized annealing conditions involve maintaining at 700 °C for 30 min in an ambient of hydrogen. This step is necessary to prevent template mask cracks and detachment during the selective MOCVD growth.
After regrowth, the oxide mask is etched away with buffered hydrofluoric (BHF) acid and the sample is inspected by scanning electron microscopy and atomic force microscopy. Scanning electron microscope images of the regrown quantum dots at two magnifications (100k and 500k) are shown Figure 3.26. The DBC pattern is locally uniform, but does not possess any long range order as depicted by the overlayed domain boundaries on Figure 3.26 (a). In each domain, the dots are arranged in a hexagonal pattern. Dots are absent in some apertures. This is probably due to the incomplete opening of the oxide aperture as even a few ångströms of oxide prevent dot nucleations in the aperture.

![Scanning electron microscope scans of regrown quantum dots.](a) Low resolution scan showing domain boundaries, (b) high resolution scan showing typical size and spacing.

Figure 3.26 (b) shows the arrangement within a single domain. On average the dots are 17 nm in diameter with a center to center spacing of 37 nm. This corresponds to an area density of $8.4 \times 10^{10}$ cm$^{-2}$. This is the highest density of patterned quantum dots reported to date and is better than the best results obtained from self-assembly which yields densities around $5 \times 10^{10}$ cm$^{-2}$. The theoretical limit for a single layer of dots with a diameter of 20 nm in a closely packed structure is $3 \times 10^{11}$ cm$^{-2}$ [51].

61
An atomic force microscope image of the quantum dots is shown in Figure 3.27. The average height was found to be 7 nm including the height of the GaAs wetting and cap layers. TEM analysis would be necessary to accurately determine the final shape and size of the dots after capping. In addition to areal density, another important consideration of patterned quantum dots is the uniformity. High resolution SEM images (e.g. Figure 3.28 (a)) of the dots were analyzed with an image processing algorithm to extract the distribution in sizes of the grown dots. Figure 3.28 (b) shows the outline of each dot in Figure 3.28 (a).

Figure 3.27: 1 μm by 1 μm AFM scan of DBC patterned InAs quantum dots.

Figure 3.28: (a) Scanning electron microscope scan of regrown quantum dots and (b) corresponding edges of quantum dots after image processing.
The dots were fitted to ellipses to extract the area, perimeter, aspect ratio and angle. The corresponding distributions are plotted in Figure 3.29. The dots are strongly elliptical, which is expected due to the directional preferences inherent to the SAE process. The average diameter is around 18 nm with a standard deviation of 1 nm. This degree of uniformity is a big improvement compared to unpatterned self-assembled quantum dots which exhibit variations as high as ±10 nm from the mean value [52].

Figure 3.29: Distribution of area, diameter, angle and aspect ratio of regrown quantum dots.
CHAPTER 4 – NANOPORE BAND STRUCTURE

4.1 Nanopore geometry

In its most general form, the two-dimensional unit cell of a nanopore lattice can be taken as a parallelogram with an embedded ellipse signifying the pore (Figure 4.1). The defining parameters of the parallelogram are the dimensions $L_u$ and $L_v$ and the angle between the two principle axes $\theta$. The pore is characterized by the major and minor axes lengths ($a$ and $b$) and the angle ($\phi$) made by the major axes with the x-axis. Of the various possible lattices, two of the most common are the square lattice ($L_u = L_v$ and $\theta = 90^\circ$) and the hexagonal lattice ($L_u = L_v$ and $\theta = 60^\circ$).

Figure 4.1: Schematic of generalized nanopore unit cell showing the relevant geometrical parameters.

The Cartesian co-ordinates can be expressed in terms of the in-plane generalized co-ordinates as

$$\hat{u} = x \quad \text{and} \quad \hat{v} = x \cos \theta + y \sin \theta$$

(4.1)
Assuming the $z$ axis as the third axis, the reciprocal lattice becomes

$$\vec{u} = \hat{x} \sin \theta - \hat{y} \cos \theta$$

and

$$\vec{v} = \hat{y}$$

(4.2)

For a hexagonal lattice, the first Brillouin zone is also a hexagon while for a square lattice, it is a square as shown in Figure 4.2. From symmetry considerations the band structure calculation can be restricted over the irreducible Brillouin zones for the hexagonal and square unit cells as illustrated in Figure 4.2.

![Brillouin zones](image)

Figure 4.2: Brillouin zones of (a) hexagonal and (b) square lattices with the corresponding irreducible sections shaded gray.

The three high symmetry points for the square Brillouin zone are $\Gamma$, $M$ and $K$ corresponding to $(0,0)$, $(\pi/L,0)$ and $(\pi/L,\pi/L)$. In the case of the hexagonal lattice, the critical points are $\Gamma$, $M$ and $K$ corresponding to $(0,0)$, $(\pi/L,0)$ and $(4\pi/3L,2\pi/3L)$. It is useful to note here that the nanopore band structure calculations are the electronic analog of optical mode calculations in 2-D photonic crystals for the TM-polarization with the potential arising from band gap offset playing the part of the dielectric constant.
4.2 Strain theory

The lattice mismatch between the InGaAs and GaAs layers results in non-uniform strain in the nanopore layer. It is important to evaluate the effect of the strain distribution on nanopore spectral properties. The effects of strain are three-fold: change in band shape which affects the effective mass, splitting of band degeneracies which affects the scattering dynamics and shifts in the energy level which alter the barrier height and hence the emission wavelength.

Assuming linear elastic theory, Hooke’s law is valid and the stress ($\sigma_{ij}$) is proportional to strain ($\epsilon_{\alpha\beta}$) with the proportional constant being the elastic stiffness tensor.

$$\sigma_{ij} = \sum_{\alpha\beta} C_{ij\alpha\beta} \epsilon_{\alpha\beta} \tag{4.3}$$

In cubic crystals due to the high symmetry, only three independent variables ($C_{11}$, $C_{12}$ and $C_{44}$) remain and the stiffness tensor takes the form

$$C_{ij\alpha\beta} = \begin{pmatrix}
C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\
C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\
C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\
0 & 0 & 0 & C_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & C_{44} & 0 \\
0 & 0 & 0 & 0 & 0 & C_{44}
\end{pmatrix} \tag{4.4}$$

where $C_{11}$ depends upon the elastic constant of the material, $C_{12}$ depends upon the Poisson’s ratio and $C_{44}$ represents the shear elastic constant.

In the case of a biaxially strained quantum well, the strain components $\epsilon_{xx}$ and $\epsilon_{yy}$ are

$$\epsilon_{xx} = \epsilon_{yy} = \frac{a_0 - a}{a} \tag{4.5}$$

This equation holds true even for a symmetric nanopore structure when $L_u = L_v$. However, it becomes very difficult to perform these calculations analytically due to the complexity of the structure. In addition to the top and bottom InGaAs-GaAs interfaces which introduce biaxial
compressive strain, additional compressive strain in the growth direction is introduced due to the requirement for lattice matching at the pore interface. This strain further couples to the in-plane strain components and creates a non-uniform strain profile.

The z-component of the strain is related to the x-y components by

\[ \epsilon_{zz} = -2 \frac{\sigma}{1 - \sigma} \epsilon_{xx} = -2 \frac{C_{12}}{C_{11}} \epsilon_{xx} \]  

(4.6)

where \( \sigma \) is the Poisson’s ratio.

An arbitrary strain can be decomposed into three separate strain tensors

\[
\begin{pmatrix}
\epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\
\epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\
\epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz}
\end{pmatrix}
= \frac{1}{3}
\begin{pmatrix}
\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz} & 0 & 0 \\
0 & \epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz} & 0 \\
0 & 0 & \epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}
\end{pmatrix}
\]

\[ + \frac{1}{3}
\begin{pmatrix}
2\epsilon_{xx} - (\epsilon_{yy} + \epsilon_{zz}) & 0 & 0 \\
0 & 2\epsilon_{yy} - (\epsilon_{xx} + \epsilon_{zz}) & 0 \\
0 & 0 & 2\epsilon_{zz} - (\epsilon_{xx} + \epsilon_{yy})
\end{pmatrix}
\]

(4.7)

where the first tensor accounts for the volume change, and the latter two tensors account for the shape change of the unit cell. Correspondingly, the first tensor represents a hydrostatic strain, and the latter two tensors describe the effect of shear strain. The second tensor is related to the change in length along the three axes while the third tensor describes the rotation of the axes. A cubic crystal under biaxial stress does not have any shear stress components and the third tensor can thus be neglected.

Since the crystal symmetry is not altered under hydrostatic strain, it has no effect on lifting the band degeneracy and only results in a shift of the bandgap. It can be understood by
considering that if two atoms are brought close, the interatomic interaction increases resulting in a wider separation between the bonding and antibonding energy states. In the case of biaxial stress, the strain elongates one axis with respect to the other two and thus changes the cubic symmetry into tetragonal symmetry. This causes a splitting of the degeneracy of the heavy hole (HH) and light hole (LH) valence bands.

The change in the conduction band edge is given by

$$\delta E_C = a_c (\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz})$$  \hspace{1cm} (4.8)

and the valence bands are shifted by

$$\delta E_{HH} = -a_v (\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) - \frac{b}{2} (\epsilon_{xx} + \epsilon_{yy} - 2\epsilon_{zz})$$ \hspace{1cm} (4.9)

$$\delta E_{LH} = -a_v (\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) + \frac{b}{2} (\epsilon_{xx} + \epsilon_{yy} - 2\epsilon_{zz})$$

To simulate the lattice mismatch between the InGaAs nanopore layer and the GaAs barrier layers, a thermal expansion coefficient equal to the mismatch is applied in the subdomain composed of InGaAs and the temperature is raised by 1 K. This results in a thermal expansion which acts as the source of stress in the finite element model.

