Substrate and Channel Engineering for Improving Performance of Strained-SiGe MOSFETs

by

Saurabh Gupta

Bachelor of Technology in Materials Science and Metallurgical Engineering
Indian Institute of Technology, Kanpur, 2002

Submitted to the Department of Materials Science and Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Electronic, Photonic, and Magnetic Materials

at the
Massachusetts Institute of Technology

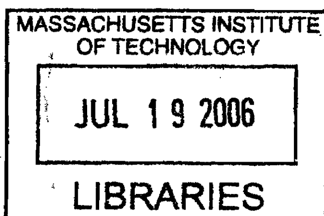
June 2006

© 2006 Massachusetts Institute of Technology
All rights reserved.

Signature of Author: ______________________________
Department of Materials Science and Engineering
May 10, 2006

Certified by: ______________________________
Eugene A. Fitzgerald
Professor of Materials Science and Engineering
Thesis Supervisor

Accepted by: ______________________________
Samuel M. Allen
POSCO Professor of Physical Metallurgy
Chair, Departmental Committee on Graduate Students


MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUL 19 2006
LIBRARIES

ARCHIVES

Substrate and channel engineering for improving performance of strained-SiGe MOSFETs

by

Saurabh Gupta

Submitted to the Department of Materials Science and Engineering
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electronic, Photonic, and Magnetic Materials

ABSTRACT

With VLSI technology moving closer towards fundamental physical limits, a way to further improve the transistor drive current for superior circuit performance is enhancing the average velocity of carriers in the channel. In order to enable technology boosters like strained channels on the commercially developed Si platform, a low threading dislocation density SiGe platform is required. A relaxed $Si_{1-x}Ge_x$ graded buffer creates a larger lattice constant on a Si substrate while providing low threading dislocation densities (TDD) on the order of 10^5 cm^{-2}. For many III-V devices on Si, such as high-efficiency multi-junction solar cells, the SiGe graded buffer approach is hindered somewhat by the fact that relatively thick buffers must be grown to relieve the strain effectively.

A way to reduce the thickness of SiGe graded buffer is to use larger mismatch (10% Ge difference) at each step. Growing a larger mismatch layer at high temperature leads to a higher threading dislocation density (TDD). In this work, the rate of plastic strain relaxation in a mismatched layer and misfit dislocation distribution in the layer are identified to be important parameters which affect the TDD in a mismatched SiGe layer. Reducing the TDD in a mismatched SiGe layer on Si requires slow plastic strain relaxation which can be achieved by growing layers at slower growth rates. Using lower growth temperatures to reduce the growth rates, however, leads to a poorly relaxed layer. A Time-Temperature-Threading Dislocation Density diagram has been developed in this work and explains that a high temperature anneal for relaxing the layer leads to undesirable TDD escalation. A gradual increase in the anneal temperature results in slow plastic relaxation leading to a reduced TDD. The remaining elastic strain can be completely removed by using strain relieving layer on top. Repetition of this minimized 10% mismatched layer achieves high Ge lattice constant in less than half the thickness of a conventional graded buffer.

Using the conventional SiGe relaxed buffers, strained layer heterostructures which provide enhanced hole and electron mobilities can be fabricated. A dual channel

heterostructure consisting of strained-Si (ε-Si) / strained-$Si_{1-y}Ge_y$ (ε-$Si_{1-y}Ge_y$) on a relaxed $Si_{1-x}Ge_x$ buffer ($y>x$), provides a platform for fabricating MOSFETs with high hole mobility (μ_{eff}). The impressive hole μ_{eff} enhancements obtained on dual channel heterostructures depend greatly on the Ge concentration and the strain[1] in the ε-$Si_{1-y}Ge_y$ layer. Ge out-diffusion during high temperature processing steps from the ε-$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ buffer reduces the Ge concentration and the strain in the layer leading to reduced hole μ_{eff} in these heterostructures[2].

For estimating the thermal budgets of the strained layer heterostructures, a diffusion coefficient for Ge in crystalline SiGe has been formulated. We have also implemented a novel, more accurate finite difference scheme for estimating thermal budgets of strained SiGe heterostructure with steeply varying concentrations. It is found that due to the very small diffusivity of Ge in Si, the out-diffusion of Ge occurs primarily from the ε-$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ buffer with much less diffusion into the ε-Si[3]. This motivated the inclusion of a ε-Si layer between the ε-$Si_{1-y}Ge_y$ and relaxed $Si_{1-x}Ge_x$ buffer for reducing the Ge out-flux. The resulting heterostructure is referred to as tri-layer heterostructure[4].

Based on our numerical investigations and annealing experiments regimes in temperature and concentration where the tri-layer is expected to be thermally more stable as compared to corresponding dual channels have been outlined. Hole μ_{eff} enhancements are retained to a much higher extent in a tri-layer heterostructure after high temperature processing as compared to a dual channel heterostructure. Improved hole mobilities in this heterostructure are also observed over similar dual channel heterostructures. The improved thermal stability and hole μ_{eff} of a tri-layer heterostructure makes it an ideal platform for fabricating high μ_{eff} MOSFETs that can be processed over higher temperatures without significant losses in hole μ_{eff}.

Thesis Supervisor: Eugene A. Fitzgerald
Title: Merton C. Flemings-SMA Professor of Materials Science and Engineering

TABLE OF CONTENTS

TABLE OF CONTENTS **6**

LIST OF FIGURES **9**

LIST OF TABLES **17**

ACKNOWLEDGEMENTS **18**

Chapter 1. Introduction **20**

1.1. Motivation *21*

1.2. Scope of the thesis *24*

Chapter 2. Ultra-thin SiGe buffers **26**

2.1. The Relaxed SiGe Integration Platform *27*

2.2. Lattice Mismatched epitaxy *29*
2.2.1 Introduction 29
2.2.2 Lattice mismatch and critical thickness 31

2.3. Experimental Techniques *36*
2.3.1 SiGe Growth 36
2.3.2 Characterization Techniques 39
2.3.2.1 Transmission Electron Microscopy 39
2.3.2.2 Etch-Pit Density Measurements 41
2.3.2.3 Atomic Force Microscopy 43
2.3.2.4 Triple-Axis X-ray Diffraction 44
2.3.2.5 Nomarski contrast microscopy 45

2.4. Threading dislocation densities in a single mismatched layer *46*
2.4.1 Variables affecting TDD in mismatched SiGe layer: Strain rate 47
2.4.2 Surface roughness evolution in a mismatched SiGe layer 51
2.4.3 Variables affecting TDD in mismatched SiGe layer: misfit dislocation spacing 53

2.5. Compositionally Graded Buffers *60*

2.6. Controlling strain rates for reducing the dislocation densities by reducing growth rates *62*
2.6.1 Determining growth rates for singly mismatched SiGe layers for simulating the strain rates of graded buffer 63
2.6.2 Achieving low growth rates for reducing TDD 65

2.7. Pulsed layer growth *66*

2.8. Low temperature growth *70*
2.8.1 Metastability in layers grown at lower temperatures. 72
2.8.1.1 Relaxation through annealing 73

2.9. Time-Temperature-Threading Dislocation Density Diagram (TTTDD Diagram) *77*

2.10. Achieving thin 10% completely relaxed layer *78*
2.10.1 Strategy for achieving thin graded buffer 78
2.10.2 Metastability in 10% layer after annealing 80
2.10.3 Strain relieving layer 82
2.10.4 Relaxation of the metastable layer by using a 12% strain relieving layer 83
2.10.5 Phenomenon of relaxation of the 10% layer 84

2.10.6 Designing the strain relieving layer 87

2.11. Ultra-Thin Buffer *88*
2.11.1 Repetition of the 10% for getting TRUT 88
2.11.2 Characteristics of a 17% TRUT buffer and comparison with conventional graded buffer 89

2.12. Summary *92*

Chapter 3. SiGe Based MOSFETs 94

3.1. Motivation for SiGe-based Devices *95*
3.1.1 New Materials Technologies 96

3.2. Short-flow MOSFET process *101*
3.2.1 Short Flow MOSFET Fabrication Sequence 104
3.2.2 Transistor layout and mobility extraction 107

3.3. Summary *108*

Chapter 4. Diffusion coefficient formulation for estimating diffusion of Ge in SiGe single crystals 109

4.1. Interdiffusion observations in SiGe heterostructures *111*

4.2. Diffusion Model formulation *117*
4.2.1 Diffusion Mechanisms 117
4.2.2 Formulating Ge diffusion coefficient in SiGe single crystal structures 120

4.3. Summary *127*

Chapter 5. Tri-Layer Heterostructure 128

5.1. Dual channel Heterostructures *130*
5.1.1 Motivation for Dual channel 130
5.1.2 Effect of Ge concentration of the ε-$Si_{1-y}Ge_y$ layer on the hole mobility 134
5.1.3 Effect of Ge concentration in the relaxed $Si_{1-x}Ge_x$ 135
5.1.4 Effect of thickness of the ε-$Si_{1-y}Ge_y$ layer 135

5.2. Thermal stability of dual channel heterostructures *138*

5.3. Tri-layer heterostructure *146*
5.3.1 Growth of tri-layer heterostructures 147

5.4. Thermal stability of Tri-Layer heterostructures *148*

5.5. Improved hole mobilities in tri-layer heterostructures *152*

5.6. Summary *155*

Chapter 6. Summary and Future Work 157

6.1. Summary *158*

6.2. Suggestions for Future work *165*
6.2.1 Extending the TRUT buffer concept 165
6.2.1.1 Optimizing annealing times for TRUT buffers 165
6.2.1.2 Higher Ge content buffers 167
6.2.2 Improving the Relaxation-Nucleation model 170
6.2.3 Advanced devices on tri-layer heterostructure 170

APPENDIX 172

Appendix 1. Script for Critical Thickness *172*

Appendix 2. Elastic strain remaining with thickness *175*

Appendix 3. *For Dislocation Density with Ge concentration* *177*
Appendix 4. *Growth rates for Pulsed layer* *180*
Appendix 5. *TDD and strain evolution with time during an isothermal anneal* *182*
Appendix 6. *Smoothing XRD data for strain determination* *184*
Appendix 7. *Diffusion in strained SiGe heterostructures* *187*

BIBLIOGRAPHY **195**

LIST OF FIGURES

Figure 2-1:The Si-Ge Binary Phase Diagram, adapted from Gandhi[24] Si and Ge are completely miscible, enabling SiGe films of arbitrary composition to be grown without phase segregation. 29

Figure 2-2: Very high TDD (>10^9 cm^{-2}) in a Ge deposited on Si substrate. Image courtesy of Tom Langdo 30

Figure 2-3 Schematic of pseudomorphic growth and the onset of dislocation formation in lattice-mismatched heteroepitaxy. (a) Initially, the in-plane film lattice constant is constrained to that of the substrate, as shown on the left. (b) Above a critical thickness, dislocation formation becomes energetically favorable and the film relaxes, as shown on the right. Image courtesy of Arthur Pitera. 32

Figure 2-4: Critical thickness (h_c) of $Si_{1-x}Ge_x$ alloys on Si (001) versus Ge content (x). Critical thickness drops rapidly with increasing lattice mismatch. 34

Figure 2-5: Strain remaining in the film plotted with thickness 35

Figure 2-6: Misfit dislocations spacing is not uniform but has a certain distribution. 36

Figure 2-7: Schematic of the Fitzgerald Group UHVCVD system 38

Figure 2-8: Plan view TEM showing a dislocation density of about $1.5\times10^7 cm^{-2}$. 40

Figure 2-9: Plan view TEM used for calculating the relaxation in the layer by measuring the spacing between the misfit dislocations 41

Figure 2-10: DIC micrograph of SiGe layer after selective etching, revealing threading dislocations. Image Courtesy of Chris Leitz 42

Figure 2-11: Example of an AFM micrograph obtained from a 10% SiGe layer grown on Si. The scan area is 10μm×10μm 43

Figure 2-12 Example of an (004) reciprocal space map obtained from a compositionally graded SiGe buffer graded to a final Ge composition of 50%. Image courtesy of C.W. Leitz 44

Figure 2-13: Figures explaining XRD data smoothing using Savitzky-Golay filter. Figure (a) plots the as acquired data while Figure (b) shows the data after subjecting it to the filter. The random noise in the data is removed and a better idea of peak position in the omega axis is achieved. In addition a peak finding subroutine is implemented to get the exact peak 45

Figure 2-14: TDD evolution with Ge concentration in a mismatched SiGe layer grown at 900°C...... 47

Figure 2-15: Experimental Matthews-Blakeslee curve for SiGe mismatched layer grown at 900°C.. 48

Figure 2-16: TDD evolution with Ge concentration in a mismatched SiGe layer grown at 900°C compared with the expected TDD distribution based on strain rate effects. 51

Figure 2-17: RMS roughness evolution with Ge concentration in a mismatched SiGe layer grown at 900°C. The scan area of all the samples is 10μm×10μm. 52

Figure 2-18: AFM images of the layer depicting the roughness increase as the Ge concentration increases in the mismatched layer. The scan area of all the samples is 10μm×10μm. 52

Figure 2-19: Illustration of a dislocation blocking mechanism in compositionally graded buffers. Local undulations of the surface coincide with the strain field generated by underlying misfit dislocations, posing an obstacle to dislocation glide. Figure adapted from Samavedam. [30] 53

Figure 2-20: Plan-view of cellular structure emerging from dislocation pile-up formation. Dislocation pile-ups form boundaries that restrict gliding threading dislocation motion[34]. 55

Figure 2-21: Misfit dislocations are spaced not uniformly but follow a log-normal distribution. The fraction of misfit dislocations that are very close to each other prevent the dislocations from gliding necessitating further nucleation events. Dotted line represents the log-normal distribution while histogram is experimentally measured spacing... 56

Figure 2-22: Plan-view optical micrograph of a selectively etched SiGe surface revealing widely spaced dislocations in the field and a large threading dislocation pile-up. Dislocations trapped in pile-ups cannot relieve mismatch strain, forcing nucleation of additional dislocations in the field and causing an escalation in overall dislocation density. Image Courtesy of Chris Leitz.. 57

Figure 2-23: The calculated dislocation density matches the observed dislocation density implying that strain rate and the surface roughness are two important factors affecting the dislocation densities in a mismatched layer. 59

Figure 2-24: (a) Schematic of Compositionally graded buffer. (b) Cross-sectional transmission electron micrograph of a compositionally graded SiGe region. Image courtesy of Thomas Langdo. 61

Figure 2-25: AFM images of a 20% graded buffer (left) compared with a 20% singly mismatched layer. The RMS roughness of the graded buffer is 3.8nm while the 20% singly mismatched layer has 60nm RMS roughness. 62

Figure 2-26: Growth rates (in angstroms per second) plotted with Ge composition in order to maintain a strain rate equivalent to a low TDD 2% layer. It shows that as the mismatch increases the layer needs to be grown at increasingly lower growth rates. 65

Figure 2-27: Using Pulsed layers Ge poor films are obtained as compared to layers grown without the pulsed 68

Figure 2-28: The roughness of the pulsed layers becomes very bad after about 9% Ge concentration. Such a high roughness will lead to a high TDD. The AFM images of these pulsed layer specimen are shown in Figure 2-29. 68

Figure 2-29: Roughness evolution in pulsed layers as the Ge content is increased. The pictures 1,2,3,4,5 corresponds to a Ge concentration of 3%, 4%, 7%, 9%, 20% respectively. The scan area of all the AFM is 10μm×10μm. 69

Figure 2-30: Pulsed layer is very metastable after growth and does not relax efficiently. 70

Figure 2-31: Growth rates vary over by many orders of magnitude with temperature changes. Such a range is required for controlling the strain rate in the layer 71

Figure 2-32: Metastability occurs in layers grown at lower temperatures[40]. 72

Figure 2-33: Annealing the layer reduces the strain remaining in the layer. The dynamics of strain relaxation is much faster at higher temperatures as compared to lower temperatures[41]. 73

Figure 2-34: Two different anneals are engineered in order to look into the strain relaxation behavior and the dislocation density trends in a 10% SiGe layer grown at 650°C based on different anneal procedures. 'A' is combination of step anneals from 650°C to 900°C over a long time which 'B' is a step anneal to 900°C. 75

Figure 2-35: Strain relaxation and TDD evolution with time in a 10% SiGe layer subjected to a gradual step anneal described in Figure 2-34. The dislocation density after the gradual anneal is very low. 76

Figure 2-36: Strain relaxation and TDD evolution with time in a 10% SiGe layer subjected to a steep step anneal described in Figure 2-34. The dislocation density after the steep anneal is higher than in the gradually annealed sample (Figure 2-35). 76

Figure 2-37: Time-Temperature-Threading Dislocation Density (TTTDD) diagram for a metastable 10% layer which is annealed isothermally. The diagram is helpful in determining the annealing path for a low TDD state in a layer. 78

Figure 2-38: Step-anneal profile used for relaxing the 10%, 1500Å layer grown at 650°C. 81

Figure 2-39: The 10% layer grown at 650°C and annealed using a step anneal profile from 650-900°C retains a lot of strain. The misfit spacing expected is about 40nm while the average spacing that we get in the layer is 180nm. 81

Figure 2-40: Relaxation of the 10% layer with thickness of 12% layer. If the 12% layer is 1800Å thick the 10% layer has already relaxed to its equilibrium lattice constant of 2.5%. 84

Figure 2-41: XTEM of the stack of layer showing the 10% layer and the 12% layer which completely relaxes the underlying 10% layer. 85

Figure 2-42: The above figure explains the mechanism of relaxation of the 10% layer by the use of a strain relieving layer. (A) the 10% layer after deposition and anneal is very metastable, (B) As the 12% layer is growing it is providing driving force to the underlying 10% layer. When the thickness of the 12% layer is below its critical thickness, it aids in the relaxation of 10% layer. .. (C) When the thickness of the 12% layer crosses its critical thickness the 10% layer is already very relaxed leading to a low mismatch between the 12% and the 10% layer. In addition the 10% layer seeds 12% with threading dislocations leading to a low TDD in the final structure. 86

Figure 2-43: A 17% TRUT buffer. The first layer $Si_{0.90}Ge_{0.10}$ is grown at 650°C and then annealed using step anneal from 650-900°C. The $Si_{0.88}Ge_{0.12}$ layer is then grown at 900°C to relax the 10% layer. Similarly, a $Si_{0.83}Ge_{0.17}$ layer is grown at 650°C and then annealed using step anneal from 650-900°C. The $Si_{0.81}Ge_{0.19}$ layer is then grown at 900°C to relax the 17% layer. 89

Figure 2-44: The TRUT buffer compared to the conventional graded buffer. Both of these pictures are at same scale. The TRUT buffer developed in this work can be more than 2 times thinner than the conventional graded buffer 91

Figure 3-1: Universal electron and hole mobility relations for bulk Si MOSFETs. At low fields, carrier mobility is limited mainly by Coulomb scattering. At moderate vertical fields, carrier mobility is limited by phonon scattering. At high vertical fields, surface roughness scattering limits carrier mobility. Image courtesy of Minjoo L. Lee 98

Figure 3-2: The various directions where uniaxial stress can be applied on a MOSFET for change in transport properties[48]. 100

Figure 3-3: Bandgap variation in $Si_{1-x}Ge_x$ alloys. The conduction band minima crosses over from X to L at $x \sim 0.8$[50]. The strained SiGe band gap corresponds to SiGe layer grown on a Si substrate. 101

Figure 3-4: Short-flow MOSFET process flow 102

Figure 3-5: SEM picture showing the oxide undercut of 102.4 nm. Image courtesy: Jongwan Jung 103

Figure 3-6: SEM picture showing a short between the Source and Gate. 104

Figure 3-7: Planar view of MOSFETs measured in this thesis 107

Figure 4-1: Meyer-Neldel equation in SiGe system. The E_A were fitted linearly and the D_0 were obtained from the Meyer-Neldel equation. The solid line shows the agreement of the measured values of pre-exponential factors with the derived values is good. 121

Figure 4-2: The formulated diffusion coefficient in this work matches well with the literature values[58]. 122

Figure 4-3: Error in the present finite difference scheme as compared to the standard finite difference scheme[73]. As the spacing of the nodes increases the error increases 126

Figure 4-4: The profile for simulation used is shown below. Initial interfaces were smoothened with a hyperbolic tangent function. Value of Q' = 23 gives a good fit with the experimental profiles. 127

Figure 5-1: Band alignment of strained Si grown on relaxed $Si_{1-x}Ge_x$ 131

Figure 5-2:Schematic band alignment of a dual channel heterostructure. Higher *y* in the compressive layer leads to a deeper well for holes. In *p*-MOSFETs based on these structures, large gate overdrives force holes towards the SiO_2/Si interface. 132

Figure 5-3:Hole mobility enhancement over Cz-Si at E_{eff} = 0.6 MV/cm for single channel and dual channel devices grown on the same relaxed buffer compositions. The buried layer Ge content in all of the dual channel devices is 30% higher than the relaxed buffer. Even at high E_{eff}, dual channel devices demonstrate far larger enhancements than single channel devices. Image Courtesy: Minjoo L. Lee 133

Figure 5-4:Effective hole mobilities of dual channel heterostructure *p*-MOSFETs[85]. The difference in Ge content between the virtual substrate and compressive layer, *y-x*, is held constant at 0.3. 135

Figure 5-5: Hole effective mobility vs N_{inv}. The use of a thicker ε-Ge layer gave rise to a 20% increase in μ_{eff} for the samples grown on $Si_{0.3}Ge_{0.7}$ and an 80% increase for the samples grown on $Si_{0.5}Ge_{0.5}$ [16] 138

Figure 5-6: The Si-Ge Binary Phase Diagram, adapted from Gandhi.[24] Si and Ge are completely miscible, enabling SiGe films of arbitrary composition to be grown without phase segregation. 139

Figure 5-7: Schematic layer sequence and band diagram for a dual channel heterostructure 139

Figure 5-8: Ge concentration in the dual channel heterostructure is reduced after annealing at a high temperature 140

Figure 5-9: Hole mobilities reduce after annealing at 800°C for 30 minutes. For a 70 on 40 dual channels the mobility drops from 2.9 times to 1.56X, while it drops from 3.8X to 2.2X for an 80 on 50 channel[2]. 141

Figure 5-10: Peak decrease in a dual channel heterostructure with increasing concentration of Ge in the ε-$Si_{1-y}Ge_y$ layer. The Ge concentration in the relaxed buffer is maintained at 30% below the Ge concentration in the ε-$Si_{1-y}Ge_y$ layer 142

Figure 5-11: Peak decrease with an increasing temperature of processing. The time of processing is one hour. The peak decrease of different dual channels have been shown. 143

Figure 5-12: Strain reduction in a Ge on 70 dual channel heterostructure[17]. As shown in this work almost all the reduction is through the Ge outdiffusion from the compressively strained layers. 144

Figure 5-13: SIMS profile of a Ge buried channel heterostructure. The outdiffusion of Ge in the strained Si cap is very less and majority of diffusion occurs into the relaxed SiGe buffer. 145

Figure 5-14: XTEM of a tri-layer heterostructure. Corresponding band diagram shows a quantum well for holes in the ε–$Si_{1-y}Ge_y$ layer and a high valence band offset ΔE_v due to underlying ε-Si layer preventing the holes from tunneling into $Si_{0.8}Ge_{0.2}$. 147

Figure 5-15: The peak decrease of the dual channel and the tri-layer with varying concentration of Ge in the buried layer and the temperature is shown. The tri-layer has much improved thermal budget over the dual channels when the concentration in the strained-$Si_{1-y}Ge_y$ layer is higher. The peak decreases are in reasonable agreements with the experimental values. 149

Figure 5-16: Comparison of SIMS profile of the ε-Si/ε-$Si_{0.5}Ge_{0.5}$ heterostructure on relaxed $Si_{0.80}Ge_{0.20}$ with ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si on relaxed $Si_{0.8}Ge_{0.2}$ after annealing for 30 minutes at 850°C. The peak decreases to 39% from 46% in case of dual channel and to 41% from 46% in case of tri-layer. 150

Figure 5-17: Comparison of SIMS profile of ε-Si/ε-Ge heterostructure on relaxed $Si_{0.5}Ge_{0.5}$ with ε-Si/ε-Ge/ε-Si on relaxed $Si_{0.5}Ge_{0.5}$ after annealing for 60 minutes at 650°C. The peak decreases by 10% in case of the dual channel and 3% in case of tri-layer. 151

Figure 5-18: Higher hole mobilities are retained in a ε-Si/ε-$Si_{0.3}Ge_{0.7}$/ε-Si on relaxed $Si_{0.6}Ge_{0.4}$ heterostructure as compared to a ε-Si/ε-$Si_{0.3}Ge_{0.7}$ on relaxed $Si_{0.6}Ge_{0.4}$ heterostructure after processing it at 850°C, showing the improved thermal stability of the tri-layer. 152

Figure 5-19: Hole effective mobility vs N_{inv} obtained from a PMOS fabricated on a ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si structure grown on relaxed $Si_{0.8}Ge_{0.2}$. The mobility enhancement for the tri-layer is 20% higher than corresponding ε-Si/ε-$Si_{0.5}Ge_{0.5}$ dual channel heterostructure grown on relaxed $Si_{0.8}Ge_{0.2}$ at $N_{inv} = 1\times10^{13}/cm^2$. 153

Figure 5-20: Mobility enhancement in a PMOS fabricated on a tri-layer heterostructure (50 on 20), and electron mobility enhancements from NMOS fabricated on a 20% relaxed SiGe buffer.

Since, the NMOS behavior is not expected to be very different from a single channel heterostructure, a possibility of symmetric hole and electron enhancement exist on a ε-Si|ε-$Si_{0.5}Ge_{0.5}$|ε-Si on relaxed $Si_{0.8}Ge_{0.2}$ 155

Figure 6-1: Step-anneal profile used for relaxing the 10%, 1500Å layer grown at 650°C. While longer times at lower temperatures are justified, longer times at higher temperatures are not necessary since the dislocation dynamics is very fast at these temperatures. 166

Figure 0-1: The above figure explains the usefulness of Savitzky-Golay smoothing filter for X-ray data[93]. While the intensity of narrow peaks remains, the high frequency random noise has been filtered out. Compare this method to the normal data averaging method for narrower peaks. 185

Figure 0-2: Figures explaining XRD data smoothing using Savitzky-Golay filter. Figure (a) plots the as acquired data while Figure (b) shows the data after subjecting it to the filter. The random noise in the data is removed and a better idea of peak position in the omega axis is achieved. Peak is then obtained using a peak finding routine. 186

LIST OF TABLES

Table 3-1: Electron and hole mobilities in bulk Si and Ge[46] 97
Table 3-2: Desired uniaxial stress state for improved carrier mobilities[48]. 99
Table 3-3: MOSFET Fabrication Sequence 105
Table 3-4: Relevant dimensions of MOSFETs characterized in this work 107

ACKNOWLEDGEMENTS

As I complete this thesis, I feel a greatest sense of thankfulness to Gene. Not only did he give me the freedom to explore whatever path I felt interesting, he also supported me enthusiastically in that. I always feel that I have been lucky to find him as an advisor and mentor. I attribute my advances, both scientific and professional, to him and believe that his training and education would be very helpful in my future life.

I am thankful to Prof. Sam Allen for his helpful discussion about the diffusion model and his very valuable inputs on improving my thesis. I would like to thank Prof. Carl Thompson for his critical suggestions on improving and enriching my dissertation and defense.

I would like to thank the Fitzgerald group for their support and help during the highs and lows of research and life. David Isaacson was with me, be it the innumerable nachos (or shall I say towers?) escapades at CBC, or during our research/work together during the last four years. I must say that his optimism is infectious and he infuses an enthusiasm in the co-workers. During my final days here, I really appreciate his help with my research. Arthur Pitera and Larry Lee have been great collaborators and friends. Larry was the person who introduced me to the research and the life in Fitzgerald group. He was almost a co-advisor to me during my first year at the grad school. I am very thankful to him as he was always very willing to help me in anything ranging from research to making presentations and writing papers. I wish him good luck for his future endeavors. Arthur Pitera has been a great friend and I have enjoyed too many outdoor adventures with him including Crystal Method concert, tasting delicacies at north end, discussing research or the arm-wrestling at B-side, gym trips, or the JD party. He is awesomely gifted and I admire his ingenious approaches to spot and solve problems. Carl was the first new graduate student whom I met. For a long time Carl's puzzles of the day was the highlight of the day. I am amazed by his comprehensive understanding of theories and his creative approaches to solve problems. I can never forget the marathon effort that we put while working on the modulator project. I wish him best of luck for the remaining graduate school and for any future auditions of "Beauty and the Geek". Bai Yu has been a great co-worker. He may be a recent graduate student in the lab, however, his attitude in sharing responsibilities and helping people has certainly won him many friends. I especially like his positive view towards life and research. With his hard working nature and intelligence, I am sure that he will make a great researcher and a succesful entrepreneur. Gianni Taraschi is a cool guy. It was his initiatives that got me started on the diffusion project. It was great to talk with him on the future of CMOS or the future of finance or "Gold". I really enjoyed his humor and wish him all the best in his future endeavors as a patent lawyer. Mayank, the entrepreneur, is one of the most helpful people in the group. I am a great admirer of his "no-nonsense" approach which has helped me numerous times for refining my presentations. I would like to thank him for his special effort in getting me interviews during my job search days. I cannot emphasize enough the help that I have received from Mike in the many aspects of my graduate life, be it getting the class notes of many classes that we took together or the help on computer administration stuff in the first 2 years, or for providing very constructive suggestions on my presentations and would like to thank him for that. Nate is never short of enthusiasm, patience, cheerfulness and positive attitude. I greatly admire his multi-talented personality which helped him to almost single-handedly manage an entire lab and pursue his hobbies alongside. His patience and help on XRD was very important for my work and I sincerely thank him for that.

It was a great pleasure working, chatting and hanging out with Nava. I will remember Lisa Mcgill for her discussions about Karate and the liquid nitrogen ice-cream which she made during many of the group meetings. I admire Ken for his in-depth knowledge in so many different fields especially in Physics and optics. I am sure that he will do a great job in the hard project that he has undertaken. I wish all the best to Kamesh for his future endeavors.

Anabela Afonso has been very helpful to me, whether it be in helping with the conference stuff, or helping me select the chocolates for my sister, or giving me extensions in UHVCVD accounting, or for securing an extra piece of the great cookies that she prepares. It was great to talk with her on matters ranging from food, movies and even the pains of research.

I would also like to thank Prof. Dimitri Antoniadis for providing suggestions which were very vital for solving some of the research problems and for making the resources of his measurement lab available to me. Prof. Judy Hoyt has provided lot of advices on research and understanding of the subject. I am very thankful to Jongwan Jung, who was there to help me during processing and device measurements and to clear up my concepts on MOS. I would like to thank Andy, Ingvar, Cait, Lee Wee and other collaborators in course 6 who helped me in innumerable ways including all the help that they gave me in device fabrication and measurements. I should express my gratitude to the MTL staff, CMSE staff, and other people who have maintained the facilities so that the research runs smoothly. Thanks is also due to all my friends at MIT and elsewhere for making my stay at MIT enjoyable.

Last but not the least, I would like to thank Saumya who came as a new life to me just when I was facing the difficult times of the grad school. Her support and care helped me to overcome the trying times of the grad school. My parents and my sister have always supported me and my gratitude towards them cannot be expressed in words. This thesis was not possible without the love and support of my family. I would like to dedicate this thesis to my parents for all their support.

Chapter 1. Introduction

1.1. Motivation

Historically, performance improvements in metal-oxide-semiconductor field-effect transistors (MOSFETs) have been attained by shrinking device dimensions, such as the gate length and oxide thickness. In addition to enhancing performance of Si based integrated circuits, it has also led to a reduction in the cost per function; thereby promoting market growth of integrated circuits. However, the practical benefit of scaling is declining as physical and economic limits are approached, and novel solutions are increasingly being sought. The ITRS roadmap puts together the challenges in the present technical capabilities that need to be developed in order to stay on Moore's Law. Since 2001, scaling has reached the point where the horizon of the roadmap challenges the most optimistic projections for continued scaling of complementary-metal-oxide-semiconductor (CMOS) transistors. International Technology Roadmap for Semiconductors (ITRS) 2001 reported on the presence of a potential "Red Brick Wall" around 2005 to 2006, indicating where there are no "known manufacturable solutions" to continued scaling in some aspect of the semiconductor technology that can block the further scaling envisaged by Moore's Law[5].

Thus, there are numerous technological challenges to be overcome to achieve further scaling and the continued growth of the semiconductor industry. The industry and academic communities are pursuing two avenues to meeting these challenges—new transistor structures and new materials. Incorporating strain and higher mobility materials in the channel can improve carrier transport properties. In addition, integrating many of

the III-V compounds, which can have direct bandgaps and high carrier mobility, with silicon utilizes existing mature Si processing infrastructure, thus preserving its economics of scale while expanding the functionality of the Si platform.

Relaxed $Si_{1-x}Ge_x$ virtual substrates create a larger lattice constant on a Si substrate and can be used as an epitaxial template for incorporating alternate materials such as SiGe, Ge, or III-V compound semiconductors and/or incorporating strain in the channel. Many other concepts for attaining relaxed $Si_{1-x}Ge_x$ buffers on Si wafers have been proposed and demonstrated over the years.[6,7] To date, compositional grading [8,9] remains the only established technique for attaining low-defect density, fully relaxed $Si_{1-x}Ge_x$ alloys on a Si substrate. In some cases, however, the presence of greater buffer layer thickness is deleterious. For many III-V devices on Si, the SiGe graded buffer approach is hindered by the relatively thick buffers. Due to high growth temperatures and the significant thermal expansion coefficient difference[10] between Si and Ge, the tensile strain that develops upon cool-down of graded buffer limits the amount of III-V material that can be deposited on structure without formation of micro-cracks[11]. The low thermal conductance that relaxed graded SiGe buffer platform provides, produces self-heating in high power devices and can lead to a significant reduction in both mobility and drain current[12].

As devices based on III-V materials are typically much thicker than Si-based devices, it will be highly beneficial to reduce the thickness of the SiGe buffer. Additionally, reducing the thickness of the buffer will also serve to improve the thermal conductivity of the buffer. One of the motivations for this work is studying the

dislocation dynamics of single mismatched layer in order to arrive at a technique for achieving ultra-thin buffers with a low threading dislocation densities.

Once a relaxed low threading dislocation density (TDD) buffer is achieved, some of the configurations for incorporating strained materials such as single channel and dual channel heterostructures[1,13,14] can be implemented. A dual channel heterostructure, consisting of a compressively strained $Si_{1-y}Ge_y$ (ε-$Si_{1-y}Ge_y$) layer grown on top of relaxed $Si_{1-x}Ge_x$($y>x$) and capped with a tensilely strained Si (ε-Si) layer, provides a platform for fabricating metal-oxide-semiconductor field-effect transistors (MOSFETs) with very high hole and electron mobility (μ_{eff}) enhancements over Czochralski grown Si[1,15,16]. The impressive hole μ_{eff} enhancements obtained on dual channel heterostructures depend greatly on the Ge concentration and the strain in the ε-$Si_{1-y}Ge_y$ layer[1].

Out-diffusion of Ge from the ε-$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ buffer and the top ε-Si cap can occur during high temperature processing. This reduces Ge concentration and, thus, the strain in the ε-$Si_{1-y}Ge_y$ layer leading to a reduced hole μ_{eff} in these heterostructures[2,17]. Development of a predictive tool for estimating the thermal budget of dual channel heterostructures is essential for outlining the design space of the processing temperatures for a particular dual channel heterostructure. In the second part of the thesis, the diffusion coefficient for Ge interdiffusion in a SiGe single crystal has been formulated. The out-diffusion of Ge from the ε-$Si_{1-y}Ge_y$ layer into the neighboring relaxed $Si_{1-x}Ge_x$ buffer and the strained Si has been analyzed. This analysis will help us to arrive at a solution for improving the thermal budgets of these strained layer

heterostructures. Improving the thermal budgets of these heterostructures will help in a smooth incorporation of these heterostructures in a commercial CMOS node.

1.2. Scope of the thesis

The thesis focuses on the physics of relaxation of SiGe layers and its applications for engineering substrates with desired properties. Once an engineered substrate has been achieved, strained channels can be grown on them for CMOS applications. This thesis will yield valuable insights into the thermal budgets of such strained layer heterostructures. Using a diffusion coefficient, which will be formulated in this thesis, solutions for improving the thermal budgets of these heterostructures will be suggested.

Chapters 2 reviews the physics of SiGe mismatched epitaxy and uses it to build a practical technique of achieving an ultra-thin buffer. It starts by reviewing the basics of lattice mismatched epitaxy and arrives at the important parameters affecting the TDD in mismatched SiGe layers. Reducing the TDD in a mismatched SiGe layer on Si requires slow plastic strain relaxation which can be achieved by growing layers at slower growth rates which can be typically achieved by using reduced growth temperatures in CVD systems. Using lower growth temperatures to reduce the growth rates, however, leads to a poorly relaxed layer. A Time-Temperature-Threading Dislocation Density diagram has been developed in this work and explains that a high temperature anneal for relaxing the layer leads to undesirable TDD escalation. Strategies for improving the relaxation of a metastable layer without any undesirable threading dislocation nucleation are described.

The chapter concludes with the very important realization of an ultra-thin buffer with a low TDD which can directly be applicable in CMOS and has important implications for optoelectronic integration on Si platform.

In Chapter 3 we will describe the fabrication techniques of MOSFETs used in this thesis and describe the procedures for extracting the mobilities of the carriers. Chapter 4 will focus on formulating a diffusion coefficient for use in strained and unstrained SiGe layers. Chapter 5 will discuss the thermal stability of the dual channel heterostructure and utilize the diffusion coefficient formulation in estimating the thermal budget of a dual channel quantitatively. Lastly the tri-layer heterostructure is introduced and important enhancements in thermal stability and electrical properties as a result of incorporating the tri-layer heterostructure have been described.

In Chapter 6, we will summarize the results and suggest directions for future work.

Chapter 2. Ultra-thin SiGe buffers

Producing relaxed defect-free SiGe alloys has been difficult due to the 4% mismatch between Si and Ge. For example, a thick uniform $Si_{0.8}Ge_{0.2}$ buffer layer deposited directly on Si contains a threading dislocation density (TDD) on the order of 10^7 cm^{-2}. Relaxed $Si_{1-x}Ge_x$ graded buffer creates a larger lattice constant on a Si substrate while providing low TDD on the order of 10^5 cm^{-2}. Many other concepts for attaining relaxed $Si_{1-x}Ge_x$ buffers on Si wafers have been demonstrated over the years[18]. However, compositional grading remains the most established technique for attaining low-dislocation density, fully relaxed $Si_{1-x}Ge_x$ alloys on a Si substrate.

In this chapter we will attempt to build from basics the concepts for achieving a completely relaxed low threading dislocation density mismatched layer. Important variables which affect the dislocation density in a mismatched layer will be identified and engineered using different techniques to achieve ultra-thin SiGe buffers. Various experimental and theoretical limits for engineering such buffers will also be considered.

