

**ANALYTICAL DESIGN AND NUMERICAL VERIFICATION OF  
P-CHANNEL STRAINED SILICON-GERMANIUM HETERO MOSFET**

**by**

**Mohan Krishnan Gopal**

Copyright © Mohan Krishnan Gopal 2008

A Dissertation Submitted to the Faculty of the  
**DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING**

In Partial Fulfillment of the Requirements

For the Degree of

**DOCTOR OF PHILOSOPHY**

In the Graduate College

**THE UNIVERSITY OF ARIZONA**

2008

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Mohan Krishnan Gopal

entitled “**Analytical Design and Numerical Verification of p-Channel Strained Silicon-Germanium Hetero MOSFET**”

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of **Doctor of Philosophy**

\_\_\_\_\_ Date: **November 25, 2008**  
**John R. Brews, Ph.D.**

\_\_\_\_\_ Date: **August 1, 2008**  
**Harold G. Parks, Ph.D.**

\_\_\_\_\_ Date: **August 1, 2008**  
**Olgierd A. Palusinski, Ph.D.**

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

\_\_\_\_\_ Date: **November 25, 2008**  
**Dissertation Director: John R. Brews, Ph.D.**

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED -----

Mohan Krishnan Gopal

## ACKNOWLEDGEMENTS

My profuse thanks go to Dr. John R. Brews for his tremendous patience with my intermittent progress. Apart from that I am extremely thankful to Prof. Brews for his technical advice, technical help and personal concern during the latter days of my dissertation. My sincere thanks go to Dr. Harold G. Parks and Dr. Olgierd A. Palusinski for reviewing my dissertation and agreeing to be members of my examination committee. I am very grateful to Dr. Kelly S. Potter for agreeing to be in the examination committee representing College of Optical Sciences, my minor department. I am sincerely thankful to Dr. Carl F. Maes of the College of Optical Sciences for his sympathetic attitude and all the help he has offered. I am very thankful to Dr. Thomas C. Cetas for making me part of the Microwave Hyperthermic Oncology research group and supporting me for many years. I would like to thank Dr. Mark A. Neifeld for his technical advice and support while doing research with amorphous silicon for few years. I am grateful to Dr. Jeffrey J. Rodriguez for his encouragement, co-operation and sympathetic attitude. My thanks go to Dr. Hal S. Tharp for being helpful and considerate about my study room accommodation at the department. One particular person I would like to thank very much is Mr. Leo Enfield. He was always there to help me with the simulation program and PC support. I am thankful to Mr. Lewis Dupont for being always responsive about my study room accommodation. I would like to thank Ms. Tami Whalen profusely for being there always willing to help with academic matters. I am thankful to Ms. Gail Varin of Optical Sciences and Ms. Nancy Lindsay of the graduate college for all their help.

**DEDICATED**  
**TO THE LOVING MEMORY OF**

My mother: *C. Ponnamma*

My father: *M. Gopalan Nair*

My uncle: *C. Rajappan Nair*

My brother: *C. Chandrasekharan Nair*

My favorite professor: *M. R. Harsha Jyothi*

Long time family friend: *A. V. Divakaran Pillay*

My good friend: *Philip Thomas*

## TABLE OF CONTENTS

<b>LIST OF FIGURES .....</b>	<b>11</b>
<b>LIST OF TABLES .....</b>	<b>23</b>
<b>ABSTRACT .....</b>	<b>24</b>
<b>1. INTRODUCTION.....</b>	<b>26</b>
1.1 Short History of Silicon CMOS Devices .....	27
1.1.1 Technology node .....	28
1.1.2 Operating Speed.....	30
1.1.3 Power Dissipation.....	30
1.1.4 Low Voltage Systems.....	31
1.2 Current Status of Silicon CMOS.....	32
1.3 Future of Silicon CMOS .....	34
1.4 Introduction to Silicon Germanium ( $\text{Si}_{1-x}\text{Ge}_x$ ).....	35
1.4.1 Advantages of $\text{Si}_{1-x}\text{Ge}_x$ .....	38
1.4.2 Silicon Germanium in all Silicon Technology .....	39
1.4.3 Current Status of $\text{Si}_{1-x}\text{Ge}_x$ Technology.....	41
1.5 Limitations of Si CMOS .....	43
1.5.1 Si/SiO <sub>2</sub> Interface .....	46
1.5.2. Perpendicular Fields .....	46
1.5.3. Substrate Doping .....	48
1.5.4. Improving Silicon CMOS Performance .....	49
1.6 Objective and Organization of the Project .....	51
1.6.1 Objective.....	51
1.6.2 Procedure .....	54
1.6.3 Organization of chapters.....	58
<b>2. MODERN SHORT CHANNEL PMOS.....</b>	<b>61</b>
2.1 Scaling.....	62
2.2 Short Channel Effects.....	64
2.2.1 Velocity Saturation .....	64
2.2.2 Pinch off, Saturation and Channel Length Modulation .....	64
2.2.3 Drain Induced Barrier Lowering (DIBL) .....	65
2.2.4 Velocity Overshoot and Hot Carriers .....	66
2.2.5 Substrate Current Induced Body Effect.....	67
2.2.6 Gate Tunnel Current .....	67
2.2.7 Gate Oxide Breakdown.....	68
2.2.8 Sub-surface Punch Through .....	68
2.2.9 Random Dopant Effects.....	68
2.2.10 Ballistic Transport .....	68