The strain calculations are then coupled to a Schrödinger equation solver to quantify the role of strain in changing nanopore electronic properties. In the effective mass approximation, the most important parameters for calculating nanopore band structure are the band offsets and the effective mass.

The three-dimensional nature of the nanopore structure makes it necessary to perform the band structure calculations on a full three-dimensional model. However, for nanopore patterns with pitches of the range of 80-100 nm, this requirement can be relaxed. Since the thickness of
the nanopore pattern (≈10 nm) is about an order of magnitude smaller than the pitch, the nanopore wavefunction can be split by separation of variables into

$$\Psi_{lmn}(u,v,z) = \Phi_{ln}(u,v) \times Z_n(z)$$  \hspace{1cm} (4.10)

From the Bloch theorem, the in-plane states of the nanopore can be written as

$$\Phi_{ln}(u,v) = \exp(ik_lu) \exp(ik_mv) \psi_{lm}(u,v)$$  \hspace{1cm} (4.11)

where $k_l$ and $k_m$ are the Bloch wave-vectors and $\psi_{lm}$ is the periodic function characteristic of the $l,m$ point in k-space.

The periodic boundary conditions for the square unit cell are

$$\Phi_{lm} \left( \frac{L_u}{2}, v \right) = \exp(ik_mL_u) \Phi_{lm} \left( -\frac{L_u}{2}, v \right)$$
$$\Phi_{lm} \left( u, \frac{L_u}{2} \right) = \exp(ik_nL_v) \Phi_{lm} \left( u, -\frac{L_v}{2} \right)$$  \hspace{1cm} (4.12)

In the natural co-ordinates for the hexagonal lattice, i.e. ($\bar{x}$ and $\bar{x}/2 + \sqrt{3} \bar{y}/2$), the periodic boundary conditions are the same.

The $Z(z)$ function takes the form of the confined states in a finite quantum well

$$Z_n(z) = \begin{cases} 
A_{-1} \exp(\alpha_n z) & \text{for } z < -L_z/2 \\
A_0 \cos(k_n z) & \text{for } -L_z/2 < z < L_z/2 \\
A_1 \exp(-\alpha_n z) & \text{for } z > L_z/2 
\end{cases}$$  \hspace{1cm} (4.13)

This assumption is validated by performing a band structure simulation with a three-dimensional model and comparing it to the results obtained from a two-dimensional model.

Periodic boundary conditions are imposed on the boundaries of the unit cell (Figure 4.3) and the simulation is performed along the high symmetry directions, i.e. $\Gamma \Rightarrow M \Rightarrow K \Rightarrow \Gamma$, to compute the nanopore band structure.
Figure 4.3: Schematic of nanopore unit cell in 3D showing the application of boundary conditions to simulate the periodicity.

4.3 Nanopore strain distribution

Figure 4.4 shows an X-Z slice of the strain distribution in a square nanopore pattern with 100 nm pitch and 50 nm pore diameter. It is instructive to study the strain at the pore-InGaAs interface to understand the effect of the pore on strain distribution and hence the emission properties of the nanopore.
The vertical GaAs-InGaAs interface introduces compressive strain in the nanopore layer in the z-direction, which in turn creates localized tensile strain in the x- and y-directions. Due to the circular symmetry of the pore, the strain fields arising due to the pore can be written in cylindrical co-ordinates (i.e. in terms of $\epsilon_{rr}$ and $\epsilon_{\theta\theta}$). The x-strain $\epsilon_{xx}$ is related to $\epsilon_{rr}$ by $\cos^2 \phi$ while the y-strain is related by $\sin^2 \phi$. Hence, along the x-axis a clear variation in $\epsilon_{xx}$ is observed while no variation is observed in $\epsilon_{yy}$. The minor deviation in $\epsilon_{yy}$ at the pore interface is due to coupling to shear stresses.

Linear plots of the strain profile along the x-axis at $z = 0$ are shown in Figure 4.5. Far away from the GaAs pore, $\epsilon_{xx}$ is compressive. As one approaches the pore, the compressive strain
reduces due to the compensating tensile strain at the pore interface. Due to the continuity boundary conditions imposed at the subdomain boundaries, the GaAs pore is compressively strained in the immediate vicinity of the InGaAs layer. This non-uniform distribution of strain at the interface results in a modification of the effective potential barrier felt by electrons and holes and thus is expected to affect the calculated band structure. The $\epsilon_{zz}$ profile at $z = 0$ shows a behavior that is exactly opposite to that of the $\epsilon_{xx}$ profile. The tensile strain in the InGaAs layer decreases as one approaches the pore interface. The GaAs in the pore is slightly tensile strained due to the InGaAs interface.

### 4.4 Effect of strain on band structure

The strain effects induce an extra potential field which alters the band structure and optical properties of the nanopore. Figure 4.4 shows the effect of strain on the conduction band energy (CBE). Note that this figure only plots the effect of strain and does not take into account the effect of compositional variation on the CBE.

The CBE in the nanopore layer is very similar to a reference quantum well far away from the GaAs pore. Due to the interface strain, the CBE in the InGaAs decreases further due to the compressive strain while that in the GaAs increases due to the tensile strain. Hence the conduction band discontinuity is 0.1205 eV when considering non-uniform strain as opposed to 0.109 eV when assuming uniform biaxial strain in the InGaAs nanopore layer. Thus, the inclusion of inhomogeneous strain effects results in only a minor correction to the band offsets calculated using the uniform strain model.

If the nanopore were to be scaled equally in all three dimensions, the strain distribution is expected to remain identical due to the linearity of the elasticity theory that was assumed. However, if the pitch is reduced with the thickness of the layers remaining constant, the strain
distribution changes. Qualitatively, it can be observed that more sharing of strain occurs between the InGaAs layer and the GaAs pore. Figure 4.5 shows the strain distribution in a 20 nm pitch nanopore pattern with a 10 nm pore diameter.

Figure 4.5: (a) Strain distribution in x-z plane of 20 nm pitch nanopore structure. (b) Profile of strain along x-axis for z = 0.
Comparing the line plots with Figure 4.4, it can be observed that the InGaAs layer is less strained and the GaAs pore is more strained than a nanopore pattern with greater pitch. This can be understood from the fact that the fraction of InGaAs in the entire structure has increased and hence the strain distributes more equally between the GaAs and the InGaAs domains. The calculated conduction band discontinuity for the 20 nm pitch nanopore with a 10 nm pore diameter is 0.112 eV which is very close to the value obtained under the uniform strain assumption (0.109 eV). Thus, it can be concluded that the effect of strain inhomogeneity in the nanopore introduces only a small correction to the band offsets and can thus be neglected in the calculation of the final band structure.

4.5 3D vs. 2D band structure simulation

To verify the validity of the separation of variables method for the nanopore wavefunctions, band structure calculations were performed for a square nanopore lattice with a 100 nm pitch and 50 nm pore diameter. Figure 4.6 shows the calculated band structure for both the full three-dimensional model and the two-dimensional x-y model of the nanopore. Excellent agreement is observed in the band structure between the two models. The full three-dimensional model is offset in energy by 32 meV, which corresponds to the quantized energy of the first confined z-state in the quantum well.
Figure 4.6: Energy dispersion of 100 nm pitch nanopore pattern in (a) 2D mode and (b) 3D mode.

Thus in the case of nanopore patterns in which the pitch is much greater than the thickness of the active layer, the separation of variables assumption holds true and a two-dimensional simulation is sufficient to determine the nanopore in-plane band structure.

However, moving to smaller nanopore patterns in which the pitch is of the order of the layer thickness, the in-plane quantization effects become as significant as the z-confinement and the results from the three-dimensional and two-dimensional calculations differ. Figure 4.7 shows the simulated band structure for a 20 nm pitch nanopore pattern with a 10 nm thick active layer. The pore diameter is chosen as 10 nm. The results for the 2D simulation of the 20 nm nanopore appear as scaled versions of the simulation for the 100 nm nanopore with a widening of the gaps as expected due to an increase in the confinement effects. The dispersion relations for the lower subbands are very similar between the 3D and 2D simulations. This is understandable because, for lower energies, the wavefunction is almost totally confined to the InGaAs layer with minimal penetration into the GaAs pore (Figure 4.8).
Figure 4.7: Energy dispersion of 20 nm pitch nanopore pattern in (a) 2D mode and (b) 3D mode.

Figure 4.8: Nanopore wavefunction along X-Z plane showing near complete confinement to quantum well section.

As a result, the 2D assumption holds true at these lower subbands. For higher order conduction subbands, we see a divergence between the results. The higher subbands penetrate significantly into the GaAs pore because of the weak confinement (163 meV) and hence behave very differently in the 2D and 3D simulations.

4.6 Wavefunction overlaps

When considering the nanopore for device applications, the ability to create designer band structures and precise bandgaps is the most interesting and useful. The nanopore band structure is closely linked to the geometry and material composition of the constituent structures.
Figure 4.9: Energy dispersion for hexagonal nanopore with 100 nm pitch.

Figure 4.9 shows the first six subbands in the conduction band for a hexagonal nanopore pattern with a pitch of 100 nm and a pore diameter of 50 nm. Comparing with Figure 4.6 (a), which shows the band structure for a similar sized square nanopore, we can see that there the bandgap between the first and second subbands disappears in the rhombic lattice. A bandgap appears between the second and third subbands. Also, the third subband is relatively flat indicating a high density of states. This has been observed experimentally in nanopore lasers that exhibit clear subband structures in the emission spectrum and the evidence of a gap between the subbands. For the rhombic lattice, the minima of the bands are at the $K$ point while in the case of the square lattice the subband minima are either at the $\Gamma$ or $K$ points.