2.1. The Relaxed SiGe Integration Platform

The silicon-germanium (SiGe) materials system possesses several attractive properties making it a natural choice for extending the performance of the Si microelectronics platform[19,20]. One of those is the ability to access a lattice constant that varies linearly between 5.43 Å - 5.66 Å by growing relaxed $Si_{1-x}Ge_x$ films of any composition from pure Si to Ge. Since, Si and Ge are completely miscible over the entire

composition range (Figure 2-1), high quality SiGe films can be grown over all composition ranges without the problem of phase segregation.

Such a template can also be used to grow thin strained films of SiGe with any concentration. Some of these include high mobility strained Si and strained SiGe layers for improved MOSFET performance. Different alloy compositions of SiGe possess very different chemical properties making them excellent etch-stop for use in MEMS and on-insulator platform[21]. SiGe alloys have great potential for optoelectronic integration with CMOS because of lattice matching of Ge alloys to GaAs and AlAs. The successful integration of high quality III-V material[22] on Si wafers holds great promise for future circuit technologies where optical devices may be monolithically integrated[23] with the most advanced VLSI logic technology.

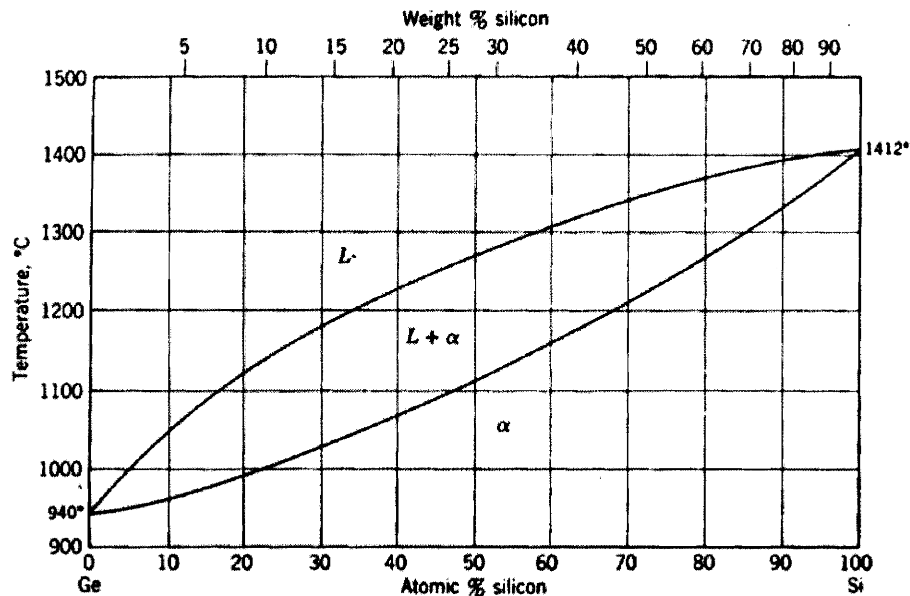


Figure 2-1:The Si-Ge Binary Phase Diagram, adapted from Gandhi[24] Si and Ge are completely miscible, enabling SiGe films of arbitrary composition to be grown without phase segregation.

2.2. Lattice Mismatched epitaxy

2.2.1 Introduction

Many electronic devices, particularly minority carrier devices are very sensitive to defect density. Accessing the improved functionality offered by silicon-germanium alloys requires relaxed films with a low defect density.[25]. SiGe wafers cannot be made from the melt using the CZ method since they will have non-uniform Ge concentration both radially and also along the boule. Thus, the most feasible method for accessing the wide

range of $Si_{1-x}Ge_x$ lattice constants is through epitaxial growth of $Si_{1-x}Ge_x$ buffers on Si wafers.

However, because of the 4% lattice mismatch between Si (a = 5.431 Å) and Ge (a = 5.658 Å), direct deposition of a SiGe film onto a Si wafer will not accomplish the requirements of low threading dislocation densities. Growth of a thick relaxed film lattice mismatched with its nearest substrate results in the formation of misfit dislocations at the growth interface. Above the critical thickness, the formation of misfit dislocations is energetically favored to reduce the strain energy in the film. Since a dislocation cannot terminate inside the crystal, it rises up to the film surface as threading segments. A highly mismatched layer grown on top of Si leads to a very high dislocation density (Figure 2-2). All defect reduction techniques rely on controlling the threading dislocation density.

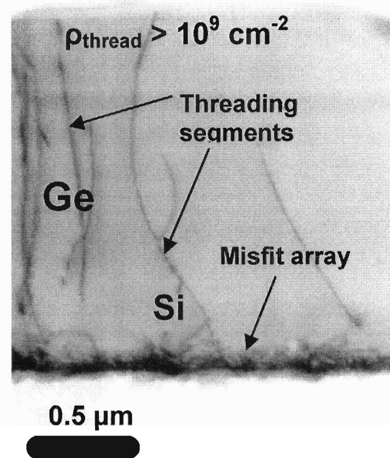


Figure 2-2: Very high TDD (>10^9 cm^{-2}) in a Ge deposited on Si substrate. Image courtesy of Tom Langdo

2.2.2 Lattice mismatch and critical thickness

An epitaxial film with a lattice constant, a_f grown on a substrate with a lattice constant, a_s will experience a lattice mismatch, f given by Equation 2-1.

$$f = \frac{a_s - a_f}{a_f}$$ **Equation 2-1**

At small film thickness, the energy required for dislocations to be present and relieve mismatch strain is less than the strain energy accumulated in the film. As a result, the SiGe layer grows pseudomorphically[26], with an in-plane lattice constant equal to that of the Si substrate. The out-of-plane lattice constant also deforms according to the Poisson ratio of the film, leading to a tetragonal distortion of the epitaxial layer. As the thickness of the layer increases, it will accumulate elastic strain energy until a critical thickness, h_c. At larger film thickness, the strain energy in the film exceeds the energy necessary for dislocations to be present, leading to dislocation formation as shown in Figure 2-3.

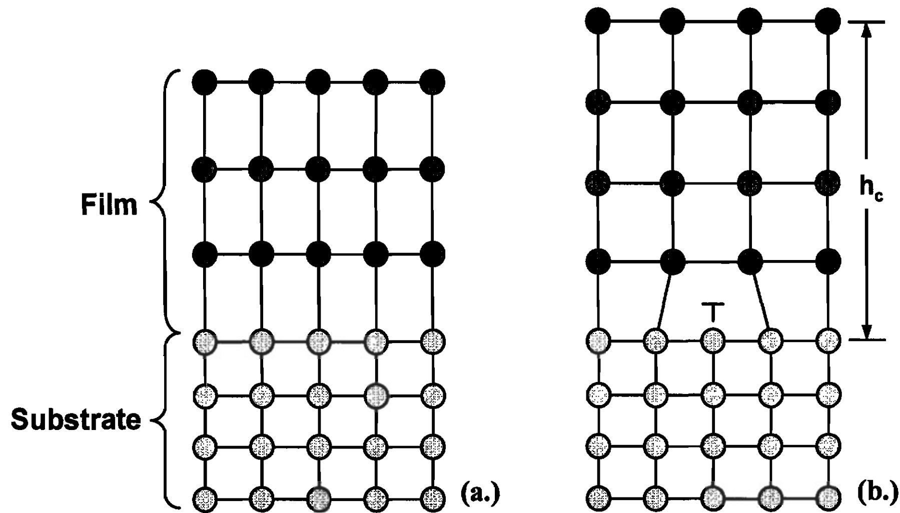


Figure 2-3 Schematic of pseudomorphic growth and the onset of dislocation formation in lattice-mismatched heteroepitaxy. (a) Initially, the in-plane film lattice constant is constrained to that of the substrate, as shown on the left. (b) Above a critical thickness, dislocation formation becomes energetically favorable and the film relaxes, as shown on the right. Image courtesy of Arthur Pitera.

Misfit dislocations form at the expense of introducing line energy to the crystal. The thickness at which the strain energy in the film exceeds the energy necessary for dislocations to be present, it becomes thermodynamically favorable to nucleate misfit dislocations. This can be expressed mathematically by minimizing the total energy of the system, leading to the classic Matthews-Blakeslee equation. This thickness is referred to as critical thickness and is an equilibrium parameter. Metastable films can be grown at lower temperatures which remain dislocation-free well past the critical thickness.

The equilibrium critical thickness, h_c, for a lattice-mismatched layer is given by

$$h_c = \frac{D(1-\nu\cos^2\alpha)(b/b_{eff})\left[\ln\left(h_c/b\right)+1\right]}{2Yf}$$ **Equation 2-2**

where D is the average shear modulus at the interface, ν is Poisson's ratio, α is the angle between the dislocation line direction and Burgers vector b, h is the film thickness, b_{eff} is the interfacial component of the Burgers vector, Y is the Young's modulus of the film, and f is the mismatch between film and substrate[8].

The lattice constant of a relaxed SiGe layer is given by Vegard's Law

$$a_{SiGe} = a_{Si}(1-x) + a_{Ge}(x)$$ **Equation 2-3**

for a given Ge atomic fraction (x). As mentioned earlier, the lattice constants for Si and Ge are 5.431 Å and 5.658 Å, respectively. By interpolating elastic constants, calculating mismatch for $Si_{1-x}Ge_x$ layers grown on Si substrates, and assuming $b_{eff}=a_{SiGe}/2$, the equilibrium critical thickness for $Si_{1-x}Ge_x$ layers grown on Si can be computed, as given in Figure 2-4 below.

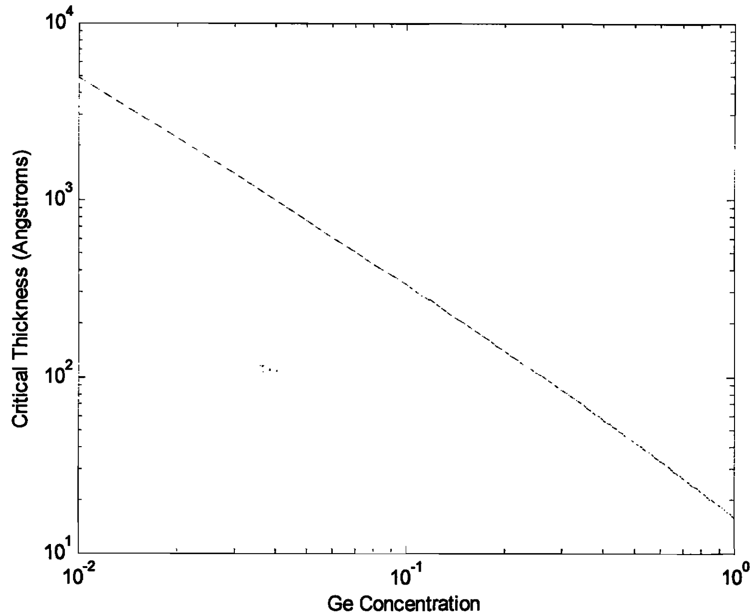


Figure 2-4: Critical thickness (h_c) of $Si_{1-x}Ge_x$ alloys on Si (001) versus Ge content (x). Critical thickness drops rapidly with increasing lattice mismatch.

A very important way to look at Equation 2-2 is given as following

$$\varepsilon = \frac{D\left(1 - v\cos^2\alpha\right)\left(b/b_{eff}\right)\left[\ln\left(h/b\right)+1\right]}{2Yh} \quad \textbf{Equation 2-4}$$

The different terms are the same as in Equation 2-1 and a plot of strain remaining in the film with thickness is given in below (Figure 2-5).

If f is the mismatch of a layer grown on Si and its thickness is h, assuming that there are no barriers to relaxation, it will relax from mismatch of f to a value of strain

given from Figure 2-5. The amount of plastic strain relaxed will be referred to as δ, while the strain remaining in the film is referred to as ε.

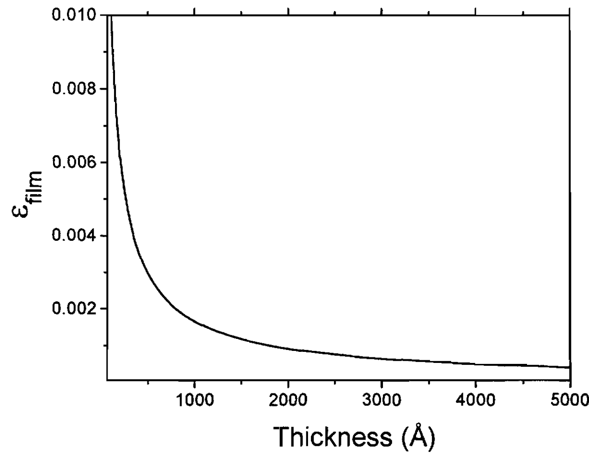


Figure 2-5: Strain remaining in the film plotted with thickness.

Thus,

$$f = \delta + \varepsilon \quad \textbf{Equation 2-5}$$

If a layer of 10% Ge concentration is grown to be of thickness of about 1000Å, assuming that it is free to relax, it will relax to just about 4% of Ge concentration. Its initial mismatch is 0.00418 (f=0.10×0.0418), the strain remaining in the layer corresponds to ε=1.672×10^{-3} (Figure 2-5) and the plastic strain relaxed is equivalent to δ which corresponds to ($f-\varepsilon$) (Equation 2-5).

The misfit dislocation spacing, S, is related to plastic strain relaxation, δ by Equation 2-6.

$$\delta = \frac{b_{eff}}{\text{mean}(\text{Spacing})}$$ **Equation 2-6**

A fully relaxed (*i.e.* $\delta = f$) 10% SiGe layer grown on Si requires a mean misfit dislocation spacing of 45 nm, assuming 60° dislocations. An important aspect to note is that all the misfits dislocations are not spaced by an equal value but have a distribution of spacing. While calculating relaxation, mean of the spacing should be taken into account.

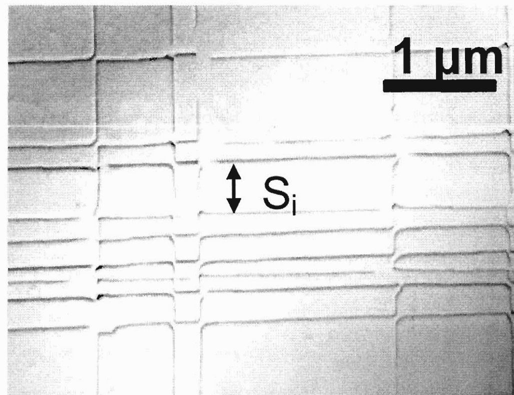


Figure 2-6: Misfit dislocations spacing is not uniform but has a certain distribution.

2.3. Experimental Techniques

In this section we will describe the various experimental techniques used for the growth and characterization of SiGe material.

2.3.1 SiGe Growth

All material for this thesis was grown in the Fitzgerald group ultra-high vacuum chemical vapor deposition (UHVCVD) system. A schematic of the UHVCVD system employed in this study is given in Figure 2-7 below.

The UHVCVD system used in this study employs SiH_4 and GeH_4 source gases. Dopants are supplied via 1% B_2H_6 in H_2 and 1% PH_3 in H_2 gases, which can be further diluted with either H_2 or Ar via two dilution stages. The reactor is a hot-walled, load-locked system with a quartz growth tube, and up to 10×6" wafers can be loaded simultaneously. Temperatures and growth pressures can be varied from 900°C to 350°C and 1-30 mTorr respectively. The upper limit for temperature is set by de-vitrification of the quartz tube and cooling of the o-rings at the tube base, while the lower limit is set by the thermally activated decomposition of source gases. Both the growth chamber and loadlock are evacuated by turbo-molecular pumps that are, in turn, backed by a roots-trivac combination pump. The seal between the quartz tube and the stainless steel chamber is accomplished by concentric o-rings. The space between the o-rings is pumped by a separate, smaller turbo-molecular pump so that each o-ring supports a vacuum of about 10^{-6} Torr. This scheme, called differential o-ring pumping, allows base pressures of 1×10^{-9} Torr to be attained at 900°C. Growth rates of around 10Å s^{-1} are typical for temperatures from 800°C to 900°C

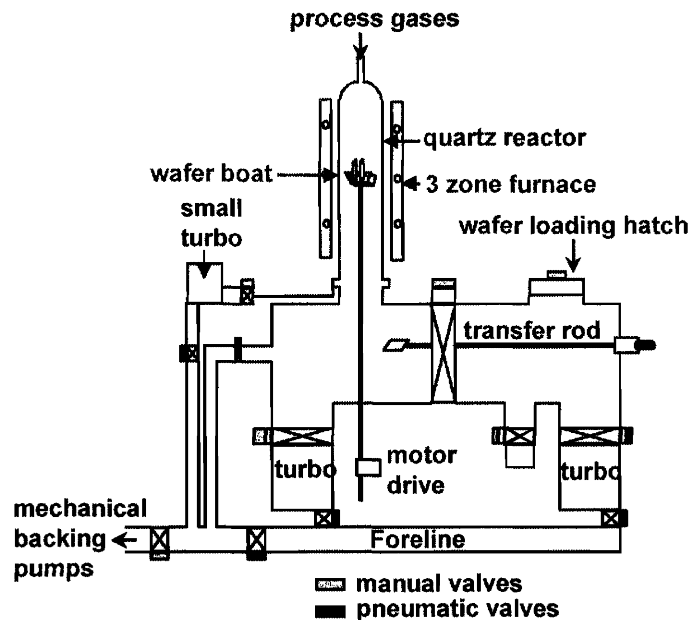


Figure 2-7: Schematic of the Fitzgerald Group UHVCVD system

Prior to growth, Si wafers are cleaned by a two-step procedure consisting of a 10 minute clean in 3:1 H_2SO_4:H_2O_2 (piranha clean) followed by de-ionized water (DI) rinsing and a 1 minute dip in 10:1 H_2O:HF. In addition to Si wafers, this procedure can be applied to virtual substrates containing up to 50% Ge. The final dilute hydrofluoric acid step serves to strip the chemical oxide that grows during piranha cleaning as well as terminating the surface with H atoms to prevent the subsequent formation of a native oxide in atmosphere.

Wafers are typically pumped down in the loadlock chamber to a pressure of ~1×10^{-7} Torr before introduction into the growth chamber. The motor driven pedestal shown in Figure 2-7 is then used to raise the wafers into an area just below the furnace. This

step is intended to serve as a low-temperature desorption to encourage volatilization of water molecules and other atmospheric contaminants. Finally, just before introduction of source gases, the wafers are held at 900°C for several minutes in order to volatilize any native oxide that may have formed after cleaning.

2.3.2 Characterization Techniques

Multiple characterization techniques were undertaken to measure defect density, composition, strain, and surface morphology. This section describes the techniques commonly employed in this thesis.

2.3.2.1 Transmission Electron Microscopy

Transmission electron microscopy (TEM) provides information pertaining to film and defect morphology by passing a monochromatic electron beam through a thin sample foil, either in cross-section or plan-view, to produce a high-resolution image. The interaction of the beam and the crystalline foil produces multiple diffracted beams that can be used to create an image of the internal structure of the sample. The variation in the intensity of these diffracted beams leads to diffraction contrast that is particularly useful for imaging features, such as defect and strain, which cause distortion in the sample lattice. The samples are manually ground and polished to a thickness of approximately 10 μm and then Ar-ion milled to render them electron-transparent. The usefulness of TEM lies in determining film morphology and thickness. A typical low-magnification plan view TEM (PVTEM) image for characterizing TDD is shown in Figure 2-8. Since dislocations may be directly imaged using this technique, TEM can also

be used to unambiguously measure dislocation densities under the correct contrast condition. However, because the sampling area in TEM is very small, in practice cross-sectional TEM has a threading dislocation detection lower limit of roughly 10^8 cm^{-2} while plan-view TEM has a threading dislocation detection lower limit of roughly 10^6 cm^{-2}. Furthermore, since dislocation pile-ups are spaced very far apart, TEM is unsuitable for accurately measuring dislocation pile-up densities. In general, then, TEM measurements of dislocation densities in low defect density material should always be confirmed with other wide-area techniques. Using PVTEM, misfit dislocation spacings can be determined and relaxation of the layer can be obtained by using Equation 2-6 (Figure 2-9).

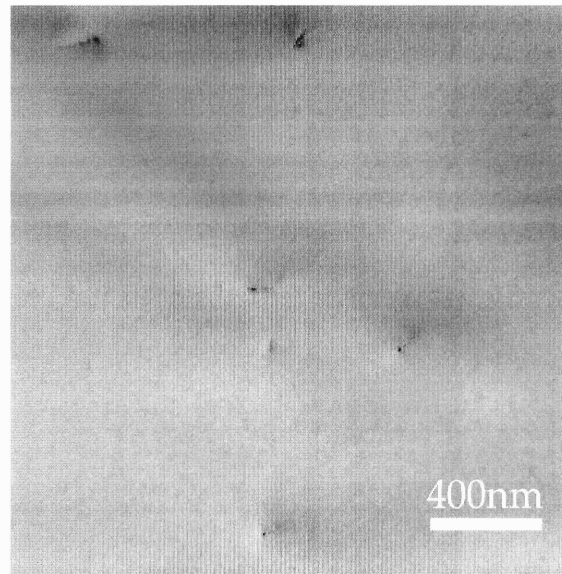


Figure 2-8: Plan view TEM showing a dislocation density of about $1.5\times10^7 cm^{-2}$.

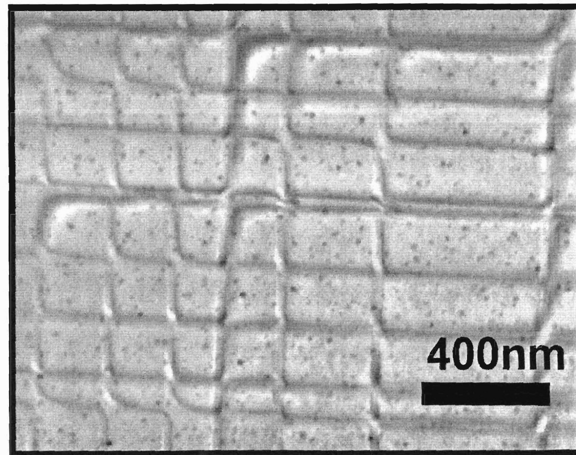


Figure 2-9: Plan view TEM used for calculating the relaxation in the layer by measuring the spacing between the misfit dislocations.

2.3.2.2 Etch-Pit Density Measurements

A common and dependable method for determining dislocation density in lattice-mismatched layers is direct imaging of the defects with plan-view TEM. However, the maximum imageable area using electron microscopy is typically smaller than 10×10 μm. Therefore, a sample containing 10^5 cm^{-2} threading dislocations will require hundreds of TEM micrographs to yield a meaningful threading dislocation density (TDD) value. An alternative method for revealing dislocations is selective etching of the sample and subsequent etch-pit density (EPD) measurements. Chemical etching reveals dislocations by selectively attacking their highly strained cores. EPD measurements are suitable for samples containing low ($<10^6$ cm^{-2}) threading dislocations since the etch-pits can be imaged over a large surface area using differential interference contrast (DIC)

microscopy. The diameter of the etch-pit is much larger than the diameter of the corresponding dislocation. Consequently EPD is not applicable to samples containing large defect densities where overlap of the etch pits seriously underestimates TDD values. Detailed correlation of etch-pit density measurements with other techniques, such as plan-view TEM, is necessary for accurate determination of dislocation densities.

For SiGe alloys, standard dislocation etches utilized are chromic acid-based etches. The standard chromic acid etch, composed of 4 parts 0.3 molar aqueous CrO_3 to 5 parts concentrated HF, is commonly known as the Schimmel etch[27] and is useful for revealing dislocations in SiGe layers for Ge compositions less than 70%.[28] It was required to calibrate the time for each sample as different times to reveal the etch pits were required for different Ge concentrations. Normally a time between 10 to 40 seconds worked for samples between 5% and 25% Ge concentrations.

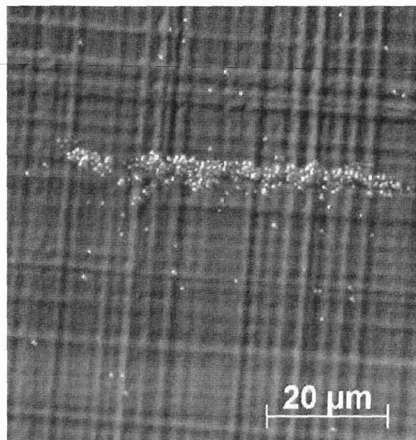


Figure 2-10: DIC micrograph of SiGe layer after selective etching, revealing threading dislocations. Image Courtesy of Chris Leitz

2.3.2.3 Atomic Force Microscopy

Atomic force microscopy (AFM) is useful diagnostic tool for characterizing the surface morphology of deposited films. The underlying mechanism for forming an AFM image is a silicon cantilever mounted on a piezoelectric transducer. The tip of the cantilever has a sharp point which is rastered across the sample surface. A constant force is maintained between the cantilever and the sample surface by application of a voltage to the piezoelectric transducer thus generating a digital image of the surface topography. The AFM method employed in the characterization of semiconductor surfaces in this work is called tapping mode which utilizes a vibrating cantilever to minimize electrostatic interaction between the tip and the sample surface. In addition to providing a qualitative description of surface morphology, AFM is useful for measuring the root-mean-square (RMS) surface roughness.

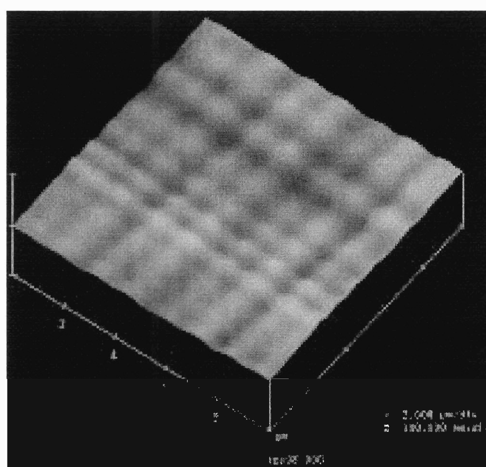


Figure 2-11: Example of an AFM micrograph obtained from a 10% SiGe layer grown on Si. The scan area is 10μm×10μm

2.3.2.4 Triple-Axis X-ray Diffraction

Triple-axis x-ray diffraction (TAXRD) is a high resolution x-ray diffraction technique capable of accurately resolving strain and composition in single mismatched layers and compositionally graded buffers. In TAXRD, the x-ray beam is conditioned by crystals before and after diffracting off the sample surface, producing a highly monochromatic beam and filtering out intensity resulting from sample bow. Triple-axis reciprocal space maps, in which a series of $\theta/2\theta$ scans are taken at various θ values, are useful for determining strain and composition and are in general more accurate than $\theta/2\theta$ scans. An example of a reciprocal space map obtained from a compositionally graded SiGe buffer layer is given in below.

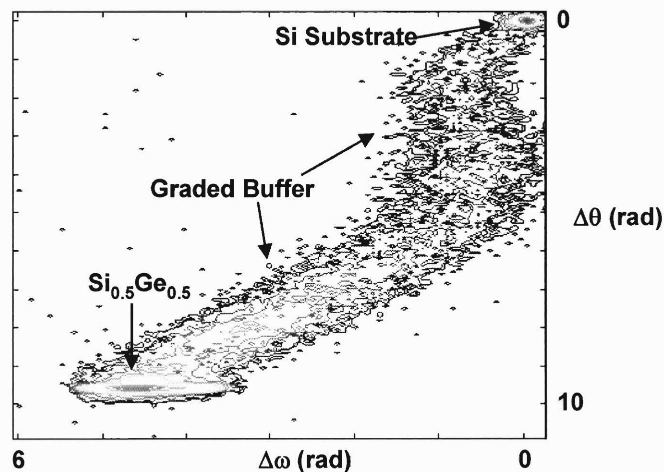


Figure 2-12 Example of an (004) reciprocal space map obtained from a compositionally graded SiGe buffer graded to a final Ge composition of 50%. Image courtesy of C.W. Leitz

Depending on the sample, one or both of TAXRD and PVTEM were used for determining strain and composition. While determining strain in the samples it was required to have a extremely accurate readings and therefore a noise removal subroutine (using Savitzky-Golay Filter, see Appendix 6) was also implemented which removes the noise from the sample to obtain the desired peak.

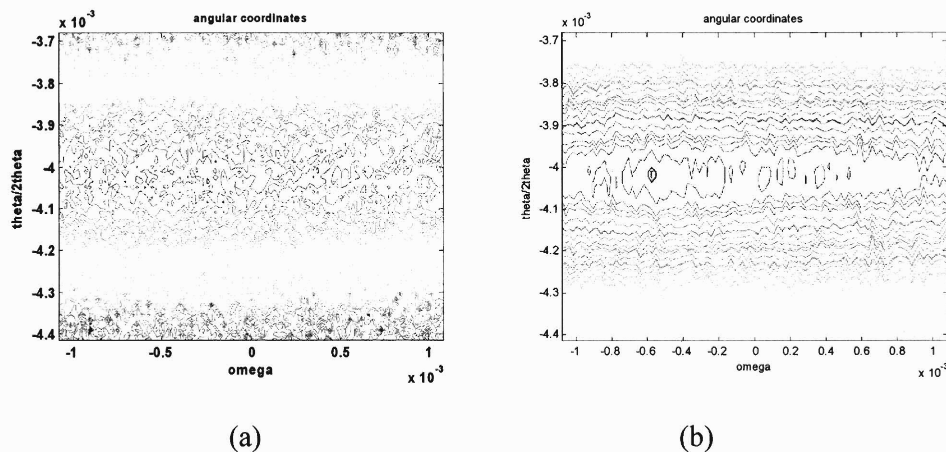


Figure 2-13: Figures explaining XRD data smoothing using Savitzky-Golay filter. Figure (a) plots the as acquired data while Figure (b) shows the data after subjecting it to the filter. The random noise in the data is removed and a better idea of peak position in the omega axis is achieved. In addition a peak finding subroutine is implemented to get the exact peak

2.3.2.5 Nomarski contrast microscopy

Nomarski microscopy or the differential interference contrast (DIC) technique was originally developed for rendering contrast in transparent specimens. Conventional optical microscopy relies on variations in surface reflectivity to create contrast. However,

the reflectivity of a blanket semiconductor film has little variation, yielding poor contrast using standard optical techniques. The DIC microscope utilizes a beam-shearing prism to produce an illumination source containing two optical paths that are slightly out of phase with each other. Contrast is produced by optical path gradients caused by variations of refractive index or topography of the sample. This technique greatly improves image contrast, particularly in samples which contain steep optical path gradients such as those present with crosshatched surfaces. These elements combine to make Nomarski microscopy ideal for imaging etch-pits for dislocation density measurements, as described in the next section.

2.4. Threading dislocation densities in a single mismatched layer

In this section we will determine the parameters that affect the threading dislocation densities in a single mismatched layer grown on Si. Identification of these variables would be essential in controlling the threading dislocation densities in a single mismatched layer. 1-1.5 micron thick relaxed single mismatched layers of varying Ge composition were grown on Si wafers using UHVCVD at 900°C. Then, the threading dislocation densities were characterized using EPD measurements and PVTEM. Roughness of the layers was characterized using atomic force microscopy. The TDD evolution with Ge concentration is shown below.

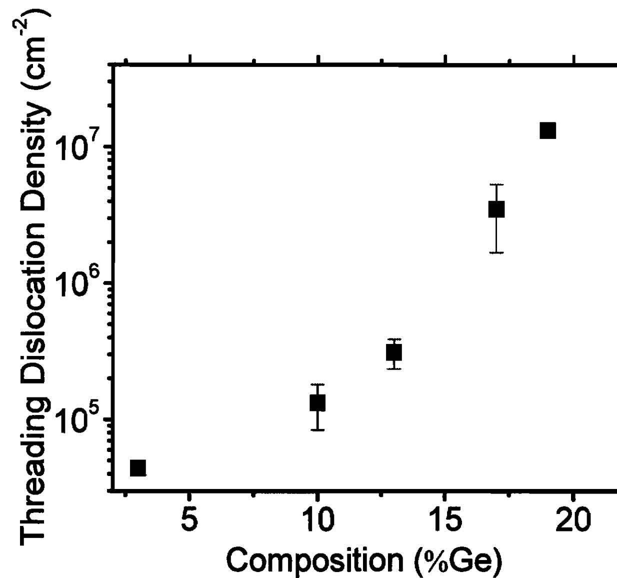


Figure 2-14: TDD evolution with Ge concentration in a mismatched SiGe layer grown at 900°C

2.4.1 Variables affecting TDD in mismatched SiGe layer: Strain rate

As a singly mismatched layer grows beyond its critical thickness, excess strain builds up in the layer leading to dislocation nucleation. The excess strain is given by the difference of strain in the layer and the equilibrium strain that should occur in the layer. Typically the dislocation nucleation does not start until the layer has surpassed the critical thickness to such an extent that the excess strain developed is large enough for it to nucleate dislocations. Thus, the relaxation of a single layer normally moves away from the ideal Matthews-Blakeslee criterion. However, since nucleation is thermally activated, a mismatched layer grown at a high temperature follows the ideal Matthews-Blakeslee curve very closely. Thus, if the layer follows equilibrium, as its growth continues, it will

follow the strain with thickness curve as shown in Figure 2-5. The ideal Matthews-Blakeslee curve is give by the Equation 2-4. The Matthews-Blakeslee curve for layers grown at 900°C has been shown in Figure 2-15.

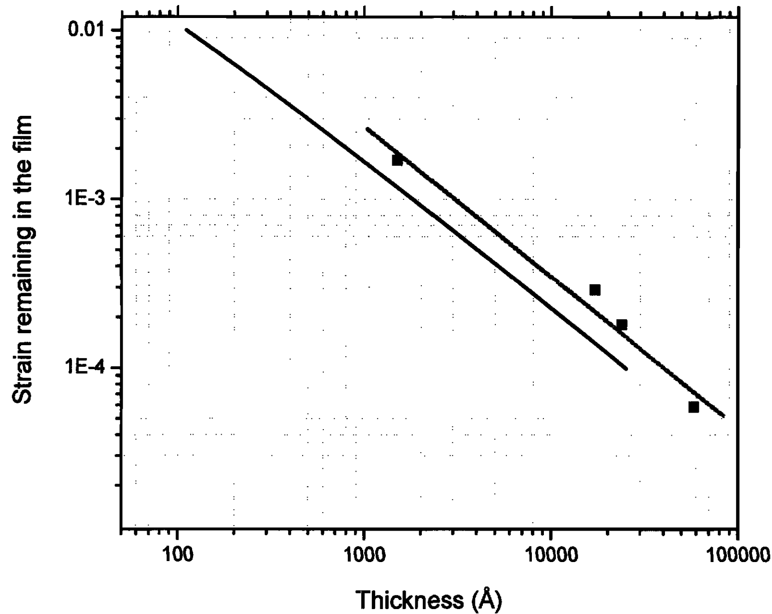


Figure 2-15: Experimental Matthews-Blakeslee curve for SiGe mismatched layer grown at 900°C

If a layer is deposited at high temperature, where nucleation of dislocation is easy, the growing mismatched layer closely follows the ideal Matthews-Blakeslee curve. As the layer exceeds the critical thickness it needs to relax the excess strain. If the growth rate of the layer is high, the amount of strain that it needs to relax is also high as more excess strain builds up per unit time. We define rate of plastic strain relaxation as the amount of strain that needs to be relaxed per second. The rate of plastic strain relaxation

then depends on number of threading dislocations participating in glide and the glide velocities of these threading dislocations[9].

$$\rho_x = \frac{\dot{\delta}_x}{b_{eff,x} v_x} = \frac{\left(\frac{\partial \varepsilon}{\partial h}\right)_x \times \left(\frac{\partial h}{\partial t}\right)_x}{b_{eff,x} v_x}$$ **Equation 2-7**

ρ_x is the threading dislocation densities in a layer with Ge concentration, x, $b_{eff,x}$ is the effective Burgers vector in the layer, v_x is the glide velocity of the threading dislocation in the layer, $\dot{\delta}_x$ is the rate of relaxation of plastic strain. It is to be noted that the plastic strain relaxation rate of the layer can be decomposed into a product of the slope of the Matthews-Blakeslee strain curve $\left(\partial \varepsilon / \partial h\right)_x$, and growth rate of the layer $\left(\partial h / \partial t\right)_x$.

From this equation we see that the number of dislocations will increase as the Ge content in the layer increases. This is intuitive since larger numbers of dislocations are required for relaxing the higher strain buildup during the growth of a high mismatch layer. The glide velocities of the dislocations are also dependent on the strain in the layer[29] amongst other factors. Rearranging the equation 2-7 and comparing it with a particular form of equation 2-7 for a 2% layer, the TDD in a mismatched layer having a Ge fraction, x, can be related to the measured value of TDD in a 2% layer as below

$$\rho_{xge,final} = \left(\rho_{2\%} \times \frac{\dot{\delta}_{xge}}{b_{xge} v_{xge}} \times \frac{b_{2\%} v_{2\%}}{\dot{\delta}_{2\%}} \right) \textbf{Equation 2-8}$$

The growth rates of $Si_{1-x}Ge_x$ (x<0.20) layers at 900°C are similar; the difference among the layers is that the slope of the strain rate curve is higher at high Ge concentration and the difference in the glide velocities of the dislocations. Using Figure 2-5, we can calculate the slope of the strain curve with thickness. This equation helps us to predict the TDD at a higher Ge concentration. The measured values of TDD with increasing Ge concentration is shown in Figure 2-16, which shows that while strain rate effect predictions match the observed TDD quite well until about 12-13% Ge, the observed TDD increases much faster with increasing Ge concentration after 12% implying that additional factors are responsible. The additional factor which affects the TDD can be the roughness of the layer.

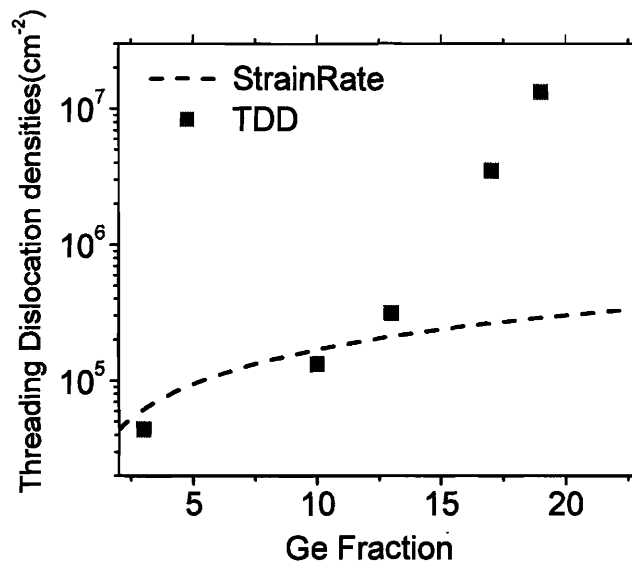


Figure 2-16: TDD evolution with Ge concentration in a mismatched SiGe layer grown at 900°C compared with the expected TDD distribution based on strain rate effects.