## TABLE OF CONTENTS - Continued

2.2.11 Velocity Overshoot.....	69
2.3. Nano Scale Device Design Aspects .....	69
2.4 Analysis of Short Channel MOSFET .....	72
2.4.1 Drain Current .....	72
2.4.2 Transconductance and Transition Frequency .....	73
2.4.3 Scaling Limits and Reliability .....	74
<b>3. PRELIMINARY STUDIES OF A QW MOSFET .....</b>	<b>75</b>
3.1 Introduction .....	75
3.2 Solution of One Dimensional Poisson Equation .....	77
3.3 Field at the front of the QW .....	84
3.4 The Gate Voltage .....	85
3.5 Simulation Results.....	86
3.5.1 Goal Seek to Set QW Width and Gate Voltage.....	87
3.5.2 Excel Program and Device Details .....	88
3.5.3 ON State Hole Sheet Density in the QW .....	91
3.5.4 OFF State Hole Sheet Density in the QW .....	93
3.5.5 Hole Density (ON/OFF) Ratio in the QW .....	94
3.5.6 Parasitic Channel .....	97
3.5.7 Sub-threshold Slope.....	98
3.6 Trade Off Studies .....	102
3.6.1 Trade off: p-delta Doping Density.....	102
3.6.1.1 The ON State.....	102
3.6.1.2 Conclusion for the ON state.....	106
3.6.1.3 Off state.....	107
3.6.1.4 Conclusion for the OFF state .....	112
3.6.1.5 Conclusion for p-delta Doping Simulations.....	112
3.6.2 Tradeoff: QW Height.....	113
3.6.2.1 Optimization of QW Height.....	115
3.6.2.2 Conclusion for QW height Simulations .....	117
3.6.3 Trade off: QW width .....	117
3.6.3.1 Simulation Results .....	118
3.6.3.2 Conclusion .....	120
3.6.4 Trade off: Substrate Doping Density .....	121
3.6.4.1 Simulation set up.....	121
3.6.4.2 Potential plots.....	122
3.6.4.3 Simulation results.....	125
3.6.4.4 Conclusions .....	128
3.6.4.5 Overall Discussion .....	129
3.7 Lessons Learned from the Above Simulations.....	132
3.8 Summary .....	135

## TABLE OF CONTENTS - Continued

<b>4. EFFECTS OF STRAIN ON <math>\text{Si}_{1-x}\text{Ge}_x</math> AND HMOSFET.....</b>	<b>137</b>
4.1 Properties of $\text{Si}_{1-x}\text{Ge}_x$ .....	138
4.2 Strained $\text{Si}_{1-x}\text{Ge}_x$ .....	140
4.2.1 Pseudomorphic Growth of $\text{Si}_{1-x}\text{Ge}_x$ .....	140
4.2.2 Critical Thickness .....	142
4.3 Compressive Strain Induced Effects in $\text{Si}_{1-x}\text{Ge}_x$ .....	144
4.3.1 Band Gap Engineering Using $\text{Si}_{1-x}\text{Ge}_x$ .....	144
4.3.2 Band Edge Alignment in $\text{Si}/\text{Si}_{1-x}\text{Ge}_x$ .....	147
4.3.3 Transport Properties.....	149
4.3.3.1 Effective Hole Mass .....	149
4.3.3.2 Hole Mobility .....	151
4.3.3.3 Velocity Saturation.....	154
4.3.3.4 Valence Band Density of States .....	156
4.3.4 Summary of Strain Induced Effects.....	157
4.4 Introduction to p-HMOSFET .....	158
4.5 Survey of p channel HMOSFETs.....	162
4.5.1. Main p-HMOSFET Structures.....	162
4.5.2. Improved p-channel HMOSFET Structures .....	167
4.6 Generalized Structure for this Project .....	169
4.7 Summary .....	173
<b>5. DESIGN CONSIDERATIONS OF A P-CHANNEL HMOSFET.....</b>	<b>175</b>
5.1. Introduction .....	175
5.2. Material Considerations .....	178
5.3. p-channel HMOSFET Structural Details .....	178
5.3.1. The Substrate .....	179
5.3.2 The p-doped Si Layer .....	181
5.3.3 The Undoped Si Buffer Layer .....	182
5.3.4 Strained $\text{Si}_{1-x}\text{Ge}_x$ Quantum Well Channel Layer.....	184
5.3.5 The Undoped Si Cap Layer .....	185
5.3.6 Gate Oxide .....	187
5.3.7 Gate Material .....	187
5.3.8 Fixed Parameters and Design Variables .....	187
5.3.9 Device Notation .....	188
5.4 Summary .....	189
<b>6. ANALYTICAL ANALYSIS AND DESIGN PROGRAM .....</b>	<b>190</b>
6.1 Solution of One Dimensional Poisson Equation .....	191
6.2 Analytical Analysis and Design Program .....	200
6.2.1 Program Inputs and Outputs .....	201
6.2.2 Dissection of the Program .....	203
6.3 Verification of the Analytical Program .....	207

## TABLE OF CONTENTS - Continued

6.3.1 Potential functions .....	208
6.3.2 Hole Densities.....	214
6.4 Summary and Conclusions.....	214
<b>7. LONG CHANNEL DEVICE SIMULATIONS.....</b>	<b>216</b>
7.1 Parametric Simulations of Design Variables .....	216
7.1.1 Ge fraction, $x$ , of the SiGe QW Channel .....	218
7.1.2 Substrate Doping Density, $N_{\text{sub}}$ .....	224
7.1.3 p-doped Layer Thickness, $X_p$ .....	230
7.1.4 Buffer Layer Thickness, $X_{\text{buf}}$ .....	235
7.2 Long Channel Device Design Windows .....	237
7.3 An Optimized Device.....	243
7.4 Summary of Long Channel Device Simulations.....	245
<b>8. SHORT CHANNEL DEVICE SIMULATIONS.....</b>	<b>247</b>
8.1 Design Program for MOSFET with $N_{\text{sub}}$ as Parameter .....	251
8.2 Modified Analytical Design Program for Short Channel HMOSFET .....	257
8.3 Short Channel simulations .....	258
8.3.1 Short channel comparison.....	260
8.3.2 Sub-threshold Comparison .....	264
8.3.2.1 ON Current Calculation .....	269
8.3.2.2 Riddle of the large threshold voltage of HMOSFET .....	279
8.4 Scalability of $\text{Si}_{1-x}\text{Ge}_x$ HMOSFET .....	285
8.5 Summary of Short Channel Simulations.....	286
<b>9. OPTIMIZATION OF THE P-HMOSFET STRUCTURE .....</b>	<b>288</b>
9.1 Limitations of the p-HMOSFET Structure.....	288
9.2 A Modified Structure .....	289
9.3 A Feasibility Analysis of the Modified Structure .....	290
9.4 Analysis Using the Analytical HMOSFET design Program.....	291
9.5 Analysis Using Numeric Simulator.....	294
9.6 Conclusion.....	298
<b>10. CONCLUSION .....</b>	<b>300</b>
10.1 Contributions of this Dissertation .....	301
10.2 Outline of the Tool .....	306
10.3 Summary .....	309
10.4 Limitations .....	309
10.5 Recommendations for Future Work .....	310