The formation of the nanopore introduces in-plane quantization into a quantum well due to the GaAs pore. In the GaAs-InGaAs-AlGaAs system, the GaAs barrier can be replaced by AlGaAs to increase the quantization effects. However, since the pitch is large compared to the electron wavelength ($\lambda_e \sim 10$ nm), the wavefunctions penetrate little into the barriers. Hence for nanopore patterns with a pitch greater than 50 nm, the barrier has almost no effect on the lower subband dispersion relation.
Similarly, increasing the indium content in the active medium has little to no effect on the band structure. However, it should be noted that as the indium content increases, the strain in the quantum well increases and the thickness that can be grown becomes limited by the onset of dislocations.

Photoluminescence characterization of the nanopore has shown enhanced emission from excited states relative to unpatterned quantum wells. Non-equilibrium carriers in higher subbands can relax either by phonon emission to the lower subbands or by recombining radiatively with holes in the valence band. Modification of the in-plane wavefunction from the free-carrier plane wave in the quantum well results in a reduction in the wavefunction overlap. This decreases the probability of carriers relaxing to the ground state and increases the probability of finding carriers in excited states.

Figure 4.10 shows the electron density in the ground and first excited states of a square and hexagonal nanopore pattern. It can be seen that there exists a significant overlap between the electron density in the ground and first excited states for the hexagonal nanopore. Hence the transition is more probable. On the other hand, in the square nanopore there is no overlap between the first and second subband states. Hence, in the case of the square nanopore, a transition from the second to the first excited state is highly improbable. The ability to design a nanopore pattern with a definite subband structure and to control the wavefunction overlaps makes the nanopore a candidate for application to intersubband devices.
4.7 Nanopore subbands and the free electron model

In the free electron model, the energy dispersion can be expressed as

\[ E = \frac{\hbar^2}{2m} |k + G|^2 \]  

(4.14)

where \( k \) is the wavevector and \( G \) is the reciprocal lattice vector. The reciprocal lattice of a 2D square real space lattice, whose lattice constant is \( a \), is also a square lattice of spacing \( 2\pi/a \).

The reciprocal lattice vector \( G \) can be written as

\[ G = \frac{2\pi}{a} (ui_x + vi_y) \]

where \( u \) and \( v \) are integers.

By adjusting the units, Eq. 4.15 can be simplified to
\[ E = (k_x + G_x)^2 + (k_y + G_y)^2 \] \hspace{1cm} (4.15)

where \( k_x \) and \( k_y \) range between \(-\frac{\pi}{a}\) and \(\frac{\pi}{a}\), and \(G_x\) and \(G_y\) are integer multiples of \(\frac{2\pi}{a}\).

Based on this equation the energy bands in any arbitrary direction can be computed. Due to the periodicity of the reciprocal lattice, it is sufficient to compute the band structure over the irreducible first Brillouin zone (BZ) from \(-\frac{\pi}{a}\) to \(\frac{\pi}{a}\).

Along the \langle 10 \rangle direction, \( k_y = 0 \) and

\[ E = (k_x + G_x)^2 + G_y^2 \] \hspace{1cm} (4.16)

The lowest energy state is located at \( G_x = G_y = 0 \). Thus the lowest branch of the E-k dispersion is a parabola \( E = k_x^2 \) as shown in Figure 4.11. The extended portion of the \( E = k_x^2 \) parabola in the second BZ can be wrapped back into the first BZ to yield the second nanopore subband.

Alternatively, this second subband can also be written as \( E = (k_x - \frac{2\pi}{a})^2 \). In the absence of a nanopore lattice, the first and second subbands are connected at the \( \left( \frac{\pi}{a}, 0, 0 \right) \) point. The third and fourth subbands are doubly degenerate and can be expressed as \( E = k_x + \left( \pm \frac{2\pi}{a} \right)^2 \). Since the dispersion is in the (100) direction, the two \( G_y \) reciprocal lattice points yield identical energies and hence the degeneracy. In the same vein, higher order subbands can be expressed in terms of Eq. 4.17 for different values of \( G_y \) as shown Figure 4.11.
Figure 4.11: Energy dispersion for square nanopore lattice in the [100] direction showing the correlation between the subbands and the free-electron bands.

It is important to note that the color coding does not strictly define a specific band as band crossing complicates the standard procedure of identifying a band based on sorting of eigenvalues in increasing order.

In the nearly free-electron model, the opening of the gap between the first and second subbands ($\Delta E_{12}$) is proportional to the strength of the perturbation, which in this case is the wide bandgap pore etched into the quantum well. For larger nanopore patterns (pitch $\approx$ 80 nm), the wavefunction penetration into the pore is negligible and hence $\Delta E_{12}$ is almost independent of the
pore material. As the pore diameter increases for a fixed pitch, the perturbation seen by the in-plane wavefunctions increases and $\Delta E_{12}$ widens (Figure 4.12 (a)).

![Figure 4.12](image1.png)

(a)  

![Figure 4.12](image2.png)

(b)  

![Figure 4.12](image3.png)

(c)  

Figure 4.12: Bandgap between first and second subbands as a function of (a) pore diameter for 80 nm pitch nanopore array, (b) band gap offset between quantum well and etched pores for 80 nm pitch array with 40 nm pore diameter and (c) pitch. The pore diameter is fixed at 50% of the pitch.

The widening gap can also be explained by the increasing confinement of carriers in the section of the quantum well between the surrounding etched pores as illustrated in Figure 3.2 (b). Shown in Figure 4.12 (b) is the variation in $\Delta E_{12}$ between the first and second subband for an 80
nm pitch nanopore array with a 40 nm pore diameter. For lower band offsets, $\Delta E_{12}$ increases linearly and saturates at higher offset values.

The change in $\Delta E_{12}$ with array pitch is shown in Figure 4.12 (c). The diameter for each pattern is assumed to be half of the pitch. The observed trend is analogous to that observed in the fundamental bandgap between the conduction and valence bands for elements in the periodic table. As the lattice constant decreases, adjacent cells are more strongly coupled and this increases the splitting between the bonding and antibonding states at the Bragg point.

4.8 Nanopore stochastic variations

To achieve significant electron confinement effects, the nanopore lattice constant should be of the order of the electron wavelength, which is about 10 nm. A lack of high uniformity fabrication techniques restricts patterns in the sub-50nm range. At these length scales, nanopore spectral characteristics are intricately linked to the pattern geometry and composition of the quantum well and barrier layers. Fabrication tolerances are of central importance since they determine the degree to which the optical gain is inhomogeneously broadened. It may so happen that realizing an intersubband device using the nanopore is theoretically feasible, but highly unrealistic in terms of the required fabrication tolerances.

The nanopore array is fabricated on an 8.7 nm In$_{0.31}$Ga$_{0.69}$As quantum well. Silicon dioxide which acts as the etch mask is patterned by electron beam lithography (EBL) and a FREON-23 (CHF$_3$) based dry etch. The oxide pattern is transferred to the quantum well by a calibrated phosphoric acid-hydrogen peroxide etch. Figure 4.13 shows an SEM image of the etched nanopore sample. Pore-to-pore variations can arise due to the EBL or the oxide dry etching or the quantum well wet etching. The PMMA patterns after electron beam lithography are very regular due to the averaging introduced by the proximity dosing effect. The anisotropy of the dry
etching ensures that oxide mask pattern after etching is identical to the PMMA resist pattern. Hence, wet etching is expected to be the prime contributor to pore-to-pore size variations.

Figure 4.13: Scanning electron micrograph of 80 nm pitch array after wet etching of the quantum well.

The nanopore geometry is defined by the pore diameter and pitch. To study the effect of pore shape and size, the extracted pore profiles were fitted to ellipses. As shown in Figure 4.14, an ellipse is characterized by three main parameters: the major axis, the minor axis and the angle of the ellipse. Alternatively, the major and minor axes can be uniquely described by the area and aspect ratio values. The aspect ratio is defined as the ratio of the major axis to the minor axis and gives a measure of the circularity of the pores. The extracted perimeter gives a measure for the edge roughness in the pore pattern.

Figure 4.14: Model for fabricated nanopores and its fundamental parameters.
Figure 4.15: Distribution of (a) area, (b) diameter, (c) angle and (d) aspect ratio for 80 nm pitch nanopore array.

The statistics for an 80 nm pitch etched quantum well pattern are shown in Figure 4.15. The mean area is 1360 nm$^2$ with a standard deviation of 50 nm$^2$. Assuming a circular pore, this corresponds to a diameter distribution of 42 ± 1 nm, i.e. $\sigma_d/d = 0.0025$. The uniformity obtained coupled with the regular periodicity represents a significant improvement compared to self-assembled dots [53, 54]. Assuming a circular pattern, this corresponds to a diameter distribution of 42±1 nm. The aspect ratio has a mean value of 1.3 which indicates that the pores are nearly circular. The ellipses are mostly aligned at 10 degrees to the x-axis which could be an artifact of the misalignment of the pores to the SEM scan.

The energy dispersion along the principal symmetry directions ($\Gamma \rightarrow M \rightarrow K \rightarrow \Gamma$) for an $L = 80$ nm, $diam = 40$ nm array is shown in Figure 4.16 (a). Error bars for each subband corresponding to the ±1 nm variation in $diam$ are overlayed on the dispersion for the ideal
pattern. The effect of the diam variation is more prominent on the higher order subbands, which is consistent with established theory for quantum size structures. Figure 4.16 (b) shows the corresponding effect of a ±1 nm variation in diam on an L = 20 nm, diam = 10 nm array. Due to the greater confinement, the subband separation is increased. However, the error bars increase proportionately relative to the L = 80 nm case. Performing the same calculations assuming a ±2 nm variation in diam yields the energy dispersions shown in Figure 4.17.

For the hexagonal lattice, there exists a gap between the second and third subbands (ΔE_{32}). ΔE_{32} is inversely proportional to L, decreasing from 2.2 meV for an L = 40 nm pitch, diam = 20 nm array to 0.7 meV for an L = 100 nm, diam = 50 nm array (Figure 4.18 (a)). In the L = 80 nm pattern, ΔE_{32} decreases from 0.92 meV to 0.72 meV due to the ±1 nm variation in diam. Assuming a ±2 nm variation in diam, the ΔE_{32} vanishes as shown in Figure 4.17 (a) and (b).