2.4.2 Surface roughness evolution in a mismatched SiGe layer

The roughness of the layers with varying Ge fraction is shown below. It is to be noted from Figure 2-14 that the increase in the RMS roughness occurs at the same point in Ge concentration where the TDD starts to increase. This leads us to believe that the roughness affects the dislocation density in some way.

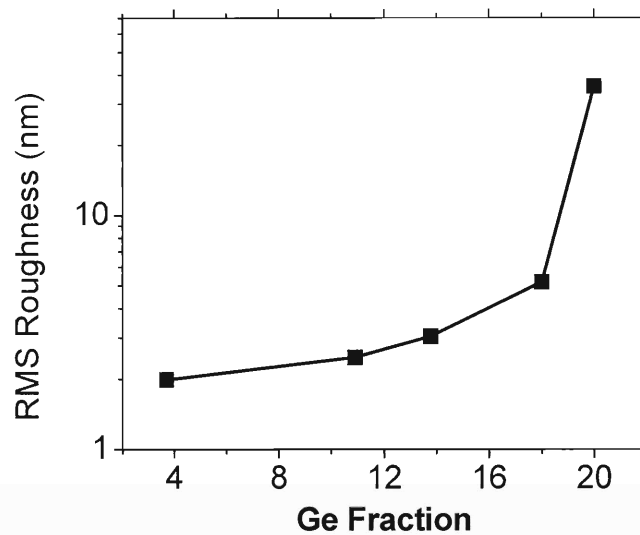


Figure 2-17: RMS roughness evolution with Ge concentration in a mismatched SiGe layer grown at 900°C. The scan area of all the samples is 10μm×10μm.

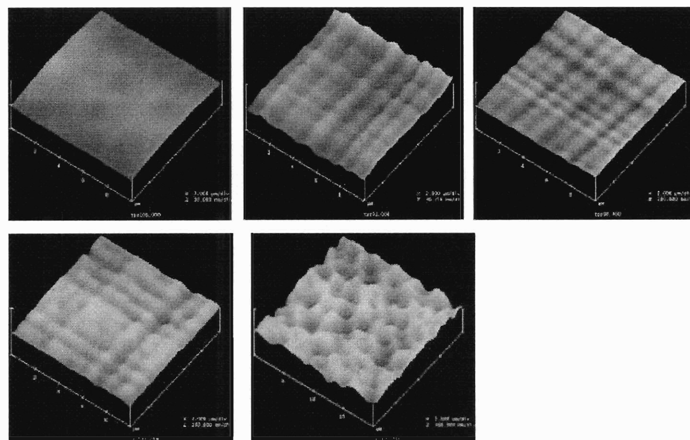


Figure 2-18: AFM images of the layer depicting the roughness increase as the Ge concentration increases in the mismatched layer. The scan area of all the samples is 10μm×10μm.

The relaxation of the layer is achieved through the glide of the threading dislocations which increases the length of the misfit dislocation. Threading dislocations can encounter obstacles inhibiting glide and thus rendering them ineffective for relaxing strain. This necessitates nucleation of a new threading dislocation leading to an increase in TDD. One blocking mechanism that is common in mismatched layers is illustrated in Figure 2-19.[30]

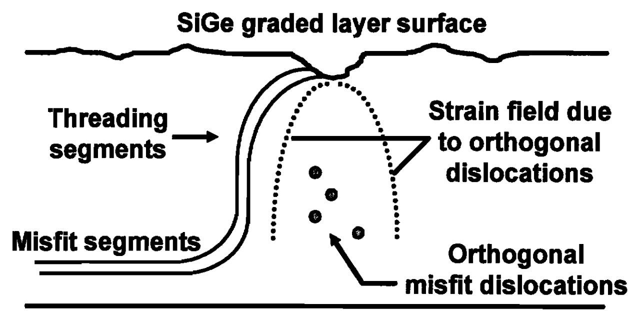


Figure 2-19: Illustration of a dislocation blocking mechanism in compositionally graded buffers. Local undulations of the surface coincide with the strain field generated by underlying misfit dislocations, posing an obstacle to dislocation glide. Figure adapted from Samavedam.[30]

2.4.3 Variables affecting TDD in mismatched SiGe layer: misfit dislocation spacing

Relaxation of the layer is achieved through the glide of threading dislocations which increases the length of the misfit dislocation. When there are no impediments to dislocation glide, the final threading dislocation density (TDD) of a compositionally graded structure should be governed by the rate of plastic strain relaxation. In practice, threading dislocations can encounter obstacles inhibiting glide and thus rendering them

ineffective for relaxing strain. This necessitates nucleation of a new threading dislocation leading to an increase in TDD. In the $Si_{1-x}Ge_x$/Si material system, the combination of surface roughness[30] and strain fields[31] of underlying misfit dislocations can trap the threading dislocations.

One of the possible impediments to the glide of threading dislocations is the interaction with dislocations on other glide planes. Relaxation of a mismatched SiGe layer occurs through formation of a misfit dislocation network. The excess strain in the layer provides a force helping the glide of threading dislocations. The stress fields of misfit dislocations can completely negate the traction due to excess strain forcing the threading dislocation to glide through a "channel" of width less than the layer thickness for bypassing the misfit dislocation. As the Ge concentration in the layer increase the average misfit dislocation spacing in the interface reduces. When the misfit dislocations get very near to each other their stress fields can overlap strongly and completely negate the driving force because of excess strain; thereby obstructing it completely in these places.[32].

Other factors like surface trenches can aid such blocking. The strain fields associated with misfit dislocations lead to the characteristic cross hatch pattern on the epilayer surface in lattice mismatched heteroepitaxy. When the misfit dislocations get very near to each other their strain fields can overlap strongly leading to deep cross hatch. For (001) epitaxy, the cross hatch pattern occurs in form of trenches and ridges aligned along the two in-plane <110> directions. Threading dislocations get blocked at the deeper

trenches since the total driving force on a threading dislocation is relatively lesser in these areas[30].

By applying both a dislocation blocking criterion and surface roughness to graded Si-Ge/Si(001) structures, a qualitative model for predicting the formation of dislocation pile-ups has been given in literature[33]. However, in this work we will develop a semi-quantitative model for explaining the TDD evolution in the layer. This model takes into account both the strain rate dependence and the misfit dislocation spacing to estimate the threading dislocation densities.

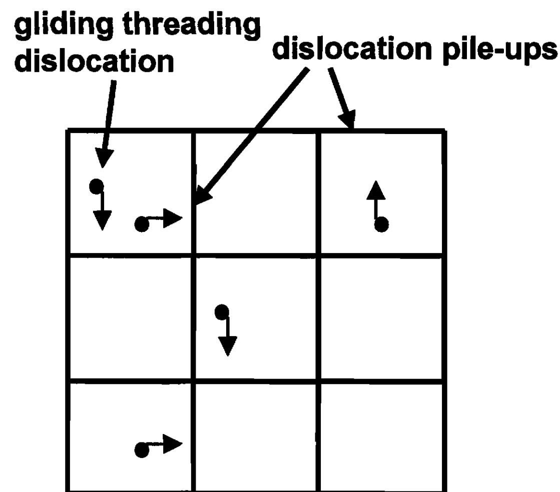


Figure 2-20: Plan-view of cellular structure emerging from dislocation pile-up formation. Dislocation pile-ups form boundaries that restrict gliding threading dislocation motion[34].

The primary premise of the model is that the plastic strain that is relaxed in the layer is relieved through misfit dislocations which are distributed in the layer in a log-

normal distribution (Figure 2-21). The fraction of the misfit dislocations which are very near to each other would have overlapping strain fields which would also lead to deep cross hatch combined with strong strain fields. Obstruction of these threading dislocations by the two mechanisms described above can therefore be related solely to misfit dislocation spacing. The fraction of the misfit dislocations which are below a critical spacing will prevent the dislocations from gliding and necessitate an increase in the dislocation density. Once we have an estimate of the fraction of misfit dislocations which will prevent the threading dislocation glide, we can better explain TDD evolution.

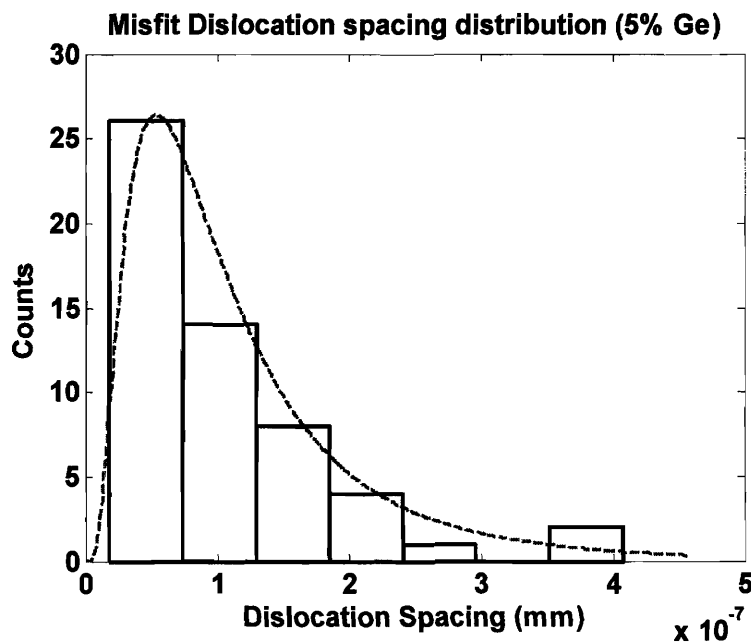


Figure 2-21: Misfit dislocations are spaced not uniformly but follow a log-normal distribution. The fraction of misfit dislocations that are very close to each other prevent the dislocations from gliding necessitating further nucleation events. Dotted line represents the log-normal distribution while histogram is experimentally measured spacing

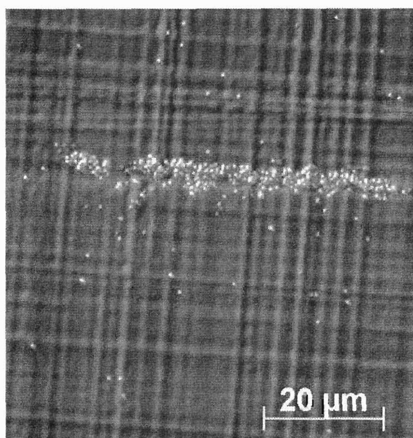


Figure 2-22: Plan-view optical micrograph of a selectively etched SiGe surface revealing widely spaced dislocations in the field and a large threading dislocation pile-up. Dislocations trapped in pile-ups cannot relieve mismatch strain, forcing nucleation of additional dislocations in the field and causing an escalation in overall dislocation density. Image Courtesy of Chris Leitz

In order to determine the extent of TDD escalation, it is essential to know the fraction of the misfits that prevent dislocation glide. The fraction that prevents dislocation glide can be incorporated in Equation 2-8, so that it includes both the strain rate dependence and the effect of dislocation glide prevention. The newer equation is given below:

$$\rho_{xge,final} = \left(\sqrt{\rho_{2\%} \times \frac{\dot{\delta}_{xge}}{b_{xge} v_{xge}} \times \frac{b_{2\%} v_{2\%}}{\dot{\delta}_{2\%}}} + 2 \times \frac{\delta}{b} \times fraction \right)^2 \quad \textbf{Equation 2-9}$$

The fraction that is referred to in this equation refers to the fraction of misfit dislocations whose spacing is below the critical spacing. The critical spacing is a fitting

parameter and is determined to be about 22nm. Thus, the fraction of misfits which are below 22nm will lead to prevention of dislocation glide and cause TDD escalation.

The calculated TDD matches the observed TDD very well as shown in Figure 2-23. An important implication from the above model is the insight that misfit dislocation spacing needs to be very uniform in order to reduce the dislocation densities. If the nucleation of misfit dislocation occurs from the edge of the wafer, the resulting distribution is very spread out. In a uniform distribution, a small fraction of misfit dislocations is closer than the critical spacing and hence their effect on escalating the threading dislocation density is smaller. Thus, if there is a seed layer which has a relatively uniform threading dislocation and a mismatched layer is deposited, the resulting misfit dislocation will be uniformly spaced leading to less dislocation density escalation as compared to if the misfit dislocation nucleation was a random process. This is indeed the case as is demonstrated in the literature[35].

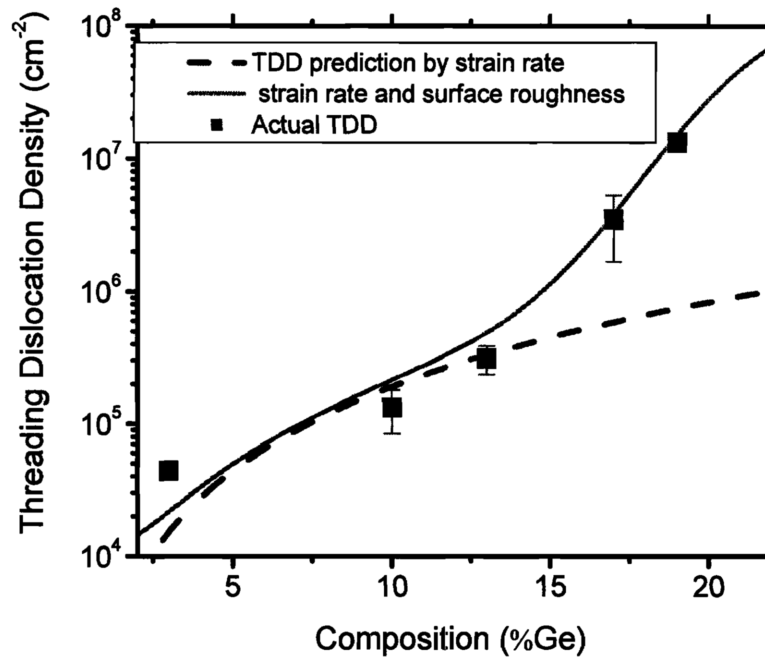


Figure 2-23: The calculated dislocation density matches the observed dislocation density implying that strain rate and the surface roughness are two important factors affecting the dislocation densities in a mismatched layer.

Another implication of this formulation is that a mismatched layer which has a high misfit density cannot have a low TDD. This is because, regardless of how uniformly spaced the misfit dislocations are, a significant fraction of them will still be more closely spaced than the critical spacing. A 20% layer has a misfit spacing of about 23nm, implying that regardless of how uniformly the misfits are spaced in a 20% layer, it will always have a high TDD since the strain fields will be very strong to prevent the dislocation glide.

2.5. Compositionally Graded Buffers

Numerous approaches have been investigated to reduce the threading dislocation density in lattice mismatched systems including growth of thick layers[36], thermal cycling[37], reduced area growth, epitaxial lateral overgrowth techniques[38], and graded buffers[39]. Of these, the most successful technique for the fabrication of low threading dislocation density epitaxial layers is the relaxed graded buffer which maintains a constant low strain rate to minimize dislocation nucleation while maximizing dislocation glide. Many intermediate layers of low mismatch are deposited in order to maintain a low strain rate. Each layer relaxes to its intermediate lattice constant and the process continues until the desired highly mismatched layer is reached. The threading dislocations from previous layers are reused at each subsequent lattice-mismatched interface as they glide to create misfit dislocations (Figure 2-24). Also, the deposition is carried out at a high growth temperature in order to maximize the glide length of threading dislocations at each interface, minimizing the need to nucleate additional dislocations for strain relief. The low density of misfit dislocations at each interface minimizes detrimental dislocation-dislocation interactions. Finally, the small lattice mismatch at each layer minimizes dislocation nucleation. Additionally, the glide length of misfit dislocation is very large which leads to reduced threading dislocation densities.

Thus, the substrate is transformed into a $Si_{1-x}Ge_x$ "virtual substrate" with a low threading dislocation density (TDD). While the bulk of the wafer is Si, the lattice constant of the top surface is that of a relaxed $Si_{1-x}Ge_x$ alloy.[39].

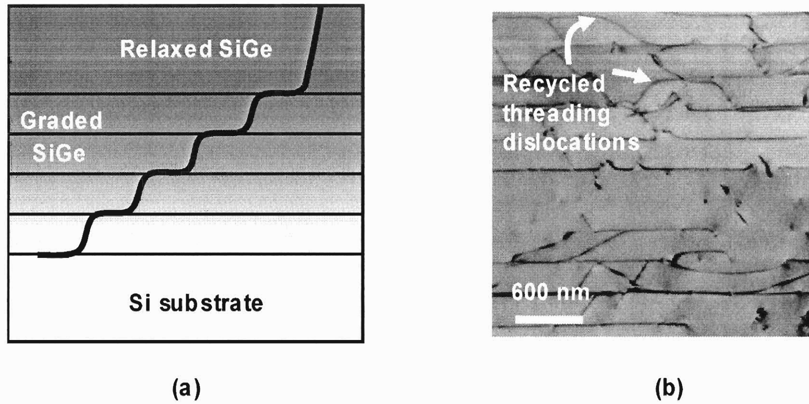


Figure 2-24: (a) Schematic of Compositionally graded buffer. (b) Cross-sectional transmission electron micrograph of a compositionally graded SiGe region. Image courtesy of Thomas Langdo.

Based on the earlier discussion in which factors affecting the dislocation densities were identified as misfit dislocation spacing (and concomitantly surface roughness) and strain rates at which the layer is grown, we will like to position the graded buffer in the roughness and the strain rate space as compared to a singly mismatched layer. While the surface roughness is very low in graded buffer as compared to a uniform layer (Figure 2-25), the misfit dislocation spacing is high leading to a lesser misfit-threading interactions. Thus, the instances where the threading dislocations get stuck are relatively less and hence the escalation of the TDD due to dislocation glide prevention does not occur.

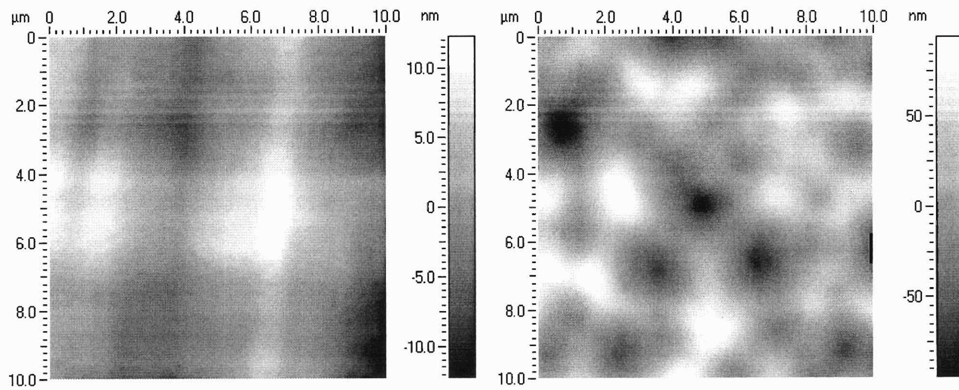


Figure 2-25: AFM images of a 20% graded buffer (left) compared with a 20% singly mismatched layer. The RMS roughness of the graded buffer is 3.8nm while the 20% singly mismatched layer has 60nm RMS roughness.

2.6. Controlling strain rates for reducing the dislocation densities by reducing growth rates

As discussed in the previous section, growth of a highly mismatched layer will lead to development of a high excess strain and if the layer has to follow equilibrium, a high rate of plastic strain relaxation. Since the glide velocity of the threading dislocations is fixed at a particular temperature, relieving the plastic strain at a high rate requires a high TDD in the layers. From Equation 2-7, we realize that the strain rate is directly proportional to the growth rate of the layer. Thus, by growing the layers very slowly the rate of development of the excess strain in the layer can be reduced to such an extent that the developed strain can be relaxed with the glide of relatively few dislocations. In order to have a TDD of similar level as in a graded buffer, the growth rates should be reduced to such an extent that the rate of the plastic strain relaxation of the mismatched layer is equal to that of a graded buffer. Estimating the required growth rate by the means of

equating the strain rate, however, yields a conservative estimate of the growth rates. This is because the glide velocities are higher in a more mismatched layer, and a higher rate of development of plastic strain can be relaxed with a fewer number of dislocations. In the following section we try to develop an analytical expression of growth rates at which the layers should be grown in order of it to have the same amount of strain as in the case of graded buffers grown with 2% step.

2.6.1 Determining growth rates for singly mismatched SiGe layers for simulating the strain rates of graded buffer

The strain remaining in the layer of thickness h, if it follows the Matthews-Blakeslee equilibrium curve is as below (Equation 2-4).

$$\varepsilon = \frac{D(1 - v\cos^2\alpha)\left(b/b_{eff}\right)\left[\ln\left(h/b\right)+1\right]}{2Yh}$$

where D is the average shear modulus at the interface, v is Poisson's ratio, α is the angle between the dislocation line direction and Burgers vector b, h is the film thickness, b_{eff} is the interfacial component of the Burgers vector, Y is the Young's modulus of the film, and f is the mismatch between film and substrate[8].

From Equation 2-7, the rate of plastic strain relaxation can be decomposed into the slope of the strain curve and the growth rate of the layer. If the plastic strain relaxation rate of a mismatched layer with a mismatch of f has to be equal to the strain rate of a 2% layer, the resulting equation is given below (Equation 2-10). The growth rate of a 2% layer is known from experiments and the slope of the strain curve for both a 2% layer and a layer with a mismatch of f can be determined from the Matthews-Blakeslee equilibrium curve.

$$\left(\frac{\partial \varepsilon}{\partial h}\right)_f \times \left(\frac{\partial h}{\partial t}\right)_f = \left(\frac{\partial \varepsilon}{\partial h}\right)_{0.02} \times \left(\frac{\partial h}{\partial t}\right)_{0.02}$$

Equation 2-10

Thus, the growth rate of a layer so that it has the strain rate equivalent to a 2% layer is given in Equation 2-11.

$$\left(\frac{\partial h}{\partial t}\right)_f = \left(\frac{h_{c,f}}{h_{c,0.02}}\right)^2 \times \left(\frac{\ln(h_{c,0.02}) - \ln(b)}{\ln(h_{c,f}) - \ln(b)}\right)_{0.02} \times \left(\frac{\partial h}{\partial t}\right)_{0.02}$$

Equation 2-11

An important point to notice is that the $h_{c,f}$ and $h_{c,0.02}$ are given by another transcendental equation that is derived in literature. Thus, the dependence of growth rate with composition is more complicated than is evident from the equation. Figure 2-26 shows the growth rate variation with composition near the critical thickness.

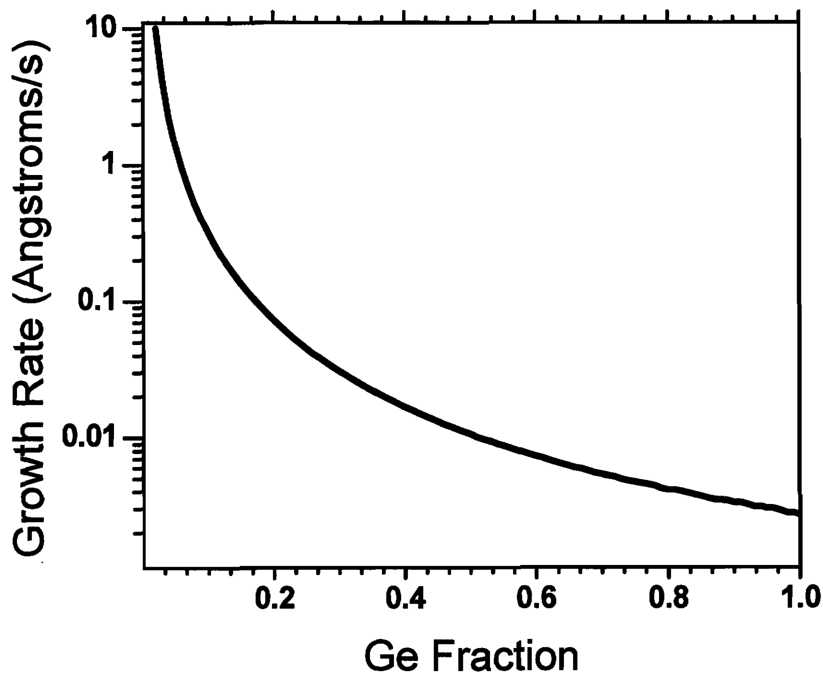


Figure 2-26: Growth rates (in angstroms per second) plotted with Ge composition in order to maintain a strain rate equivalent to a low TDD 2% layer. It shows that as the mismatch increases the layer needs to be grown at increasingly lower growth rates.

2.6.2 Achieving low growth rates for reducing TDD

Figure 2-26 describes that as the mismatch increases the layer needs to be grown at very low growth rates. Such growth rates can be achieved only at very low temperatures using CVD. At these temperatures the glide velocities of the dislocations are very small. Thus, a fundamental limitation exists since extremely slow growth rates and high glide velocities cannot be achieved simultaneously.

In order to circumvent this problem we have invented a pulsed growth technique for growing a highly mismatched layer. In this technique the growth is done in UHVCVD reactor at a high temperature of 900°C. This is useful for obtaining high glide velocities, the growth rates of the layer is very high (~1.5 nm/s). In order to reduce the effective growth rates, we grow for a very short time (henceforth called as pulse time) and then wait for a certain amount of time (henceforth called as the delay time). The effective growth rate obtained can thus be varied by varying the delay times.

A low growth rate regime at lower temperatures has also been explored. Since such a procedure will lead to a metastable layer, efforts has been put to determine techniques using which the layer can be relaxed without a drastic increase in the dislocation densities. Both the pulsed layer technique and low temperature growth are described in the following sections.

2.7. Pulsed layer growth

We define a pulsed layer technique in which gas flow is intermittent at 900°C, thus achieving low effective growth rates. By varying the delay time and the flow time we can vary widely the effective growth rates. This technique is required as CVD control via mass flow controllers introduces a minimum growth rate. It also allows the decoupling of growth rate and growth temperature inherent to CVD growth. While the lower growth rates achieve a low strain rate, the ability to have these lower growth rates at higher temperature improves the kinetics of relaxation.

Because enough pressure cannot be built in the pulsed layer growth, recalibration of composition is required for the growth of SiGe layers. The composition of the SiGe layer using the pulse technique at 900°C is Ge poor at the same gas flows as compared to the uniform layers. A plot of the composition of the pulsed flow is compared with the regular growth flows (Figure 2-27).

While the lower growth rates obtained by employing this pulsed layer technique helps us to get a low strain rate, the higher temperature improves the kinetics of relaxation. However, high adatom mobility during this process results in an increased surface roughness (Figure 2-28 and Figure 2-29). Thus, lower growth rates at high temperature inherently couples surface roughness to lower growth rate. The growth rate variable and dislocation velocity in Equation 2-7 are not independent under these conditions, and lower threading dislocation densities can not be achieved.

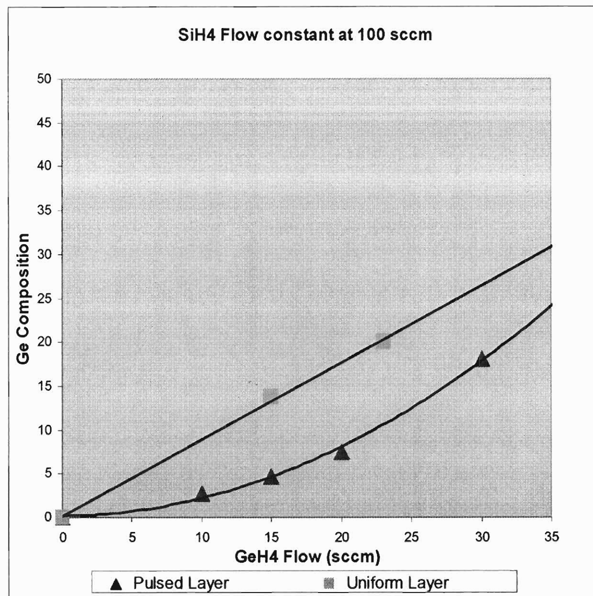


Figure 2-27: Using Pulsed layers Ge poor films are obtained as compared to layers grown without the pulsed

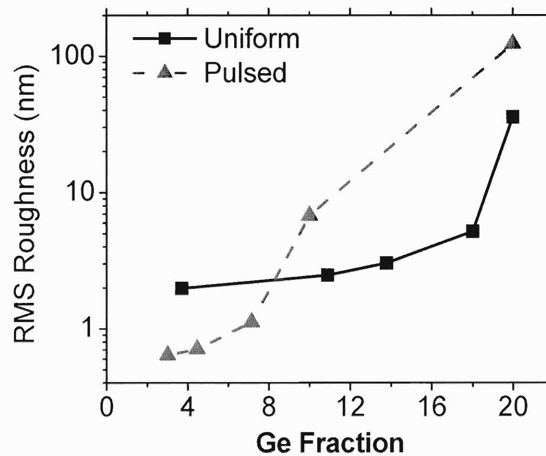


Figure 2-28: The roughness of the pulsed layers becomes very bad after about 9% Ge concentration. Such a high roughness will lead to a high TDD. The AFM images of these pulsed layer specimen are shown in Figure 2-29.

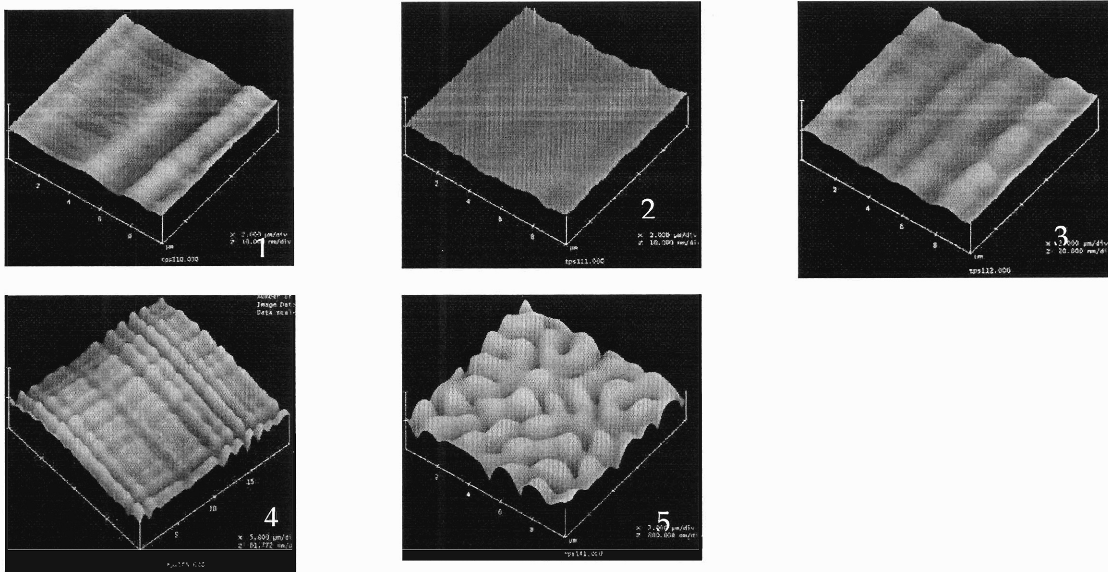


Figure 2-29: Roughness evolution in pulsed layers as the Ge content is increased. The pictures 1,2,3,4,5 corresponds to a Ge concentration of 3%, 4%, 7%, 9%, 20% respectively. The scan area of all the AFM is 10μm×10μm.

In addition to the roughness problem, the pulsed layer also has a metastability problem. When a pulsed layer is grown thick it is still far from the equilibrium curve for relaxation. The primary reason behind the fact that the pulsed layer does not relax efficiently is based on the way that the relaxation has to occur during this mode of growth. A very thin layer is deposited which gives the excess strain in the layer. This excess strain is the driving force for dislocation motion. The strain is relaxed as the dislocations glide and this results in a reduced driving force for dislocation. Thus, the dislocations are not able to relax the layer completely after one pulse of growth. Over many pulses, the layer follows a different curve for relaxation which is much higher than the equilibrium curve (Figure 2-30).

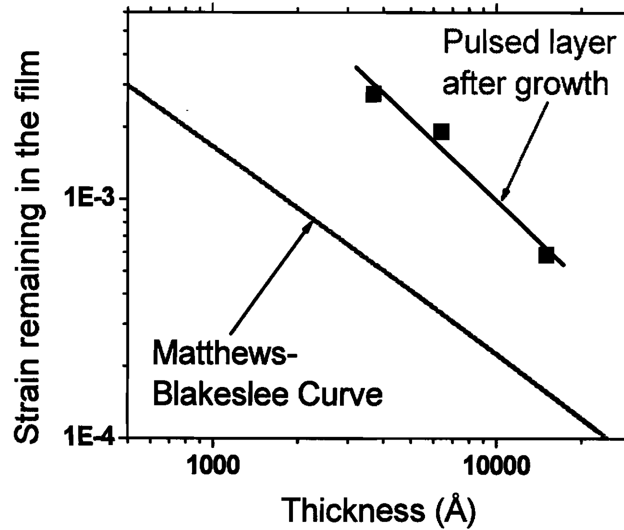


Figure 2-30: Pulsed layer is very metastable after growth and does not relax efficiently.

The high surface roughness and metastability which occur on using the pulsed layer technique, severely limit its usefulness. Therefore alternate ways to reduce the growth rates have to be found.

2.8. Low temperature growth

The traditional means to reduce the growth rate is to lower the growth temperature (Figure 2-31). Since the surface mobility of the atoms is reduced, the resulting layer has less surface roughness. However, since dislocation velocity is

exponentially dependent on temperature, decreasing growth temperature will increase threading dislocation density via Equation 2-7. Fortunately, dislocation nucleation for layers on (001) surfaces is difficult and there is a window in which decreased growth temperatures does not increase threading dislocation density. Instead, a metastable strain is introduced, thus preventing complete relaxation of the strained layer. Therefore, achieving a completely relaxed low TDD layer requires the removal of the metastable strain without increasing dislocation nucleation.

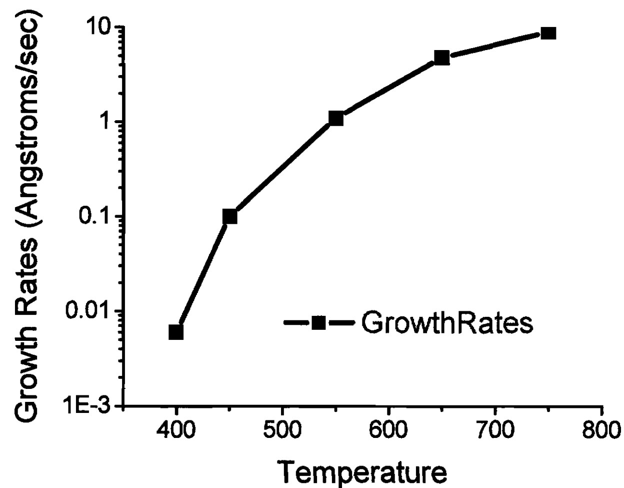


Figure 2-31: Growth rates vary over by many orders of magnitude with temperature changes. Such a range is required for controlling the strain rate in the layer

2.8.1 Metastability in layers grown at lower temperatures.

As discussed earlier, the dislocation velocity is exponentially dependent on temperature and therefore the glide of the dislocations is severely depressed at low temperatures. This causes a metastability to set in the layer (Figure 2-32). However, since the dislocation nucleation is also suppressed, this metastability does not cause excess dislocation nucleation.

One possible way for removing the metastability of this layer is by annealing of the layer at high temperature. The constraint on the relaxation scheme is that excess nucleation of dislocations should not take place.

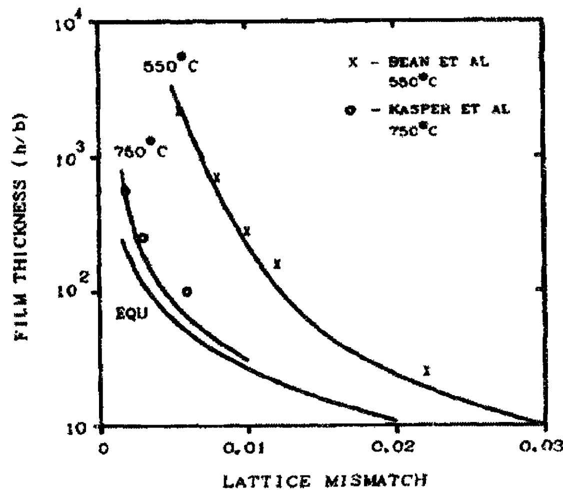


Figure 2-32: Metastability occurs in layers grown at lower temperatures[40].

2.8.1.1 Relaxation through annealing

Relaxation of the layer can be achieved by high temperature annealing. This is because the glide of dislocations is thermally activated. Some amount of residual strain remains after annealing because the strain is not enough for the dislocations to glide. Since dislocation glide is thermally activated, the dynamics of relaxation are much faster at higher temperatures and the residual strain remaining in the layer is lower at higher temperatures. (Figure 2-33).

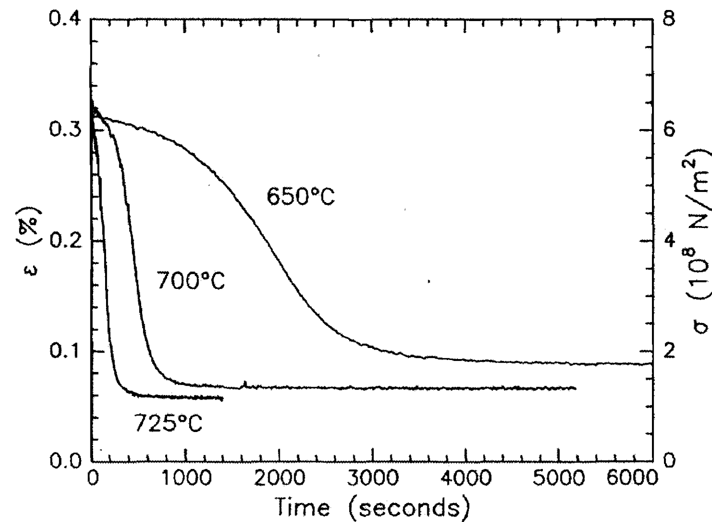


Figure 2-33: Annealing the layer reduces the strain remaining in the layer. The dynamics of strain relaxation is much faster at higher temperatures as compared to lower temperatures[41].

It is to be noted that during annealing, in addition to the dislocation glide, dislocation nucleation is also happening with a smaller rate due to normally higher activation energies of nucleation. However, as temperatures are increased during anneals, the rate of nucleation also increases. Thus a high temperature anneal will help in

relaxation of the layer but will be detrimental because the dislocation densities will also increasing.

Ideally, an optimized time-temperature anneal procedure is required in order to relax the strain without any significant increase in the TDD. In order to arrive at such an optimized procedure, we need to look closely at the kinetics of nucleation and relaxation in the layer.