**TABLE OF CONTENTS - Continued**  
**INDEXES**

<b>A: REVIEW OF CONVENTIONAL MOSFET .....</b>	<b>312</b>
<b>B: EXCEL SPREAD SHEET – BAND BENDING WITH P-DELTA .....</b>	<b>328</b>
<b>C: ONE DIMENSIONAL POISSON EQUATION SOLUTION .....</b>	<b>335</b>
<b>D: EXCEL ANALYTICAL P-HMOSFET DESIGN PROGRAM .....</b>	<b>343</b>
<b>E: A HIGH FREQUENCY STRUCTURE OF P-HMOSFET .....</b>	<b>359</b>
<b>REFERENCES.....</b>	<b>361</b>

## LIST OF FIGURES

Figure 1-1. Sketch of a p-MOSFET.....	44
Figure 1-2. Band diagram of the p-MOSFET at high $V_{gs}$ .....	45
Figure 1-3. Hole mobility vs effective perpendicular field, $E_{perp}$ . The upper and lower curves corresponds to doping densities of $3.9 \times 10^{15}$ and $3 \times 10^{17}$ respectively.....	47
Figure 1-4. Hole mobility vs substrate doping density $N_a$ .....	49
Figure 1-5. A simple SiGe MOSFET. It has a SiGe surface channel. ....	55
Figure 1-6. A recessed SiGe channel MOSFET. It has an undoped Si cap layer. ....	55
Figure 1-7. A recessed SiGe channel MOSFET with additional p+Si layer and an undoped Si buffer layer.....	56
Figure 2-1. Cross sectional view of a basic long channel p-MOSFET.....	62
Figure 2-2. Potential plot from source to drain. As the gate length is reduced the potential barrier at the source junction reduces and carrier injection from source to channel occurs.....	66
Figure 2-3. Sketch of a modern nanometer scale p-MOSFET. Drawing is not to scale. Source/Drain regions are enlarged to show details. The device has source/drain extensions (SDE) and a deep high doped Retrograde layer. ....	70
Figure 3-1. A modified structure of a MOSFET with the introduction of a QW. A p-delta layer is also introduced for additional reasons mentioned in the text.....	78
Figure 3-2. Valence band sketch of an ideal device in the ON state. The bulk potential, $u_{nB}$ , ( $u_B - u_W$ ) is assumed. $u_W$ and $Q_d$ are input parameters. ....	79
Figure 3-3. ON and OFF potentials with a substrate doping of $5 \times 10^{18}/\text{cm}^3$ .....	89
Figure 3-4. Display panel for Figure 3-3. ....	89
Figure 3-5. Valence band diagram at zero gate bias. Notice that the bulk potential drop is not zero due to the negative ions in the p-delta layer. The QW tilts downwards indicating that the ions in the p-delta layer is stronger than the positive ions in the substrate. ....	90
Figure 3-6. Hole sheet density dependence on QW width. $n_B = 5 \times 10^{18}/\text{cm}^3$ , .....	91
Figure 3-7. Hole sheet density dependence on QW height. $n_B = 5 \times 10^{18}/\text{cm}^3$ ,.....	92
Figure 3-8. Hole sheet density dependence on p-delta doping. $n_B = 5 \times 10^{18}/\text{cm}^3$ ,.....	92

### LIST OF FIGURES - Continued

Figure 3-9. Valence band plots for the ON and OFF states with a QW height of 4 kT/q. The marker just to the right of the QW is the top of the cap layer. This point is lying above the QW and forms a parasitic surface channel. ....	93
Figure 3-10. Valence band plots for ON and OFF states with a QW height of 10 kT/q. It can be seen that the top of cap layer (marker to the right of the QW) is lying below the top of the QW. Now the QW has much higher carrier population compared to the cap layer. This figure is the same as Figure 3-3.....	94
Figure 3-11. Valence band sketch of the device in the ON state.....	95
Figure 3-12. Hole sheet density ON/OFF ratio dependence on QW width.....	96
Figure 3-13. Hole density ON/OFF ratio vs QW height. $n_B = 5e18/cm^3$ ,.....	96
Figure 3-14. Hole sheet density ON/OFF ratio vs p-delta doping.....	97
Figure 3-15. Log(hole sheet density) vs $V_{gs}$ , sub-threshold plot for the p-MOSFET device. Since hole sheet density is proportional to the drain current the ratio of hole sheet densities in the ON and OFF states can be used instead of $I_{ON}/I_{OFF}$ .....	100
Figure 3-16. Sub-threshold slope dependence on QW width, $x_W.u_W = 10$ kT/q,.....	100
Figure 3-17. Sub-threshold parameter, S, dependence on QW height, $u_W$ .....	101
Figure 3-18. Sub-threshold slope dependence on p-delta doping. $n_B = 5e18/cm^3$ ,.....	101
Figure 3-19. ON state sketch. Negative p-delta ions are less than the positive bulk depletion ions. The field in the QW is positive since the positive ions is stronger. ....	104
Figure 3-20. ON state sketch . p-delta negative ions is equal to the bulk depletion positive ions. Since the ions cancel each other the field at the back of the QW is zero and the QW is flat at the left side.....	105
Figure 3-21. ON state sketch. p-delta negative ions are more than the bulk depletion positive ions. So the field in the QW is negative and the band bends downwards at the back of the QW.....	106
Figure 3-22. Simulated potentials, using the Excel program, in the ON state for a device with $u_W = 10$ kT/q, $x_W = 5$ nm, $n_B = 5e18/cm^3$ . $u_{nB}$ is the bulk depletion potential, $u_{cap}$ is the potential at the top of the cap layer, $u_{on}$ is the potential at the top of the QW in the ON state or $V_{gs} = 1.5$ V.....	107
Figure 3-23. OFF state sketch. p-delta ions were less than depletion ions in the ON state. But now p-delta ions may be stronger. Since bulk depletion is caused by p-delta ions the depletion will be weak, $u_B$ will be low and QW will not be lifted much compared to the next two cases.....	108