![Figure 4.16: Effect of a ±1 nm variation in the pore diameter on the energy dispersion of an (a) 80 nm and (b) 40 nm pitch nanopore array.](image)
For lower pitch patterns, the increase in $\Delta E_{32}$ is accompanied by a greater sensitivity of the absolute band position to variations in $\text{diam}$. Figure 4.18 (a) shows the change in the position of the second ($\delta E_2$) and third subbands ($\delta E_3$) due to a $\pm 1$ nm variation in $\text{diam}$ for arrays with pitches from 40 to 100 nm. As $L$ decreases, a 1 nm change in $\text{diam}$ causes a larger change in $\delta E_2$ and $\delta E_3$ which partially offsets the ability to create wider band gaps. Calculations for a $\pm 2$ nm variation in $\text{diam}$ reveal that the gap between the second and third subbands is completely smeared out by the geometry-induced inhomogeneous broadening (Figure 4.17 (a) and (b) and 4.18 (b)).

The dependence of the subband gap ($\Delta E_{32}$) on pitch for $\pm 1$ nm and a $\pm 2$ nm variation in the pore diameter is shown in Figure 4.18. It can be seen that for all pitches, the stochastic variation needs to be maintained under 2 nm to prevent the gap from disappearing due to the inhomogeneous broadening.

In addition to the size of the pore, the aspect ratio (i.e. ratio of major axes to minor axes) and the angle are also sources of variation in the fabricated structures. Figure 4.19 (a) shows a
schematic of an elliptical pore that is aligned ideally at 90° to the x-axes with distributions on either side. The corresponding effect on the energy dispersion is shown in Figure 4.19 (b). Since the overall area of the pore is unaffected, no change is observed at the Γ point.

Figure 4.18: Dependence of ΔE_{32} on array pitch. Pore diameter is set to half of the pitch. δE_2 and δE_3 are the errors in the position of the second and third subbands due to a (a) ±1 nm and (b) ±2 nm variation in the pore diameter.

Figure 4.19: (a) Schematic of simulation setup showing idealized elliptical pore (red) and rotated versions (blue and green dotted). (b) Energy dispersion showing the variation for a ±40° rotation of the ellipse.
Rotation of the pore perturbs the energy band structure in the K direction. This can be explained by the fact that the wave-vectors around the K point correspond to traveling waves along the diagonal of the rhombic unit cell which are most affected by the rotation.

Figure 4.20: (a) Energy dispersion showing the variation in the energy levels for a constant area pore with aspect ratio changed from 0.8 to 1.2. (b) Energy dispersion for a pore with constant x length (40 nm) and aspect ratio changed from 0.8 to 1.2.

Figure 4.20 (a) shows the energy dispersion and the associated error bars for an $L = 80$ nm pitch pattern for which the area of the pore has been fixed and the aspect ratio varied from 0.8 to 1.2. No change in $\Delta E_{32}$ is observed in this case as the effect of the pore ellipticity is not manifest on the $\Gamma$ point at which the bandgap exists. If the same calculation is performed by fixing the length of one axis and changing the aspect ratio, a drastic change is observed in $\Delta E_{32}$ (Figure 4.20 (b)). This is due to a change in the effective percentage of GaAs in the nanopore layer.
CHAPTER 5 – CARRIER RELAXATION DYNAMICS

The non-equilibrium dynamics of particles in semiconductors are dominated by random scattering events due to impurities, lattice vibrations, etc., which randomize the momentum and energy of charged particles in time. Based on first order, time-dependent perturbation theory, the transition rate from an initial state $\Psi_i$ to a final state $\Psi_f$ is given by Fermi’s golden rule

$$\Gamma(\Psi_i;\Psi_f) = \frac{2\pi}{\hbar} \left| \langle \Psi_f | H' | \Psi_i \rangle \right|^2 \delta(E_f - E_i \pm \hbar \omega)$$

(5.1)

where $H'$ is the scattering potential of this process, $E_i$ and $E_f$ are the initial and final state energies of the particle. The delta function describes conservation of energy, valid for long times after the collision is over, with $\hbar \omega$ being the energy emitted (upper sign) or absorbed (lower sign) during the process. The total rate of scattering for a particle in an initial state $\Psi_i$ to any possible final state is given by an integral (sum) over all final states

$$\Gamma(\Psi_i) = \frac{2\pi}{\hbar} \sum_{\Psi_f} \left| \langle \Psi_f | H' | \Psi_i \rangle \right|^2 \delta(E_f - E_i \pm \hbar \omega)$$

(5.2)

The scattering mechanisms are roughly divided into scattering due to crystal defects, which is primarily elastic in nature; scattering between electrons and lattice vibrations or phonons, which is inelastic; and finally scattering between the particles themselves, including both single particle and collective type excitations (Figure 5.1). Phonon scattering involves different modes of vibration, either acoustic or optical, as well as both transverse and longitudinal modes. Carriers may either emit or absorb quanta of energy from the lattice, in the form of phonons, in individual scattering events.
Because of the spherical symmetry of the wave function at \( \Gamma \), optical deformation potential scattering is zero in the lowest GaAs minimum. The effective mass in this minimum is very small \( m^* = 0.067m_0 \). Therefore acoustic phonon scattering contributes little. These two facts are the reason why GaAs exhibits high electron mobility compared to silicon. The most important scattering mechanisms for GaAs are polar optical phonon scattering, ionized impurity scattering if the density of impurities is significant (\( \gtrsim 10^{17} \text{ cm}^{-3} \) at room temperature), and intervalley scattering in a very high electric field. The maxima of the valence band for GaAs are at \( \Gamma \), and the wave function does not have spherical symmetry. Therefore optical deformation potential scattering is important for holes in GaAs. Also, acoustic phonon scattering and impurity scattering are important in the valence band. The scattering mechanisms are further enhanced by the presence of roughness at interfaces.

At low temperatures the dominant process of energy relaxation by hot electrons in semiconductors is by acoustic phonon scattering. The rate of energy loss due to electromagnetic radiation is absolutely tiny by comparison. At higher electron temperatures optical phonon emission take over. The changeover from acoustic to optical phonon emission occurs at about 36
meV and the changeover carrier temperature is between 30 and 50 K. At room temperature optical phonon emission is dominant. The relaxation pathways for high energy electrons by phonon scattering processes are sketched in Figure 5.2. Considering a typical GaAs-Al₃₀Ga₇₀As QW with 100 Å thickness, we see that optical phonon emission will be the dominant mechanism for carrier relaxation. There is now agreement that the intersubband relaxation time is of the order of 0.3 to 0.6 ps for GaAs QWs with $E_{21}$ of the order of 100 meV.

![Diagram of energy relaxation](image)

**Figure 5.2:** Energy relaxation of hot carriers for (a) intersubband spacing greater than optical phonon energy and (b) intersubband spacing lesser than optical phonon energy.

The expressions for scattering rate for relevant scattering processes which can be applied to the nanopore are derived in the following sections. Acoustic and polar optical phonon scattering for intravalley scattering is reviewed. The use of InGaAs as the active material makes it necessary to compute the alloy scattering rate and determine its relative importance.
5.1 Acoustic phonon scattering

Acoustic phonons can be viewed as perturbations that cause local lattice compression or dilation. The local homogeneous compression or dilation can be described as a local relative volume change \( \nabla . U(r) = \frac{\Delta V}{V} \). An electron feels this deformation as a local change of the lattice constant which shifts the single particle energy. For a simple nondegenerate band, the electron-phonon interaction can be written as

\[
H_{e-\text{ac}} = \Xi_{ac} \nabla . U(r,t) = \Xi_{ac} q . u(r,t)
\]  

(5.3)

where \( \Xi_{ac} \) is the acoustic deformation constant and \( u(r,t) \) is the displacement given by

\[
u(r) = \sum_q \sqrt{\frac{\hbar}{2 \rho V_c \omega_q}} \left( a_q e^{i qr} + a_q^\dagger e^{-i qr} \right) \bar{e}_q
\]  

(5.4)

where \( \rho \) is the crystal density, \( V_c \) the unit cell volume, \( \omega_q \) the phonon frequency, \( a_q \) and \( a_q^\dagger \) are respectively the annihilation and creation operators for phonon mode \( q \) and \( \bar{e}_q \) is the polarization unit vector. The above expression is a result of the definition of the displacement operator in terms of the annihilation and creation operations \( a_q \) and \( a_q^\dagger \).

The acoustic phonon perturbing potential becomes

\[
H_{ac} = \Xi_{ac} \sum_q \sqrt{\frac{\hbar}{2 \rho V_c \omega_q}} \left( a_q e^{i qr} + a_q^\dagger e^{-i qr} \right) \bar{e}_q \bar{q}
\]  

(5.5)

The matrix element for scattering from \( \Psi_i \) to \( \Psi_f \) is given by

\[
M_{ac} = \langle \Psi_f | H_{ac} | \Psi_i \rangle = \Xi_{ac} \sqrt{\frac{\hbar}{2 \rho V_c \omega_q}} q \left( N_q + \frac{1}{2} \pm \frac{1}{2} \right)^{1/2} I_{if}
\]  

(5.6)

where \( N_q \) is the number of phonons given by the Bose-Einstein distribution function.
\[ N_q \approx N_q + 1 = \left( e^{\frac{\hbar \omega}{k_B T}} - 1 \right)^{-1} \quad (5.7) \]

and \( I_{gf} \) is the overlap integral

\[ I_{gf} = \iiint \Psi_f e^{iq \cdot r} \Psi_e dxdydz \quad (5.8) \]

In the equipartition approximation

\[ N_q = \frac{1}{e^{\frac{\hbar \omega}{k_B T}} - 1} = \frac{1}{1 + \frac{\hbar \omega}{k_B T} + \ldots - 1} = \frac{k_B T}{\hbar \omega} \quad (5.9) \]

and in the linear dispersion regime for acoustic phonons defined by \( \omega = \nu_s q \), we get

\[ |M_{ac}|^2 = \Xi_{ac}^2 \frac{\hbar q^2}{2 \rho V_c \nu_s q} \frac{k_B T}{\hbar \nu_s q} I_{gf}^2 = \Xi_{ac}^2 \frac{k_B T}{2 \rho V_c \nu_s^2} I_{gf}^2 \quad (5.10) \]

where \( \nu_s \) is the sound velocity in the crystal.