The strain relaxed by the layer in time *dt* is given by

$$d\varepsilon_1 = -\rho_0 b v_0 (\varepsilon_1 - \varepsilon_{eq})^m e^{\frac{-Q_v}{kT}} \times dt \quad \textbf{Equation 2-12}$$

While the layer is getting relaxed, the dislocation density is increasing as well because it is also thermally activated and is strongly dependent on the strain as well.

$$d\rho = \xi_0 (\varepsilon_1 - \varepsilon_{eq})^m e^{\frac{-Q_p}{kT}} \times dt \qquad \textbf{Equation 2-13}$$

Where, $d\varepsilon$ is the strain relaxed in time dt, $\varepsilon(t)$ is the strain remaining in the layer at a time t, b_{eff} is the effective burgers vector, ε_{eq} is the Matthews-Blakeslee equilibrium strain, Q_v and Q_p are the activation energies for the velocity and the threading dislocation densities respectively, $\rho(t)$ is the mobile threading dislocation density, v_0 and ξ_0 are material constants. Both the equations are coupled since strain relaxation will lead to reduction in the dislocation nucleation rate which will in turn lead to slower strain relaxation. The parameters used for these equations have been obtained from literature[42].

In order to see what the effect of the anneal treatment is on the strain and the dislocation density of the structure we propose two different anneals, and model the trends in strain relaxation and the dislocation density. One is a slower anneal and is referred to as gradual step anneal (Figure 2-34). The other anneal is a sudden high temperature anneal. The resulting strain relaxation and TDD for slower anneal is described in Figure 2-35, while for the faster anneal is described in Figure 2-36.

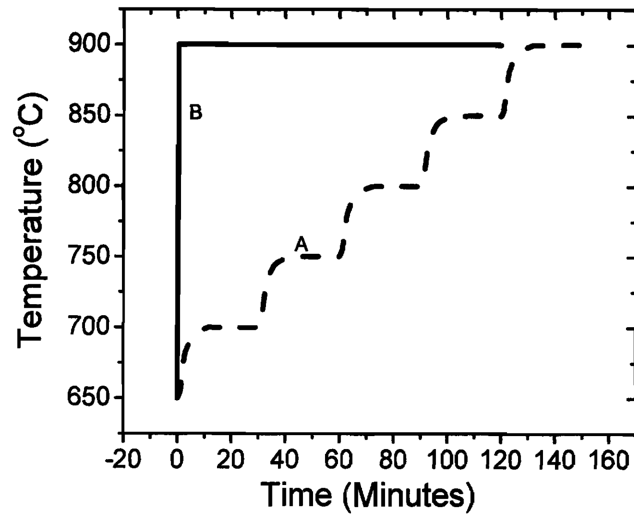


Figure 2-34: Two different anneals are engineered in order to look into the strain relaxation behavior and the dislocation density trends in a 10% SiGe layer grown at 650°C based on different anneal procedures. 'A' is combination of step anneals from 650°C to 900°C over a long time which 'B' is a step anneal to 900°C.

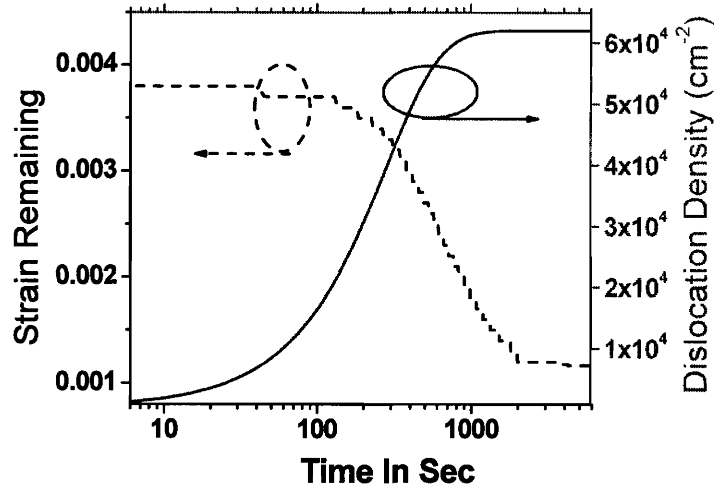


Figure 2-35: Strain relaxation and TDD evolution with time in a 10% SiGe layer subjected to a gradual step anneal described in Figure 2-34. The dislocation density after the gradual anneal is very low.

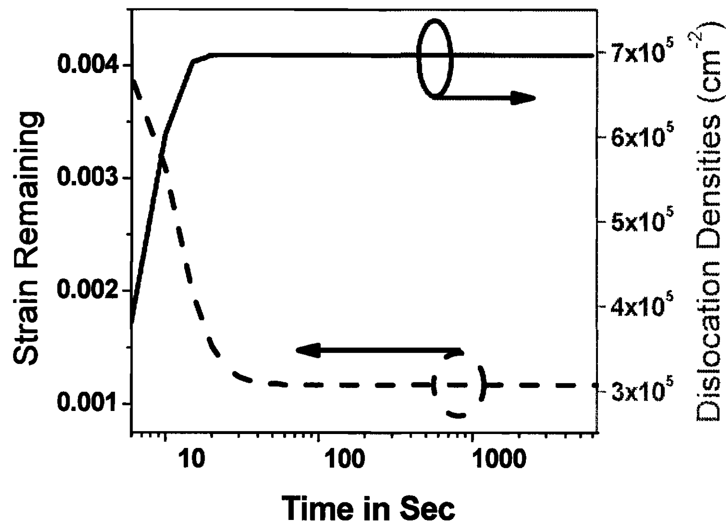


Figure 2-36: Strain relaxation and TDD evolution with time in a 10% SiGe layer subjected to a steep step anneal described in Figure 2-34. The dislocation density after the steep anneal is higher than in the gradually annealed sample (Figure 2-35).

From the above numerical exercise we arrive at the fact that a sudden high temperature anneal is detrimental to the TDD in a metastable layer because it leads to faster strain relaxation and therefore higher TDD. Thus, gradual anneal, in which the temperature is increased gradually from a lower temperature, is a way to achieve relaxation in the layer with lower TDD. This procedure works because dislocation nucleation is dependent on strain and temperature. The rate of dislocation nucleation at a low temperature is very low, but the glide of the dislocation is still happening at that temperature. After some time, the metastable strain in the layer has reduced to such an extent that increasing the temperature to a higher value would not increase the dislocation nucleation. However, at a higher temperature, the glide velocities of the dislocation are higher leading to improved kinetics of relaxation with suppressed nucleation. Thus, a gradual anneal could lead to a slow strain relaxation in the layer and smaller value of TDD.

2.9. Time-Temperature-Threading Dislocation Density Diagram (TTTDD Diagram)

Using the model developed earlier, we can plot threading dislocation density evolution with time at a constant temperature. If we plot iso-TDD contours we can get the time-temperature-threading dislocation density diagram, which will be referred to as TTTDD diagram. Using this diagram isothermal anneal pathways which lead to a low

TDD can be determined. The TTTDD is determined by using the nucleation-relaxation model that has been developed earlier.

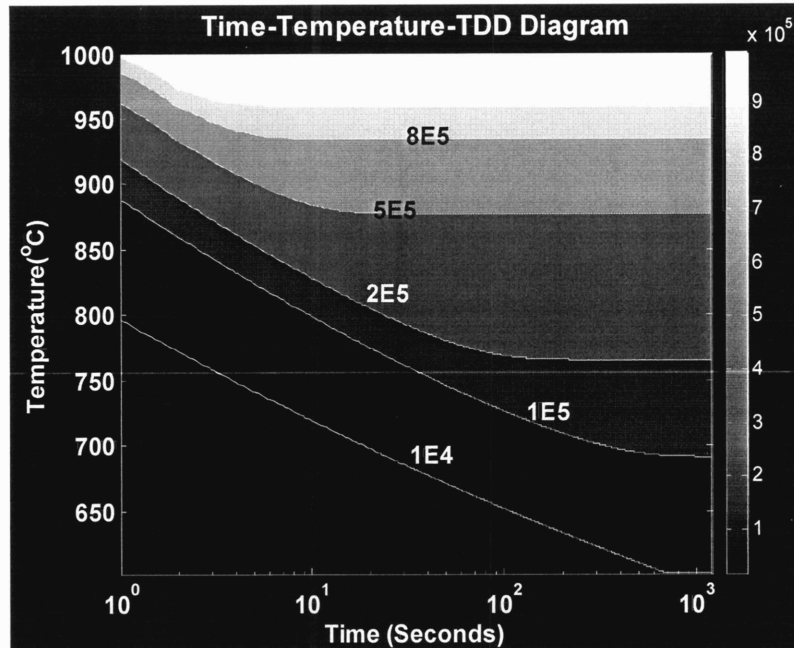


Figure 2-37: Time-Temperature-Threading Dislocation Density (TTTDD) diagram for a metastable 10% layer which is annealed isothermally. The diagram is helpful in determining the annealing path for a low TDD state in a layer.

2.10. Achieving thin 10% completely relaxed layer

2.10.1 Strategy for achieving thin graded buffer

While undertaking the study of TDD evolution and dislocation glide and relaxation in a mismatched layer, our aim was to achieve a thin completely relaxed

platform with a low TDD. If we take the strategy of manipulating the graded buffer kinetics that will suggest increasing the grading rate and reducing the growth rate so that the net TDD remains similar to a conventional graded buffer. However, in order to make a significant difference, we pursue the strategy of optimizing a single relaxed high mismatched (~10% Ge) layer. Once a completely relaxed high mismatch layer has been fabricated with low TDD it can be repeated to get an ultra-thin buffer. In order to decide on the Ge concentration of this high mismatched layer we will look at the formulation which leads to estimating the TDD evolution in the mismatched layer. It has also been shown in Section 2-4 that a relaxed high mismatched layer will have misfit dislocation very close to each other and these groups of closely spaced misfit dislocations will provide impedance to the gliding threading dislocations, leading to an escalation of the TDD.

From the TDD formulation, we see that the misfit dislocations start to affect the TDD from about 12-13% Ge concentration. Therefore, a Ge concentration of 10% is our target mismatch for achieving a thin relaxed low TDD platform. We cannot grow this layer at a high temperature since the plastic rate of relaxation for this layer is very high and will lead to a high TDD. Therefore, in order to achieve lower strain rate we will grow the 10% layer ~1500Å thin at 650°C. This layer will be very metastable and we will need to improve the relaxation of this layer through step annealing procedures as described in Section 2.8.1.1. It was seen from the model developed that using such a procedure the TDD is expected to remain at a low level of about 7×10^{4} cm^{-2}.

An experimental study was undertaken for determining the extent of relaxation in such a layer after annealing.

2.10.2 Metastability in 10% layer after annealing

A 10% Ge layer 1500Å thick was grown at 650°C and was annealed using step anneals. Since the temperatures in the UHVCVD system do not increase very fast, the steps of the Time-Temperature anneal profiles are not very steep (Figure 2-38). Once annealed, the extent of relaxation was determined using plan view TEM by measuring misfit dislocation spacing. According to the model discussed earlier, such time-temperature anneals are expected to relax the 10% Ge layer to about 7% Ge equivalent lattice constant. However, as measured from the plan-view TEM, such a time-temperature anneal relaxes the layer only to about 2.5% Ge concentration lattice constant (Figure 2-39).

In keeping with our aims of completely relaxing and maintaining the low TDD of this 10% layer, we need to improve its relaxation. Increasing the final annealing step to a higher temperature is not an option as the highest temperature that we can achieve in our reactor is 900°C. Therefore, an alternative procedure for relaxing the layer is required.

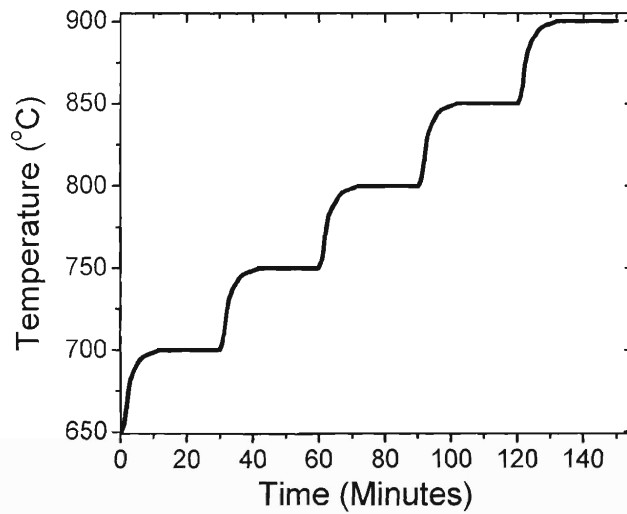


Figure 2-38: Step-anneal profile used for relaxing the 10%, 1500Å layer grown at 650°C.

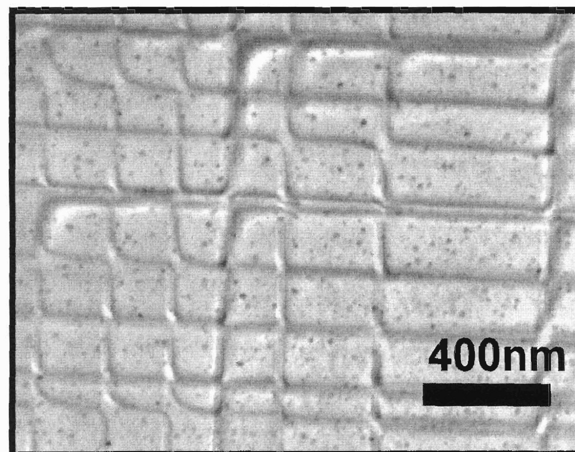


Figure 2-39: The 10% layer grown at 650°C and annealed using a step anneal profile from 650-900°C retains a lot of strain. The misfit spacing expected is about 40nm while the average spacing that we get in the layer is 180nm.

2.10.3 Strain relieving layer

The reason for the remaining metastability in the layer despite the time-temperature anneal, is primarily because the driving force for dislocation glide is reduced to an extent that the relaxation through the glide does not occur. In addition, since the layer does not have enough residual strain, the nucleation of the layer does not occur also. A high temperature anneal can nucleate more dislocations but the reduction in the metastability of the 10% layer is not expected to be removed significantly. Therefore, a procedure needs to be derived in which additional strain can be provided to these dislocations so that they can move and relax the layer.

This additional strain can be provided by growing the 10% layer thick so that the layer can achieve excess strain for gliding the dislocation. This way to provide excess strain is not pursued because the thickness of the layer required to relax the 10% layer significantly is large and the aim of achieving a thin layer is defeated.

In order to improve the relaxation of the 10% layer, we have looked at the mechanism of relaxation in conventional graded buffers. In a graded buffer, layers are relaxed to a level much more than what is predicted by the Matthews-Blakeslee equilibrium. This happens because the next step of the graded buffer relaxes the previous layer. Therefore, an alternate way of relaxing the metastable 10% layer consists of depositing a mismatched layer on top of the 10% layer for relaxing the 10% layer. This layer will provide excess strain to the dislocations in order to make them glide. By

modifying the composition and the thickness of this layer, we can change the amount of driving force that we provide to the dislocations in the 10% layer and help it relax to the level that is desired. We will henceforth refer to this layer as the strain relieving layer.

2.10.4 Relaxation of the metastable layer by using a 12% strain relieving layer

A 12% layer is used for reducing the metastability of the 10% layer. As described above, a 10% layer is grown at first with a thickness of about 1500Å and is subjected to a step anneal procedure. This relieves only slightly the metastability present in the layer. The final step is a 30 minute anneal at 900°C. At this temperature, a 12% layer is grown, which provides the necessary driving force to the dislocations for gliding.

By varying the thickness of this 12% layer we can change the amount of relaxation of this 10% layer. As described earlier, the amount of relaxation of the 10% layer has been determined based on the PVTEM measurements of the misfit dislocation spacing.

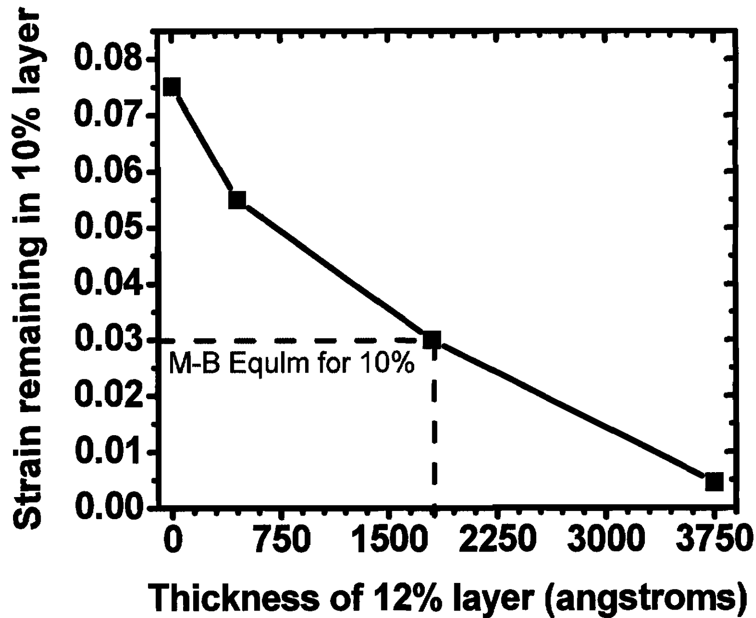


Figure 2-40: Relaxation of the 10% layer with thickness of 12% layer. If the 12% layer is 1800Å thick the 10% layer has already relaxed to its equilibrium lattice constant of 2.5%.

2.10.5 Phenomenon of relaxation of the 10% layer

As described above, a 10% layer grown at 650°C and subjected to step anneals from 650-900°C does not relax to its equilibrium value of an equivalent lattice constant of 7.5%. However, growing a strain relieving layer relaxes the 10% layer completely. A thick 12% layer will actually relax the layer to a higher extent than predicted by the equilibrium. To the author's knowledge, this is the first report of a layer relaxing below its Matthews-Blakeslee equilibrium value. The phenomenon of complete relaxation of the layer introduces many new important results which will not be explored in this thesis.

The phenomenon of the relaxation of the layer is described in the following figure. When the 12% layer starts to deposit on the 10% layer, which is relaxed to a lattice constant of 2.5% Ge, it starts to provide a driving force to the dislocations which then start to glide. The driving force keeps on increasing as the thickness of the 12% layer increases leading to an increasing relaxation of the 10% layer. As the underlying lattice constant continuously increases, the mismatch between the metastable 10% layer and the 12% layer decreases, making the critical thickness for the 12% layer larger. This phenomenon has been described in the figure below (Figure 2-42).

If the thickness of the 12% layer is larger than about 1800Å (Figure 2-40), it relaxes the layer to a lattice constant more than its equilibrium value. Shown is the XTEM of a layer which has been relaxed completely by growing a 3800Å, 12% layer (Figure 2-41).

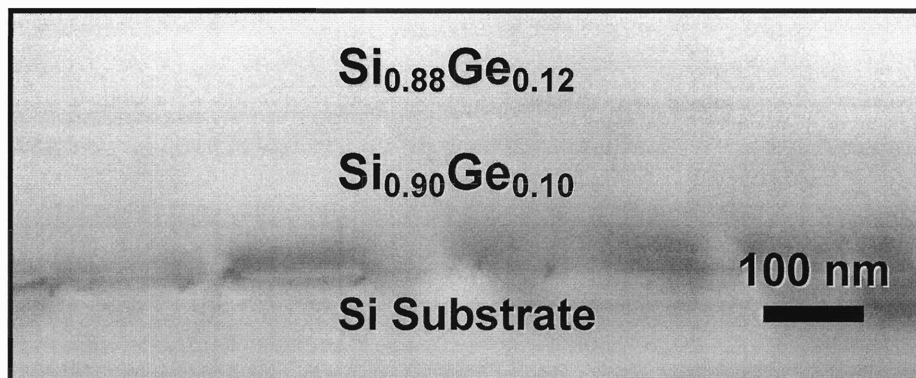


Figure 2-41: XTEM of the stack of layer showing the 10% layer and the 12% layer which completely relaxes the underlying 10% layer.

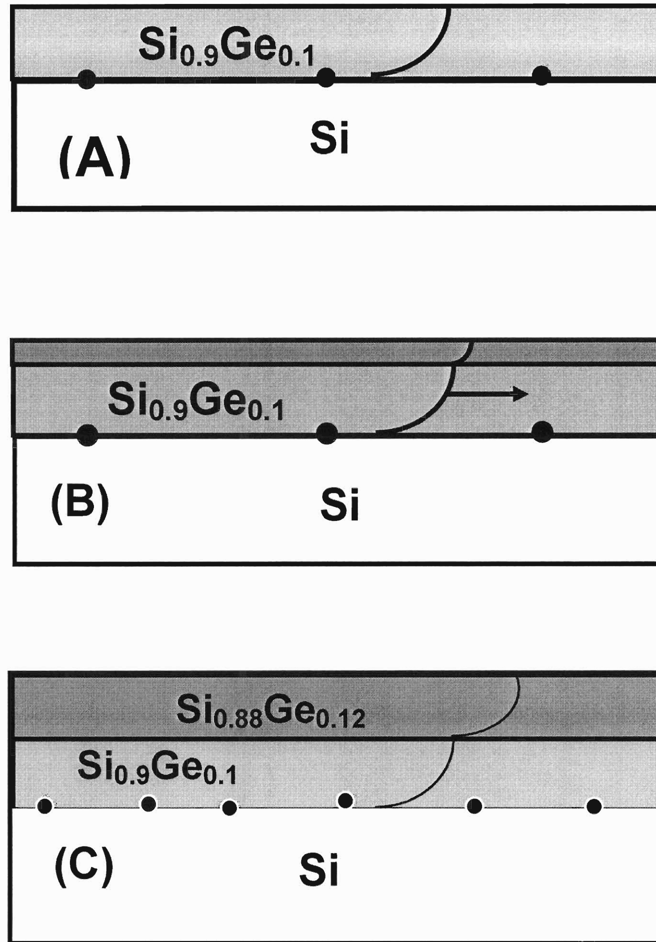


Figure 2-42: The above figure explains the mechanism of relaxation of the 10% layer by the use of a strain relieving layer. (A) the 10% layer after deposition and anneal is very metastable, (B) As the 12% layer is growing it is providing driving force to the underlying 10% layer. When the thickness of the 12% layer is below its critical thickness, it aids in the relaxation of 10% layer. (C) When the thickness of the 12% layer crosses its critical thickness the 10% layer is already very relaxed leading to a low mismatch between the 12% and the 10% layer. In addition the 10% layer seeds 12% with threading dislocations leading to a low TDD in the final structure.

2.10.6 Designing the strain relieving layer

The purpose of the strain relieving layer is to provide a large driving force to the dislocations in the underlying 10% Ge layer to help them move while maintaining a low threading dislocation density in the layer. In accordance with the aims of producing a thin buffer, a minimum possible thickness of the strain relieving layer is desired.

In order to obtain low threading dislocation density in the strain relieving layer, additional nucleation in the strain relieving layer should be prevented by keeping the strain relieving layer below its critical thickness. A high Ge concentration layer has a smaller critical thickness and, while it is growing on top of the 10% Ge layer, can exceed its critical thickness before relaxing the underlying 10% Ge layer to a sufficient extent. For example, a 20% Ge layer exceeds its nominal critical thickness when the underlying 10% Ge has barely relaxed to an equivalent lattice constant of 3% Ge. When a high Ge content layer exceeds its critical thickness, it will nucleate a higher threading dislocation density making it undesirable for the purpose of a low threading dislocation density buffer. For example, a thick 12% layer completely relaxes the underlying 10% layer while maintaining a low threading dislocation density of $2.5\times10^4 cm^{-2}$ as compared to a higher threading dislocation density of about $7.7\times10^4 cm^{-2}$ which is developed on using a thick 15% layer as the strain relieving layer. The above arguments suggest using a lower Ge concentration for the strain relieving layer in order to maintain a low threading dislocation density.

However, if we use a minimally mismatched layer (such as a 10% Ge layer) as the strain relieving layer, the thickness that we need to grow in order to provide the required amount of strain in the layer is very large. For example, if we use a 10% Ge layer as the strain relieving layer, the thickness that we need to grow it in order to completely relax the underlying 10% layer is infinite. A 12% Ge layer is a good choice for a strain relieving layer as it can relax a 10% Ge layer to a lattice equivalent of 7% Ge in a reasonable thickness of about 1800Å (Figure 2-40) and retains a low threading dislocation density.

2.11. Ultra-Thin Buffer

2.11.1 Repetition of the 10% for getting TRUT

Our experiments suggest a means to engineer graded buffers with significantly reduced thickness. The 10% optimized stack as demonstrated in the previous sections, can be repeated two times to achieve a very thin relaxed buffer with a very low thread count. This is because a relaxed 10% layer (relaxed to an extent of ~7%) can be achieved in about 330 nm thickness which is about 21% per micron grading. This is about 2.1 times the grading of a standard graded buffer virtual substrate. Since these buffers utilize temperatures for relaxation, they are referred to as Thermally Relaxed Ultra-Thin (TRUT) buffers. A XTEM of a 17% TRUT buffer is seen below. The important characteristics of this TRUT buffer are described in the subsequent section.

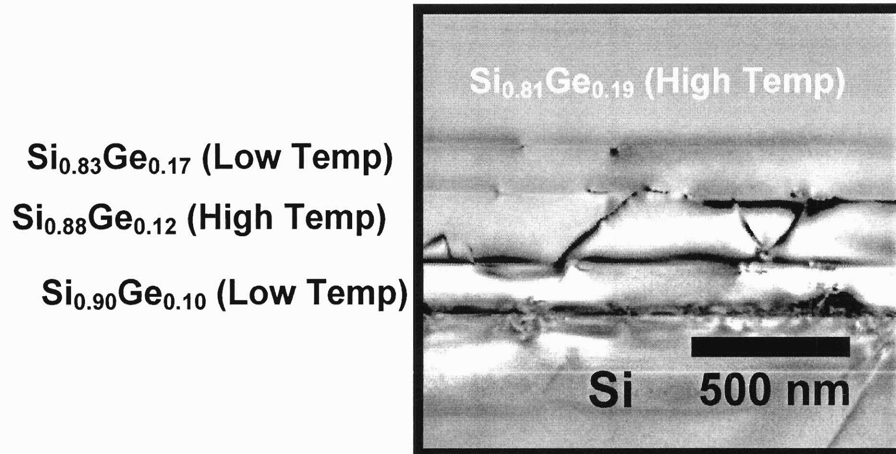


Figure 2-43: A 17% TRUT buffer. The first layer $Si_{0.90}Ge_{0.10}$ is grown at 650°C and then annealed using step anneal from 650-900°C. The $Si_{0.88}Ge_{0.12}$ layer is then grown at 900°C to relax the 10% layer. Similarly, a $Si_{0.83}Ge_{0.17}$ layer is grown at 650°C and then annealed using step anneal from 650-900°C. The $Si_{0.81}Ge_{0.19}$ layer is then grown at 900°C to relax the 17% layer.

2.11.2 Characteristics of a 17% TRUT buffer and comparison with conventional graded buffer

Our results reveal more insight into the need for compositionally-graded buffers to achieve low TDD, relaxed layers. The structure described above is essentially a graded composition buffer; we were forced back to that profile in order to achieve high dislocation velocities and prevent additional dislocation nucleation. However, from the earlier experiments we can engineer graded buffers with reduced thickness. In this

technique, different characteristics of the TRUT buffer are compared with the conventional graded buffer so that the strengths and weaknesses of the TRUT buffers are established.

The first characteristic that we will like to compare the two buffers is the thickness. In some cases the presence of greater buffer layer thickness is deleterious. For many III-V devices on Si, SiGe graded buffer approach is hindered by the relatively thick buffers. Due to high growth temperatures and significant thermal expansion coefficient difference between Si and Ge, the tensile strain that develops upon cool down of graded buffer limits the amount of III-V material that can be deposited on structure without formation of micro-cracks. In this work, our results demonstrate that, for the 10% Ge layer used in these experiments, we can reduce the buffer thickness 2.1X as compared to the conventional graded buffer. Thus, this development will be very helpful in applications requiring extremely thick graded buffers, as is the case for Ge on Si.

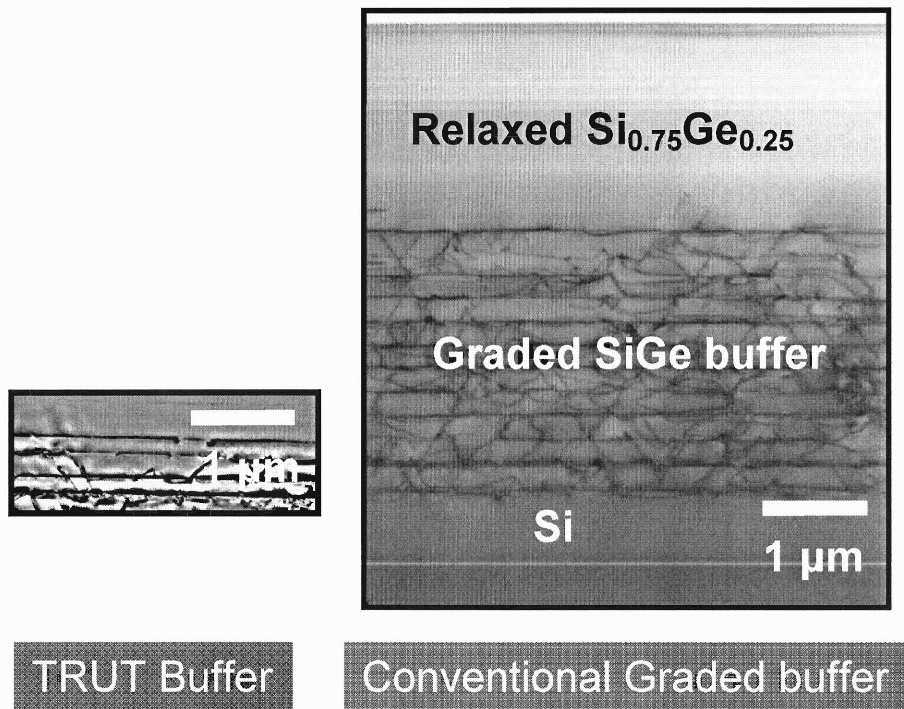


Figure 2-44: The TRUT buffer compared to the conventional graded buffer. Both of these pictures are at same scale. The TRUT buffer developed in this work can be more than 2 times thinner than the conventional graded buffer

The TDD of the TRUT buffer is very comparable to the conventional graded buffer. For a TRUT buffer of about 17% Ge composition, the TDD is measured to be about $8.8\times10^4 \pm 3.5\times10^4$ per cm^2. The conventional graded buffer of a similar composition has the TDD of about 1.4×10^5 per cm^2.

Compared to the 90 minutes required for the growth of the conventional graded buffer of 20% Ge concentration, the TRUT buffers require a total time of about 9 hours. However, the total volume of gas used in growth of TRUT buffer is much less than that used by the conventional graded buffers. This makes the technique of TRUT buffer

attractive for use in molecular beam epitaxy (MBE) systems. The longer time steps in the TRUT buffers are due to annealing steps and the temperature drop in the high heat capacity UHVCVD reactor. If this procedure was carried out in rapid-thermal chemical vapor deposition system, the time required for the complete process could be reduced to about 6 hours. This is still a low throughput process as compared to a simple high temperature graded composition layer process. However, the time of anneal for the 10% layer can also be optimized leading to a reduction in total time to such an extent that the TRUT buffers may become commercially promising.

In addition, this work sheds new light on the efficacy of graded composition layers to form relaxed lattices on common substrates. Our experiments set out to use graded-layer kinetics on a uniform composition film. In doing so, we were required to employ compositions and thicknesses that forced the structure back to an overall graded-composition structure very similar to a graded composition layer. Therefore, this work connects the physics behind the dislocation kinetics directly to the dual requirements of complete relaxation and low threading dislocation density, revealing the universal nature of graded composition layers in lattice engineering.

2.12. Summary

In this chapter, we explore the detailed relationship between strain rate and threading dislocation density in uniform composition SiGe films on silicon substrates. We find that the rate of plastic strain relaxation and misfit dislocation distribution on the interface in a mismatch layer are the important parameters which affect threading

dislocation density in a mismatched SiGe layer. 10% Ge is determined to be the highest Ge concentration layer in which threading dislocation flow without relative impedance from the underlying misfit stress fields. Reducing the threading dislocation density in this 10% Ge layer on Si then requires slow plastic strain relaxation which could be achieved by growing the layer at a slower growth rate. If the slower growth is achieved at high growth temperatures the concomitant surface-induced roughness limits the threading dislocation motion, thereby necessitating increased threading dislocation nucleation for relaxing the film. Using lower growth temperatures to reduce the growth rates, however, leads to a poorly relaxed 10% Ge layer. Since, the window to achieve relaxation and low threading dislocation density is narrow, in order to aid with the visualization of this window; we have developed a Time-Temperature-Threading Dislocation Density diagram. The only possible kinetic path suggested is a gradual increase in the anneal temperature which results in a slow rate of relaxation leading to a low threading dislocation density with fair degree of relaxation. The remaining elastic strain can be completely removed by using a strain relieving layer on top which drives the threading dislocation with a larger force than is possible with an equilibrium single-layer system.

Repetition of this couple leads to completely relaxed layers with very low threading dislocation densities. This structure is essentially a graded composition region with an absolute maximum in the rate of composition increase with thickness. The results shed increased understanding on creating highly deformed thin films with low threading dislocation densities, as well as providing guidance on graded composition buffers that can be less than 50% of the thickness of conventional graded composition buffers.

Chapter 3. SiGe Based MOSFETs

The scaling in MOSFET has been the strategy to meet the ever increasing demands for performance in logic devices as well as memory devices. However, the present momentum for scaling cannot be maintained without innovations in technology. The present chapter describes the motivation for SiGe-based devices to improve MOSFET performance. Short flow MOSFET process, which is used for fabricating the long channel MOSFETs described in this work, is explained in detail. In addition, techniques used for experimentally measuring the MOSFET mobilities of the various devices discussed in this thesis are also described.

3.1. Motivation for SiGe-based Devices

Scaling, which is the process of reducing the sizes of the transistor, improves the device packing density and circuit speed. Some simple general approaches to scaling include constant field scaling, constant voltage scaling and quasi-constant voltage scaling[43]. Since 2001, scaling has reached the point where the horizon of the roadmap challenges the most optimistic projections for continued scaling of CMOS. The primary limitations of scaling has been in fabrication technology and device reliability as seen in less aggressive scaling of metal line width as compared to gate length of active devices. Limits to scaling include thermodynamic limits which are doping concentrations in source and drain, physical limits exemplified by tunneling through gate oxide and statistical fluctuation of body doping. The factory cost also increases exponentially with scaling. The most severe limitation occurs in the form of gate oxide thickness which

cannot be reduced below 15Å. The mobility of the carriers is rapidly degraded with scaling because of higher field required for scaling.

In the saturation mode

$$I_D = \frac{Z\mu_{eff}C_0}{L_{eff}}\left[(V_G - V_T)V_{DSat} - \frac{V_{DSat}^2}{2}\right]$$ **Equation 3-1**

Improving I_d can be achieved by either improving the transport characterized by μ_{eff} or by scaling which can change both the effective gate length and the capacitance. Because of the technological and the fundamental limits, the ability to maintain the performance trend can only be achieved through innovations in lithography, device structures and new materials.

3.1.1 New Materials Technologies

One option that remains is available for improving and maintaining performance trends in CMOS is the use of new materials in the channel. This focuses on improvement of the transistor drive current for improved circuit performance by enhancing the carrier mobilities and the saturation velocities. Mobility determines carrier velocity in an applied electric field. The parameter, μ, originates from the band structure of the material, and thus changes to the band structure can significantly change μ. Drift mobility is given by:[44]

$$\mu = \frac{e\tau}{m^*}$$ **Equation 3-2**

where, e is the elementary charge, τ is the scattering time, and m^* is the carrier effective mass. The low-field electron and hole mobilities in lightly doped bulk Si and Ge are shown in Table 3.1.[45]

Material	Electron Mobility (cm^2/Vs)	Hole Mobility (cm^2/Vs)
Si	1450	450
Ge	3900	1900

Table 3-1: Electron and hole mobilities in bulk Si and Ge[46]

The relationship between effective mobility and effective vertical field in bulk Si MOSFETs follows a universal relation regardless of substrate doping.[47] A schematic of the universal mobility curve for both electrons and holes is given in Figure 3-1 below.

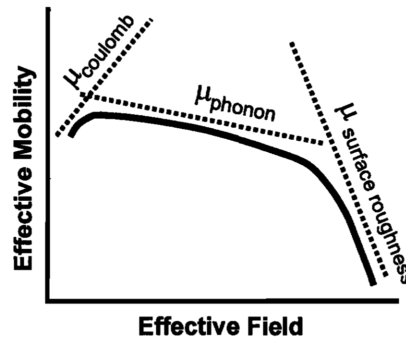


Figure 3-1:Universal electron and hole mobility relations for bulk Si MOSFETs. At low fields, carrier mobility is limited mainly by Coulomb scattering. At moderate vertical fields, carrier mobility is limited by phonon scattering. At high vertical fields, surface roughness scattering limits carrier mobility. Image courtesy of Minjoo L. Lee

At low vertical fields, where the density of carriers in the inversion layer is low, carrier mobility is controlled by Coulomb scattering, which becomes screened out at high vertical fields (where there is a high density of carriers in the inversion layer). At moderate vertical fields, carrier mobility is controlled by phonon scattering. Finally, at high vertical fields, carrier mobility is degraded by surface roughness scattering. The evolution of effective mobility versus effective vertical field is a key performance metric for SiGe-based heterostructure MOSFETs.

Carrier mobility is of fundamental importance in MOSFET design, since it determines the die size needed to extract a given drive current from a MOSFET. Since hole mobility in bulk Si is much lower than electron mobility, *p*-MOSFETs must be made larger than *n*-MOSFETs to obtain the same drive current, consuming valuable chip real estate. Moreover, the mismatch between *n*- and *p*-MOSFET size increases overall circuit

capacitance, decreasing the operating speed of logic elements. An approach to reduce the design complexity for VLSI designers will be to enhance the mobilities to symmetric values. Another approach can be to improve the mobility of both the carriers by equal factor resulting in the use of present technology but with improved performance[4].

Carrier transport can be modified using stress in one or more of the x-y and z directions. The stress can be localized (uni-axial) or biaxial. A uni-axial compressive strain on the channel can be produced by shallow trench isolation (STI) or shallow trench isolation in conjunction with SiGe source drains. Since the stress profile in the wafer will change corresponding to stress being applied at other places via STI or SiGe source drains, these methods of producing stress lead to a complex distribution of stress profile over the whole wafer.

	Transverse	Longitudinal	Z
p type-MOSFET	Tensile	Compressive	Tensile
n type-MOSFET	Tensile	Tensile	Compressive

Table 3-2: Desired uniaxial stress state for improved carrier mobilities[48].