### LIST OF FIGURES - Continued

Figure 3-24. OFF state sketch. p-delta ions were equal to the depletion ions in the ON state. But now the p-delta ions are stronger. This can increase $u_B$ and lift the QW .....	109
Figure 3-25. OFF state sketch. p-delta ions were more than bulk depletion ions in the ON state. Now p-delta ions are very high leading to high $u_B$ and lifts the QW. This case may have the highest OFF state hole density .....	109
Figure 3-26. ON and OFF state potentials for a device with $u_W = 10 \text{ kT/q}$ , .....	110
Figure 3-27. ON and OFF state potentials for a device with $u_W = 10 \text{ kT/q}$ , .....	111
Figure 3-28. Simulated potentials in the OFF state for a device with $u_W = 10 \text{ kT/q}$ , .....	111
Figure 3-29. Valence band plots for ON and OFF states with a QW height of $10 \text{ kT/q}$ . It can be seen that the top of cap layer (marker to the right of the QW) is lying below the top of the QW. Now the QW has much higher carrier population compared to the cap layer. ....	114
Figure 3-30. ( $u_{on} - u_{cap}$ ) vs QW height. A difference of 2.3 corresponds to the QW hole sheet density being ten times the hole sheet density in the cap layer. The value 2.3 is plotted with the dashed line. Substrate doping $n_B = 5e18/cm^3$ , QW width $x_W = 5 \text{ nm}$ and p-delta doping density $Q_d = 1.25e12/cm^2$ . ....	115
Figure 3-31. Sub-threshold slope parameter $S_t$ vs QW height. $n_B = 5e18/cm^3$ , .....	116
Figure 3-32. Log(hole sheet density ON/OFF) vs QW height. $n_B = 5e18/cm^3$ , $x_W = 5 \text{ nm}$ , $Q_d = 1.25e12/cm^2$ . ....	116
Figure 3-33. Variation of sub-threshold parameter, $S$ , with variation of QW width, $x_W$ . $u_W = 10 \text{ kT/q}$ , $Q_d = 2 \times 10^{12}/cm^2$ , $n_B = 5 \times 10^{18}/cm^3$ .....	118
Figure 3-34. Variation of QW hole sheet density with variation of QW width, $x_W$ .....	119
Figure 3-35. Hole sheet density ON/OFF ratio vs QW width. $u_W$ is $10 \text{ kT/q}$ , .....	119
Figure 3-36. Difference between QW potential and cap layer potential. As discussed in previous sub-sections a minimum difference of 2.3 is needed to keep the hole sheet density in the cap layer less than a tenth of the QW hole sheet density. ....	120
Figure 3-37. Potential plots for the ON and the OFF state for a device with .....	122
Figure 3-38. Potential plots for the ON and the OFF states for a device with .....	123
Figure 3-39. Potential plots for the ON and the OFF state for a device with $u_W = 10 \text{ kT/q}$ , $x_W = 5 \text{ nm}$ , $Q_d = 8 \times 10^{10}$ , $V_{gs} = 1.5 \text{ V}$ , $n_B = 1 \times 10^{18}$ .....	123

### LIST OF FIGURES - Continued

Figure 3-40. Potential plots for the ON and the OFF state for a device with $u_W = 10 \text{ kT/q}$ , $x_W = 5 \text{ nm}$ , $Q_d = 8 \times 10^{10}$ , $V_{gs} = 1.5 \text{ V}$ , $n_B = 5 \times 10^{18}$ .....	124
Figure 3-41. Potential plots for the ON and the OFF state for a device with $u_W = 10 \text{ kT/q}$ , $x_W = 5 \text{ nm}$ , $Q_d = 8 \times 10^{10}$ , $V_{gs} = 1.5 \text{ V}$ , $n_B = 1 \times 10^{19}$ .....	124
Figure 3-42. QW hole sheet density vs substrate doping, $n_B$ . $u_W = 10 \text{ kT/q}$ , .....	126
Figure 3-43. Substrate potential drop vs substrate doping for the device with $u_W = 10 \text{ kT/q}$ , $x_W = 5 \text{ nm}$ , $Q_d = 8 \times 10^{10}$ , $V_{gs} = 1.5 \text{ V}$ . .....	126
Figure 3-44. Log(hole density ON/OFF ratio) vs substrate doping for the device with $u_W = 10 \text{ kT/q}$ , $x_W = 5 \text{ nm}$ , $Q_d = 8 \times 10^{10}$ , $V_{gs} = 1.5 \text{ V}$ . .....	127
Figure 3-45. Sub-threshold slope parameter, $S$ , vs. substrate doping density for the device with $u_W = 10 \text{ kT/q}$ , $x_W = 5 \text{ nm}$ , $Q_d = 8 \times 10^{10}$ . .....	127
Figure 3-46. The difference between the front side QW potential and the potential at the top of the cap layer vs substrate doping density. $V_{gs} = -1.5 \text{ V}$ .....	128
Figure 3-47. OFF state peak QW potential vs. substrate doping. This potential is nearly the same throughout the doping density range considered. ....	129
Figure 3-48. Potential plot of an ideal device. $t_{OX}$ , $X_{cap}$ , $X_{QW}$ and $X_{buf}$ are 3 nm and the substrate n-type doping is $4 \times 10^{18}/\text{cm}^3$ . There is no p-layer. So the potential keeps rising. The diamond marks are on Fermi level and dots represent potentials including the effects of holes in the QW.....	131
Figure 3-49. Potential plot of an ideal device with the inclusion of a p-doped $\delta$ -layer with doping density of $5.64 \times 10^{12}/\text{cm}^2$ . The layer dimensions and doping are the same as before. $t_{OX}$ , $X_{cap}$ , $X_{QW}$ , $X_{buf}$ are 3 nm and the n-substrate doping is $4 \times 10^{18}/\text{cm}^3$ . Notice that the p-layer doping is selected such that the field in the buffer layer is zero. ....	132
Figure 4-1. Lattice matched growth of $\text{Si}_{1-x}\text{Ge}_x$ over silicon substrate. The overgrown $\text{Si}_{1-x}\text{Ge}_x$ lattice is compressed in the plane of growth leading to a tetragonal distortion .....	141
Figure 4-2. Critical thickness $hc$ as a function of Ge fraction $x$ . Solid line is from Mathews - Blakeslee theory. The dashed line is from equation ( 4.2).....	143
Figure 4-3. Band gap of $\text{Si}_{1-x}\text{Ge}_x$ as a function of the Ge fraction $x$ . The top line is for bulk or relaxed $\text{Si}_{1-x}\text{Ge}_x$ . For thin strained $\text{Si}_{1-x}\text{Ge}_x$ films (bottom two curves) the band gap reduction is more pronounced [ 114].....	148
Figure 4-4. Band gap $E_g(\text{Si}_{1-x}\text{Ge}_x)$ as a function of Ge fraction $x$ . Using the linear interpolation used in ISE-TCAD (solid), Using the linear approximation of equation (4.5) (dashed) .....	148