This matrix element is then incorporated into the transition probability given by Eq. 5.2. In the case of a three-dimensionally confined system such as a quantum dot or a nanopore, the sum in Eq. 5.2 remains a sum and needs to be numerically computed keeping the energy conservation in mind. Since in an elastic process no distinction is made between final states attained by means of absorption or emission processes, the total transition rate can be taken to be twice the absorption or emission rates. Hence

\[ |M_{ac}|^2 = \Xi_{ac}^2 \frac{k_B T}{\rho V_c \nu_s^2} I_{gf}^2 \quad (5.11) \]

The relevant parameters for GaAs and InAs are given in Table 5.1.
Table 5.1: Scattering parameters for silicon, gallium arsenide and indium arsenide.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Si</th>
<th>GaAs</th>
<th>InAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac. def. pot. $\Xi_{ac}$</td>
<td>9 eV $^a$</td>
<td>-7.17 eV $^a$</td>
<td>-5.08 eV $^a$</td>
</tr>
<tr>
<td>Opt. def. pot. $\Xi_{opt}$</td>
<td>80 $^b$</td>
<td>31 $^b$</td>
<td>37 $^b$</td>
</tr>
<tr>
<td>Sound velocity $v_s$</td>
<td>$9 \times 10^3 m/s$ $^c$</td>
<td>$4.73 \times 10^3 m/s$ $^c$</td>
<td>$3.83 \times 10^3 m/s$ $^c$</td>
</tr>
<tr>
<td>Density $\rho$</td>
<td>2.33 g/cm$^3$ $^c$</td>
<td>5.317 g/cm$^3$ $^c$</td>
<td>5.68 g/cm$^3$ $^c$</td>
</tr>
<tr>
<td>Diel. const $\epsilon_0$</td>
<td>11.9 $^b$</td>
<td>13.1 $^b$</td>
<td>15.15 $^b$</td>
</tr>
<tr>
<td>High freq. diel. const. $\epsilon_\infty$</td>
<td>--</td>
<td>10.89 $^c$</td>
<td>12.3 $^c$</td>
</tr>
</tbody>
</table>

$^a$ – Reference [55]
$^b$ – Reference [56]

The above deformation potential values are only approximate and correct to within an order of magnitude.

In general, acoustic phonon scattering is assumed to be elastic with no change in energy between initial and final states. The acoustic phonon dispersion is assumed to be a straight line with the slope proportional to the sound velocity in the medium. This is the Debye approximation which is used in conjunction with the equipartition principle to simplify the calculation of transition rates due to acoustic phonons.

The assumption that acoustic phonon scattering is an elastic process is based on the fact that, at the length scales involved in periodic heterostructures, the relevant acoustic phonons have very low energies compared to the carrier energy. However, in situations where acoustic
scattering is the dominant scattering mechanism, the energy of the phonon, however low, should be considered in the interaction for accurate simulation of the energy dissipation process by acoustic phonons. In such a scenario it becomes necessary to compute the scattering from an initial to a final state considering only those phonons that possess energies corresponding to the energy difference between the two states. Also, to ensure that dissipation occurs with a higher probability than energy accumulation, the phonon emission process should have a higher probability that the phonon absorption process. Thus the energy conservation equation is written as

\[ E_f = E_i \pm \hbar q \nu \]  

(5.12)

### 5.2 Optical phonon scattering

In crystals with two or more basis atoms per unit cell, a long-wavelength optical phonon involves relative displacements of atoms within the primitive unit cell. Unlike acoustic phonons, a long-wavelength optical phonon does not involve a macroscopic strain of the crystal, since there are no macroscopic distortions of the lattice. Instead, optical phonons can be regarded as “microscopic distortions” within a primitive unit cell. Optical phonons can change the energy of an electronic band in two ways. In non-polar crystals, optical phonons alter the electronic energies by changing the bond lengths and/or bond angles. This electron-optical phonon interaction is the analog of the deformation potential interaction of acoustic phonons and is known as the deformation-potential electron-optical phonon interaction.

Although the deformation-potential optical phonon matrix element is almost identical to the acoustic one, its effect is different. In case of optical phonons, the two sublattices vibrate out of phase with each other. Hence the scattering process is sensitive to the symmetry of the particular band involved. If the electron is scattered close to the \( \Gamma \) minimum and if the wavefunction has
spherical symmetry, the matrix element vanishes. Hence, optical deformation potential scattering is forbidden in the conduction band of GaAs and this is the cause for the high electron mobility in GaAs.

In polar crystals, a long-wavelength longitudinal optical (LO) phonon involves uniform displacements of the charged atoms within the primitive cell. Such relative displacements of the charged atoms generate a macroscopic electric field. This electric field can then interact with electrons in a process known as the Fröhlich interaction.

The Hamiltonian describing the interaction of an electron with a potential field \( \phi(r) \) is

\[
H_{\text{pop}} = \int \rho(r') \phi(r') d^3r'
\]

(5.13)

where \( \rho(r) \) is the electron charge density. From Poisson’s equation \( \nabla \cdot E = \rho/\epsilon \)

\[
H_{\text{pop}} = \int \epsilon \nabla E_c(r') \phi(r') d^3r'
\]

(5.14)

where \( E_c(r) \) is the electric field associated with the electron charge. Integrating by parts yields

\[
H_{\text{pop}} = -\epsilon \int E_c(r') \nabla \phi(r') d^3r' = \epsilon \int E_c(r') E_p(r') d^3r'
\]

(5.15)

where \( E_p(r) \) is the phonon electric field. The electric field due to an electron assuming screening is given by

\[
E_c(r) = \frac{e}{4\pi\epsilon} \nabla_r \left[ \frac{1}{|r - r'|} \exp\left\{-q_0 |r - r'| \right\} \right]
\]

(5.16)

where \( q_0 \) is the inverse screening length. The screening length is a function of the carrier density \( n \) and the dielectric permittivity \( \epsilon \). In the Debye formulation \( q_0 \) is given by

\[
q_0 = \sqrt{\frac{e^2 n}{ek_B T}}
\]

(5.17)
The polarization field due to optical phonons is

\[ P(r) = e^* u(r)/V_c \]  

(5.18)

where \( V \) is the unit cell volume and \( e^* \) is the effective charge on the atoms. If a low frequency electric field is applied to a material, both the charged atoms and the electrons contribute to the polarization. At high enough frequencies, the atoms being heavier do not contribute and only the electron contribution remains. Thus by taking the difference between the high and low frequency polarizations, the effective charge on the atoms is calculated to be

\[ \left[ e^* \right]^2 = e_0 \bar{M} \omega_{op}^2 V \left[ \frac{1}{\epsilon_r(\infty)} - \frac{1}{\epsilon_r(0)} \right] \]  

(5.19)

where \( \bar{M} \) is the reduced mass of the atoms \( \bar{M} = [1/M_1 + 1/M_2]^{-1} \), \( \omega_{op} \) is the optical phonon angular frequency and \( \epsilon_r(\infty) \) and \( \epsilon_r(0) \) are the high frequency and low frequency dielectric permittivities. The electric field produced by this polarization is

\[ E_p(r) = \frac{1}{\epsilon_0} P(r) = \frac{1}{\epsilon_0} e^* u(r)/V_c \]  

(5.20)

Since

\[ u(r) = \sum_q \sqrt{\frac{\hbar}{2\rho Vq \omega_q}} (a_q e^{iq \cdot r} + a_q^\dagger e^{-iq \cdot r}) \tilde{e}_q \]  

(5.21)

the polar optical phonon interaction Hamiltonian \( H_{pop} \) then becomes

\[ H_{pop} = \frac{e}{4\pi \epsilon V} \left[ \sum_q \sqrt{\frac{\hbar}{2\rho Vq \omega_q}} (a_q e^{iq \cdot r} + a_q^\dagger e^{-iq \cdot r}) \right] \left[ \int \nabla' \left( \frac{1}{|r - r'|} e^{-i|q| r'} \right) e^{iq \cdot r'} d^3 r' \right] \]  

(5.22)

Here, it is assumed that optical phonons are dispersionless with \( \omega_q = \omega_0 \), where \( \omega_0 \) is the bulk optical phonon frequency. Making the substitution \( r' - r = s \) and assuming that \( q \) is in the
z-direction (valid assumption due to the circular symmetry of potential), we can perform a conversion to polar co-ordinates giving

\[
I = \int \nabla_r \left( \frac{1}{|r-r'|} e^{-q_0 r - r|} \right) e^{iq \cdot r'} d^3 r'
\]
\[
= e^{iq \cdot r} i q \int \frac{1}{s} e^{-q_0 s} e^{iq \cdot s} d^3 s
\]
\[
= e^{iq \cdot r} i q 2\pi \int_0^\infty s^2 ds \int_{-1}^1 d(cos \theta) e^{iq s \cos \theta} \frac{e^{-q_0 s}}{s}
\]
\[
= e^{iq \cdot r} i q 2\pi \int_0^\infty s^2 ds \left[ \frac{e^{iq s} - e^{-iq s}}{iq s} \right] \frac{e^{-q_0 s}}{s}
\]
\[
= e^{iq \cdot r} i q \frac{4\pi}{q} \int_0^\infty \sin(qs) e^{-q_0 s} ds
\]
\[
= e^{iq \cdot r} i q \frac{4\pi}{q} \frac{q}{q^2 + q_0^2}
\]
\[
= e^{iq \cdot r} i q \frac{4\pi}{q^2 + q_0^2}
\]