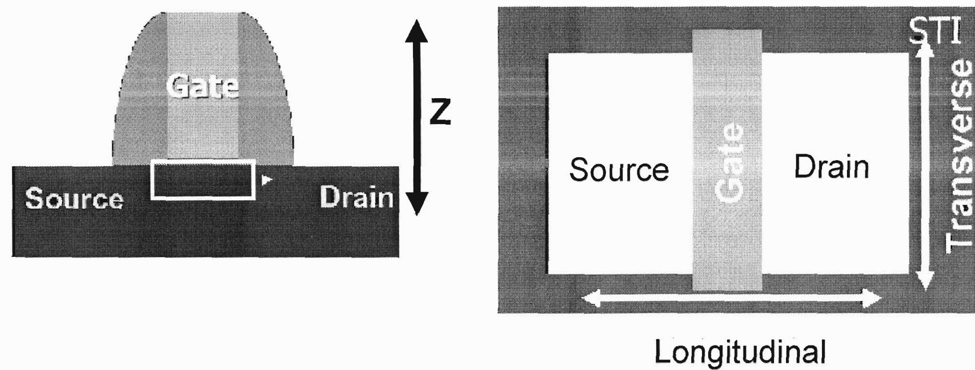


Figure 3-2: The various directions where uniaxial stress can be applied on a MOSFET for change in transport properties[48].

A much better way to apply strain in the device is biaxial strain by means of relaxed SiGe layers. The defect concentration in these layers is reduced to benign levels by the use of virtual substrates. A detailed description on virtual substrates has been given in Chapter 2. With a relaxed $Si_{1-x}Ge_x$ template, with x varying anywhere from 0 to 1, we can provide a range of lattice constant from Si to Ge.

Figure 3-3 plots the variation in bandgap, E_g, for $Si_{1-x}Ge_x$ alloys versus x. As the Ge content in the alloy is increased, E_g drops slowly, remaining substantially Si-like (X-valley minima) until $x \sim 0.8$, after which the conduction band minima crosses over to the L-valley. For $x > 0.8$, E_g drops rapidly to the value of bulk Ge ($E_{g,Ge} = 0.66$ eV).[49]

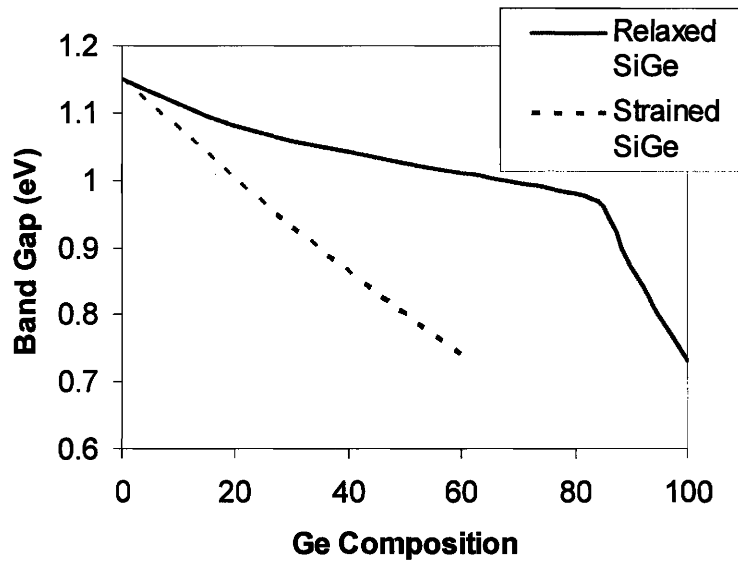


Figure 3-3: Bandgap variation in $Si_{1-x}Ge_x$ alloys. The conduction band minima crosses over from X to L at $x \sim 0.8$[50]. The strained SiGe band gap corresponds to SiGe layer grown on a Si substrate.

3.2. Short-flow MOSFET process

Mobilities in heterostructures are often deduced from Hall-effect measurements. In these experiments, modulation doping is generally employed in order to maximize the mobility in the quantum well, and care must be taken to avoid parallel conduction through layers outside of the region of interest. However, the Hall mobility is not always indicative of the mobilities that will be encountered in MOS devices with large carrier densities, oxide semiconductor interface and large band-bending near the oxide-semiconductor interface.

Therefore, in this thesis MOSFETs were fabricated using short-flow process to explore carrier transport in $Si_{1-x}Ge_x$ heterostructures. Instead of Hall-effect measurements, room-temperature electrical measurements using standard semiconductor parameter analyzers were performed. Despite the large MOSFET geometries, the short-flow process is useful for extracting data relevant to short channel devices. Specifically, the MOSFETs can be biased in the gate region with a very high vertical field, thus simulating the nature of the channel in a short channel device.

The gate lengths of the transistors vary from 50 to 200 μm and have a ring-shaped geometry. A schematic of the processing steps of the short-flow MOSFET is shown in Figure 3-4, and a full outline of the process flow can be found in Table 3-3. The mobility extraction is explained in Section 3.2.2.

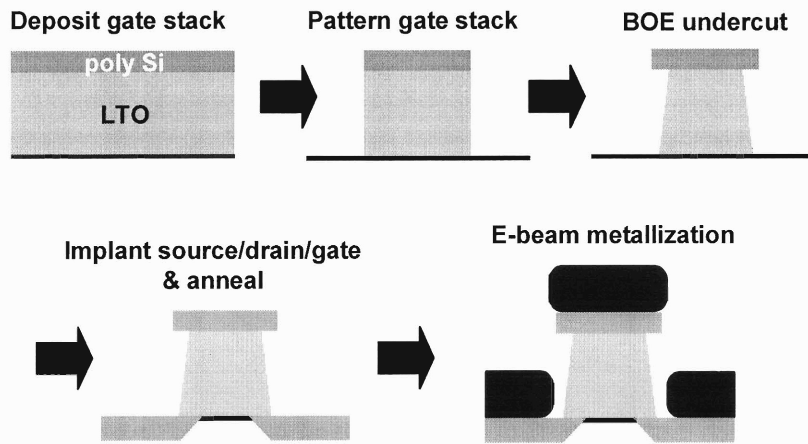


Figure 3-4: Short-flow MOSFET process flow

One of the critical step in the short-flow MOSFET process is blanket Ti/Al metallization performed via e-beam deposition at perpendicular incidence. Due to the

extreme geometry of the "T-gate" FET structure, breaks occur in the metal which isolate the source, gate, and drain regions without further lithography, much like a traditional liftoff process. BOE undercut is very important step of the MOSFET fabrication process.

The undercut helps the lift-off process and the metal breaks up after the metallization. The following SEM pictures show the profiles of the oxide after the under-cut and an instance where such a short occurs in case of the gate and source.

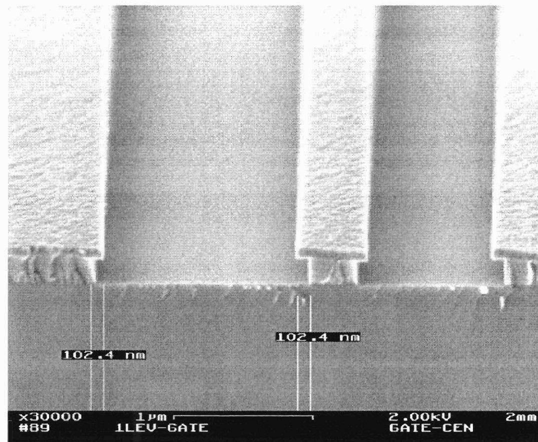


Figure 3-5: SEM picture showing the oxide undercut of 102.4 nm. Image courtesy: Jongwan Jung

Shorting of the gate and source/drain can occur if the undercut is not proper or the amount of metal deposited is more than critical. Such a short will render the device useless. Thus, extreme care should be taken during metallization to prevent such shorting.

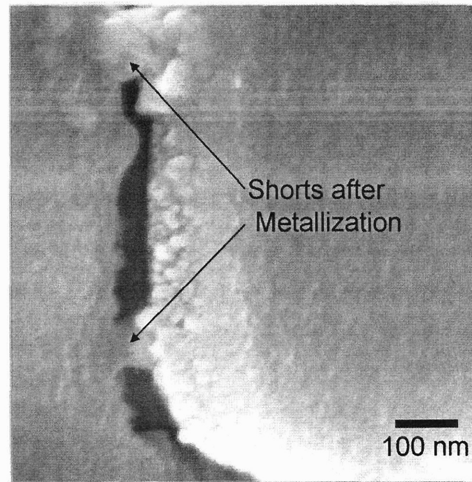


Figure 3-6: SEM picture showing a short between the Source and Gate.

The full MOSFET run can be completed in much less time as compared to full MOSFET runs which can take a month to complete. Essentially, the short flow MOSFET process allows the expedient exploration of Si/SiGe heterostructure channels under large vertical electric fields.

3.2.1 Short Flow MOSFET Fabrication Sequence

The following section outlines the MTL machines and recipes used in short flow MOSFET fabrication. Both MIT's Integrated Circuits Laboratory (ICL) and Technology Research Laboratory (TRL) were used in this process. In cases where more than one machine can be utilized for a given step, the preferred machine is italicized.

Table 3-3: MOSFET Fabrication Sequence

Process Module	Machine	Description
Gate Stack Deposition	ICL RCA	Modified RCA Clean (5-10 min. piranha clean, 15 s 50:1 H_2O:HF dip, 2.5-15 min. SC-2 clean)
Gate Stack Deposition	ICL tube 6C	Deposition of 3000 Å LTO (recipe 45Amin, 400°C; average deposition rate ~ 50 Å/min.) Total Time: 5 hours
Gate Stack Deposition	ICL tube 6B	Deposition of 1000 Å poly-Si (recipe 620DEP, 600°C; average deposition rate ~100 Å/min.) Total Time: 2 Hours 30 min
Backside Clear	ICL Coater6	Coat wafers with HMDS with a subsequent coat of 1 µm positive resist and post bake (Recipe T1HMDS/ Saur1)
Backside Clear*	ICL oxide	Remove backside native oxide with 5 s BOE dip
Backside Clear+	ICL *AME5000* or ICL LAM 490B	Remove backside poly-Si (Recipe on AME 5000: Duheon CP, with Main etch 35" and other etch 45"). On LAM 490B (Recipe: Backside Poly, etch time:3300 Å/min, etch time:30")
Backside Clear+	ICL oxide	Remove backside LTO with 7:1 BOE dip for 90 seconds.
Backside Clear	ICL *pre-metal* or *ICL* asher	Remove photoresist with 10 min. piranha clean (blue piranha only) or standard asher recipe for 3 minutes. If using asher for PR removal, before the next step do green piranha (8 minutes) followed by blue piranha (8 minutes) and 50:1 HF dip (30 seconds) to get rid of oxide formed in asher.
Frontside Patterning	ICL Coater 6	Coat wafers with HMDS with a subsequent coat of 1 µm positive resist and post bake (Recipe T1HMDS/ Saur1)
Frontside Patterning	ICL I-stepper	MASK: SiGe-6inch-1level.gds Recipe: JONGWAN-ONEL Step Distance: X:22500, Y:11500 X_m: -11000, X_p: 11000, Y_m:0, Y_p: 11000
Frontside Patterning	ICL coater6	Develop and Post-bake photoresist. Recipe: DEV6.
Frontside Patterning*	ICL oxide	Remove frontside native oxide in exposed regions with 5 s BOE dip
Frontside Patterning+	ICL AME5000	Remove frontside poly-Si (Recipe: Duheon CP, with ME 35", OE: 15". OE:10").
Frontside Patterning+	ICL AME5000	Etch frontside LTO in exposed regions with Jongwan LTO, leaving ~400 Å LTO remaining. Etch time: 62"
Frontside	ICL UV1280	Use ellipsometry to verify pre- and post-etch LTO

Patterning		thicknesses and calibrate LTO etch rates
Frontside Patterning	ICL asher	Remove photoresist with standard asher recipe
Frontside Patterning	ICL Premetal-Piranha	Blue Piranha (7 minutes). Complete removal of PR.
Frontside Patterning	ICL Premetal-Piranha	LTO under-cut. Calibrate etch rate of 50:1 HF dip using dummy wafers. Undercut around 500-1000 Å. Estimated time for undercut: 2 minutes 20 seconds.
Implant	Outside vendor	BF_2 or As, 30 keV, 4×10^{15} ions cm^{-2} consisting of 4 identical implants at 90° rotation. Tilt: 10° - 20°.
Implant	ICL Premetal-Piranha	Post-implant clean; 7 min. Blue piranha, 7 min. Green piranha, 30s 50:1 HF dip.
Implant	ICL RCA	Modified RCA. Piranha:5 min., SC-2: 7 min. No HF dip.
Implant	ICL 5B	40 min. anneal at 600°C in N_2 ambient. Recipe: 2A600. Actual temperature is 590°C. Total process time: 2 hours 10 minutes
Metallization	ICL pre-metal	Pre-metal clean; 5 minute Green piranha clean, 20 s 50:1 H_2O:HF dip
Metallization	ICL e-beam CMOS	Deposit 500 Å Ti + 1000 Å Al on wafer fronts. Ti: Deposition rate: 2 Å/s, Tooling Fac.: 160% Al: Deposition rate: 3 Å/s, Tooling Fac.: 190% Thicker deposition can lead to short circuit between gate and S/D. Normal Incident planetary: 4 wafers /cycle.
Metallization	ICL e-beam CMOS	Deposit 5000 Å Al on wafer backsides. Normal planetaries. Tooling Factor: Same as in Bindfile.
Sinter	TRL tubeB3	40 min. anneal at 400°C in N_2/H_2 ambient

*Optional

+Measurements needed to confirm step.

3.2.2 Transistor layout and mobility extraction

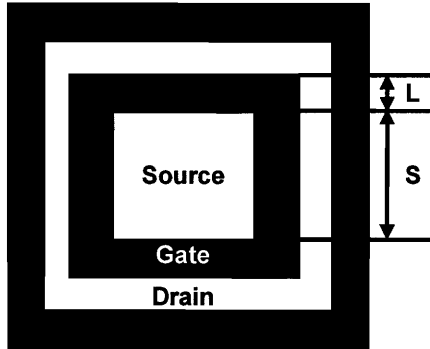


Figure 3-7: Planar view of MOSFETs measured in this thesis

L (μm)	200
S (μm)	250
A_{gate} (μm^2)	360,000
G	0.138

Table 3-4: Relevant dimensions of MOSFETs characterized in this work.

Effective carrier mobilities were extracted from measurements of the drain current in the linear regime, given by

$$I_D = \frac{W}{L}\mu_{eff}C_{ox}(V_{GS} - V_T)V_{DS} \qquad \textbf{Equation 3-3}$$

where W/L is replaced by the inverse of the geometry factor, G. $|V_{DS}|$ was set to 0.1 V for all measurements. A more detailed description of the derivation of G was given by Armstrong[51].

3.3. Summary

The present chapter describes the motivation for SiGe based devices to improve MOSFET performance, and the effect of strain on the band structure of SiGe crystals. Short flow MOSFET sequence, which is used to fabricate the MOSFETs described in this thesis, has been outlined. Mobility extraction procedures which use linear regime I_D-V_G characteristics have also been described.

Chapter 4. Diffusion coefficient formulation for estimating diffusion of Ge in SiGe single crystals

Mono-crystalline layers of relaxed $Si_{1-x}Ge_x$ can be grown epitaxially over the entire composition range of $0 \le x \le 1$. It is very important to know the layer structure and quality of material after high temperature treatment of strained heterostructures. One important parameter to model is the Ge fraction in layers after the temperature treatment of a heterostructures. In view of the various limitations in experimentally observing Ge fraction, and also the tedious, expensive and time consuming nature of experiments, it is essential to have a diffusion model which will help to estimate the diffusion characteristics of Ge in SiGe single crystal. As ITRS 2003 reiterates "In view of the need to increase carrier mobilities in the channel, the modeling of stress and strain and their influence on diffusion and activation has become vital, especially for strained Si, SiGe, and for SOI structures[5]."

This chapter focuses on this problem by formulating a diffusion coefficient which incorporates the complex Ge diffusion behavior and also strain effects on Ge diffusion. Such a model will be very helpful in thermal stability analysis of heterostructures as well as other heterostructures such as etch stop layers. We first start with different techniques to measure interdiffusion followed by the diffusion mechanisms for Ge diffusion in SiGe single crystals and then move on to describing the diffusion model itself. The effect of including strain in the diffusion is also discussed.

4.1. Interdiffusion observations in SiGe heterostructures

Accurately measuring the thickness and Ge concentration for each of the epitaxial Si layers in various devices such as single channel and dual channel MOSFETs and HBTs is essential for effective process control of these devices. For example, the most significant measurement parameters in the case of dual channel and single channel heterostructures are the thickness and the concentration of Ge in the active layers which can change during high temperature processing. These measurements are necessary because both parameters have a large influence on the device properties. Interdiffusion observations and/or diffusion coefficient measurements can be achieved from either obtaining a composition profile after thermal processing (e.g. SIMS) or, relating some observed property to the diffusion coefficient such as X-ray intensity from superlattice peaks.

Normally the composition measurements in a thin film are done using X-ray diffraction (XRD). However, in the case of heterostructures the X-ray intensities are very much reduced because of very small thickness (~50-100Å). The intensities can be strengthened by providing a multilayer stack structure which builds up the intensities or by using a high-intensity x-ray source such as a synchrotron. However, such a layer will not be useful as a MOSFET heterostructure, and the XRD will also not be able to provide a depth profile of the composition. Hence, we have to look at other methods in order to determine the composition profile of the structure. In the following section we will

outline many major methods for the determination of composition in SiGe heterostructures.

Quantum wells and barriers normally arise in different devices depending on the Ge concentration in adjacent layers and the layer itself. Figure 3-3 plotted the variation in band gap, E_g, for $Si_{1-x}Ge_x$ alloys versus x. As the Ge content in the alloy is increased, E_g drops slowly, remaining substantially Si-like (X-valley minima) until $x \sim 0.8$, after which the conduction band minima crosses over to the L-valley. For $x > 0.8$, E_g drops rapidly to the value of bulk Ge ($E_{g,Ge} = 0.66$ eV).[49] A technique which uses emitted frequency of light to determine the band gap and hence, the composition is photoluminescence (PL). PL is a powerful technique for detecting the interdiffusion in quantum well (QW) which is reflected in the form of a PL peak energy shift. Interdiffusion in a single quantum well (SQW) causes an increase in the SQW width and hence, reduces the quantum confinements. Thus, a blue shift is observed in PL measurements because of interdiffusion. The diffusion coefficients can be derived from luminescence peak energy blue shifts observed in annealed samples. In addition, PL measurements will show the interface quality by the PL intensity versus the photon energy measurements. A high quality of crystallinity is required for efficient luminescence. Also, if the interface quality is worsened, it shows up in the PL measurements as perturbations[52-54]. XTEM can also be used in conjunction with the PL to observe the increase in the SQW width.

However, if the SiGe layers are not relaxed, their band structures and band gaps do not match relaxed SiGe layers of the same composition and hence, effects of strain on

the band gaps should be taken into account. This is a complicated effect and it limits the use of PL in strained heterostructures. In addition to X-ray comparison of in plane and out-of plane lattice constant to determine the amount of biaxial strain, XTEM can also be used to qualitatively observe the strain. The strain is manifested as contrast difference in a XTEM image by using a <004> diffraction condition for <100> wafers.

Contrast can also be seen in TEM between different layers of the heterostructure based on their ability to stop electrons. Thus, a heavier material (e.g. Ge) is more effective in stopping electrons as compared to a lighter material (e.g. Si) and hence, Si appears to be lighter as compared to Ge. Thus, an abrupt contrast difference which occurs in the adjacent layers because of different Ge concentration is seen to reduce to more graded contrast after thermal annealing. This happens because of diffusion of Ge.

Secondary ion mass spectrometry (SIMS) is an offline analytical technique that can provide independent measurements of film thickness and Ge concentration on each layer with a high degree of sensitivity and accuracy. SIMS is often considered to be the benchmark technique for thickness and elemental composition measurements. A typical scan, however, can take an hour or more and results in the destruction of the device, making the technique inappropriate for production monitoring. Further, SIMS suffers from measurement artifacts, such as enhanced count rates at interfaces, which can distort the true SiGe concentration profile of a multilayer film stack. While low concentration dopants are easy to quantify in Si using SIMS, high concentration elements are subjected to "matrix effect". The matrix effect[55] refers to the phenomenon in which the intensity of

sputtered atoms is not linearly proportional to the concentration. SiGe layers with high Ge concentration display the matrix effect, leading to an underestimation of the Ge concentration. Thus, a process like SIMS lends itself to doubt in many cases and should be calibrated with other techniques, such as RBS before actual measurements.

X-ray techniques have been used for many years to characterize semiconductor materials in a laboratory environment. These techniques have excellent intrinsic measurement capability but are limited in providing high throughput, ease of use, low cost and a small spot size. The primary X-ray methods can be broadly divided into three different techniques: X-ray diffraction (XRD), X-ray reflectometry (XRR) and X-ray fluorescence (XRF)[56].

XRD provides the most comprehensive information on film characteristics since stress, thickness and Ge concentration can be theoretically extracted from diffraction spectra. For thickness measurements, XTEM is the most accurate method to be used in conjunction with other methods. Small-spot systems are only now becoming available due to recent advances in X-ray hardware and algorithm software. But these systems have a high cost of ownership and have not been fully proven in a production environment.

XRD is a very sensitive technique for measuring the inter-diffusivity in compositionally modulated thin films. To an approximation the integrated intensity of the successive low angle reflection is proportional to the square of the corresponding Fourier

coefficient of the composition profile. Thus, the interdiffusion in a modulated structure of wavelength λ can be obtained by monitoring the intensity as a function of annealing time.

$$\frac{d}{dt}\ln\left(\frac{I(t)}{I(0)}\right)=\frac{-8\pi^2}{\lambda^2}D_\lambda \qquad \text{Equation 4-1}$$

Since, we will be able to strengthen the intensity of a particular wavelength by growing a superlattice; X-ray diffraction is very sensitive in superlattice structures. If there is an arbitrary composition, then Fourier analysis should be considered. The diffusivity can be modified by both strain and composition.

XRF can be used to measure Si and Ge composition individually. However, the technique is limited to measurements on single layers on un-patterned monitor wafers. This is because X-rays penetrate many microns into the substrate, in turn causing the Ge signal that is collected from all of the layers to be counted as a single, "lumped" Ge signal.

Finally, XRR can provide information on multilayer thickness, density and interfacial and surface roughness, both independently and simultaneously. But XRR suffers from a poor signal-to-noise ratio, since the electron density contrast between the various SiGe layers is low. This results in poor sensitivity to small changes in layer thickness.

Another way to observe interdiffusion is a radiotracer technique in which radioactive atoms of ^{31}Si or ^{71}Ge are implanted. They are then isothermally diffusion annealed. Subsequently, serial sectioning is done using ion beam sputtering and radioactivities are then measured. A concentration profile can thus be constructed by determining the radioactivities which are directly proportional to the isotope concentrations. Stoichiometry can be checked by Rutherford Backscattering spectroscopy (RBS). Appropriate solutions of the diffusion equation are fitted and the diffusion coefficient is thus determined[57].

Spectroscopic ellipsometry (SE) is the most widely used method for determination of the thickness and concentrations in present industrial setups. However, it has poor sensitivity to SiGe stacks. With improved measurement calibrations, SiGe stack metrology systems based on spectroscopic ellipsometry can now achieve good precision.

MOSFET heterostructures, which will be discussed later, are typically very thin and hence can be quite difficult to analyze. This is because, while RBS cannot resolve layers which are of the order of the typical layer thickness, the limitations of SIMS have been discussed earlier. Also, since the layers are thin, the X-ray intensities are too low to be differentiated from noise. Most of the heterostructures mentioned in this thesis, have very steep concentration gradients and are not well resolved by techniques such as SIMS. A coefficient incorporating strain and composition parameter is formulated below.

4.2. Diffusion Model formulation

4.2.1 Diffusion Mechanisms

Solid Si and Ge occur in the diamond cubic lattice structure. There is a 4% lattice difference in between the two materials. They are both iso-electronic. Hence, both of them are highly soluble in each other and have a lens shaped phase diagram as shown in Figure 2-1.

The diffusion of Ge can occur both through point defects and lattice site exchanges. Below 925°C point defect dominated diffusion takes place because of its low activation energies[53]. Thus, below 925°C, they can diffuse with both interstitial and vacancy diffusion mechanism (each having its own pre-factor and activation energy) and the other methods of diffusion are not important.

Ge self diffusion in Ge occurs through the vacancy mechanism, while in Si, self diffusion occurs both through interstitial related and vacancy mechanism. The net diffusion coefficient taking into account both the interstitial and the vacancy contribution[58] is given as:

$$D = D_i \exp\left(\frac{-E_i}{kT}\right) + D_v \exp\left(\frac{-E_v}{kT}\right) \qquad \text{Equation 4-2}$$

Where,

$E_i = H_f^i + H_m^i$

$E_v = H_f^v + H_m^v$

and,

H_f^v = Formation energy of the vacancy

H_m^v = Migration energy for the vacancy etc.

The ratio of interstitial to vacancy diffusion is important for determining the total vacancy contribution to diffusion. Ge has 0.6-0.8 vacancy contribution in Si, Si has 0.4 vacancy contribution in Si. However in Ge, both of them have vacancy diffusion and hence their E_a is approximately the same[58]. As atom size increases the fractional vacancy contribution increases because the interstitials become increasingly harder to form.

Several experiments have been performed to determine the activation energies and the pre-exponential factors but no consensus has generally been obtained. However, it is believed that the activation energy for Ge interdiffusion decreases with increasing Ge concentration from 4.7eV in pure Si to 3.2eV in 50% SiGe. Thereafter, the activation energy appears to remain constant with further increases in Ge concentration.

Real diffusion data can always be fitted with $D = D_0 \exp\left(\frac{-E}{kT}\right)$, which will include the effect from both the terms for the interstitial and the vacancy mechanisms. Hence, the pre-exponential factor D_0 as well as activation energy E will be somewhere between the values for interstitial and vacancy.

We will describe the inter-diffusivity later in this chapter. Considering the two diffusion mechanisms, the inter-diffusivity in SiGe will also have to be written as the sum of separate terms for the two-point defect mechanisms. The degree to which each mechanism operates depends on concentration.

Since the vacancy mechanism dominates, interdiffusion in Ge-rich alloys and in Si, the interstitial mechanism is not taken into account. In addition, the interstitial mechanism is much slower than the vacancy mechanism and hence the assumption is strengthened[59]. Therefore, we accept Cowern's assumption that interstitial effects are negligible in low as well as high Ge concentration films[60].

Diffusion in the polycrystalline state is much higher than in the monocrystalline state, because grain boundaries serve as a fast path for diffusion. Thus, Zangenberg[58] measures the activation energy of 3.1eV at 50% Ge concentration which remains constant with increasing Ge concentration while McVay[61] seems to get this activation energy at around 30% Ge content. This implies that the polycrystalline has many other fast diffusion paths and hence, diffusion occurs faster than the mono-crystalline case. However, for our case we are interested in the single crystal structures and hence, use Zangenberg's data for the diffusivity.

4.2.2 Formulating Ge diffusion coefficient in SiGe single crystal structures

Many groups have looked at both the self diffusion and the interdiffusivity of Ge in SiGe single crystal, employing a variety of experimental techniques including SIMS[58], XRD[59], tracer diffusion[62,63] and TEM[64]. In addition, efforts have also been made to physically describe the process of diffusion and to arrive at the diffusivity through first principles calculations[65]. Even though the values of E_a and D_0 differ from the other values in the literature, the values of diffusivity are similar to the other reported values in the literature.

Among the different measurements of diffusivity taken in literature, focus has been on small ranges of Ge concentration[66] and higher temperatures[58,62]. For using the diffusivity to determine the thermal budget of a strained heterostructure containing layers with varying Ge concentration, diffusivity is needed for the full range of the Ge concentration and for temperatures as low as 600°C. Strohm's work evaluates the D_0 and the E_A for the diffusivity of Ge over the whole range of Ge concentration[63].

We found that the values of E_a and D_0 for self-diffusion in SiGe[63] follow Meyer–Neldel's equation. The presence of Meyer-Neldel rule in the SiGe system was seen by Dunstan[67] and Zangenberg[68] and the theoretical form of the rule (Equation 4-3) was derived by Khait[69].

$$\ln(D_0) = \left(\ln\left(\frac{\alpha^2}{6\Delta\tau}\right) - \frac{\Delta S_m}{k} - 3\ln(A) \right) + \frac{E_A}{kT_m}$$ **Equation 4-3**

where, D_0 is pre-exponential factor, α is the diffusion step, $\Delta\tau$ is the time scale, T_m is the absolute melting temperature, A is a dimensionless length scale fitting parameter, E_A is the activation energy and ΔS_m is the entropy of melting. The values of the T_m and α have been linearly interpolated between values for Si and Ge. The values for the factors used are the same as Zangenberg[68].

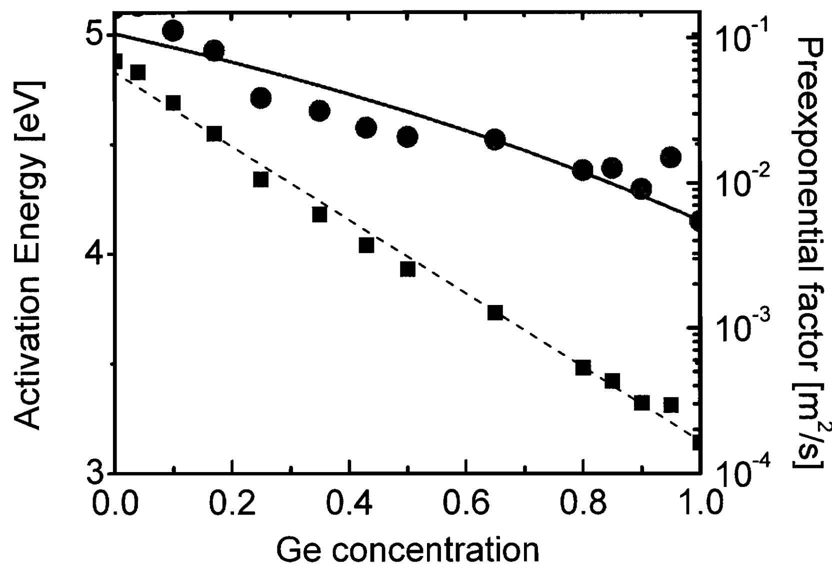


Figure 4-1:Meyer-Neldel equation in SiGe system. The E_A were fitted linearly and the D_0 were obtained from the Meyer-Neldel equation. The solid line shows the agreement of the measured values of pre-exponential factors with the derived values is good.

Figure 4-2 shows values of the D_0 obtained by using the data for E_A matches well with the experimental values. The diffusion coefficients obtained from these values of the D_0 and the E_A agree well with the other values from the literature. Interdiffusivity is given by including a Darken's term in the diffusion coefficient[59]. The Kirkendall effects are neglected[59] resulting in relatively simple modification of the self diffusivity of Ge in SiGe to interdiffusivity.

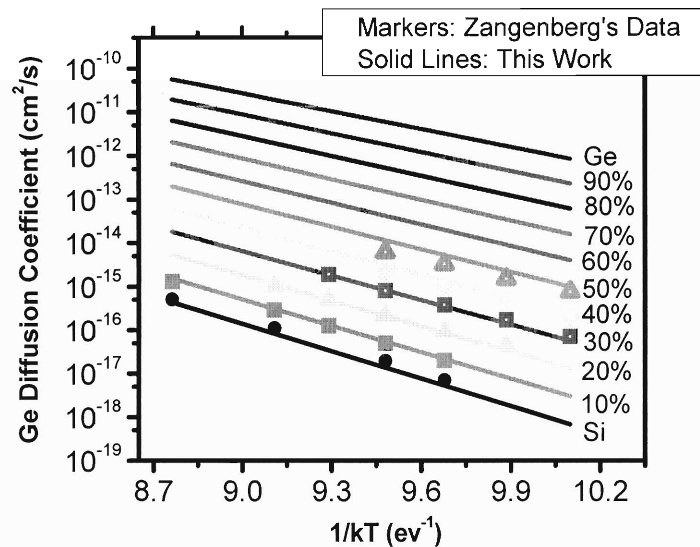


Figure 4-2: The formulated diffusion coefficient in this work matches well with the literature values[58].

For interdiffusion in SiGe interfaces there are 2 driving forces for diffusion[64], concentration difference and strain potential gradient. The effect of concentration on diffusivity has been described earlier. Strain affects the diffusion by providing an additional potential gradient which affects the flux[64], in addition to changing the activation energy of formation of point defects[60,70]. The interdiffusion is initially faster

when both the concentration and the strain potential gradients are higher. In addition, since the diffusion is point defect mediated, modification of the activation energy will increase or decrease the concentration of point defects and hence modifies the diffusion coefficient. Since such an effect also enhances interdiffusion in the early stages it is hard to decouple these two effects of strain. Thus, a factor called coupling strength (Q') has been introduced in literature can account for both these effects[59]. An important disagreement occurs in the literature on the value of coupling strength[66] Q', which determines the contribution of strain on E_A. Q' measures the apparent change in the activation energy with strain at a constant composition.

The factor Q' results from separating the effect of biaxial strain from the composition method. Atomistic calculations using molecular statics or dynamics have provided activation volumes under hydrostatic conditions[70]. Biaxial stress experiments can also address the same atomistic parameters if the effect of composition at constant strain is controlled for[71]. However, both the hydrostatic and biaxial experimental results can be given a single, consistent theoretical framework so that they may be compared directly with each other[72].

The comparison between biaxial and hydrostatic stress states proceeds readily because the activation volume generalizes, for non-hydrostatic stresses, to the activation strain tensor. In essence, the ''shape'' of the activation volume is relevant when stresses are non-hydrostatic. The hydrostatic and the biaxial stress effects on diffusion can be

related to each other[72]. Therefore, the biaxial factor Q' can be derived from the hydrostatic components.

The diffusion characteristics of dual channels are very sensitive to the value of Q' since the layers have different states and magnitude of strain. Calculating the profiles of different strained SiGe heterostructures and then comparing with the experimentally measured profiles, we find that a value of 23 eV/unit strain for Q' results in good fit.

The time evolution of the concentration profile is given by Fick's Law. The analytical solutions of Fick's law are intractable in strained channel heterostructures because of empirical nature of diffusion coefficient and complicated initial boundary conditions, leading us to resort to finite difference methods. The standard finite difference scheme[73] has inaccuracies when applied to a system with steeply varying concentration gradients because of exponential dependence of diffusivity on concentration. Hence, in this work, we have implemented a more accurate finite difference scheme, for use in strained SiGe heterostructure with steeply varying concentrations.

With the diffusivity varying with the concentration and strain in the layer the Fick's law in explicit finite difference scheme can be written as Equation 4-4.

$$\frac{c_{i,j+1}-c_{i,j}}{\Delta t}=\frac{D\left(c_{i+1/2,j}\right)\left(\frac{c_{i+1,j}-c_{i,j}}{\Delta x}\right)-D\left(c_{i-1/2,j}\right)\left(\frac{c_{i,j}-c_{i-1,j}}{\Delta x}\right)}{\Delta x} \quad \textbf{Equation 4-4}$$

Where, $c_{i,j}$ is the concentration of Ge in the $Si_{1-y}Ge_y$ film at the node i at the jth time step , $D(c_{i,j})$ is the diffusivity of Ge at the node i at the jth time step, and depends on concentration and the strain at that point, t is the diffusion time and x is the coordinate axis.

In the standard finite difference scheme, the diffusion coefficient at node $i+1/2$ is approximated to be equal to the average of the diffusivity at node i and node $i+1$ and results in a finite difference scheme as described elsewhere[73]. This approximation has some inaccuracies because since the diffusion is exponential, this averaging scheme will lead to an overestimation of the diffusion coefficient. We can improve this scheme by evaluating the diffusion coefficient of the averaged concentration at $(i+1/2)^{th}$ node as given in Equation 4-5. The improved scheme for evaluating Fick's law in finite difference is given below as

$$\frac{c_{i,j+1}-c_{i,j}}{\Delta t}=\left(\frac{D\left(\frac{c_{i-1,j}+c_{i,j}}{2}\right)c_{i-1,j}-\left(D\left(\frac{c_{i-1,j}+c_{i,j}}{2}\right)+D\left(\frac{c_{i+1,j}+c_{i,j}}{2}\right)\right)c_{i,j}+D\left(\frac{c_{i+1,j}+c_{i,j}}{2}\right)c_{i+1,j}}{\Delta x^2}\right)$$

Equation 4-5

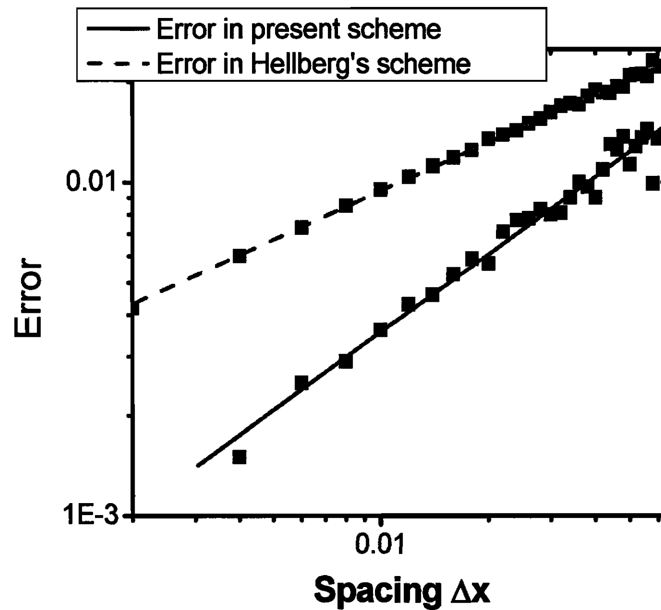


Figure 4-3: Error in the present finite difference scheme as compared to the standard finite difference scheme[73]. As the spacing of the nodes increases the error increases.

Using the formulated diffusion coefficient and the proposed finite difference method, we have developed a platform for evaluating the thermal budget of the structure. Simulations match with the experimental results (Figure 4-4) in numerous situations dealing with different concentration and temperatures, confirming the validity of our analytical tool.