## LIST OF FIGURES - Continued

Figure 4-5. Band edge alignment diagram for a strained $\text{Si}_{1-x}\text{Ge}_x$ film grown over (100) silicon for an $x$ value of 0.3. The alignment is of type I with a valence band discontinuity of approximately 222 meV and a negligible conduction band discontinuity which is less than 20 meV.....	149
Figure 4-6. The valence band diagrams for unstrained (left) and strained (right) $\text{Si}_{1-x}\text{Ge}_x$ grown over (100) silicon substrate. Notice the warping of the heavy hole band and the lowering of the light hole band with strain [ 111]......	151
Figure 4-7. Hole mobility as a function of the electric field along the direction of transport. The field dependence is computed using Caughey-Thomas formula. The mobilities are for low doped silicon MOS and strained $\text{Si}_{1-x}\text{Ge}_x$ with $x$ values of 0.3, 0.4 and 0.5. ....	154
Figure 4-8. Sketch of a GaAs MODFET .....	161
Figure 4-9. Band diagram for the MODFET .....	161
Figure 4-10. Layer sequence for some earlier SiGe p-HMOSFET structures. It has a p+ layer on top of the SiGe QW channel layer. This is an adaptation from the III-V MODFET.....	164
Figure 4-11. Strained silicon channel p-MOS .....	164
Figure 4-12. A simple structure of p-HMOSFET. It has a recessed SiGe channel .....	166
Figure 4-13. A p-HMOSFET structure that seems to have promise for general acceptance. It has an additional p-Si layer that can be used to adjust $V_T$ and thus increase the hole density in the QW. ....	167
Figure 4-14. A p-HMOSFET structure with graded SiGe channel .....	168
Figure 4-15. Recessed SiGe channel HMOSFET in SOI process. ....	169
Figure 4-16. Double gated p-HMOSFET. The channel is at the center, away from interfaces. ....	170
Figure 4-17. Structure of the proposed p-HMOS. It has an undoped silicon cap layer, undoped $\text{Si}_{1-x}\text{Ge}_x$ quantum well layer, undoped silicon buffer layer and a variable thickness p-doped layer followed by an n-Si substrate.....	171
Figure 4-18. Potential (top two curves) and valence band edge (bottom two curves) of a p-HMOSFET for $V_{gs} = 0$ (outer curves) and $V_{gs} = -1.5$ V (inner curves). The straight line is the Fermi level.....	172
Figure 5-1. Sketch of the generalized p-HMOSFET used in this study. It consists of a .....	180
Figure 5-2. (a) Band diagrams from source to drain through p-Si layer through the section A-A (b) through Si substrate through the section BB of .....	181

## LIST OF FIGURES - Continued

Figure 5-3. Source to p-layer barrier at the source junction. Higher p-layer doping reduces this barrier .....	182
Figure 5-4. Valence band diagram of p-HMOSFET at high $V_{gs}$ . The n-Si substrate parabolically bends the band upwards and the p-Si layer within its width bends it downwards. Both the buffer layer and the cap layer have no charges. Hence the transition through them is linear. The QW is nonlinear since the hole distribution within it and the potential are coupled.....	183
Figure 5-5. Valence band plot at high $V_{gs}$ value. The thick cap layer leads to inversion near Si/SiO <sub>2</sub> interface.....	186
Figure 6-1. Sketch of a p-HMOSFET showing the different regions, dimensions and.....	192
Figure 6-2. Flow chart of the analytical program showing the computation of hole density $p_H$ in the QW.....	206
Figure 6-3. Potential variation with depth for the device 3,5,5,5/0.3/5e18,5e18 for $V_{gs} = -0.5, -1.2, -1.5$ . The device has a $V_T$ of $-1.25$ . The match is very good for .....	209
Figure 6-4. An enlarged view of Figure 6-3. Both low and high $V_{gs}$ simulations match extremely well but a slight split is obvious for $V_{gs}$ near $V_{gs} = -1.2$ V. This is explained in Figure 6-3. ....	210
Figure 6-5. Comparison of potential plots from analytical program and numerical simulator. $V_{gs} = -1.2$ V. Device is 3,5,5,5/0.3/5e18,5e18.....	211
Figure 6-6. Comparison of potential plots from analytical program and numerical simulator. $V_{gs} = -1.5$ V. Device is 3,5,5,5/0.3/5e18,5e18.....	211
Figure 6-7. Potential at the top of the QW for the device 3,5,5,5/0.3/5e18,5e18 as a function of $V_{gs}$ . Three numerical simulated values are shown by asterisks.....	212
Figure 6-8. Hole density at the top of the QW $p_H$ for the device 3,5,5,5/0.3/5e18,5e18 as a function of $V_{gs}$ . Numerical simulation values are shown by asteriks for $V_{gs} = -0.5, -1.0$ and $-1.5$ V.....	212
Figure 6-9. Hole density at the top of the cap layer for the device 3,5,5,5/0.3/5e18,5e18 as a function of $V_{gs}$ . Three numerical simulation values are shown at three points by asteriks .....	213
Figure 6-10. Integrated hole sheet density in the QW for the device 3,5,5,5/0.3/5e18,5e18 as a function of $V_{gs}$ . Numerical simulator outputs are shown by asteriks .....	213