The Hamiltonian for this interaction is given by

\[
H_{pop} = \frac{e e^*}{4\pi e V_c} \left[ \sum_q \sqrt{\frac{\hbar}{2\rho V_c \omega_0}} \frac{4\pi}{q^2 + q_0^2} \left( a_q + a^*_q \right) e^{iq \cdot r} \hat{e}_q \hat{q} \right]
\]
\[
= \frac{ee^*}{e V_c} \sqrt{\frac{\hbar}{2\rho V_c \omega_0}} \left[ \sum_q \frac{1}{q^2 + q_0^2} \left( a_q + a^*_q \right) e^{iq \cdot r} \hat{e}_q \hat{q} \right]
\]
\[
= \sqrt{\frac{\hbar \omega_0}{2V_c}} \left[ \frac{1}{e(\infty)} - \frac{1}{e(0)} \right] \left[ \sum_q \frac{1}{q^2 + q_0^2} \left( a_q + a^*_q \right) e^{iq \cdot r} \hat{e}_q \hat{q} \right]
\]
\[
(5.24)
\]

The polar optical phonon matrix element \(M_{pop}\) becomes

\[
M_{pop} = \frac{C_{pop}}{q^2 + q_0^2} \left( a_q + a^*_q \right) \hat{e}_q \hat{q} \langle \Psi_f | e^{iq \cdot r} | \Psi_i \rangle
\]
\[
(5.25)
\]

where
\[ C_{\text{pop}} = \sqrt{\frac{\epsilon_0^2}{2V} \left[ \frac{1}{\epsilon(\infty)} - \frac{1}{\epsilon(0)} \right]} \]  

(5.26)

Since the nanopore wavefunctions are not in an analytical form, the sum over \( q \) cannot be converted to an integral and has to remain a sum. The final form of the matrix element then is

\[ M_{\text{pop}} = \frac{C_{\text{pop}}}{q^2 + q_0} q I_g(q) \left( N_0 + \frac{1}{2} \pm \frac{1}{2} \right) \]  

(5.27)

with \( I_g(q) \) denoting the Fourier transform of the overlap of the initial and final wavefunctions and \( N_0 \) being the number of phonons having an energy of \( \hbar \omega_{\text{op}} \), which is given by

\[ N_0 = \frac{1}{\frac{\hbar \omega_{\text{op}}}{k_B T} - 1} \]  

(5.28)

The Fröhlich interaction depends upon the phonon wave vector as \( q^2 \). If screening is neglected, it diverges, as \( q \) approaches zero. This is not possible in intraband electron-LO-phonon scattering because the LO phonon frequency is nonzero even at \( q = 0 \). Energy and momentum conservation prevents electrons from undergoing intraband scattering via \( q \equiv 0 \) optical phonons. Nevertheless, depending on the band dispersion, \( q \) can be quite small, and this scattering mechanism can dominate at temperatures where a significant number of LO phonons are excited (i.e. \( k_B T \geq \hbar \omega_{\text{LO}} \)). The matrix element calculation follows closely the method used for intraband acoustic phonon scattering with the substitution of \( H_{\text{ac}} \) by \( H_{\text{pop}} \) in Eq. 5.6.

### 5.3 Multi-phonon scattering

In quantum wells, the dominant relaxation process is longitudinal-optical (LO) phonon emission via the Frohlich interaction, with subpicosecond relaxation times. In a quantum dot, the process is forbidden due to the very discrete nature of the levels, unless the level separation equals the LO phonon energy \( \hbar \omega_{\text{LO}} \). Deformation potential interaction with longitudinal-
acoustic (LA) phonons, which is already weak in the bulk compared to the Frohlich interaction, becomes even weaker due to the decreasing form factor. The form factor decreases rapidly with wavevector $q$ and therefore with the energy of the LA phonon emitted. Hence, relaxation via LA phonon emission in a dot is slower than in the bulk by many orders of magnitude. This slowing down is more significant for electrons than holes, since the electron effective mass is usually 1/10 of the hole mass, and the energy separation is correspondingly larger.

In nanostructures the emission of acoustic phonon should be strongly inhibited as soon as the energy separation between the levels is higher than a few meV, while LO-phonon emission should not be possible unless the intersublevel energy separation matches the LO-phonon energy. Experimental studies have demonstrated that the carrier relaxation in SAQDs occurs on a timescale of tens of picoseconds. This was explained in terms of anharmonicity of the polarons formed by the interaction between electrons and longitudinal optical phonons. The anharmonic terms of the lattice vibrational potential energy then trigger the polaron disintegration into a two-phonon state. The various one- and two-phonon processes that contribute to dissipation are shown in Figure 5.3. Due to energy conservation, each of these processes is possible in a certain limited range of level spacing $\Delta E$. For each decay channel, the relaxation rate strongly decreases with decreasing intersublevel energy. Three factors mainly contribute to this behavior: the reduction of the available density of two-acoustic-phonon states, the decreasing strength of the anharmonic couplings involved and the increasing detuning with respect to the longitudinal optical phonon energy.
Figure 5.3: Calculated relaxation time in InGaAs quantum dots showing the effect of single-phonon and multi-phonon processes on intersubband decay time [57, 58].

5.4 Alloy scattering

In a compound semiconductor consisting of three or more elements, each of the three elements is expected to be aperiodic in the crystal. For example, a ternary alloy such as $A_xB_{1-x}C$ usually consists of $(AC)_x$ and $(BC)_{1-x}$ on average in the ratio $x:(1-x)$. Under this assumption the energy band structure of $A_xB_{1-x}C$ is calculated with the average lattice constant and average pseudopotentials. This assumption is called the virtual-crystal approximation. This approximation is based on the assumption that the atoms A and B are distributed uniformly in the ratio $x:(1-x)$ around the cation C. In real alloys this distribution is not uniform. This non-uniformity results in a local variation of the periodic potential and thus in electron scattering due to the non-uniform potential. This scattering is called alloy scattering.

The alloy potential is given as
The Fourier coefficient $V_{alloy}(q)$ is taken to be the root-mean-square of the shift from the average energy and assumed to be independent of $q$. Defining the Fourier coefficients for atoms A and B by $V_a$ and $V_b$, the average potential of the Fourier coefficients is given by

$$V_0 = V_a x + V_b (1 - x)$$

(5.30)

When the occupation of atom A is changed from $x$ to $x'$, the change in the potential is given by

$$V' - V_0 = (V_a - V_b)(x' - x)$$

(5.31)

Therefore the root-mean-square value of the potential difference is given by

$$\left| \langle V' - V_0 \rangle \right| = \left| V_a - V_b \right| \left[ \frac{x(1-x)}{N_c} \right]^{1/2}$$

(5.32)

where $N_c$ is the number of cations C and corresponds to the unit cell number. The matrix element for the scattering is then

$$\langle k' | H' | k \rangle = \left| V_a - V_b \right| \left[ \frac{x(1-x)}{N_c} \right]^{1/2} \delta_{kzq,k'}$$

(5.33)

and the scattering rate is

$$W_{alloy}(k) = \frac{2\pi}{\hbar} \left( V_a - V_b \right)^2 x(1-x)D(E) \times \Omega$$

(5.34)

where $\Omega$ is the unit cell volume and $D(E)$ is the density of states at $k$ [59].

The $V_a$ and $V_b$ parameters could be taken as the energy band gaps for the two components of the alloy. A more reasonable approach would be to take the vacuum level as the reference energy level and consider the electron affinities of the two materials instead [60]. The significant difference in electron affinities between GaAs (4.07 eV) and InAs (4.90 eV) indicates that alloy
scattering could potentially be significant. It is also important to note that alloy scattering is an elastic scattering process involving no change in the electron energy and only randomizes the electron momentum; hence, it does not contribute directly to energy dissipation.

5.5 Multiple exciton generation

The formation of multiple electron-hole pairs per absorbed photon in photoexcited semiconductors is typically modeled as an impact ionization process. Figure 5.4 depicts the mechanisms for multiple exciton generation in bulk direct and indirect bandgap semiconductors when compared to nanostructures. In this process, an electron or hole with kinetic energy greater than the band gap produces one or more additional electron-hole pairs. The impact ionization rate, which originates from the Coulomb interaction between carriers, is a critical factor affecting the relaxation dynamics of carriers far above the band minima.

In a bulk semiconductor, when the absorbed photon energy is greater than the band gap, the excess energy between the electron and hole is given by

$$\Delta E_e = \left( h \nu - E_g \right) \left[ 1 + m_e/m_{hh} \right]^{-1}$$

$$\Delta E_{hh} = \left( h \nu - E_g \right) - \Delta E_e$$

(5.35)

where $m_e$ and $m_{hh}$ are the effective masses of the electrons and holes, respectively. This excess energy manifests as kinetic energy. These non-equilibrium carriers seek and establish equilibrium over a time period depending upon the various equilibration pathways.

The first step towards achieving equilibrium is for the carriers to interact among themselves by carrier-carrier collisions and intervalley scattering to form separate Boltzmann distributions for electrons and holes. The two Boltzmann distributions can be individually assigned a temperature that reflects the kinetic energy distribution of the respective population. If the excess energies are at least $kT$ above the conduction and valence bands, then both initial carrier
temperatures are above the lattice temperature and the carriers are termed as hot carriers. This first stage of relaxation or equilibration occurs very rapidly (<100 fs) and is referred to as carrier thermalization.