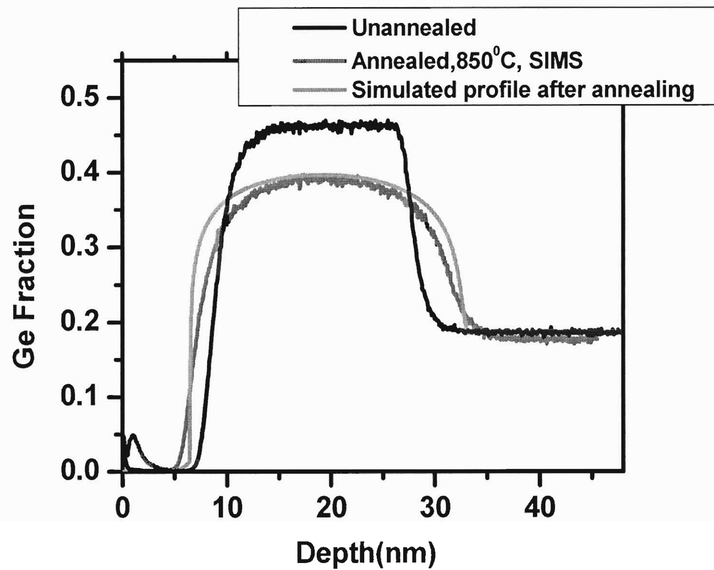


Figure 4-4: The profile for simulation used is shown below. Initial interfaces were smoothened with a hyperbolic tangent function. Value of Q' = 23 gives a good fit with the experimental profiles.

4.3. Summary

This chapter outlines the methods to determine experimentally and mathematically the diffusion of Ge in a SiGe heterostructure. A diffusion coefficient is developed for estimating the Ge diffusion coefficient. This is later used in developing a platform which implements a novel, improved finite difference scheme for analyzing and engineering thermal budgets of strained SiGe heterostructures.

Chapter 5. Tri-Layer Heterostructure

Dual channel heterostructures provide very high carrier mobility heterostructures for use in various microelectronic applications. Different parameters such as the thickness of compressively strained SiGe, tensilely strained Si, and the Ge composition in the compressively strained SiGe, and the relaxed SiGe buffer, in such heterostructures provide a handle for exactly modifying the heterostructure to match the requirements for the mobility over a large range. However, as we will see in this chapter, a tradeoff exists between the achievable mobility enhancements and the thermal stability of the heterostructure.

Dual channel heterostructures have been shown to have a tradeoff between thermal stability and achievable mobility because while the thermal budget reduces with increasing Ge concentration in the buried layer, the hole mobility increases. The mobility enhancements, which depend on the Ge concentration present in the buried layer, will decrease with such an out-diffusion since it will lead to a decrease in the Ge concentration in the buried layer.

The thermal stability of a dual channel heterostructure is extensively discussed by utilizing the diffusion coefficient and the finite difference scheme built earlier. Effect of different parameters, including Ge concentrations in the compressively strained layer, temperature of processing, is considered based on mathematical arguments. A background of the proposed tri-layer heterostructure is established based on such analysis and the thermal analysis reveals that the tri-layer heterostructure is much more stable than a dual channel heterostructure.

We present a tri-layer heterostructure in this chapter and take a look at some of its electrical and physical properties. In this chapter we look at some experimental and numerical results that describe its improved thermal stability. Some general trends on the thermal stability are also derived using the platform built earlier for evaluating thermal budgets. The band structure of tri-layer heterostructure is also very important since it has a barrier for the hole to prevent its leakage into the underlying low mobility relaxed SiGe buffer layer. This is expected to improve the hole mobility in a PMOS fabricated on the tri-layer heterostructure as compared to a dual channel. Some results on the PMOS mobility enhancements on the tri-layer heterostructure are also described. Finally, a structure is proposed which is expected to have symmetric mobility enhancements for the carriers.

5.1. Dual channel Heterostructures

5.1.1 Motivation for Dual channel

Hole mobility enhancements in p-type MOSFETs on strained Si are typically small and dependent on the strain and the gate overdrive[74,75]. These enhancements reduce to zero beyond gate overdrives greater than 0.5 MV/cm for unclear reasons[13,20,76-78]. With a strain lower than 0.8% the ε-Si PMOSFET loses its enhancements and in some cases shows worse mobilities than Si PMOS.

Biaxial tensile strain in ε-Si splits the light-hole/heavy-hole degeneracy and reduces both the out-of-plane and in-plane effective mass of holes[75,79,80]. A biaxial tension greater than 1% in Si can cause the out of plane mobilities to become higher than the in-plane mobilities[79,80]. This leads to a hole wavefunction which is substantially spread in the vertical direction. Experimental studies show the extent of the hole wavefunction to be more than[20] 5 nm which implies that the hole wavefunction samples the valence band structure of the material below the ε–Si for small h_{si}. This implies that even at high vertical fields the properties of holes in the relaxed SiGe below the ε–Si can make a significant contribution to hole transport.

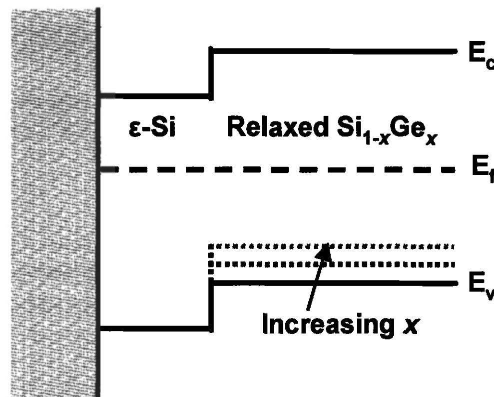


Figure 5-1: Band alignment of strained Si grown on relaxed $Si_{1-x}Ge_x$

Since the hole transport properties in the layer below the ε–Si are important, a way to improve the hole mobility can be the incorporation of a compressively strained layer below the ε–Si[1,15,81]. A compressively strained $Si_{1-y}Ge_y$ layer on top of a relaxed $Si_{1-x}Ge_x$ layer (y>x) provides a high mobility channel for holes. The Si capping layer on top of the

compressively strained $Si_{1-y}Ge_y$ layer provides compatibility with conventional CMOS processing and a high quality interface with a standard SiO_2 gate dielectric. This structure, known as a dual channel heterostructure, provides a platform for fabricating both p-type and n-type MOSFETs, with very high hole and electron mobility enhancements over CZ-Si[14,82].

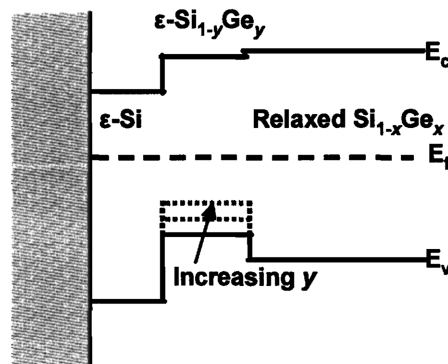


Figure 5-2:Schematic band alignment of a dual channel heterostructure. Higher *y* in the compressive layer leads to a deeper well for holes. In *p*-MOSFETs based on these structures, large gate overdrives force holes towards the SiO_2/Si interface.

In this heterostructure, the compressive strain in the $Si_{1-y}Ge_y$ layer reduces the inter-valley scattering due to breaking of the valence band degeneracy and also reduces the in-plane and out-of-plane effective masses in a manner analogous to tensile strain, leading to increased hole mobilities[81,83]. Also, the resulting band structure has a valence band offset which confines the holes in the high mobility compressively strained $Si_{1-y}Ge_y$ layer. The hole mobilities are further enhanced when the band structure crosses over from the Si-like X-valley to the Ge-like L-valley at y~0.8 and thus, the ε-$Si_{1-y}Ge_y$ starts exhibiting Ge-like mobilities[49,83] ($\mu_{h,Ge}$=1900 cm^2/Vs compared to $\mu_{h,Si}$ = 450 cm^2/Vs).

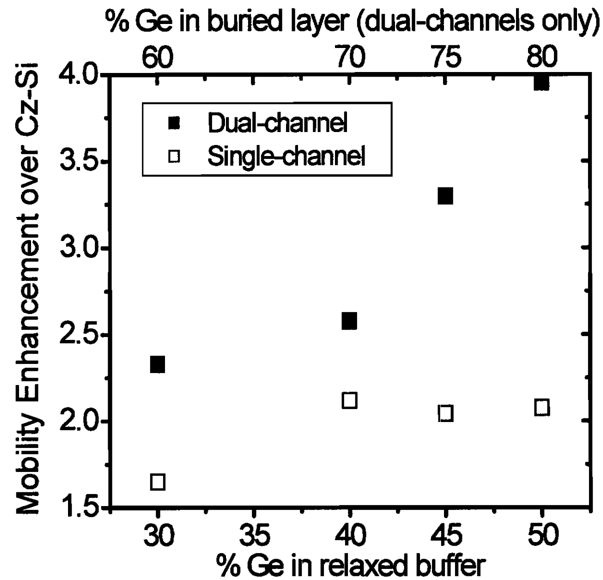


Figure 5-3:Hole mobility enhancement over Cz-Si at E_{eff} = 0.6 MV/cm for single channel and dual channel devices grown on the same relaxed buffer compositions. The buried layer Ge content in all of the dual channel devices is 30% higher than the relaxed buffer. Even at high E_{eff}, dual channel devices demonstrate far larger enhancements than single channel devices. Image Courtesy: Minjoo L. Lee

The improved hole mobilities of the dual channel heterostructures (Figure 5-7) as compared to a single channel heterostructure can be seen in Figure 5-3. Thus, for all the structures presented in Figure 5-3, the buried ε-$Si_{1-y}Ge_y$ layer has a Ge-content 30% higher than the relaxed $Si_{1-x}Ge_x$. A single channel heterostructure refers to a structure consisting of a strained Si over relaxed buffer, the Ge concentration of which is being varied. Thus, we can see that for a constant strain level, an increase in Ge concentration increases the mobilities even at high fields of 0.6 MV/cm. Thus it can be expected that the holes are populating the high mobility compressive Ge-rich layer even at higher fields. The ε-Si cap is 7-9 nm and can be reduced so that the holes occupy the bottom compressive layer even at higher fields.

The hole mobility enhancement in the dual channel heterostructure depends on many parameters which includes (i) the Ge concentration in the relaxed $Si_{1-x}Ge_x$, (ii) the Ge concentration in the compressively strained $Si_{1-y}Ge_y$ layer, (iii) the thickness of the top ε-Si layer and (iv) the thickness of the ε-$Si_{1-y}Ge_y$ layer. These effects have been described elsewhere[1,13,16,84] and only a brief summary is being presented here.

5.1.2 Effect of Ge concentration of the ε-$Si_{1-y}Ge_y$ layer on the hole mobility

The Ge content in the ε-$Si_{1-y}Ge_y$ layer directly influences (i) the intrinsic mobility of hole in the channel, (ii) the valence band offset between the relaxed $Si_{1-x}Ge_x$ buffer and the ε-$Si_{1-y}Ge_y$ layer and (iii) the strain and hence, the amount of inter-valley splitting and the effective masses in different directions. Taking the relaxed $Si_{1-x}Ge_x$ composition constant, the increase in the Ge concentration in the ε-$Si_{1-y}Ge_y$ layer will lead to a higher compressive strain. The higher strain leads to a large decrease in the effective mass and an increase in the inter-valley splitting leading to increased mobilities due to both these effects. In addition, the intrinsic hole mobility increases as the Ge fraction is increased, with further enhancements as the Ge fraction becomes more than 0.8, since the band structure becomes Ge-like. The alloy scattering phenomenon is not seen to be dominant as verified experimentally[13]. The valence band offset also increases with increasing y and helps in preventing the hole wavefunction from leaking in the low mobility relaxed $Si_{1-x}Ge_x$.

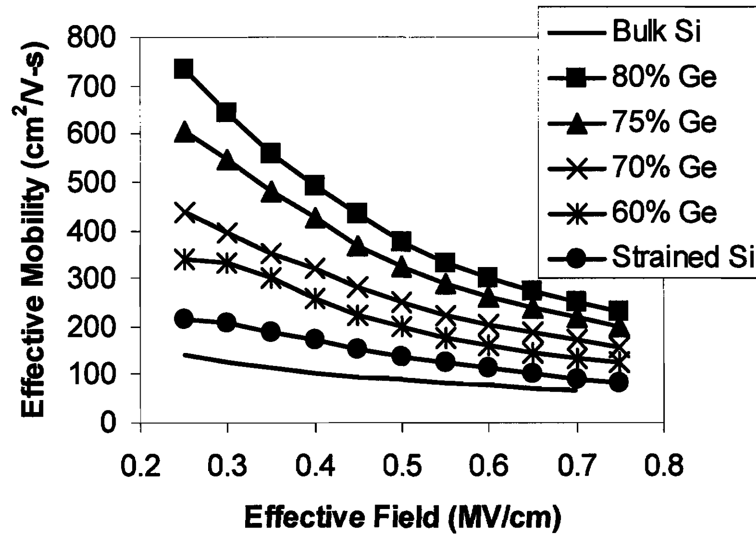


Figure 5-4:Effective hole mobilities of dual channel heterostructure *p*-MOSFETs[85]. The difference in Ge content between the virtual substrate and compressive layer, *y*-*x*, is held constant at 0.3.

5.1.3 Effect of Ge concentration in the relaxed $Si_{1-x}Ge_x$

A reduction in Ge fraction in the relaxed $Si_{1-x}Ge_x$ increases the strain in the compressively strained $Si_{1-y}Ge_y$ layer and the valence band offset between the two layers. Ideally, higher strain and higher offset will lead to enhanced mobility and hence, reduction of Ge concentration in the relaxed $Si_{1-x}Ge_x$ is desired. A lower Ge concentration in the relaxed buffer also allows higher degree of compatibility with standard Si CMOS processing.

5.1.4 Effect of thickness of the ε-$Si_{1-y}Ge_y$ layer

In ε-$Si_{1-y}Ge_y$, compressive strain greatly reduces the vertical effective mass of holes, and the out-of-plane mobility is predicted to exceed the already large in-plane mobility. Thus, the hole wave function can spread into the ε-Si layer above and the relaxed $Si_{1-x}Ge_x$ below, despite the deep potential well for holes that forms in the ε-$Si_{1-y}Ge_y$. The penetration of the hole wavefunction into the relaxed $Si_{1-x}Ge_x$ below is not desirable because it provides a very low mobility channel for holes and hence reduces the overall mobility[75,77,79]. Increasing compressive strain in the compressively strained $Si_{1-y}Ge_y$ layer can therefore, serve to *decrease* the overall μ_{eff} of the inversion layer, if the ε-$Si_{1-y}Ge_y$ is not thick enough to confine the hole wave function vertically[86].

To study the effect of ε-Ge thickness on hole mobility, h_{Ge} was varied while holding the strain level and $h_{Si\ cap}$ constant. Theory suggests that a ε-Ge layer grown on $Si_{0.5}Ge_{0.5}$ will exhibit higher mobility than one grown on $Si_{0.3}Ge_{0.7}$, both because of the lowered in-plane effective mass[11] and the larger valence band offset (type-I band alignment) between the ε-Ge and relaxed $Si_{1-x}Ge_x$. However, for the two devices with thin (h_{Ge} = 6 nm) ε-Ge layers, the more highly strained sample demonstrates *lower* μ_{eff} across the entire range of N_{inv}. As the out-of-plane mobility is predicted to exceed the already large in-plane mobility, the hole wave function can spread into the layers above and below. For the sample with thin ε-Ge on $Si_{0.5}Ge_{0.5}$, the 2% compressive strain and resultant high out-of-plane mobility cause the hole wave function to penetrate considerably into the ε-Si cap and the lower mobility buffer below. While the same spreading must take place in the thin ε-Ge sample on $Si_{0.3}Ge_{0.7}$, the vertical extent of the wave function is lessened due to the lower strain.

Increasing h_{Ge} gave rise to an 80% increase in μ_{eff} for the devices on $Si_{0.5}Ge_{0.5}$ and only a 20% increase for the devices grown on $Si_{0.3}Ge_{0.7}$ (Figure 5-5). μ_{eff} increases considerably for the thick ε-Ge layer on $Si_{0.5}Ge_{0.5}$ because the 12 nm ε-Ge layer allows the large hole wave function to be better confined in the ε-Ge. Conversely, it may be deduced that 6 nm is nearly sufficient to confine the hole wave function when ε-Ge is grown on $Si_{0.3}Ge_{0.7}$ based upon the relatively small gain in μ_{eff} caused by doubling thickness of Ge[16,82].

Thus, we can see that a large enough Ge thickness is required for obtaining the highest hole mobilities in the particular dual channel heterostructure. The thickness should be large enough to accommodate the vertical extent of hole wavefunction which is determined by the amount of strain in the Ge channel.

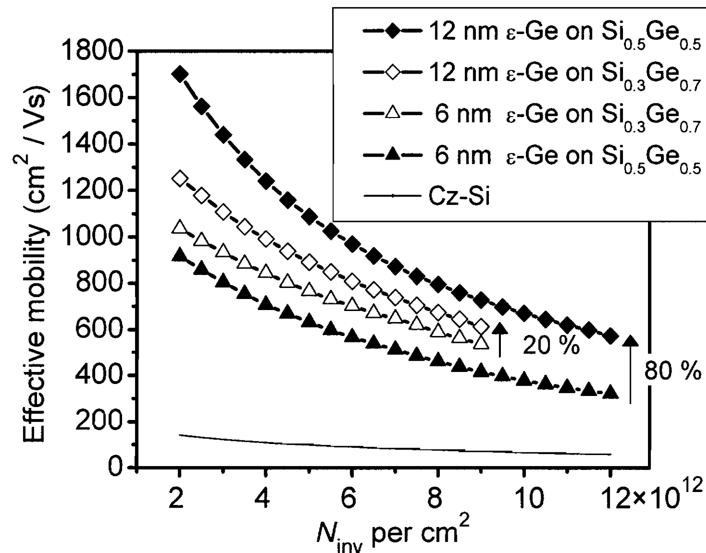


Figure 5-5: Hole effective mobility vs N_{inv}. The use of a thicker ε-Ge layer gave rise to a 20% increase in μ_{eff} for the samples grown on $Si_{0.3}Ge_{0.7}$ and an 80% increase for the samples grown on $Si_{0.5}Ge_{0.5}$ [16]

The above discussion on the effect of thickness of compressive Ge layer on hole mobility points in the direction that the hole wavefunction penetrates the underlying buffer layers as well as top Si layer. The penetration of the hole wavefunction into the relaxed $Si_{1-x}Ge_x$ below is not desirable because it provides a very low mobility channel for the holes and hence reduces the overall mobility. Thus, it is expected that higher hole mobilities can be extracted if the hole wavefunction penetration in the underlying buffer is reduced so that the hole is confined to the high mobility Ge layer. A large enough ε–$Si_{1-y}Ge_y$ thickness, which can accommodate the vertical extent of hole wavefunction, is required for obtaining the highest hole mobilities in these dual channel heterostructures.

5.2. Thermal stability of dual channel heterostructures

The phase diagram of the SiGe materials system[24] is shown in Figure 5-6. The low melting point of SiGe films present in a dual channel heterostructure necessitates a low temperature processing for the heterostructures containing such films. Low temperature processing is also desirable because the thermal budget of Ge containing heterostructures, such as dual channel heterostructures, is very low. The diffusion coefficient of Ge in a SiGe single-crystal film is high compared to Si and increases exponentially with an increase in the Ge content in the SiGe films as was shown in Chapter 4.

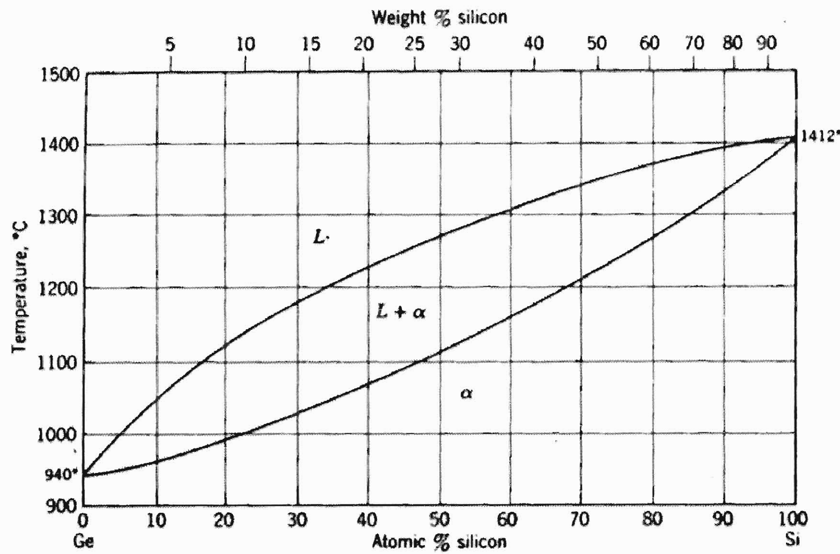


Figure 5-6: The Si-Ge Binary Phase Diagram, adapted from Gandhi.[24] Si and Ge are completely miscible, enabling SiGe films of arbitrary composition to be grown without phase segregation.

Diffusion of Ge from the ε–$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ and the strained Si layers during high temperature processing such as source-drain activation, gate oxide growth etc. and epitaxial growth will reduce the concentration and strain in the compressively strained Ge rich layer (Figure 5-8).

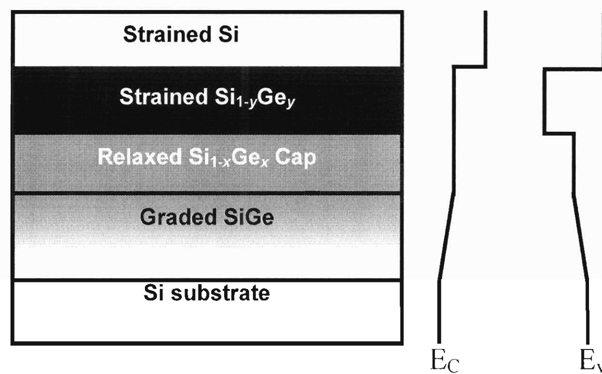


Figure 5-7: Schematic layer sequence and band diagram for a dual channel heterostructure

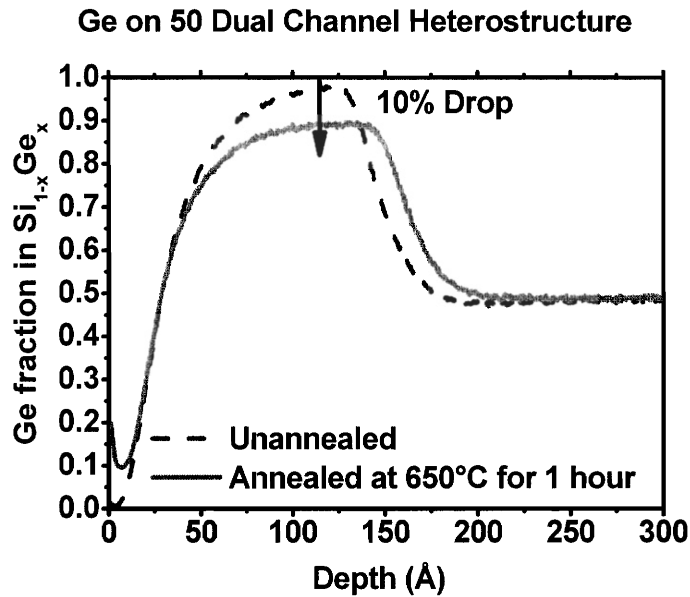


Figure 5-8: Ge concentration in the dual channel heterostructure is reduced after annealing at a high temperature

The hole μ_{eff} reduces in a dual channel heterostructure due to the outdiffusion of Ge during high temperature processing. The hole μ_{eff} reduction is related to the peak decrease (Δy) of the dual channel heterostructure, which is defined as the difference of maximum concentration of Ge in the ε-$Si_{1-y}Ge_y$ layer before and after the high temperature processing.

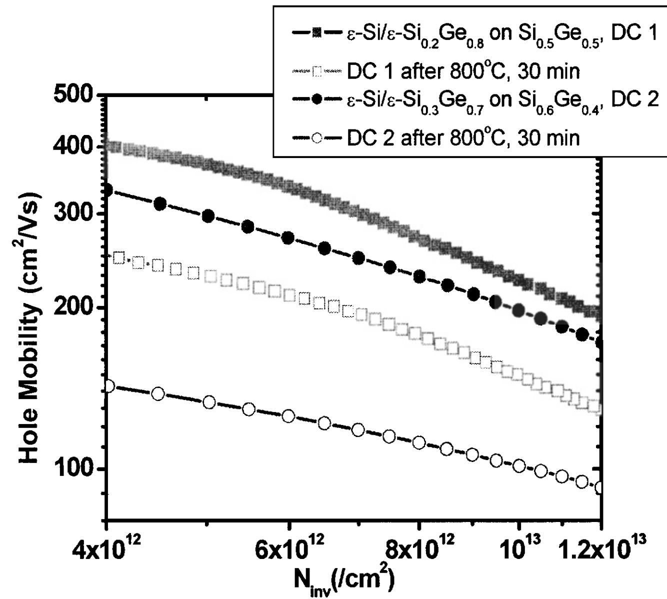


Figure 5-9: Hole mobilities reduce after annealing at 800°C for 30 minutes. For a 70 on 40 dual channels the mobility drops from 2.9 times to 1.56X, while it drops from 3.8X to 2.2X for an 80 on 50 channel[2].

Using the developed predictive scheme, we are in a position to estimate the peak decrease of a dual channel heterostructure and to evaluate the effect of different variables such as temperature and Ge concentration on the thermal budget of the structure. In order to simulate the diffusion in the dual channels, an initial concentration profile was set up and the interfaces were approximated by hyperbolic tangent. No-flux boundary conditions are used on both ends of the profile. We have found that the Δy depends superlinearly on the y in the ε-$Si_{1-y}Ge_y$ layer (Fig. 5-10).

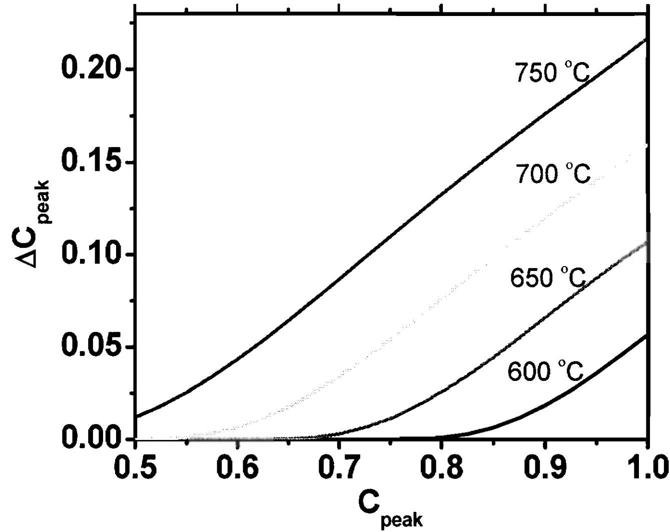


Figure 5-10: Peak decrease in a dual channel heterostructure with increasing concentration of Ge in the ε-$Si_{1-y}Ge_y$ layer. The Ge concentration in the relaxed buffer is maintained at 30% below the Ge concentration in the ε-$Si_{1-y}Ge_y$ layer

The hole μ_{eff} enhancements also increase exponentially with the y in the ε-$Si_{1-y}Ge_y$ layer[18]. Thus, the reduction in μ_{eff} after a particular high temperature processing step is the largest for highest Ge concentration in the ε-$Si_{1-y}Ge_y$ layer. This implies that the highest mobility dual channel heterostructures (Figure 5-3), which have highest Ge concentrations, are thus, most prone to the out-diffusion resulting from high temperature processing.

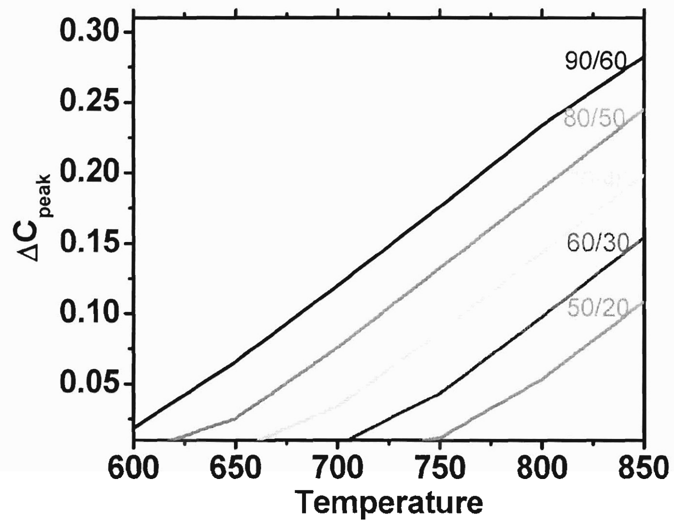


Figure 5-11: Peak decrease with an increasing temperature of processing. The time of processing is one hour. The peak decrease of different dual channels have been shown.

Because of outdiffusion of Ge from the compressively strained layer, the compressive strain in the layer is also reduced. This can also cause a reduction in the achievable hole mobilities in the layer (Figure 5-12).

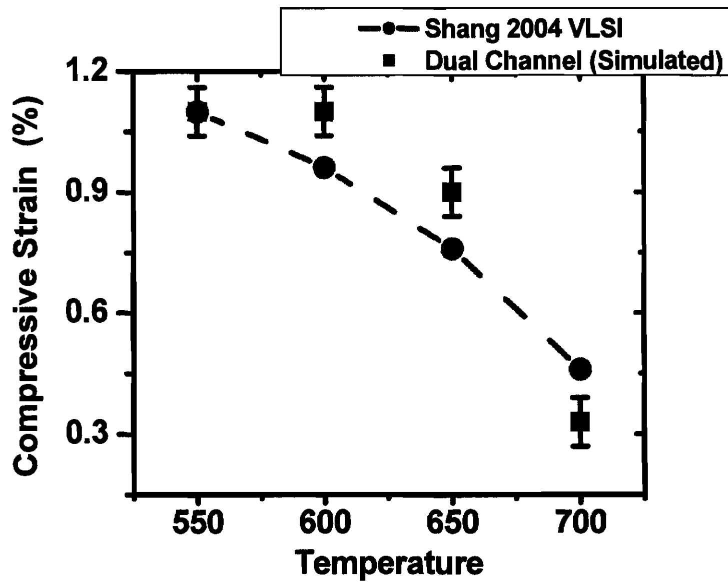


Figure 5-12: Strain reduction in a Ge on 70 dual channel heterostructure[17]. As shown in this work almost all the reduction is through the Ge outdiffusion from the compressively strained layers.

In order to obtain try to engineer structures which have a high thermal stability, a closer look at the Ge profile of a strained Ge dual channel heterostructure is warranted.

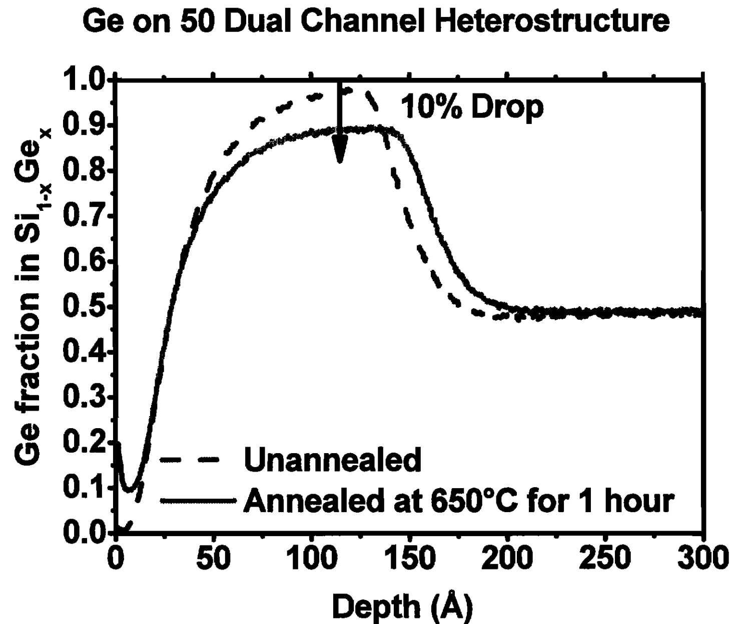


Figure 5-13: SIMS profile of a Ge buried channel heterostructure. The outdiffusion of Ge in the strained Si cap is very less and majority of diffusion occurs into the relaxed SiGe buffer.

As seen from the Figure 5-13, the amount of Ge out-diffusion into the ε-Si layer is minimal with most of the outdiffusion occurring into the relaxed SiGe layer. The diffusion coefficient for Ge interdiffusion in a SiGe single crystal increases exponentially with increasing Ge concentration and compressive strain[59,65]. This is primarily the reason why out-diffusion of Ge occurs primarily from the ε-$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ with much less diffusion into the ε-Si. The fact that little Ge out-diffusion occurs in the capping ε-Si layer provides a motivation for including a ε-Si layer between the ε-$Si_{1-y}Ge_y$ and relaxed $Si_{1-x}Ge_x$ buffer for reducing the Ge out-flux from the ε-$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ buffer. The resulting heterostructure is henceforth referred to as tri-layer heterostructure.

5.3. Tri-layer heterostructure

A heterostructure in which a ε–Si layer is present below the compressively strained $Si_{1-y}Ge_y$ layer, which in turn is capped by a tensilely strained Si layer for SiO_2 compatibility is henceforth referred to as a tri-layer heterostructure[3,4,87]. As discussed above, this structure is expected to have a higher thermal budget as compared to a conventional dual channel structure since the outdiffusion of Ge into the relaxed buffer will be reduced.

A cross-section transmission electron micrograph and schematic band structure of a ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si tri-layer heterostructure on relaxed $Si_{0.8}Ge_{0.2}$ are shown in Figure 5-14. Low temperature growth techniques (350°C $\leq$ T $\leq$ 550°C) were employed for growing these heterostructures in order to maintain planarity and abrupt Ge profiles[88]. Confinement of holes in the ε-$Si_{0.5}Ge_{0.5}$ of a tri-layer heterostructure is expected to be better than in a corresponding dual channel because of the higher valence band offset. Interface abruptness is also enhanced in a tri-layer heterostructure because the flux of Ge from ε-$Si_{0.5}Ge_{0.5}$ layer into the relaxed $Si_{0.8}Ge_{0.2}$ is higher as compared to the flux into ε-Si[3]. The improved valence band offset and interface abruptness compared to the dual channel case results in better hole confinement in the ε-$Si_{0.5}Ge_{0.5}$ layer in the tri-layer heterostructure.

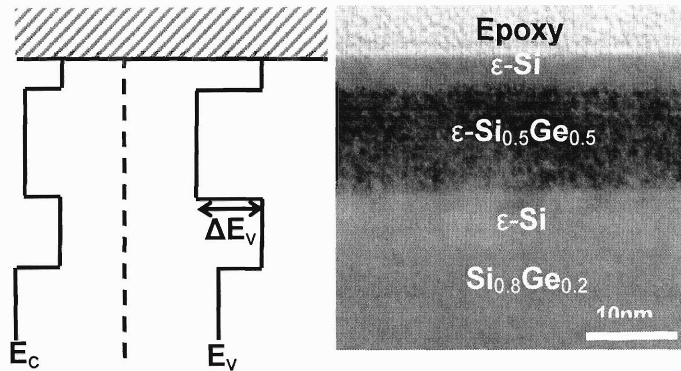


Figure 5-14: XTEM of a tri-layer heterostructure. Corresponding band diagram shows a quantum well for holes in the ε–$Si_{1-y}Ge_y$ layer and a high valence band offset ΔE_v due to underlying ε-Si layer preventing the holes from tunneling into $Si_{0.8}Ge_{0.2}$.

5.3.1 Growth of tri-layer heterostructures

This heterostructure is grown in an ultrahigh-vacuum chemical vapor deposition reactor with different parts of the structure grown at different temperatures to minimize roughening and inter-diffusion. While the relaxed graded buffer is grown at a temperature of 900°C to maximize relaxation, the heterostructure is grown at lower temperatures in order to have a smooth interface. The bottom Si is tensilely strained to a value of 0.8% and hence formation of the step type surface undulations[89] are not expected. We see from the interface in the XTEM (Figure 5-14) that it is indeed the case. However, the ε-$Si_{0.5}Ge_{0.5}$ is under compressive strain and can be expected to give undulations[88]. The growth of the ε-$Si_{0.5}Ge_{0.5}$ is done at a low temperature of 450°C and is not undulated at these temperatures. However, since the Si has extremely low growth rates at these temperatures, the temperature needs to be increased in order to get the top Si cap. An

increase in the temperature can lead to surface undulations of the compressively strained $Si_{0.5}Ge_{0.5}$ layer. Hence, a 'raise and flow' technique is utilized in which the temperature is raised from 450°C to 550°C with SiH_4 flowing during temperature increase. A thin layer of Si that forms on top of ε-$Si_{0.5}Ge_{0.5}$ helps to prevent the undulations of the ε-$Si_{0.5}Ge_{0.5}$. A detailed description of such technique for growing these structures is described elsewhere[88].

5.4. Thermal stability of Tri-Layer heterostructures

In order to quantify the achievable improvements over the dual channel heterostructures, numerical and experimental evidence on the thermal stability of tri-layer heterostructure has been collected. Numerical investigations show the peak decrease in the tri-layer is much reduced as compared to the corresponding dual channel (Fig. 5-15). Thus, not only is the Ge concentration in the tri-layer higher, higher strain is also retained as compared to the dual channel. The improvement in the thermal budget of the tri-layer over the dual channel is highest at the higher Ge fraction implying that the use of an underlying ε-Si at higher Ge concentration is much more beneficial than at the lower Ge concentrations. This result is particularly significant since the highest Ge concentration structures which have the lowest thermal budget will be most benefited by the approach of a tri-layer.

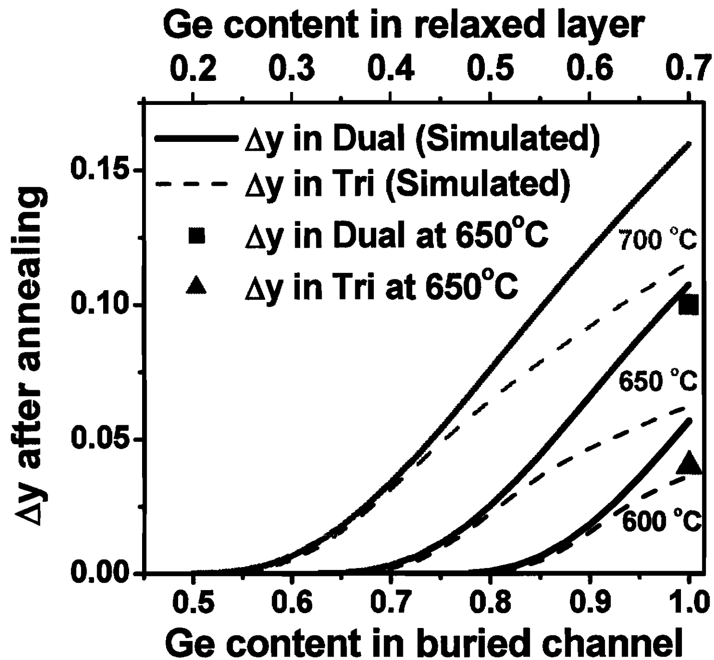


Figure 5-15: The peak decrease of the dual channel and the tri-layer with varying concentration of Ge in the buried layer and the temperature is shown. The tri-layer has much improved thermal budget over the dual channels when the concentration in the strained-$Si_{1-y}Ge_y$ layer is higher. The peak decreases are in reasonable agreements with the experimental values.