## LIST OF FIGURES - Continued

Figure 7-1. Potential and valence band plots for $V_{gs} = 0$ V and $V_{gs} = -1.5$ V. The device is 3,5,5,5,0.3/5e18,5e18. The lowermost curve and the uppermost curve corresponds to valence band edge $E_v$ and potential $\Phi(x)$ respectively for a $V_{gs} = 0$ V.....	218
Figure 7-2. Valence band plots starting from the top of the cap layer ( $x = 0$ ) for a .....	222
Figure 7-3. Valence band plots starting from the top of the cap layer ( $x = 0$ ) for a .....	222
Figure 7-4. Valence band plots starting from the top of the cap layer ( $x = 0$ ) for a Ge fraction $x = 0.5$ , for $V_{gs} = 0$ V and $V_{gs} = -1.5$ V, $V_{ds} = -0.05$ V. Device is 3,5,5,5,0.5/5e18,5e18. $L_g = 500$ nm.....	223
Figure 7-5. Hole density variation at the top of the QW with variation of Ge fraction $x$ in the QW. The device is 3,5,5,5/x/5e18,5e18 (see footnote 1), $L_g = 500$ nm, $V_{gs} = -1.5$ V, $V_{ds} = -0.05$ V.....	223
Figure 7-6. Variation of the ratio of hole density (see footnote 3) at top of QW to that at the top of the cap layer with variation in Ge fraction $x$ . $V_{gs} = -1.5$ V, $V_{ds} = -0.05$ V. Device is 3,5,5,5/x/5e18,5e18, $L_g = 500$ nm.....	224
Figure 7-7. Threshold voltage variations with variation of Ge fraction $x$ of the QW channel layer. Device is 3,5,5,5/x/5e18,5e18, $L_g = 500$ nm.....	224
Figure 7-8. Hole density at the top of the QW as a function of substrate doping for the device 3,5,5,5/0.3/5e18, $N_{sub}$ . $V_{gs} = -1.5$ V, $V_{ds} = -0.05$ V .....	226
Figure 7-9. Potential and valence band edge plots along the depth starting from the top of the cap layer ( $x = 0$ ). $V_{gs} = -1.5$ V, $V_{ds} = -0.05$ V. The device is 3,5,5,5/0.3/5e18, $N_{sub}$ . Two sets of curves are for $N_{sub}$ values of $2 \times 10^{18}/\text{cm}^3$ and $8 \times 10^{18}/\text{cm}^3$ . Solid curves are $E_v$ and $\Phi(x)$ for $2 \times 10^{18}/\text{cm}^3$ .....	227
Figure 7-10. Ratio of hole density at the top of QW to hole density at the top of the cap layer as a function of $N_{sub}$ . Device is 3,5,5,5/0.3/5e18, $N_{sub}$ . $V_{gs} = -1.5$ V, $V_{ds} = -0.05$ V,.....	227
Figure 7-11. Substrate doping dependence of $V_T$ and $V_{TS}$ for the device 3,5,5,5/0.3/5e18, $N_{sub}$ . $L_g = 500$ nm.....	228
Figure 7-12. Sub threshold parameter (S) dependence on substrate doping concentration for the device 3,5,5,5/0.3/5e18, $N_{sub}$ . $L_g = 500$ nm .....	228
Figure 7-13. Dependence of hole density at the top of the QW upon the thickness of the p-layer and the buffer layer. The device is 3,5,Xbuf,Xp/0.3/5e18,5e18 and 500nm channel length. While varying one of them the other is kept fixed at 5 nm.....	233
Figure 7-14. Valence band edge along depth starting from the top of the cap layer ( $x = 0$ ) for two p-layer thicknesses. Xbuf = 5 nm. Device is 3,5,Xbuf,Xp/0.3/5e18,5e18 and 500nm channel length .....	233