![Figure 5.4: Multiple exciton generation (MEG) process in (a) direct-gap bulk semiconductors, (b) nanostructures and (c) indirect-gap bulk semiconductors. Creation of subbands with very little dispersion relaxes the momentum conservation rules, thus potentially enhancing MEG rate.](image)

The next step in the equilibration process is for the hot carriers to equilibrate to the lattice temperature. The lattice temperature is the ambient temperature and is lower than the initial hot-electron and hot-hole temperatures. Equilibration of hot carriers can be achieved by carrier-phonon interactions whereby the excess kinetic energy of the carriers is transferred to lattice vibrations. Generally, the phonons involved in this process are longitudinal optical (LO) phonons. This may occur either by each carrier undergoing separate interactions or in an Auger process in which the excess energy is transferred from one carrier to another which in turn dissipates its energy to a phonon. Phonon emission results in carrier cooling and heating of the lattice until the carrier and lattice temperatures become equal.
The final stage of equilibration results in complete relaxation of the system. The electrons and holes can recombine, either radiatively or nonradiatively. Another possibility is for the electrons and holes to be spatially separated.

Figure 5.5: Impact ionization rate ($\Gamma_{ii}$) vs. primary electron energy for bulk GaAs (adapted from [61]).

Impact ionization has not, however, contributed meaningfully to improved quantum yield in working solar cells, primarily because the impact ionization efficiency does not reach significant values until photon energies reach the ultraviolet region of the spectrum. The rate of impact ionization must compete with the rate of energy relaxation by electron-phonon scattering. Figure 5.5 shows the impact ionization rate versus electron energy for bulk GaAs. Electron-phonon scattering is a very fast process (sub-ps) in bulk semiconductors, and the rate of impact ionization becomes competitive with phonon scattering rates only when the kinetic energy of the electron is many times the band gap energy ($E_g$). The observed transition between inefficient and efficient impact ionization occurs slowly, rising in Si from 5% at $h\nu \approx 4$ eV to 25% at $h\nu \approx 4.8$ eV. This large blue-shift of the threshold photon energy for impact ionization in semiconductors prevents bulk materials from yielding improved solar conversion efficiencies.
It has been proposed that impact ionization could be greatly enhanced in semiconductor nanostructures compared to bulk because the limitations discussed above for impact ionization in bulk semiconductors are greatly alleviated in QDs. The rate of electron relaxation through electron-phonon interactions can be significantly reduced because of the discrete character of the electron-hole spectra, and the rate of Auger processes, including the inverse Auger process of impact ionization, is greatly enhanced due to carrier confinement and the concomitantly increased electron-hole Coulomb interaction. Additionally, crystal momentum need not be conserved since momentum is not a good quantum number for three-dimensionally confined carriers.

It should be noted that MEG is not the same as impact ionization, since the electrons in higher subbands do not possess “kinetic energy,” which is a prerequisite for impact ionization. MEG is modeled as a Coulomb interaction between carriers using the well-developed theory of Auger recombination, the only difference being that the initial state in the MEG process is at a high energy while in the Auger process, the final state is at a high energy (Figure 5.6).

Figure 5.6: Comparison between multiple exciton generation and Auger recombination.
Using Fermi’s golden rule, the MEG rate per unit area rate is given as

\[ R = \frac{1}{A} \frac{4\pi}{\hbar} \sum P |M|^2 \delta(E_i + E_2 - E_1 - E'_2) \]  

(5.36)

where \( E_i(E_f) \) is the initial (final) energy of state \( i \), \( P \) is the carrier occupation number and \( M \) is the matrix element for the interaction. The above expression includes a factor of two for the two possible initial spin states. The matrix element \( M \) is given by

\[ M = \int \int \Psi^*_i(r_1) \Psi^*_2(r_2) \frac{e^2 \exp(-q_0 r)}{\epsilon|r_1 - r_2|} \Psi_1(r_1) \Psi_2(r_2) d^3r_1 d^3r_2 \]  

(5.37)

where \( \Psi_m(r) \) are the wavefunctions for the states involved in the interaction, and \( q_0 \) is the Debye wavevector that includes free-carrier screening into the interaction Hamiltonian.

To evaluate the matrix element, it is convenient to use the Fourier representation of the potential

\[ \frac{\exp(-q_0 r)}{|r_1 - r_2|} = \frac{1}{8\pi^3} \int \frac{4\pi}{q^2 + q_0^2} \exp(iq[r_1 - r_2]) d^3q \]  

(5.38)

and then

\[ M = \frac{4\pi e^2}{\epsilon} \frac{1}{8\pi^3} \int \frac{1}{q^2 + q_0^2} I_{\gamma 1}(q) I_{\gamma 2}(-q) d^3q \]  

(5.39)

where

\[ I_{mn}(q) = \int \Psi^*_m(r) \Psi_n(r) \exp(iq r) d^3q \]  

(5.40)

The MEG rate thus depends upon the transition probability between states C1 and C2 and the transition probability between states C3 and H1. Energy conservation through the delta function should also be included during the selection of states involved in the MEG process.

From a photovoltaic perspective, hot carriers can be harnessed in two ways. One way produces enhanced output voltage and the other produces enhanced output current. The former
requires that the carriers be extracted before they cool, while the latter requires the hot carriers to produce multiple electron-hole pairs by the multiple exciton generation process. To realize enhanced photovoltage, the rates of hot carrier separation, transport and capture must be faster than the carrier cooling rate. To enhance the photocurrent, the rate of multiple exciton generation should exceed the rate of carrier cooling.

5.6 Nanopore matrix elements

The nanopore wavefunctions calculated by the finite element method can be expressed by

\[ \Psi_{m,n}(x, y, z) = \Phi_m(x, y)Z_n(z) \] (5.41)

The invariance of the z-direction wavefunction for varying lateral dimensions validates the application of the separation of variables method to the nanopore wavefunction. The in-plane function \( \Phi(x, y) \) can be written in the Bloch form as

\[ \Phi_m(x, y) = \exp(ik_x x + ik_y y)u_m(x, y) \] (5.42)

where \( k_x \) and \( k_y \) are the in-plane periodic wavevectors and \( u_m(x, y) \) is the periodic envelope function corresponding to band \( m \). The periodic function can then be expanded in a Fourier series as

\[ u_m(x, y) = \int_{-\pi/L}^{\pi/L} \int_{-\pi/L}^{\pi/L} U(G_x, G_y) \exp(iG_x x)\exp(iG_y y) \] (5.43)

Since the wavefunctions are calculated numerically on a grid, a discrete Fourier transform (DFT) can be performed on the wavefunctions to compute the \( U(G_x, G_y) \) coefficients in the expansion. Numerically, a DFT expands a matrix \( u(M \times M) \) into

\[ u(M \times M) = \sum_{k=0}^{M-1} \sum_{j=0}^{M-1} U(k, l)\exp\left( i \frac{2\pi j}{M} x \right)\exp\left( i \frac{2\pi k}{M} y \right) \] (5.44)
Figure 5.7: (a) Schematic representation of DFT matrix showing the mapping from the position in the matrix to the corresponding exponential function. (b) 2D DFT of $\exp(ik_x x + ik_y y)$ yields a delta function.

A schematic view of the DFT matrix is shown in Figure 5.7 (a). Each point (j,k) in the matrix corresponds to the contribution from the $\exp(i 2\pi (j-1)x/M) \times \exp(i 2\pi (k-1)y/M)$ term in the expansion. This can be verified by taking the DFT of a complex exponential function which yields a delta function in Fourier space (Figure 5.7 (b)).

Due to the symmetry of the complex exponential functions

$$\exp\left(i \frac{2\pi}{M} \left[ \frac{M}{2} + k \right] x \right) = \exp\left(i \left[ \frac{\pi}{M} + \frac{2\pi}{M} k \right] x \right) = \exp\left(i \left[ -\pi + \frac{2\pi}{M} k \right] x \right)$$

and the sum in the Fourier expansion can be changed to

$$u(M \times M) = \sum_{k=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} U(k,l) \exp\left(i \frac{2\pi}{M} j x \right) \exp\left(i \frac{2\pi}{M} k y \right)$$

which is the discretized version of the Fourier transform. The usefulness of this equation is apparent when considering that overlap integrals required for computing the phonon scattering rate are of the form
\[
\int \Psi_1^* \times \Psi_2 \exp(iq.r) d^3r
\]  
(5.47)

which is in essence a Fourier transform of the product \( \Psi_1^* \times \Psi_2 \) and for discretized wavefunctions can be easily computed using Matlab’s in-built two-dimensional fast Fourier routine (FFT2). For optimal performance of the FFT2 routine, the data should be an order of 2 in dimension. Increasing the size of the data results in an exponential increase in computational time. The result of this computation needs to be normalized by the number of points due to the nature of the FFT computation; i.e., an \( N \times N \) FFT must be divided by \( N \).

For the hexagonal nanopore lattice, it becomes necessary to transform the lattice from its original rhombic shape to a square to facilitate computation of the DFT. The corresponding transformation is given by

\[
\begin{align*}
x_n &= x - \frac{y}{\sqrt{3}} \\
y_n &= 2 \times \frac{y}{\sqrt{3}}
\end{align*}
\]  
(5.48)

where \( x_n \) and \( y_n \) are the new co-ordinates in the square unit cell, and \( x \) and \( y \) are the original co-ordinates in the rhombic unit cell. An example of such a transformation is shown in Figure 5.8.