Improved thermal budget of the tri-layer heterostructures as compared to a otherwise equivalent dual channel heterostructure has also been confirmed by SIMS characterization of annealed and as-grown structures (Figure 5-16, Figure 5-17). Figure 5-16 shows that the Ge peaks decreases by 5% in the case of the ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si tri-layer heterostructure on relaxed $Si_{0.8}Ge_{0.2}$ and by 7% in the case corresponding to ε-Si/ε-$Si_{0.5}Ge_{0.5}$ dual channel heterostructure on relaxed $Si_{0.8}Ge_{0.2}$ after a 30 minutes 850°C anneal. As seen in Figure 5-17 after a 60 minutes, 650°C anneal, the Ge profile of a tri-layer ε-Si/ε-Ge/ε-Si heterostructure on a relaxed $Si_{0.5}Ge_{0.5}$ buffer show a 3% peak

decrease as compared to a 10% peak decrease in the corresponding dual channel ε-Si/ε-Ge heterostructure on a relaxed $Si_{0.5}Ge_{0.5}$ buffer.

The efficiency of the bottom ε-Si layer in preventing the Ge out-diffusion increases drastically with Ge concentration in the ε-$Si_{1-y}Ge_y$ layer (Figure 5-15). Thus, while the ε-Si/ε-Ge/ε-Si heterostructure on relaxed $Si_{0.5}Ge_{0.5}$ has much improved thermal stability as compared to the ε-Si/ε-Ge heterostructure on relaxed $Si_{0.5}Ge_{0.5}$ (Figure 5-17), the thermal stability of the ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si heterostructure on relaxed $Si_{0.8}Ge_{0.2}$ is only slightly larger as compared to the ε-Si/ε-$Si_{0.5}Ge_{0.5}$ heterostructure on relaxed $Si_{0.8}Ge_{0.2}$ (Figure 5-16).

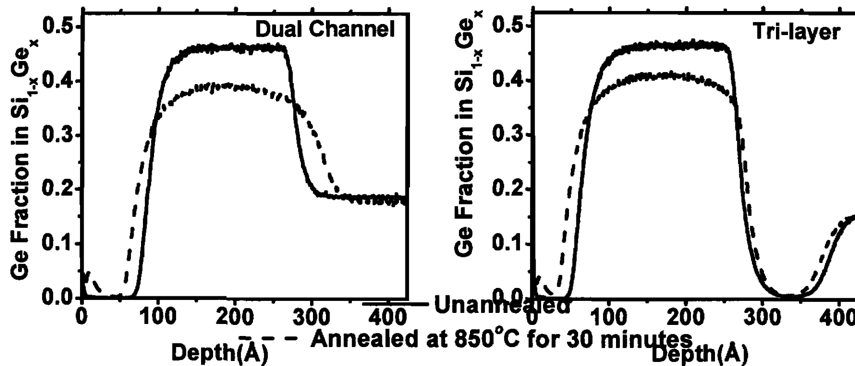


Figure 5-16: Comparison of SIMS profile of the ε-Si/ε-$Si_{0.5}Ge_{0.5}$ heterostructure on relaxed $Si_{0.80}Ge_{0.20}$ with ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si on relaxed $Si_{0.8}Ge_{0.2}$ after annealing for 30 minutes at 850°C. The peak decreases to 39% from 46% in case of dual channel and to 41% from 46% in case of tri-layer.

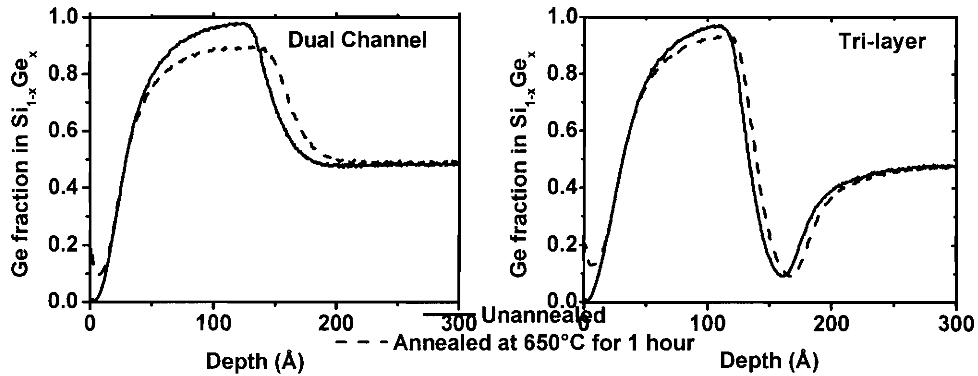


Figure 5-17: Comparison of SIMS profile of ε-Si/ε-Ge heterostructure on relaxed $Si_{0.5}Ge_{0.5}$ with ε-Si/ε-Ge/ε-Si on relaxed $Si_{0.5}Ge_{0.5}$ after annealing for 60 minutes at 650°C. The peak decreases by 10% in case of the dual channel and 3% in case of tri-layer.

In order to compare the extent of μ_{eff} reduction with thermal budget, ring shaped MOSFETs were fabricated on ε-Si/ε-$Si_{0.3}Ge_{0.7}$/ε-Si on relaxed $Si_{0.6}Ge_{0.4}$ and ε-Si/ ε - $Si_{0.3}Ge_{0.7}$ on relaxed $Si_{0.6}Ge_{0.4}$ using low temperature processing (<600°C). For activating the dopant implants, they were subjected to anneal at 600°C for 1 hour and 800°C for 30 minutes. The mobility characteristics are given in Figure 5-18 and show that while the mobilities at low temperature anneals are similar, at higher temperatures hole μ_{eff} enhancements retained in a tri-layer heterostructure are about 10% higher as compared to a dual channel heterostructure.

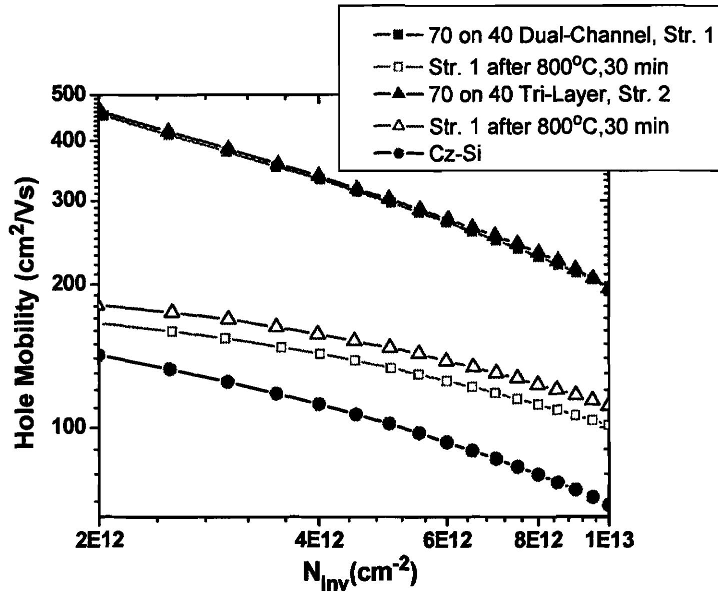


Figure 5-18: Higher hole mobilities are retained in a ε-Si/ε-$Si_{0.3}Ge_{0.7}$/ε-Si on relaxed $Si_{0.6}Ge_{0.4}$ heterostructure as compared to a ε-Si/ε-$Si_{0.3}Ge_{0.7}$ on relaxed $Si_{0.6}Ge_{0.4}$ heterostructure after processing it at 850°C, showing the improved thermal stability of the tri-layer.

5.5. Improved hole mobilities in tri-layer heterostructures

To investigate the possibility of improved hole μ_{eff} enhancements in a tri-layer heterostructure because of better hole confinement in the high μ_{eff} ε-$Si_{1-y}Ge_y$ layer (Figure 5-14), we fabricated long channel ring-type *p*-MOSFETs on a ε-Si(4nm)/ε-$Si_{0.5}Ge_{0.5}$(12nm)/ε-Si(8nm) tri-layer heterostructure on a relaxed $Si_{0.8}Ge_{0.2}$ buffer. As seen from the Figure 5-19, hole μ_{eff} enhancements of about 2 are observed for the tri-layer heterostructure, and this enhancement persists to $N_{inv} = 1\times10^{13}$ /cm^2. These enhancements

are 20% higher than a corresponding ε-Si(7nm)/ε-$Si_{0.5}Ge_{0.5}$(23nm) on $Si_{0.8}Ge_{0.2}$ dual channel device at $N_{inv} = 1\times10^{13}/cm^2$. The relatively high μ_{eff} of the tri-layer as compared to the dual channel, which has identical composition and strain to the tri-layer, implies that the valence band offset provided by the bottom ε-Si layer contributes to the improved hole transport.

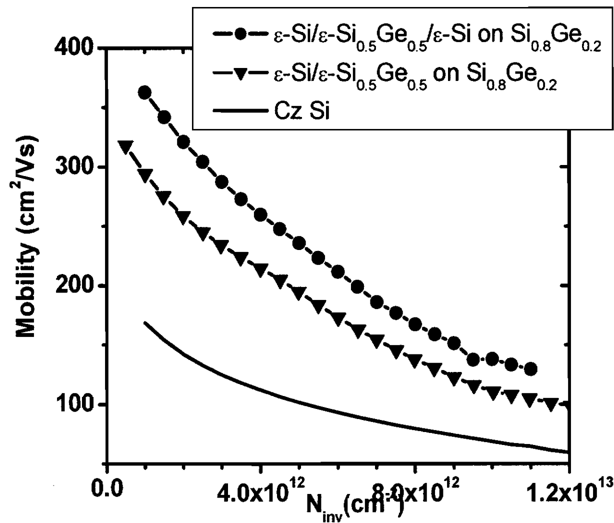


Figure 5-19: Hole effective mobility vs N_{inv} obtained from a PMOS fabricated on a ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si structure grown on relaxed $Si_{0.8}Ge_{0.2}$. The mobility enhancement for the tri-layer is 20% higher than corresponding ε-Si/ε-$Si_{0.5}Ge_{0.5}$ dual channel heterostructure grown on relaxed $Si_{0.8}Ge_{0.2}$ at $N_{inv} = 1\times10^{13}/cm^2$

These results show that the hole μ_{eff} enhancements of the order of 2X can be obtained on substrates with only moderate Ge concentrations (20% Ge in the buffer and 50% in the compressed layer, in this case). Lower Ge-content graded buffers are less challenging from the perspective of epitaxial growth, since thinner graded buffers and shorter growth times can be used, and the defect density tend to be slightly lower, as well. Lower Ge-

content compressive layers are also considerably easier to deposit, since strain-driven surface diffusion, which leads to non-planar growth, decreases with Ge content[88].

The NMOS μ_{eff} on a dual channel heterostructure with a thick Si cap ($\geq$5 nm) is similar to a single ε-Si channel on a relaxed buffer case[90], and is not expected to be different in the tri-layer heterostructure. Since a NMOS fabricated on a ε-Si layer grown on a relaxed $Si_{0.8}Ge_{0.2}$ gives a 2X enhancement (Figure 5-20)[18], both the *p*-MOSFET and the *n*-MOSFET fabricated on a ε-Si/ε-$Si_{0.5}Ge_{0.5}$/ε-Si tri-layer heterostructure on relaxed $Si_{0.8}Ge_{0.2}$ are expected to give a symmetric hole and electron μ_{eff} enhancements of 2X. These enhancements translate to similar increase in the NMOS and PMOS drive currents with the same CMOS design and drive current ratios.

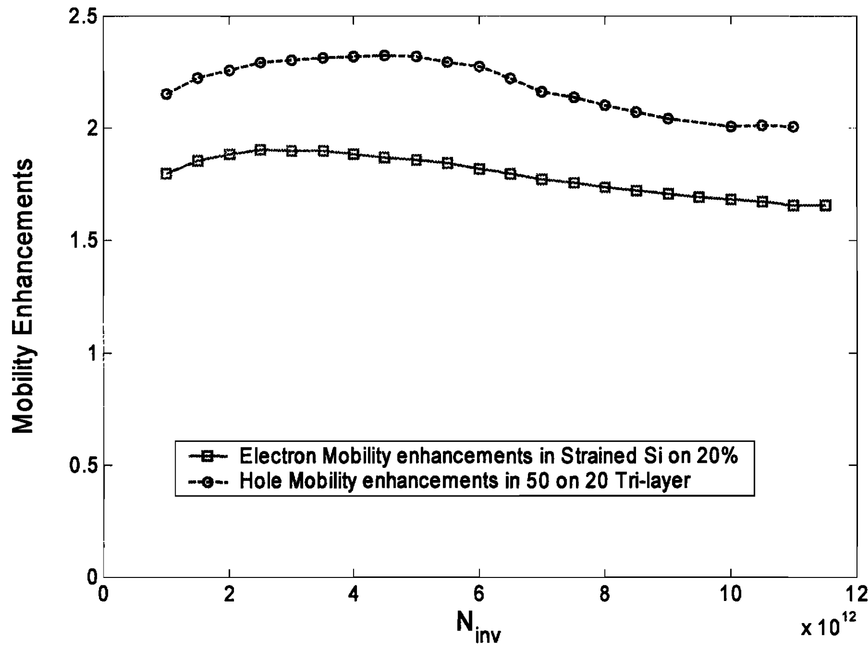


Figure 5-20: Mobility enhancement in a PMOS fabricated on a tri-layer heterostructure (50 on 20), and electron mobility enhancements from NMOS fabricated on a 20% relaxed SiGe buffer. Since, the NMOS behavior is not expected to be very different from a single channel heterostructure, a possibility of symmetric hole and electron enhancement exist on a ε-Si|ε-$Si_{0.5}Ge_{0.5}$|ε-Si on relaxed $Si_{0.8}Ge_{0.2}$

5.6. Summary

Dual channel heterostructures are summarized in this chapter with effect of various parameters on the hole mobility enhancements. Thermal stability analysis of a dual channel heterostructure is conducted based on the diffusion framework established in the previous chapter. Dual channel heterostructures have been shown to have a tradeoff between thermal stability and achievable mobility. This is because, while the thermal budget reduces with increasing amount of Ge concentration present in the buried layer, the hole mobility increases. The mobility enhancements which depend on the Ge

concentration present in the buried layer will decrease with such an out diffusion, since it will lead to decrease in the Ge concentration in the buried layer.

To address the out-diffusion problems seen in dual channel devices, we have researched a tri-layer heterostructure, which exhibits improved thermal stability as has been verified experimentally and numerically. Improved hole μ_{eff} in this heterostructure are also observed over similar dual channel heterostructures which can be a result of better hole confinement in the strained-$Si_{1-y}Ge_y$ layer of the proposed heterostructure. In addition, a tri-layer heterostructure which provides a symmetric hole and electron μ_{eff} enhancements of about 2X has also been proposed.

In conclusion, the tri-layer heterostructure described in this thesis is an excellent candidate for a high mobility and high thermal budget CMOS platform.

Chapter 6. Summary and Future Work

6.1. Summary

In this work, single mismatched layers are used to elucidate the parameters which affect the TDD in a mismatched SiGe layer. Two important variables which affect the dislocation density in a mismatched layer have been identified as the strain rate at which the layer is grown and surface roughness. Relaxation in a mismatched SiGe layer occurs through formation of a misfit dislocation network. As the spacing between misfit dislocations is reduced, their strain fields overlap strongly. These strain fields negate the driving force for threading dislocations and obstruct it in places where the misfit dislocation spacing is very less[32]. The strain fields also lead to a characteristic cross hatch pattern on the surface of the film. Threading dislocations can get blocked at the deeper trenches since the total driving force on a threading dislocation is locally reduced in these areas[30]. Obstruction of these threading dislocations by the two mechanisms described above necessitates the nucleation of more threading dislocations for relaxing the misfit strain.

Plastic relaxation of the layer occurs through glide of dislocation. It is shown in this work that at higher temperatures the mismatched layer follows the Matthews-Blakeslee curve closely. The rate of plastic relaxation in a mismatched SiGe layer then depends on the rate of material deposition (growth rates) and the composition of the SiGe layer. At a higher Ge composition where a faster rate of strain relaxation is required, a larger number of dislocation density exists.

We have developed a model which takes into account both the strain rate dependence and the misfit dislocation spacing to estimate the threading dislocation densities. An important implication from the above developed model is the insight that misfit dislocation spacing needs to be very uniform in order to reduce the dislocation densities. If the nucleation of misfit dislocation occurs from the edge of the wafer, the resulting distribution is very spread out leading to a higher fraction of misfit dislocations that are close to each other, while the other that are far apart. In a uniform distribution, lesser fraction of misfit dislocations are closer than critical spacing and hence their effect on increasing the threading dislocation density is smaller. Another important implication of this model is the result that a single relaxed high Ge composition layer (>20% Ge) grown on Si substrate is certain to be of high TDD. This is because, regardless of how uniformly spaced the misfit dislocations are, a significant fraction of them will still be more closely spaced than the critical spacing. The 20% layer has the misfit spacing of about 22nm, implying that regardless of how uniformly the misfits are spaced in the 20% layer, it will always be a high TDD layer since the strain fields will be very strong to prevent the dislocation glide.

Reducing the TDD in a mismatched SiGe layer on Si requires low strain rates which can be achieved by growing layers at slower growth rates. We have used two techniques for achieving slower growth rates: pulsed layer technique and low growth temperatures. We define a pulsed layer technique in which gas flow is intermittent at 900°C, thus achieving low growth rates. It also allows the decoupling of growth rate and growth temperature inherent to CVD growth. By varying the growth delay time and the

flow time we can vary widely the effective growth rates. While the lower growth rates achieve a low strain rate, the ability to have these lower growth rates at higher temperature improves the kinetics of relaxation. However, high ad-atom mobility during this process results in increased surface roughness. Thus, lower growth rates at high temperature inherently couples surface roughness to lower growth rate and under these conditions lower threading dislocation densities can not be achieved.

The traditional means to reduce the growth rate is to lower growth temperature. Since the surface mobility of the atoms is reduced, the resulting layer has less surface roughness. However, since dislocation velocity is exponentially dependent on temperature, decreasing growth temperature can increase threading dislocation density. Fortunately, dislocation nucleation for layers on (001) surfaces is difficult and there is a window in which decreased growth temperatures does not increase threading dislocation density. Instead, metastable strain is introduced, thus preventing complete relaxation of the strained layer. Therefore, achieving a completely relaxed low TDD layer requires the removal of the metastable strain without increasing dislocation nucleation.

Glide of misfit dislocations relieves the strain in a strained SiGe mismatched layer. The process of annealing the mismatched layer, for relaxing the strain, has been explored. One possible way for removing the metastability of this layer is by depositing strain relieving layers on top of the metastable layer and then annealing. The increased stress at the surface will increase the dislocation velocity for a given anneal temperature and will lead to complete relaxation of the layers as compared to relaxation predicted by

Matthews-Blakeslee equilibrium. We have demonstrated this technique with mismatched $Si_{0.90}Ge_{0.10}$ layers on Si since 10% Ge is determined to be the highest Ge concentration layer in which threading dislocation flow without relative impedance from the underlying misfit stress fields. Reducing the threading dislocation density in this 10% Ge layer on Si then requires slow plastic strain relaxation which could be achieved by growing the layer at a slower growth rate. If the slower growth is achieved at high growth temperatures the concomitant surface-induced roughness limits the threading dislocation motion, thereby necessitating increased threading dislocation nucleation for relaxing the film. Using lower growth temperatures to reduce the growth rates, however, leads to a poorly relaxed 10% Ge layer. Since, the window to achieve relaxation and low threading dislocation density is narrow, in order to aid with the visualization of this window; we have developed a Time-Temperature-Threading Dislocation Density diagram. The only possible kinetic path suggested is a gradual increase in the anneal temperature which results in a slow rate of relaxation leading to a low threading dislocation density with fair degree of relaxation. The remaining elastic strain can be completely removed by using a strain relieving layer on top which drives the threading dislocation with a larger force than is possible with an equilibrium single-layer system.

Repetition of this couple leads to completely relaxed layers with very low threading dislocation densities. This structure is essentially a graded composition region with an absolute maximum in the rate of composition increase with thickness. The results shed increased understanding on creating highly deformed thin films with low threading dislocation densities, as well as providing guidance on graded composition buffers that can be less than 50% of the thickness of conventional graded composition buffers.

Once a low TDD SiGe virtual substrate has been developed, it can be used for enabling technology boosters for improving transistor performance such as strained channel materials with improved carrier transport. A dual channel heterostructure consisting of strained-Si (ε-Si) / strained-$Si_{1-y}Ge_y$ (ε-$Si_{1-y}Ge_y$) on a relaxed $Si_{1-x}Ge_x$ buffer ($y>x$), provides a platform for fabricating MOSFETs with high hole mobility (μ_{eff}). The impressive hole μ_{eff} enhancements obtained on dual channel heterostructures depend greatly on the Ge concentration and the strain[1] in the ε-$Si_{1-y}Ge_y$ layer. Ge out-diffusion during high temperature processing steps from the ε-$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ buffer reduces the Ge concentration and the strain in the layer leading to reduced hole μ_{eff} in these heterostructures[2]. Our work is motivated by the desire to understand thermal stability in dual channel heterostructures and to apply this understanding to engineer heterostructures with higher thermal stability and improved electrical properties.

We show that pre-exponential factors (D_0) and the activation energies (E_A) for Ge diffusivity[63] at different Ge concentrations follow the Meyer-Neldel relationship. Using these values of D_0 and E_A we have formulated a diffusion coefficient for Ge in crystalline SiGe, which covers the entire range of Ge composition and matches well with the literature values. This diffusion coefficient will be used for analyzing the diffusion characteristics of the dual channel heterostructures in this work. The analytical solutions of Fick's law are intractable in such heterostructures because of empirical nature of diffusion coefficient and complicated initial boundary conditions, leading us to resort to finite difference methods. The standard finite difference scheme[73] has inaccuracies when applied to a system with steeply varying concentrations because of exponential dependence of diffusivity on concentration. Hence, in this work, we have implemented a

novel, more accurate finite difference scheme, for use in strained SiGe heterostructure with steeply varying concentrations.

An important disagreement occurs in the literature on the value of coupling strength[66] Q', which determines the contribution of strain on the activation energy of diffusion. The diffusion characteristics of dual channels are very sensitive to the value of Q' since the layers have different states and magnitude of strain. Calculating the profiles of different strained SiGe heterostructures and then comparing with the SIMS results we find that a value of 23 eV/unit strain for Q' results in good fit with the experimental data.

Using the formulated diffusion coefficient and our new finite difference scheme, we have evaluated the effect of different parameters such as temperature, Ge concentration and the thickness of the strained SiGe layer on the thermal budget of the structure. We have found that the peak decrease ($\Delta C_{\varepsilon\text{-SiGe}}$) depends almost exponentially on the peak concentration of the compressively strained layer. The highest mobility dual channel heterostructures, which have highest Ge concentrations, are thus most prone to this out-diffusion. Due to the very small diffusivity of Ge in Si, the out-diffusion of Ge occurs primarily from the ε-$Si_{1-y}Ge_y$ layer into the relaxed $Si_{1-x}Ge_x$ buffer with much less diffusion into the ε-Si[3]. This motivated us to include a ε-Si layer between the ε-$Si_{1-y}Ge_y$ and relaxed $Si_{1-x}Ge_x$ buffer for reducing the Ge out-flux. The resulting heterostructure is henceforth referred to as tri-layer heterostructure[4].

Based on our numerical investigations and annealing experiments we see that the thermal budget of the tri-layer heterostructure is higher over some ranges of temperatures. Regimes in temperature and concentration where the tri-layer is expected to be thermally more stable as compared to corresponding dual channels have been outlined using simple flux analysis and verified experimentally in this work. Improved hole mobilities in this heterostructure are also observed over similar dual channel heterostructures which can be a result of better hole confinement in the strained-$Si_{1-y}Ge_y$ layer of the proposed heterostructure.

Ring shaped MOSFETs were fabricated on both platforms and subjected to varying processing temperature in order to compare the extent of μ_{eff} reduction with thermal budget. Hole μ_{eff} enhancements are retained to a much higher extent in a tri-layer heterostructure after high temperature processing as compared to a dual channel heterostructure.

In conclusion, we have developed a platform which implements a novel, improved finite difference scheme, for analyzing and engineering strained SiGe heterostructures. To address the out-diffusion problems seen in dual channel devices, we have researched a novel tri-layer heterostructure, which exhibits improved thermal stability and high p-channel μ_{eff}. The improved thermal stability and hole μ_{eff} of a tri-layer heterostructure makes it an ideal platform for fabricating high μ_{eff} MOSFETs that can be processed over higher temperatures without significant losses in hole μ_{eff}.

6.2. Suggestions for Future work

6.2.1 Extending the TRUT buffer concept

Effort has been made in this thesis to completely outline the rationale behind the experimental design of TRUT buffers. Therefore, extensive experimental and modeling effort helped us to arrive at the step anneal procedure for relaxing the 10% layer. Similarly, the design of the 12% strain relaxing layer was done through extensive experimentation. However, a few aspects of TRUT buffer still remain to be explored

6.2.1.1 Optimizing annealing times for TRUT buffers

It was earlier explained that the TRUT buffers require a larger amount of time as compared to a normal graded buffer. While a significant amount of time is involved in the temperature drops in the UHVCVD chamber used in this work, a large portion of the time is for the annealing of the 10% layer. The anneal profile is reproduced here in Figure 6.1.

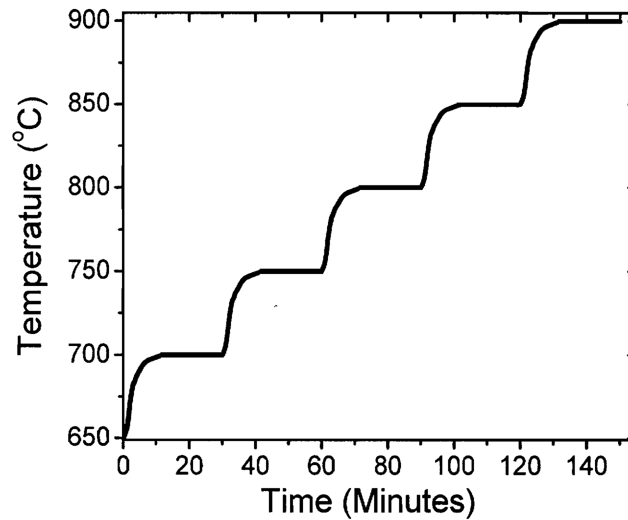


Figure 6-1: Step-anneal profile used for relaxing the 10%, 1500Å layer grown at 650°C. While longer times at lower temperatures are justified, longer times at higher temperatures are not necessary since the dislocation dynamics is very fast at these temperatures.

The times for annealing are decided on time required for relaxing the layers completely. While longer times at lower temperatures are justified on the basis of slower dislocation dynamics, at a higher temperature such long times are not needed since the kinetics of relaxation is faster. A smaller time at higher temperature will also help to prevent the surface roughness of the structure since it provides lesser time for the ad-atoms to diffuse on the surface. Therefore, in future implementation of TRUT buffers, an ideal anneal time profile should be used which has longer time at the lower temperature and smaller times at the end. Such a technique can help to achieve a similar relaxation as with step anneals at times of about 40-50 minutes (based on the nucleation-relaxation model developed earlier). Once the time for anneal has been optimized, the total time for TRUT buffers can be reduced. Similarly, a further reduction in the time for growth of the

TRUT buffers can be achieved through shifting to RTCVD system, which should be used for future TRUT buffers growth.

SiGe layers can be grown by rapid thermal chemical vapor deposition (RTCVD). For example, a commercial system can have SiH_2Cl_2 and GeH_4 as the precursors and H_2 carrier gas and can take 200 mm Si wafers as the starting substrates. Heating is provided by incandescent lamp arrays located above and below the silicon carbide coated graphite susceptor with the top bank directly illuminating onto the wafer. With a RTP system the wafer can be cooled rapidly after the growth by turning of the heating lamps. The transition from heating to cooling is limited by the thermal time constant of the wafer, typically 15ms, and the time constant for the heat source. A lamp and temperature diagnostic response time of less than 1 ms can easily be maintained which is much faster than the wafer response time of 15 ms[91]. The wafer thermal response time thus determines the turn around time. However, this time is insignificant as compared to the time required for the temperature drops in the large heat capacity UHVCVD chamber. A temperature drop of 900°C to 650°C requires a time of about 2 hours. Using an optimized anneal and a RTCVD system, the time for the growth of a TRUT buffer can be reduced to about 2.5 hours which could help to make the process commercially feasible.

6.2.1.2 Higher Ge content buffers

Despite its ability to produce lattice mismatched epitaxy of unprecedented quality, the virtual substrate approach requires growth of a thick graded buffer to ensure complete relaxation of the individual mismatched layers. In the case of Ge virtual substrates which are compositionally graded from Si to Ge, the buffer thickness is typically greater than 10 μm. Such thick layers complicate subsequent device integration with the underlying Si since the device levels are not coplanar and must be co-processed across a deep step. In addition, there are other issues with a thicker graded buffer such as thermal expansion mismatch which can lead to microcracks. The presence of cracks in thin films is not desirable for device fabrication because they can act as scattering centers for light propagation, resist in-plane electrical current flow, and introduce electrical shorting paths in vertical current flow.[92]

During hetero-epitaxial growth, one must consider the thermal mismatch effects as a high thermal mismatch can lead to high density arrays of cracks in the thin film. When the film and the substrate experience different rates of thermal expansion and contraction, the thin film will experience considerable tensile or compressive stress during a temperature cycle, leading to crack formation or delamination. For the GaAs/Si system, the GaAs film has a coefficient of thermal expansion (CTE), α, of 6.8×10^{-6} $^{o}C^{-1}$, much higher than the CTE of Si, with α of approximately 2.6×10^{-6} $^{o}C^{-1}$. The thermal expansion coefficient of the SiGe substrate is dominated by that of Si since the Si substrate is much thicker than the total thickness of the graded layers plus the top Ge layer, which is approximately 10 μm. For many III-V devices on Si, such as high-efficiency multi-junction solar cells which require thick layers, the tensile strain that develops upon cooldown limits the amount of III-V material that can be deposited on

structure without formation of micro-cracks[11]. As devices based on III-V materials are typically much thicker than Si-based devices, it will be highly beneficial to reduce the thickness of the SiGe buffer. Also since Ge is more closely matched with GaAs in terms of CTE, it is more beneficial to reduce the thickness of the Ge rich region. This is because the Si rich region of the graded-buffer is more closely matched to Si and therefore its CTE is determined by the substrate. Ge and the III-V material can be considered to be similar layer. Therefore, any amount of thickness reduction is Ge buffer directly helps to increase the thickness of GaAs that can be grown.

The TRUT buffers have a great advantage in helping to reduce the thickness of the virtual substrate. An important part of the relaxation of the layer is thermal anneals. While the thermal anneals are less effective in the Si rich region, the Ge rich regions are softer and have much higher glide velocities. Therefore, it stands to reason that the annealing procedure will be very effective in the Ge rich regions. Another useful thing in this regards is the fact that the normal growth temperatures for high Ge regions are very low, therefore, one can achieve higher relaxation by reaching higher temperatures. This was a limitation for the Si rich region since a high temperature of 900°C was not high enough for causing relaxation.

We can use the model developed earlier in order to determine the ideal time-temperature anneals for a high Ge content layers. We need to determine the activation energies for the glide and nucleation for these layers.

6.2.2 Improving the Relaxation-Nucleation model

The direction for the experiments in this thesis has been suggested in many cases by the models that are developed in this work. One such model which led us to step anneals for relaxing a mismatched layer however does not correctly predict the exact extent of relaxation (See Section 2.8.1). An important parameter that needs to be added to this relaxation-nucleation model is the resisting force by the already developed misfit dislocation densities. In most of the literature, the distribution of the misfit dislocations have been assumed to be random or uniform, we are in a better position to estimate the resisting force on the gliding dislocation. By assuming that the distribution of the misfit dislocations is log-normal, we have been able to predict correctly the TDD evolution with Ge content in a mismatched layer. Using the same log-normal distribution we can incorporate a better assumption about the resisting force to the gliding dislocations. This will help to reconcile the differences in the measured and experimentally observed relaxation of the layer.

6.2.3 Advanced devices on tri-layer heterostructure

Although improve thermal stability in a tri-layer heterostructure has already been shown, the greatest effect of improved thermal budgets in tri-layer heterostructure will be seen on high Ge concentration dual channel heterostructures. This is expected even more because the achievable hole mobilities are exponentially dependent on Ge concentration in the compressively strained SiGe layer. Devices made on a high Ge content tri-layer

heterostructure will translate the achievable hole mobilities in the layer to a higher retained hole mobilities. Even though such an effect is shown in a 70 on 40 tri-layer heterostructure, it will be more pronounced in a high Ge content tri-layer heterostructure.

APPENDIX

All the MATLAB source codes used for the calculations carried out in various part of thesis are described in this section.

Appendix 1. Script for Critical Thickness

The thickness at which the strain energy in the film exceeds the energy necessary for dislocations to be present, it becomes thermodynamically favorable to nucleate misfit dislocations. This can be expressed mathematically by minimizing the total energy of the system, leading to the classic Matthews-Blakeslee equation. This thickness is referred to as critical thickness and is an equilibrium parameter.

The equilibrium critical thickness, h_c, for a lattice-mismatched layer is given by

$$h_c = \frac{D(1-\nu\cos^2\alpha)(b/b_{eff})\left[\ln\left(h_c/b\right)+1\right]}{2Yf}$$

where D is the average shear modulus at the interface, ν is Poisson's ratio, α is the angle between the dislocation line direction and Burgers vector b, h is the film thickness, b_{eff} is the interfacial component of the Burgers vector, Y is the Young's modulus of the film, and f is the mismatch between film and substrate[8].