### LIST OF FIGURES - Continued

Figure 7-15. Dependence of the ratio of the hole density at the top of the QW to the density at the top of the cap layer with variation of p-layer and buffer layer thickness. While varying one of them the other is kept fixed at 5 nm. Device is 3,5,Xbuf,Xp/0.3/5e18,5e18 and 500nm channel length .....	234
Figure 7-16. Dependence of $V_T$ and $V_{TS}$ upon the thickness of the p-layer $X_p$ . Device is 3,5,5,Xp/0.3/5e18,5e18 and 500nm channel length.....	234
Figure 7-17. Variation of the sub threshold parameter $S$ with variation of p-layer thickness and buffer layer thickness. One is varied at a time keeping the other fixed at 5 nm. Device is 3,5,Xbuf,Xp/0.3/5e18,5e18 and 500nm channel length .....	235
Figure 7-18. Threshold voltage variation with buffer layer thickness. $X_p = 5$ nm. The device is 3, 5, $X_{buf}$ , 5/0.3/5e18, 5e18 and 500 nm channel length.....	236
Figure 7-19. $N_{sub}$ vs $x$ design space for $X_p = 2$ nm. Device is 3,5,5,2/x/5e18,5e18. Combinations of $x$ and $N_{sub}$ to get $V_T = -1$ V is given by the solid line. Above this line magnitude of $V_T > 1$ V. Combinations of $x$ and $N_{sub}$ to get $V_{TS} = -1.5$ V is given by the dashed line. Below this line magnitude of $V_{TS} < -1.5$ V. So the window lies in between the two lines. ....	239
Figure 7-20. $N_{sub}$ vs $x$ design space for $X_p = 5$ nm. Device is 3,5,5,5/x/5e18,5e18. Combinations of $x$ and $N_{sub}$ to get $V_T = -1$ V is given by the solid line. Above this line magnitude of $V_T > 1$ V. Combinations of $x$ and $N_{sub}$ to get $V_{TS} = -1.5$ V is given by the dashed line. Below this line magnitude of $V_{TS} < -1.5$ V. So the window lies in between the two lines. ....	240
Figure 7-21. $N_{sub}$ vs $x$ design space for $X_p$ of 8 nm. Device is 3,5,5,8/x/5e18,5e18. Combinations of $x$ and $N_{sub}$ to get $V_T = -1$ V is given by the solid line. Above this line magnitude of $V_T > 1$ V. Combinations of $x$ and $N_{sub}$ to get $V_{TS} = -1.5$ V is given by the dashed line. Below this line magnitude of $V_{TS} < -1.5$ V. So the window lies in between the two lines. ....	240
Figure 7-22. Design space for $I_{ON} = 100$ $\mu$ A for the device 3,5,5,Xp/x/5e18,5e18 for p-layer thickness $X_p$ of 2, 5 and 8 nm. $X_{buf} = 5$ nm. The area below the respective lines have $I_{ON}$ more than 100 $\mu$ A.....	241
Figure 7-23. Design space for $I_{ON} = 100$ $\mu$ A for the device 3,5,10,Xp/x/5e18,5e18 for p-layer thicknesses of 2, 5 and 8 nm. $X_{buf} = 10$ nm. The area below the respective lines have $I_{ON}$ more than 100 $\mu$ A.....	242
Figure 7-24. Design space for $I_{ON} = 100$ $\mu$ A for the device 5,5, $X_p$ /x/5e18, $N_{sub}$ for QW Ge fraction $x = 0.2$ and $0.5$ . $X_{buf} = 5$ nm. The area below the respective lines shown by the arrows satisfies the requirement that $I_{ON}$ is more than 100 $\mu$ A .....	243

## LIST OF FIGURES - Continued

Figure 7-25. Analytical and numerical simulation comparison of drain characteristic for an optimized device. The device is described as 3,5,5,2/0.5/5e18,3e17. It has an $I_{ON}$ of 228 $\mu A$ .....	244
Figure 8-1. $I_{Dsat}$ comparison of p-MOSFET using analytical MOSFET design program and ISE-TCAD numerical simulations. $L_g = 500$ nm, $t_{OX} = 3$ nm, $x_j = 30$ nm, p+ polysilicon gate. $V_{gs} = -1.5$ V, $V_{ds} = -1.5$ V. The Excel computed values are lower by a constant value.....	254
Figure 8-2. Estimated minimum value of $L_g$ , at the onset of short channel effects, as a function of substrate doping density for the MOSFET. junction depth $x_j = 30$ nm, .....	255
Figure 8-3. Threshold voltage $V_T$ for a MOSFET as a function of substrate doping $N_{sub}$ obtained using the numerical simulator ISE-TCAD. $L_g = 500$ nm, $tox = 3$ nm, $x_j = 30$ nm, .....	256
Figure 8-4. Off currents as a function of the substrate doping, $N_{sub}$ , obtained using the numerical simulator ISE-TCAD. This is for a MOSFET with $L_g = 500$ nm, $tox = 3$ nm, .....	256
Figure 8-5. Saturation values $I_{Dsat}$ for 500, 400, 300, 200, 150 and 100 nm channel lengths. Up to 200 nm $I_{Dsat}$ for HMOSFET shows almost linear variation but at 150 nm $I_{Dsat}$ is much higher indicative of short channel effects. In the case of MOSFET the linear variation is only upto about 300 nm. MOSFET has $tox = 3$ nm, $N_{sub} = 5e17/cm^3$ , $x_j = 30$ nm, p+ gate. HMOSFET is 3,5,5,2/0.24/5e18,5e17 with $tox = 3$ nm, .....	262
Figure 8-6. Threshold voltage $V_T$ vs channel length. Lines are from numeric simulator ISE-TCAD and the dark circles are from the analytical analysis and design program of .....	263
Figure 8-7. $I_{ds}$ for MOSFET and HMOSFET for 500 nm and 150 nm channel lengths. For 150 nm the MOSFET do not show any saturation indicating SCE. MOSFET has .....	263
Figure 8-8. $\log(I_{ds})$ vs $V_{gs}$ . MOSFET has $tox = 3$ nm, $x_j = 30$ nm, $N_{sub} = 5e17/cm^3$ , .....	265
Figure 8-9. $\log(\text{mag}(I_{ds}))$ vs $V_{gs}$ for a HMOSFET 3,5,5,2/0.24/5e18,5e17 with.....	266
Figure 8-10. Comparison of $\log(\text{mag}(I_d))$ vs $V_{gs}$ for MOSFET and HMOSFET. Both devices have the same $L_g = 500$ nm, $N_{sub} = 5 \times 10^{17}/cm^3$ , $t_{OX} = 3$ nm, $x_j = 30$ nm, p+ gate. Both have the same $I_{Dsat}$ at $V_{gs} = -1.5$ V and $V_{ds} = -1.5$ V. HMOSFET is 3,5,5,2/0.24/5e18,5e17. Notice that $I_{off}$ in the case of HMOSFET is almost seven orders of magnitude less compared to that of MOSFET. These plots are from ISE-TCAD.....	267