![Figure 5.8: Transformation of ground state of rhombic nanopore at \( \Gamma \) point from rhombic unit cell (a) to square unit cell (b).](image-url)
The z-wavefunctions in the nanopore can be approximated to first order by the states of an infinite quantum well

\[ Z_n(z) = \sqrt{\frac{2}{L}} \sin \left( \frac{n\pi x}{L} \right) \]  

(5.49)

The matrix element for the confinement direction has been calculated previously ([62])

\[
\left| \langle Z_m | e^{iq_z z} | Z_n \rangle \right|^2 = \frac{1}{4} \left[ \frac{\sin(Q + K_m + K_n)}{Q + K_m + K_n} + \frac{\sin(Q + K_m - K_n)}{Q + K_m - K_n} \right] \\
\pm \frac{\sin(Q - K_m + K_n)}{Q - K_m + K_n} \pm \frac{\sin(Q - K_m - K_n)}{Q - K_m - K_n} \right]^2
\]

(5.50)

with \( Q = q_z L/2 \), \( K_m = m\pi/2 \) and \( K_n = n\pi/2 \). The upper signs stand for the band index \( m \) and \( n \) both even or odd, and the lower signs are for one of them even and the other odd. Figure 5.9 shows the overlaps between the wavefunctions \( Z_m Z_n \) in real space (blue curves) and reciprocal space (red curves).

Figure 5.9: Confined wavefunction overlaps (a) and corresponding matrix elements as function of wavevector (b) for first and second confined states. Blue: \( Z_1 Z_1 \), Green: \( Z_1 Z_2 \), Red: \( Z_2 Z_2 \).
Considering that the matrix elements for the x-y directions are discretized, it is logical to toe the same line for the matrix element in the confined z-direction and discretize it. While sampling the z-matrix element, care should be taken to include as much information of the signal in as few samples as possible. A figure of merit in this quest for an optimal sampling strategy is Plancherel’s theorem that states that the total energy in both real and reciprocal spaces should be equal. For the cases plotted in Figure 5.9, taking only four samples leads to results that are 85 percent accurate. Henceforth, the z-direction matrix element is designated by \( M_z \).

The \( N \times N \) FFT matrix for the x-y matrix element must then be combined with \( M_z \) to yield the complete matrix element for the specific transition. This is done by multiplying each element of the \( N \times N \) matrix to every sample in \( M_z \) to yield a \( N^2N_z \) sized array. Note that \( N_z \) is the size of \( M_z \). Simultaneously, a \( q \) matrix is also composed using the relation

\[
q_{j,k,l} = \sqrt{\left(\frac{2\pi j}{M}\right)^2 + \left(\frac{2\pi k}{M}\right)^2 + k_{zl}^2}
\]  

(5.51)

The FFT and \( q \) matrices are then used to compute the total transition probability by performing a weighted sum over all the elements. The weighting factors are a function of the specific scattering mechanism considered. For example the acoustic scattering rate from an initial state \( i \) to a final state \( f \) is calculated as

\[
\Gamma_{i \rightarrow f}^{ac} = \frac{2\pi}{\hbar} \sum_q \frac{\hbar}{\rho V_s} \sqrt{q} \left(N_q + \frac{1}{2} \pm \frac{1}{2}\right) ||\psi_f(q)||^2 \delta\left(E_f - E_i \mp \hbar \omega_q\right)
\]  

(5.52)

The delta function in the above equation accounts for the energy conservation in the transition process. It works to constrict the sum to only those wave vectors for which

\[
|q| = \frac{\Delta E}{\hbar V_s}
\]  

(5.53)
Since the magnitude of the phonon wave vector is fixed by the energy conservation relation, the set of allowed $q$ points trace out a sphere with radius equal to $|q|$. In calculations performed on bulk structures, the sum over $q$ is converted into an integral over all phonon wave vectors. However, in the nanopore the sum over $q$ remains a sum. This complicates handling of the delta function. Mathematically, a delta function can be approximated by either a Gaussian or a Lorentzian as

$$
\delta(t) = \lim_{t \to 0} \frac{1}{\sqrt{2\pi \sigma}} e^{-t^2/2\sigma^2}
$$

with the $\sigma$ and $\epsilon$ parameters being a measure of the uncertainty in the phonon energy. In general, homogeneous broadening is generally modeled as a Lorentzian lineshape, which makes it a good choice for this calculation. As a rule of thumb, the broadening should equal the largest energy difference between adjacent $k$ states in any computed subband.

It should be noted that every nanopore wavefunction possesses two sets of wavevectors, i.e. the slowly varying wavevector $(k_x, k_y)$ and the cell periodic wavevectors $(2\pi j/L)$, which should both be taken into account in the momentum conservation rules. However, $k_x$ and $k_y$ can be neglected relative to $2\pi j/L$ as they are a lot longer in wavelength. As a result, only the cell periodic portion of the in-plane wavefunction is taken into account for the transition rate calculation. It should also be noted that the in-plane wavevectors are always less than the z-wavevector due to the stronger confinement in the z-direction.

The transition probability calculation discussed previously assumes a two level system with a single initial and final state. However, the scattering probability in an actual structure is scaled by the number of final states. This leads to the computation of the density of states in a nanopore
lattice. Assuming an $N \times N$ unit cell structure, the only allowed wave vectors are given total number of allowed wave vectors equal to $N^2$. This follows from the Born-von-Karman boundary conditions. The area of the sample is

$$\text{Area} = N^2 L_x L_y$$

(5.55)

Hence,

$$\text{Number of States per unit area} = N^2 / N \times N \times L_x L_y = 1 / L_x L_y$$

(5.56)

This number differs from the classical definition of density of states as this is defined only as a per unit area quantity and not in terms of per unit area per unit energy. This discrepancy makes it difficult to perform a direct comparison with the theoretical density of states of conventional structures (bulk or confined). In the asymptotic limit of small pores, the nanopore is expected to behave as a quantum well.

5.7 Phonon scattering rates

The final scattering rate between two states depends upon (a) satisfaction of energy conservation rules, (b) the matrix element between initial and final states and (c) the number of final states available. Due to the discretization scheme employed for the nanopore, the number of states for each $k$ point are constant. This does not mean that the density of states at all $k$-points is the same, as was explained at the end of the previous section. Thus the main factor influencing transition probabilities is the availability of phonons with sufficient energy and with the correct momentum to drive the transition process.
Phenomenologically, the nanopore can be compared to a quantum dot. Figure 5.10 shows the variation in the sum of square of the matrix element of a quantum dot with dot size. This result is contrary to the known fact that the oscillator strength in quantum dots increases as the confinement increases [63, 64]. In the case of interband transitions, this can be explained by the increasing mismatch between electron and hole wavefunctions with decreasing dot size. The reduction in the overlap is due to the more sensitive size dependence of the electron wavefunction compared to the hole wavefunction due to its lower effective mass. As the dot size is reduced, the electron wavefunction penetrates deeper into the barrier than the hole reducing the overlap.

Another competing effect is the increase in the energy gap between levels as the confinement is increased. This energy gap affects the transition rates through the energy conservation requirement. Thus the actual dependence between the transition rate and geometry is a tradeoff between the strength of the wavefunction overlap and the energy separation between the two states.
Figure 5.11: Energy dispersion for 20 nm pitch hexagonal nanopore with 10 nm pore diameter.

The energy dispersion for the first six subbands along the principal directions for a hexagonal nanopore lattice with a pitch of 20 nm and a pore diameter of 10 nm is shown in Figure 5.11. Due to the strong quantization, the separation between the energy levels is much greater than the acoustic phonon energy (~ 1 meV) and hence acoustic scattering is strongly suppressed between subbands. However, within a subband, electrons quickly relax from higher energies to the bottom of the band via acoustic phonon emission. Polar optical phonon emission is the dominant process for energy relaxation. Assuming an optical phonon energy of 36 meV, only certain transitions become allowed and hence, compared to a quantum well, the overall relaxation tends to be considerably slower. For nanopore patterns with larger pitch, the subband separation becomes of the order of a few meV and optical phonon emission becomes altogether prohibited due to energy conservation. The existence of a bandgap as well as the reduced overlap due to the in-plane modification of the wavefunction reduces the acoustic phonon scattering rate to about 10 ns, which is two orders less than that for a quantum well. This is in agreement with experimental observations of increased emission from excited states in the nanopore relative to a quantum well.
CHAPTER 6 – THE FUTURE

Patterned zero-dimensional nanostructures – quantum dots and inverse quantum dots (nanopore) – are promising tools for designing the next generation of optical devices. The nanopore structure is extremely flexible since it presents an opportunity to design devices covering the continuum between fully three-dimensionally confined quantum dots and one-dimensionally confined quantum wells. Patterned quantum dots permit precise control of size and composition and can be easily integrated into optical devices such as single photon emitters that contain only a few dots.

In this work, nanopore lattices have been fabricated with fabrication of ±1 nm which is a significant improvement over the uniformity obtained by self-organized growth. Photoluminescence analysis suggest a reduced intersubband scattering rate in the nanopore relative to a reference quantum well. This is understandable given that the nanopore perturbs the in-plane wavefunction, which in turn results in a reduced wavefunction overlap between states. This reduced carrier relaxation effect has the potential to enhance the probability of multiple exciton generation in the nanopore. The current body of work was restricted to the GaAs-InGaAs system which possesses a type-1 band alignment. Implementing the nanopore on type-2 band alignment such as GaSb-InAs could potentially lead to novel superlattices and applications.

A method for site-localized growth of InAs quantum dots within nano-oxide templates by selective area epitaxy was developed. A novel hybrid etch technique for maintaining critical dimensions without creating excessive damage was used. Either single or multiple clusters of dots are formed depending upon the mask dimensions and amount of deposited material. The ability to create vertically aligned stacked quantum dots was demonstrated. The fabrication method was also applied to grow InAs dots in di-block copolymer patterned substrates with
excellent uniformity over large areas. Future work could leverage the techniques developed in this thesis to create quantum dot molecules (Figure 6.1) which are essential building blocks for quantum computers. Also the ability to tailor the dot size by controlling mask geometry could enable design of broadband quantum dot based emitters as shown in Figure 6.2.

Figure 6.1: Quantum dot molecules by selective area epitaxy. (a) Schematic. (b) SEM micrograph of fabricated sample.

Figure 6.2: Variable size quantum dot arrays for superluminescent emitters. (a) Sample emission spectrum. (b) SEM micrograph of fabricated sample.
REFERENCES


122