A function which takes in the Ge concentration in the substrate and the film to determine the critical thickness of the film is outlined below.

```
function h=SiGexony(x,y)

% Materials Parameters

% Lattice Constant
ASi=5.431;
AGe=5.658;

% Elastic Constants
C11Si=16.577e11;
C12Si=6.393e11;
C44Si=7.962e11;
C11Ge=12.60e11;
C12Ge=4.40e11;
C44Ge=6.77e11;

%Ge in the substrate
xGes=y;

%Ge in the film
xGeo=x;

%Interpolated materials properties

C11o=(xGeo*C11Ge)+((1-xGeo)*C11Ge);
C12o=(xGeo*C12Ge)+((1-xGeo)*C12Ge);
C44o=(xGeo*C44Ge)+((1-xGeo)*C44Ge);
C11s=(xGes*C11Ge)+((1-xGes)*C11Si);
C12s=(xGes*C12Ge)+((1-xGes)*C12Si);
C44s=(xGes*C44Ge)+((1-xGes)*C44Si);
Go=C44o-(2*C44o+C12o-C11o)/2;
Gs=C44s-(2*C44s+C12s-C11s)/2;

Y=C11o+C12o-2*C12o^2/C11o;


%if Y==0
  %Y=C11o;
%end

  nu=C12o/(C12o+C11o);

% Interpolation constants

A=(xGeo*AGe)+((1-xGeo)*ASi);
```

```
As=(xGes*AGe)+((1-xGes)*ASi);
b=2^.5/2*A;
beff=b/2;
D=(Go*Gs*b)/(3.1428*(Go+Gs)*(1-nu));
alph=60/180*3.1428;
f=abs((As-A)/A);
```

%initial guesses

```
hc1=50;
hc2=100;
%hc2=D*(1-nu*cos(alph)^2)*(log(hc1/b)+1)/(Y*f)
```

%Iterate to find critical thickness

```
while(abs(hc1-hc2)>1)
    hc1=hc2;
    hc2=D*(1-nu*cos(alph)^2)*(log(hc1/b)+1)/(Y*f);
end

%prints answer
h=hc2
```

Appendix 2. Elastic strain remaining with thickness

If f is a mismatch of a layer grown on Si and its thickness is h, assuming that it is free to relax, it will relax from mismatch of f to a value of strain given by the following equation.

$$\varepsilon = \frac{D\left(1 - v\cos^2\alpha\right)\left(b/b_{eff}\right)\left[\ln\left(h/b\right)+1\right]}{2Yh}$$

where D is the average shear modulus at the interface, v is Poisson's ratio, α is the angle between the dislocation line direction and Burgers vector b, h is the film thickness, b_{eff} is the interfacial component of the Burgers vector, Y is the Young's modulus of the film, and f is the mismatch between film and substrate[8].

```
%plotstrain.m
%Plots the M-B curve, and uses SiGexony function (Appendix I) and StrainSiGe %(Appendix
2) function

clear

xsub=0;%Substrate Composition
xfilm=0.99;     % Film Composition


hc=SiGexony(xfilm,xsub);

        for i=1:1000
                b(i)=hc+i*5;
                epsilon(i)= strainSiGe(xfilm,xsub,b(i));
        end
relaxation= (epsilon(1)-epsilon(2));
%Plot elastic strain
plot(b,epsilon/0.042);
```

```
%function StrainSiGe
```

```
function epsilon=strainSiGe(x,y,h)

%Lattice Constants
ASi=5.431;
AGe=5.658;

%Elastic Constants

C11Si=16.577e11;
C12Si=6.393e11;
C44Si=7.962e11;
C11Ge=12.60e11;
C12Ge=4.40e11;
C44Ge=6.77e11;

%Ge in the substrate
xGes=y;
%Ge in the film
xGeo=x;

%Interpolated materials properties
C11o=(xGeo*C11Ge)+((1-xGeo)*C11Ge);
C12o=(xGeo*C12Ge)+((1-xGeo)*C12Ge);
C44o=(xGeo*C44Ge)+((1-xGeo)*C44Ge);
C11s=(xGes*C11Ge)+((1-xGes)*C11Si);
C12s=(xGes*C12Ge)+((1-xGes)*C12Si);
C44s=(xGes*C44Ge)+((1-xGes)*C44Si);
Go=C44o-(2*C44o+C12o-C11o)/2;
Gs=C44s-(2*C44s+C12s-C11s)/2;
Y=C11o+C12o-2*C12o^2/C11o;
 nu=C12o/(C12o+C11o);
% Constants
A=(xGeo*AGe)+((1-xGeo)*ASi);
As=(xGes*AGe)+((1-xGes)*ASi);
b=2^.5/2*A;
beff=b/2;
D=(Go*Gs*b)/(3.1428*(Go+Gs)*(1-nu));
alph=60/180*3.1428;
hc1=h;

%f=abs((As-A)/A);

epsilon=D*(1-nu*cos(alph)^2)*(log(hc1/b)+1)/(Y*hc1);
```

Appendix 3. For Dislocation Density with Ge concentration

Two important variables which affect the dislocation density in a mismatched layer have been identified as the strain rate at which the layer is grown and surface roughness. We have developed a model which takes into account both the strain rate dependence and the misfit dislocation spacing to estimate the threading dislocation density (TDD) evolution with Ge concentration.

$$\rho_{xge,final} = \left(\sqrt{\rho_{2\%} \times \frac{\dot{\delta}_{xge}}{b_{xge} v_{xge}} \times \frac{b_{2\%} v_{2\%}}{\dot{\delta}_{2\%}}} + 2 \times \frac{\delta}{b} \times fraction \right)^2$$

The fraction that is referred to in this equation refers to the fraction of misfit dislocations whose spacing is below the critical spacing. The critical spacing is a fitting parameter and is determined to be about 23nm. This fraction is determined in the function fractions_ln.

```
%Determines and plots the dislocation densities from the model incorporating strain
rate and %misfit dislocation spacing and compares it with the uniform layer TDD
%just use 1 mm^2 of area for TDD determination

asi=5.431*1e-7;                    %in mm
age=5.658*1e-7;                    %in mm
%crit_spacing=2.29e-5;
crit_spacing=2.145e-5;
tdd2pc=2e4; %threading dislocation density of 2%
asige2pc=(1-0.02)*asi+0.02*age;
b2pc=asige2pc*0.353; %burgers vector of 2%
```

```
d_dot2pc=relax(0,0.02);
%assume that the velocity of dislocation is similar till about 20%
m=1;

for i=21:250
   j=i-1;
   xge=0.001*(j);
   asige=(1-xge)*asi+xge*age;
   b=asige*0.353;
   d_dot(i)=relax(0,xge);
   tdd1(i)=tdd2pc*(d_dot(i)/d_dot2pc)*(b2pc/b)*((0.02/xge)^m);
   %tdd1=tdd2pc*(d_dot(i)/d_dot2pc);
   %[fraction(i) rho(i)]=fractions(xge,crit_spacing,tdd1);
   [fraction(i) rho(i)]=fractions_ln_modified(xge,crit_spacing,tdd1(i));
   %hold on;
end
i=2:0.1:25

%plot(fraction);
%semilogy(i(1:230), rho(11:240));
semilogy(i(1:230), rho(21:250));
%semilogy(i,tdd1);
hold on;
uniformlayer
```

```
%Fraction_ln
% plots the fraction of misfits that have spacing less that the critical spacing


function [fraction rho_rem]=fractions_ln(xge0,spac_crit,rho)

%xge0=0.20;
%spac_crit=1e-4;
%xge=0.001; %just using Si
xge=xge0;

%determining the spacing
asi=5.431*1e-7;%in mm
age=5.658*1e-7;%in mm

asige=(1-xge)*asi+xge*age;

b=asige*0.353;

relaxation=xge*0.0042;
delta=relaxation;
spacing= b/delta %this is the expected value for the rayleigh's distribution in mm

n_points=1/spacing%number of dislocations in 1 mm
```

```
%try to put the number of point to be less than 1e6
x=0:spacing/100:spacing*5;
%p = raylpdf(x,spacing*0.79788456);
%plot(x,p);
fraction=logncdf(spac_crit,log(spacing),600*spacing);
pfrac=logncdf(x,log(spacing),spacing);
distribution=lognpdf(x,log(spacing),spacing*600);
plot(x,distribution);
rho_rem=((sqrt(rho)+(2*n_points*fraction)))^2;
```

Appendix 4. Growth rates for Pulsed layer

Reducing the TDD in a mismatched SiGe layer on Si requires low strain rates which can be achieved by growing layers at slower growth rates. The growth rates of a layer should be reduced to such an extent that it has the strain rate equivalent to a 2% layer. The growth rates of a layer with a mismatch f is then given by,

$$\left(\frac{\partial h}{\partial t}\right)_f = \left(\frac{h_{c,f}}{h_{c,0.02}}\right)^2 \times \left(\frac{\ln(h_{c,0.02}) - \ln(b)}{\ln(h_{c,f}) - \ln(b)}\right)_{0.02} \times \left(\frac{\partial h}{\partial t}\right)_{0.02}$$

An important point to notice is that the $h_{c,f}$ and $h_{c,0.02}$ is given by another transcendental equation that is derived in literature. Thus, the dependence of growth rate with composition is more complicated that is evident from the equation.

```
%growth rate determination for same relaxation rate as 2% graded buffer
clear
% growth rate of 2% = 10A/s
xsub=0.0;
std_grrate=10;
std_rate=relax(0,0.02)*std_grrate;

%put the thickness of the layer here, critical thickness on 10% = 330A
hc=760;
%hc=330;
%film

for i=1:300
  b(i)=hc+i*10;
  gr_rate(i)=(std_rate/relax2(0,0.05,b(i)));
end

plot(b,gr_rate);
```

```
function relaxation=relax(xsubstr,xf)
```

```
xsub=xsubstr;
xfilm=xf;

hc=SiGexony(xfilm,xsub);

for i=1:100
  b(i)=hc+i*1;
  epsilon(i)= strainSiGe(xfilm,xsub,b(i));
end

relaxation= (epsilon(2)-epsilon(3));
```

```
function relaxation=relax2(xsubstr,xf,thick)

%plotstrain.m

%send the critical thickness parameters.
xsub=xsubstr;
%film
xfilm=xf;

%hc=SiGexony(xfilm,xsub);

for i=1:100
  b(i)=thick+i*1;
  epsilon(i)= strainSiGe(xfilm,xsub,b(i));
end

relaxation= (epsilon(2)-epsilon(3));
```

Appendix 5. TDD and strain evolution with time during an isothermal anneal

Utilizing the significant theoretical work done in literature, a model of dislocation dynamics has been proposed in this work.

The glide of the dislocation is thermally activated. The strain relaxed by the layer in time *dt* is given by

$$d\varepsilon_1 = -\rho_0 b v_0 (\varepsilon_1 - \varepsilon_{eq})^m e^{\frac{-Q_v}{kT}} \times dt$$

While the layer is getting relaxed, the dislocation density is increasing as well because it is also thermally activated and is strongly dependent on the strain as well.

$$d\rho = \xi_0 (\varepsilon_1 - \varepsilon_{eq})^m e^{\frac{-Q_p}{kT}} \times dt$$

Both the equations are coupled since strain relaxation will lead to reduction in the dislocation nucleation rate which will in turn lead to slower strain relaxation. The parameters used for these equations have been obtained from literature[42]. In the attached MATLAB code we calculate the evolution of TDD and the strain remaining in the layer with time.

```
clear;
%________________________________________________
rho0=2e2*1e-2;  %initial dislocation density per mm2
```

```
T=650+273;     %Temperature
time=50*60;   %In sec
thick=10000;    %in angstrom
xge=0.10;
xsubstrate=0;
epsilon_init=xge*0.0418;
%epsilon_init=0.004;

%____________________________________________________________

k=1.38e-23;     %Boltzman's Constant
v0=1.5e7;      %preexponential factor for velocity,mm per second
m= 1.1;        %exponent of the strain in the velocity
Qv=1.1*1.6e-19; %Activation energy for the velocity
kT= k*T;

asi=5.431*1e-7;%in mm
age=5.658*1e-7;%in mm
asige=(1-xge)*asi+xge*age;
b=asige*0.353;

psi=0.27e22; %/mm2.sec
Qp=2.7*1.6e-19;
n=2.8;

epsiloneq=strainSiGe(xge,xsubstrate,thick);

tsteps=5;
npoints=time/tsteps;
e1=epsilon_init;

rho_temp=rho0;

for i=1:npoints
   eq_eps=epsiloneq;
   t(i)=i*tsteps;
   e2(i)=epsiloneq+((m-1)*rho_temp*b*v0*exp(-Qv/kT)*tsteps+(e1-epsiloneq)^(-
m+1))^(1/(1-m));
   rho(i)=rho_temp+(psi*((e1-epsiloneq)^(n))*exp(-Qp/kT)*tsteps);
   rho_temp=rho(i);
   e1=e2(i);
end

e2
plot(t/60,e2)
% hold on
%plot(t, rho*100)
```

Appendix 6. Smoothing XRD data for strain determination

X-ray diffraction (XRD) is the most direct and accurate way to characterize the crystallographic quality and residual strain in a deposited semiconductor epilayer. However, in order to determine the residual strain in the layer it is important to determine the omega spread of the (224) reciprocal space map. Accurate determination of the peak position in the omega direction of the (224) can be difficult because of noise in the XRD data. In order to accurately determine the peak position some amount of data cleaning needs to be achieved in order to locate the peak position. The premise of this data smoothing is that the XRD data is both slowly varying and also corrupted by random noise. It can be useful to replace each data point by some kind of local average of surrounding data points. Since nearby points measure very nearly the same underlying value, averaging can reduce the level of noise without (much) biasing the value obtained.

Methods that can improve the detection limit of X-ray diffraction are therefore highly desirable. Savitzky-Golay smoothing which is a signal processing approach has been used to preprocess X-ray diffraction data. As described in a following figure, this approach maintains the intensity of the XRD peaks as well as removes the noise in order to better outline the peak position.

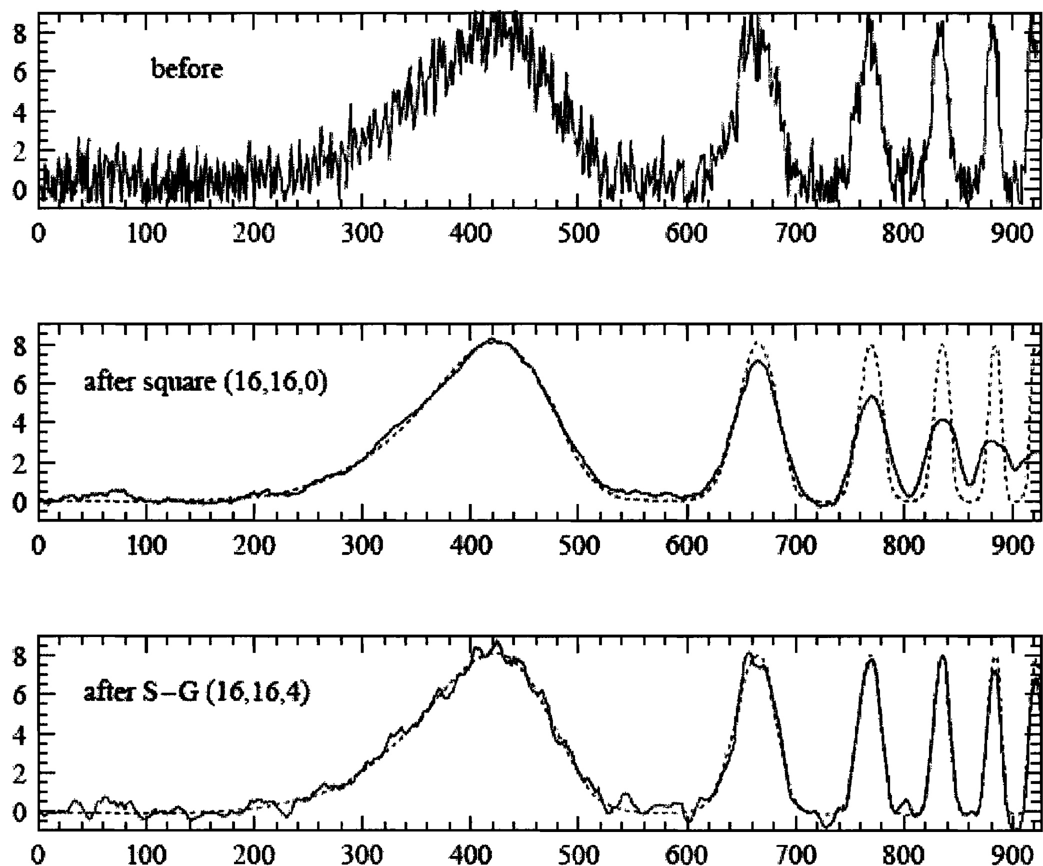
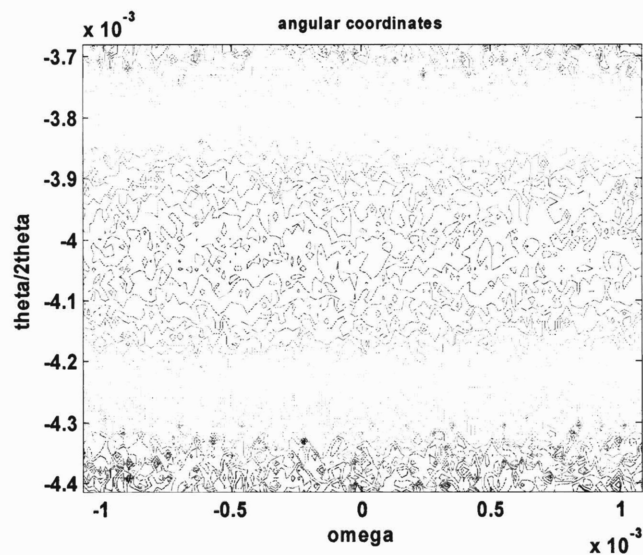
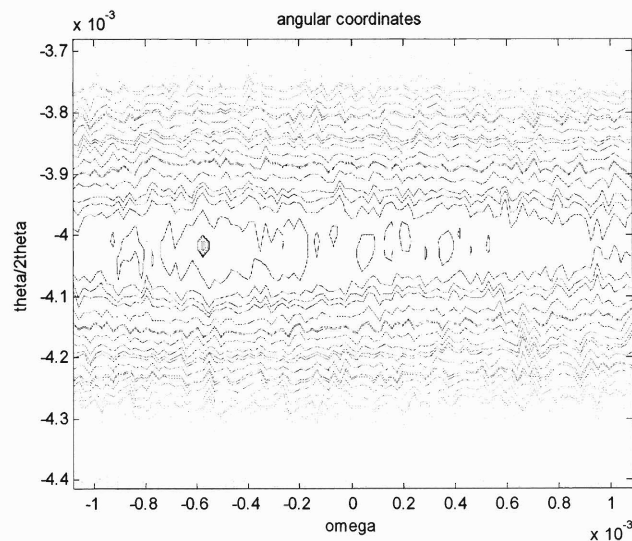


Figure 14.8.1. Top: Synthetic noisy data consisting of a sequence of progressively narrower bumps, and additive Gaussian white noise. Center: Result of smoothing the data by a simple moving window average. The window extends 16 points leftward and rightward, for a total of 33 points. Note that narrow features are broadened and suffer corresponding loss of amplitude. The dotted curve is the underlying function used to generate the synthetic data. Bottom: Result of smoothing the data by a Savitzky-Golay smoothing filter (of degree 4) using the same 33 points. While there is less smoothing of the broadest feature, narrower features have their heights and widths preserved.

Figure 0-1: The above figure explains the usefulness of Savitzky-Golay smoothing filter for X-ray data[93]. While the intensity of narrow peaks remains, the high frequency random noise has been filtered out. Compare this method to the normal data averaging method for narrower peaks.



(a)



(b)

Figure 0-2: Figures explaining XRD data smoothing using Savitzky-Golay filter. Figure (a) plots the as acquired data while Figure (b) shows the data after subjecting it to the filter. The random noise in the data is removed and a better idea of peak position in the omega axis is achieved. Peak is then obtained using a peak finding routine.

Appendix 7. Diffusion in strained SiGe heterostructures

The time evolution of the concentration profile is given by Fick's Law. The analytical solutions of Fick's law are intractable in strained channel heterostructures because of empirical nature of diffusion coefficient and complicated initial boundary conditions, leading us to resort to finite difference methods. In this work, we have implemented a more accurate finite difference scheme, for use in strained SiGe heterostructure with steeply varying concentrations.

The first step of solving such a equation is setting up of initial conditions of Ge concentration, strained SiGe thickness, temperature and time. This is done below.

```
for i=1:5
   cesige=0.97; %strained SiGe concentration
   csige=0.48; %relaxed layer composition
   Temp=600+i*50; % in degree celsius
   tesige=15; %in nm
   time=60;%in minutes
   qprime=23;
   diffusion17(cesige,csige, Temp, i, tesige, time, qprime)
end
```

The second step is setting up the initial Ge profile for the strain SiGe heterostructure. It is also important to set up the gradient for concentration change. Then, a grid is set up and its size is dependent on where the concentration is changing steeply. The gradient is modeled by hyperbolic tangent. Finally, we solve initial value problem for ordinary differential equations (ODEs). Please see MATLAB help for function ode15s for a more detailed description. We will pass the equation $c'=f(y,t)$ to the ode15s solver which will then integrate this equation to give us the time evolution of the concentration as diffusion proceeds.

```
function cpeak=diffusion17(cesige,csige, Temp, n, tesige, time,qprime)

%thickness in nanometers
tsit=5;
tsib=5;
tsige=15;
tgradient1=3;
tgradient2=3;

B=0;
    if(tsib==0)
    B=1; %B=1 dual channel
    end


%t=tsit+tgradient+tesige+tgradient+tsib+(B-1)*tgradient+tsige;

    if(B==1)
    t=tsit+tgradient1+tesige+tgradient2+tsige;
    end

    if(B==0)
    t=tsit+tgradient1+tesige+tgradient1+tsib+tgradient2+tsige;
    end

csit=0;
csib=0;

dz=0.05;
```

```
npoint=t/dz;
ngradpoint1=tgradient1/dz;
ngradpoint2=tgradient2/dz;


tempz1=[tsit:dz:tsit+tgradient1]';
tempz2=[tsit:dz:tsit+tgradient2]';


cgradpoint1=(tanh(3.0*((tempz1-(2*tsit+tgradient1)/2)/(tgradient1/2)))+0.995)*((cesige-
csit)/1.99);

cgradpoint2=(tanh(-3.0*((tempz1-(2*tsit+tgradient1)/2)/(tgradient1/2)))+0.995)*((cesige-
csib)/1.99);

cgradpoint3=(tanh(3.0*((tempz2-(2*tsit+tgradient2)/2)/(tgradient2/2)))+0.995)*((csige-
csib)/1.99);

cgradpoint4=(tanh(-3.0*((tempz2-(2*tsit+tgradient2)/2)/(tgradient2/2)))+0.995)*((cesige-
csige)/1.99)+csige;


   if(B==0)

   y=[csit*ones(tsit/dz,1);cgradpoint1;cesige*ones(tesige/dz,1);cgradpoint2;csib*ones(tsib/dz,1)
; cgradpoint3;csige*ones(tsige/dz,1)];

   end

   if(B==1)

   y=[csit*ones(tsit/dz,1);cgradpoint1;cesige*ones(tesige/dz,1);cgradpoint4;csige*ones(tsige/dz,
1)];

   end


plot((1:length(y))*dz,y);

pause(1);


%description of the concentration

[t,y]=ode15s(@odefun10,[0 time],y,[],Temp,0,csige,dz,qprime);
```

```
name='tri_varyGetime=';
name1=[name num2str(time) 'temp=' num2str(Temp) 'qprime=' num2str(qprime) 'ceSiGe=' num2str(cesige) '.mat']

save(name1)

clear;
```

In the function odefun20 we will set up the Fick's law in the form of $c'=f(y,t)$.This equation is integrated by the ode15s solver between the specified time to give the time evolution of the concentration. In the standard finite difference scheme, the diffusion coefficient at node $i+1/2$ is approximated to be equal to average of diffusivity at node i and node $i+1$ and results in a finite difference scheme as described in reference [73]. This approximation has some inaccuracies because since the diffusion is exponential, this averaging scheme will lead to an overestimation of the diffusion coefficient. We can improve this scheme by evaluating the diffusion coefficient of the averaged concentration at $(i+1/2)^{th}$ node as given in Equation 4-5. The improved scheme for evaluating Fick's law in finite difference is given below as.

$$\frac{c_{i,j+1}-c_{i,j}}{\Delta t}=\left(\frac{D\left(\frac{c_{i-1,j}+c_{i,j}}{2}\right)c_{i-1,j}-\left(D\left(\frac{c_{i-1,j}+c_{i,j}}{2}\right)+D\left(\frac{c_{i+1,j}+c_{i,j}}{2}\right)\right)c_{i,j}+D\left(\frac{c_{i+1,j}+c_{i,j}}{2}\right)c_{i+1,j}}{\Delta x^2}\right)$$

Where, $c_{i,j}$ is the concentration of Ge in the $Si_{1-y}Ge_y$ film at the node i at the jth time step , $D(c_{i,j})$ is the diffusivity of Ge at the node i at the jth time step, and depends on concentration and the strain at that point, t is the diffusion time and x is the coordinate

axis. We will also require Diffusion coefficient at the nodes and that will be calculated using the function Dgesi_Zang_saur_withstrain1. The boundary conditions are also set to be no flux boundary conditions at the boundaries.

```
function aprime = odefun20(time,a,T,asi,asige,dz,qprime)

% 1D diffusion time propagator where

% a=activity vector
% Dstr(a) is the diffusivity function passed as a string
% dz= spatial descretization in nm
% j*dz = total length of the system
% T is temperature in C
% NOTE: In this function, a(1)=constant

%T=900;


n=length(a);
%dz=1/(n-1);
Dstr='Dgesi_Zang_saur_withstrain1';



a=[asi; a ; asige];

%c_half=(a(1:n+1)+a(2:n+2))/2;
%Dp=feval(Dstr,c_half,T);
Dp=feval(Dstr,a,T,qprime,asige);

% for i=2:n+1
%
%     p1=feval(Dstr,a(i+1),T);
%     p2=feval(Dstr,a(i),T);
%     p3=feval(Dstr,a(i-1),T);
%
%     aprime(i-1)=0.5/dz^2* ( (p1+p2)*(a(i+1)-a(i)) - (p2+p3)*(a(i)-a(i-1)) );
%
% end

A=spdiags([(Dp(2:n+1)+Dp(1:n)) -(Dp(3:n+2)+2*Dp(2:n+1)+Dp(1:n))
(Dp(3:n+2)+Dp(2:n+1))]/(2*(dz^2)), -1:1,n,n);

A(1,1)=A(1,1)+Dp(1)/(dz^2);
```

```
%A(n,n)=A(n,n)+Dp(n)/(dz^2);
bc=zeros(n,1);
bc(1)=0; %Dp(1)*asi/(dz^2);
bc(n)=Dp(n+1)*asige/(dz^2);
aprime=A*a(2:n+1)+bc;
%aprime=aprime';
```

The final step for solving the Fick's law is evaluating the diffusion coefficient at each node. We found that the values of E_a and D_0 for self-diffusion in SiGe[63] follow Meyer–Neldel's equation. The presence of Meyer-Neldel rule in the SiGe system was seen by Dunstan[67] and Zangenberg[68] and the theoretical form of the rule was derived by Khait[69].

$$\ln(D_0) = \left(\ln\left(\frac{\alpha^2}{6\Delta\tau} \right) - \frac{\Delta S_m}{k} - 3\ln(A) \right) + \frac{E_A}{kT_m}$$

where, D_0 is pre-exponential factor, α is the diffusion step, $\Delta\tau$ is the time scale, T_m is the melting temperature, A is a dimensionless length scale fitting parameter, E_A is the activation energy and ΔS_m is the entropy of melting. The values of the T_m and α have been linearly interpolated between values for Si and Ge. The values for the factors used are the same as Zangenberg[68]. Interdiffusivity is given by including a Darken's term in the diffusion coefficient[59]. The Kirkendall effects are neglected[59] resulting in relatively simple modification of the self diffusivity of Ge in SiGe to interdiffusivity.

An important disagreement occurs in the literature on the value of coupling strength[66] Q', which determines the contribution of strain on E_A. The diffusion characteristics of dual channels are very sensitive to the value of Q' since the layers have different states and magnitude of strain. Calculating the profiles of different strained SiGe heterostructures and then comparing with the experimentally measured profiles, we find that a value of 23 eV/unit strain for Q' results in good fit.

```
function D = Dgesi_Zang_saur_withstrain1(xge,T,qprime,xsub)

xge0=xsub; %substrate Ge composition

%Put Qprime=0 , if unstrained case is desired
Qprime=qprime;
%D01=zeros(length(xge));

%%_______________________________________________


alpha=2.35*(1-xge)+2.45*xge;
tau=3e-13;
Sm=13.5;
A=6;
Tm=1687*(1-xge)+1211*(xge);


% ______________________________________________

%using Strohm's equation - fitting the the values of activation energy

A1=4.82862;
B1=-1.68249;
Q=A1+B1*xge;

Q1=Q+Qprime*(-0.042*(xge-xge0));

k=8.617342e-5; %in eV/K
kT=k*(273+T); %in K

% ______________________________________________

% Calculating the preexponential factor in m^2/s
```

```
a1=log(((alpha.^2)*1e-20)/(6*tau));
a2=(-Sm-3*log(A))*ones(size(xge));
a3=((Q*1.6e-19)./(1.38e-23*Tm));

D01=exp(a1+a2+a3);

D0=D01*1e18*60; % Diffusion coefficient in nm^2/min
%D0=D01; % Diffusion coefficient in m^2/s

%_______________________________________________
%calculates interdiffusivity

darken=1-(6500*xge.*((1-xge)/(8.314*(273+T))));

D=D0.*darken.*exp(-(Q1)/kT);

%D=D';
```

BIBLIOGRAPHY

1. C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, Applied Physics Letters **79,** 4246-4248 (2001).
2. J.-W. Jung, S. F. Yu, O. O. Olubuyide, J. L. Hoyt, D. A. Antoniadis, M. L. Lee, and E. A. Fitzgerald, Applied Physics Letters **84,** 3319-3321 (2004).
3. S. Gupta, M. L. Lee, D. M. Isaacson, and E. A. Fitzgerald, Materials Science and Engineering B **124-125,** 102 (2005).
4. S. Gupta, M. L. Lee, G. Taraschi, A. Pitera, and E. A. Fitzgerald, in *Tri-layer heterostructures for improved PMOS enhancements preserved over a large processing temperature range*, Honolulu, 2004 (ECS conference proceedings).
5. ITRS, International Technology Roadmap for Semiconductors **http://public.itrs.net/** (2003).
6. H. Yin, K. D. Hobart, F. J. Kub, S. R. Shieh, T. S. Duffy, and J. C. Sturm, Applied Physics Letters **82,** 3853-3855 (2003).
7. T. Tezuka, N. Sugiyama, S.-I. Takagi, and T. Kawakubo, Applied Physics Letters **80,** 3560-3562 (2002).
8. E. A. Fitzgerald, Materials Science Reports **7,** 87-142 (1991).
9. E. A. Fitzgerald, A. Y. Kim, M. T. Currie, T. A. Langdo, G. Taraschi, and M. T. Bulsara, Materials Science and Engineering **B67,** 53-61 (1999).
10. M. T. Currie, S. B. Samavedam, T. A. Langdo, C. W. Leitz, and E. A. Fitzgerald, Applied Physics Letters **72,** 1718-1720 (1998).
11. V. K. Yang, M. E. Groenert, C. W. Leitz, A. J. Pitera, M. T. Currie, and E. A. Fitzgerald, Journal of Applied Physics **93,** 3859 (2003).
12. A. L. Caviglia and A. A. Iliadis, IEEE Electron Device Letters **14,** 133 (1993).
13. C. W. Leitz, M. T. Currie, M. L. Lee, Z. Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, Journal of Applied Physics **92,** 3745-3751 (2002).
14. C. W. Leitz, Ph.D. Thesis, Massachusetts Institute of Technology, 2002.
15. G. Hock, E. Kohn, C. Rosenblad, H. von Kanel, H.-J. Herzog, and U. Konig, Applied Physics Letters **76,** 3920-3922 (2000).
16. M. L. Lee, Z.-Y. Cheng, C. W. Leitz, A. J. Pitera, T. A. Langdo, M. T. Currie, G. Taraschi, E. A. Fitzgerald, and D. A. Antoniadis, Applied Physics Letters **79,** 3344-3346 (2001).
17. H. Shang, J. O. Chu, X. Wang, P. M. Mooney, K. Lee, J. Ott, K. Rim, K. Chan, K. Guarini, and M. Ieong, Symposium on VLSI Technology, 204-205 (2004).
18. M. L. Lee, E. A. Fitzgerald, M. T. Bulsara, M. T. Currie, and A. Lochtefeld, Journal of Applied Physics **97,** 011101-1 (2005).
19. E. A. Fitzgerald, Annual Review of Materials Science **25,** 417-54 (1995).
20. M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, D. A. Antoniadis, and E. A. Fitzgerald, Journal of Vacuum Science and Technology B **19,** 2268-2279 (2001).

[21] E. A. Fitzgerald, K. C. Wu, M. T. Currie, N. Gerrish, D. Bruce, and J. Borenstein, in *Microelectromechanical Structures for Materials Research Materials Research Society Symposium Proceedings*, 1998.
[22] J. A. Carlin, S. A. Ringel, E. A. Fitzgerald, M. Bulsara, and B. M. Keyes, Applied Physics Letters **76,** 1884-1886 (2000).
[23] V. K. Yang, M. E. Groenert, G. Taraschi, C. W. Leitz, A. J. Pitera, M. T. Currie, Z.-Y. Cheng, and E. A. Fitzgerald, Journal of Materials Science: Materials in Electronics **13,** 377-380 (2002).
[24] S. K. Gandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide* (John Wiley & Sons, New York, 1994).
[25] E. A. Fitzgerald, Y.-H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, A. R. Kortan, F. A. Thiel, and B. E. Weir, Journal of Vacuum Science and Technology B **10,** 1807-1819 (1992).
[26] J. W. Matthews, in *Epitaxial Growth Part B*, edited by J. W. Matthews (Academic Press, New York, 1975), p. 559-607.
[27] D. G. Schimmel, J. Electrochem. Soc. **126,** 479-482 (1979).
[28] J. Lai, M.S. Thesis, Mass. Institute of Technology (2002).
[29] D. C. Houghton, J. Appl. Phys. **70,** 2136 (1991).
[30] S. B. Samavedam and E. A. Fitzgerald, Journal of Applied Physics **81,** 3108-3116 (1997).
[31] L. B. Freund, J. Appl. Phys. **68,** 2073 (1990).
[32] L. B. Freund, Journal of Applied Physics **68,** 2073 (1990).
[33] C. W. Leitz, M. T. Currie, A. Y. Kim, J. Lai, E. Robbins, and E. A. Fitzgerald, Journal of Applied Physics **90,** 2730-2736 (2001).
[34] C. W. Leitz, M. T. Currie, A. Y. Kim, J. Lai, E. Robbins, E. A. Fitzgerald, and M. T. Bulsara, Journal of Applied Physics **90,** 2730 (2001).
[35] M. Erdtmann, M. Currie, J. C. Woicik, and D. Black, Materials Research Society Symposia Proceedings **2005** (2005).
[36] J. S. Speck, M. A. Brewer, G. Beltz, A. E. Romanov, and W. Pompe, J. Appl. Phys. **80,** 3808-3816 (1996).
[37] H. C. Luan, D. R. Lim, K. K. Lee, K. M. Chen, J. G. Sandland, K. Wada, and L. C. Kimerling, Applied Physics Letters **75,** 2909-2911 (1999).
[38] O. Nam, M. Bremser, T. Zheleva, and R. Davis, Applied Physics Letters **71,** 2638-2640 (1997).
[39] E. A. Fitzgerald, Y.-H. Xie, M. L. Green, D. Brasen, A. R. Kortan, J. Michel, Y.-J. Mii, and B. E. Weir, Applied Physics Letters **59,** 811-813 (1991).
[40] B. W. Dodson and J. Y. Tsao, Appl. Phys. Lett. **52,** 852 (1988).
[41] Volkert, Journal of electronic materials (1991).
[42] T. C. Wang, Y. W. Zhang, and S. J. Chua, Journal of Applied Physics **89,** 6069 (2001).
[43] P. K. Ko, VLSI Electronics:Microstructure Science **18,** 1-37 (1989).
[44] M. Lundstrom, *Fundamentals of Carrier Transport*, Vol. X (Addison-Wesley Publishing Company, Reading, MA, 1990).
[45] B. G. Streetman, *Solid State Electronic Devices* (Prentice Hall, Englewood Cliffs, NJ, 1995).

[46] S. M. Sze, *Physics of Semiconductor Devices* (John Wiley and Sons, New York, 1981).
[47] S.-i. Takagi, A. Toriumi, M. Iwase, and H. Tango, IEEE Transactions on Electron Devices **41,** 2357-2362 (1994).
[48] M. Ieong, MIT MTL Seminar (2004).
[49] T. Fromherz and G. Bauer, in *Properties of strained and relaxed Silicon Germanium*; *Vol. 12*, edited by E. Kasper (INSPEC, London, U.K., 1994), p. 87-93.
[50] R. People and J. C. Bean, Applied Physics Letters **48,** 538-540 (1986).
[51] M. Armstrong, PhD Thesis, Massachusetts Institute of Technology, 1999.
[52] Sunamura, Fukatsu, Usami, and Shiraki, Applied Physics Letters **63** (1993).
[53] Zeindl, Nilsson, and Bugiel, Applied Surface Science **92,** 552-556 (1996).
[54] H. P. Zeindl, S. Nilsson, U. Jagdhold, J. Klatt, R. Kurps, D. Kroeger, and E. Bugiel, Applied Surface Science **102,** 107-111 (1996).
[55] G. Prudon, B. Gautier, J. C. Dupuy, C. Dubois, M. Bonneau, J. Delmas, J. P. Vallard, G. Bremond, and R. Brenier, Thin Solid Films **294,** 54-58 (1997).
[56] M. Narasimhan, (http://www.eetimes.com/in_focus/silicon_engineering/OEG20030822S0025).
[57] Laitnen, Strohm, Huikari, Nieminen, Voss, and Grodon, Physics review Letter **89** (2002).
[58] N. R. Zangenberg, J. L. Hansen, J. Fage-Pedersen, and A. N. Larsen, Physical Review Letters **87,** 125901-1-125901-4 (2001).
[59] D. B. Aubertine, M. A. Mander, N. Ozguven, A. F. Marshall, P. C. McIntyre, J. O. Chu, and P. M. Mooney, Journal of Applied Physics **92,** 5027-5035 (2002).
[60] N. E. B. Cowern, P. C. Zalm, P. van der Sluis, D. J. Gravesteijn, and W. B de Boer, Physical Review Letters **72,** 2585-2588 (1994).
[61] McVay and DuCharme, Physical Review B **9,** 627-631 (1974).
[62] A. Strohm, T. Voss, W. Frank, J. Raisanen, and M. Dietrich, Physica B: Condensed Matter **308,** 542-545 (2001).
[63] A. Strohm, T. Voss, W. Frank, P. Laitinen, and J. Raisanen, Zeitschrift fur Metallkunde **93,** 737 (2002).
[64] Y. S. Lim, J. Y. Lee, H. S. Kim, and D. W. Moon, Appl. Phys. Lett. **77,** 4157-4159 (2000).
[65] H. Zhu, The Electrochemical Society Proceedings, 206th Meeting, SiGe: Materials, Processing and Devices **2004 -7,** 923-934 (2004).
[66] D. B. Aubertine and P. C. McIntyre, J. Appl. Phys. **97,** 013531-10 (2005).
[67] D. J. Dunstan, Solid State Communications **107,** 159-163 (1998).
[68] N. R. Zangenberg and A. N. Larsen, Physica B **340-342,** 780-783 (2003).
[69] Y. L. Khait, R. Beserman, D. Shaw, and K. Dettmer, Phys. Rev. B **50,** 14893-14902 (1994).
[70] A. Antonelli and J. Bernholc, Phys. Rev. B **40,** 10643-10646 (1989).
[71] P. Kuo, J. Hoyt, J. F. Gibbons, J. E. Turner, R. D. Jacowitz, and T. I. Kamins, Appl. Phys. Lett. **66,** 580 (1993).
[72] M. J. Aziz, Appl. Phys. Lett. **70,** 2810 (1997).
[73] P.-E. Hellberg, S.-L. Zhang, F. M. d'Heurle, and C. S. Petersson, J. Appl. Phys. **82,** 5779-5787 (1997).

[74] D. K. Nayak, J. C. S. Woo, J. S. Park, K. L. Wang, and K. P. MacWilliams, Applied Physics Letters **62,** 2853-2855 (1993).

[75] K. Rim, J. Welser, J. L. Hoyt, and J. F. Gibbons, in *Enhanced Hole Mobilities in Surface-channel Strained-SI p-MOSFETs*, Washington, DC, 1995, p. 517-519.

[76] K. Rim, J. Chu, H. Chen, K. A. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Ieong, and H.-S. Wong, in *Characteristics and Device Design of Sub-100 nm Strained Si N- and PMOSFETs*, Honolulu, HI, 2002, p. 98-99.

[77] C. K. Maiti, L. K. Bera, S. S. Dey, D. K. Nayak, and N. B. Chakrabarti, Solid-State Electronics **41,** 1863-1869 (1997).

[78] D. K. Nayak, K. Goto, A. Yutani, J. Murota, and Y. Shiraki, IEEE Transactions on Electron Devices **43,** 1709-1716 (1996).

[79] D. K. Nayak and S. K. Chun, Applied Physics Letters **64,** 2514-2516 (1994).

[80] M. V. Fischetti and S. E. Laux, Journal of Applied Physics **80,** 2234-2252 (1996).

[81] M. Arafa, K. Ismail, P. Fay, J. Chu, B. S. Meyerson, and I. Adesida, Electronics Letters **31,** 680-681 (1995).

[82] M. L. Lee, Ph.D. Thesis, MIT, 2003.

[83] S. K. Chun and K. L. Wang, IEEE Transactions on Electron Devices **39,** 2153-2164 (1992).

[84] M. L. Lee and E. A. Fitzgerald, Journal of Applied Physics **94,** 2590-2596 (2003).

[85] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, Applied Physics Letters **79,** 4246 (2001).

[86] M. L. Lee and E. A. Fitzgerald, Applied Physics Letters **83,** 4202-4204 (2003).

[87] S. Gupta, M. L. Lee, and E. A. Fitzgerald, Appl. Phys. Lett. **86** (2005).

[88] M. L. Lee and E. A. Fitzgerald, Journal of Vacuum Science and Technology B **22,** 158-164 (2004).

[89] Y. H. Xie, G. H. Gilmer, C. Roland, P. J. Silverman, S. K. Buratto, J. Y. Cheng, E. A. Fitzgerald, A. R. Kortan, S. Schuppler, M. A. Marcus, and P. H. Citrin, Physical Review Letters **73,** 3006-3009 (1994).

[90] M. L. Lee and E. A. Fitzgerald, Journal of Applied Physics **95,** 1550-1555 (2004).

[91] S. Motamarri, http://www2.egr.uh.edu/~smotamar/RTP/RTP.html.

[92] Zamir, Meyler B., and Salzman J., Appl. Phys. Lett. **78,** 288-90 (2001).

[93] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *NUMERICAL RECIPES IN C: THE ART OF SCIENTIFIC COMPUTING* (1992).