### LIST OF FIGURES - Continued

Figure 8-11. Log( $I_{ds}$ ) vs $V_{gs}$ sketch when both MOSFET and HMOSFET have the same $V_T$ . $I_{on}$ is higher for HMOSFET at $V_{gs} = -1.5$ V and $V_{ds} = -1.5$ V. This is an approximate sketch only to convey the idea. ....	268
Figure 8-12. Log( $I_{ds}$ ) vs $V_{gs}$ sketch when the HMOSFET has a higher $V_T$ . $V_T$ for either devices is chosen such that $I_{on}$ is the same for both MOSFET and HMOSFET at .....	269
Figure 8-13. Potential plots from source to drain for 100 and 500 nm MOSFET. ....	274
Figure 8-14. Potential along the channel for 100 nm and 500 nm HMOSFET. ....	275
Figure 8-15. An enlarged view of Figure 8-14.. $V_{gs} = -1.5$ V, $V_{ds} = -1.5$ V. ....	275
Figure 8-16. Lateral electric field along the channel for long and short channel MOSFET. $V_{gs} = -1.5$ V, $V_{ds} = -1.5$ V. ....	276
Figure 8-17. Lateral electric field along the QW channel for long and short channel HMOSFET. $V_{gs} = -1.5$ V, $V_{ds} = -1.5$ V. ....	276
Figure 8-18. Lateral electric field along the channel for MOSFET and HMOSFET. ....	277
Figure 8-19. Potential contours for the 100 nm MOSFET. $V_{gs} = -1.5$ V, $V_{ds} = -1.5$ V. ....	278
Figure 8-20. Potential contours for 100 nm channel HMOSFET. $V_{gs} = -1.5$ V, ....	278
Figure 8-21. Low field hole mobility for strained $Si_{1-x}Ge_x$ vs Ge fraction .....	281
Figure 8-22. High field mobility for strained $Si_{1-x}Ge_x$ with $x = 0.24$ and silicon substrate with $N_{sub}$ of $5e17/cm^3$ .....	281
Figure 8-23. Velocity vs field for $Si_{1-x}Ge_x$ with $x = 0.24$ and silicon substrate with.....	282
Figure 8-24. Plot of hole velocity vs source-to drain distance. This is for the 500 nm channel length MOSFET. ....	284
Figure 8-25. Plot of velocity vs source-to-drain distance. This is for the 500 nm channel length HMOSFET. ....	285
Figure 9-1. Band diagram from source to drain through the p-layer. This is copied here.....	289
Figure 9-2. Structure of a modified p-HMOSFET. Here the p-layer doping is adjusted such that the field in the buffer layer goes to zero.....	290
Figure 9-3. The valence band diagram for the structure in Figure 9-2. p-layer doping is adjusted to make the field in the buffer zero.....	292

### LIST OF FIGURES - Continued

Figure 9-4. Valence band edge of a 5 nm buffer layer device. The band edge in the buffer layer below the quantum well is nearly flat. The device is 3,5,5,3/0.5/5e18,5e17.....	293
Figure 9-5. Plot of valence band edge. The buffer layer is 35 nm which puts it below the source which is 30 nm deep. p-layer doping and rest of the values are same as in.....	293
Figure 9-6. Analytically computed drain characteristic for the device in Figure 9-5. The buffer is 35 nm thick. The device is 3,5,35,3/0.5/5e18,5e17. Analytical design program used. ....	294
Figure 9-7. Valence band of a 5 nm thick buffer device. Notice the flat region over 5 nm below the quantum well. The result is from numerical simulation. The left side sharp rise is in the oxide. $V_{ds} = -1.5$ V, $V_{gs} = -1.5$ V .....	294
Figure 9-8. Numerically simulated drain characteristic of the 5 nm buffer device. ....	295
Figure 9-9. Id vs $V_{gs}$ of the 5nm buffer device from numerical computation .....	295
Figure 9-10. Valence band edge of a 35 nm thick buffer layer device. It is not flat all over the 35 nm width. It has a slight slope up to 25 nm. This plot is from numerical simulation. The left side sharp rise is in the oxide. $V_{ds} = -1.5$ V, $V_{gs} = -1.5$ V. Buffer layer extends to 43 nm. But buffer band edge is not flat below 25 nm. ....	296
Figure 9-11. Numerically simulated drain characteristic of the 35 nm buffer device. ....	297
Figure 9-12. Id vs $V_{gs}$ of the 35 nm thick buffer device. Notice the leakage at low gate voltage. This result is from numerical simulation.....	297
Figure 9-13. Two dimensional potential contours for the case of 35 nm buffer. p-layer is located where the notches are seen. It looks like there is a tendency for the notches to spread out to the left (source) if the p-layer is moved down. ....	299
Figure 10-1. Fermi distribution at 0° K (solid) and at room temperature 30° C (dash). The chain line is the Boltzman distribution. ....	314
Figure 10-2. Band diagram of n and p doped silicon. $E_g = 1.12$ eV is the band gap, $E_C$ and $E_V$ are conduction and valence band respectively, $E_{fn}$ and $E_{fp}$ are the electron quasi Fermi potential and hole quasi Fermi potential respectively and $\phi_b$ is the doping parameter .....	316
Figure 10-3. The p-MOSFET structure used in this project as a comparison reference device. It has a n+ polysilicon gate.....	317

### LIST OF FIGURES - Continued

<p>Figure 10-4. Band structure of n+ polysilicon gate, SiO<sub>2</sub> gate oxide and silicon substrate prior to forming device by metallurgically joining them. <math>\phi_G</math> and <math>\phi_{Si}</math> are work functions of n+ polysilicon and silicon substrate respectively. <math>\chi_{OX}</math> and <math>\chi_{Si}</math> are the electron affinities of oxide and silicon respectively. <math>\phi_G = \phi_{Si} = 4.05</math> eV, <math>\chi_{OX} = 0.95</math> eV and <math>\chi_{Si} = 4.05</math> eV .....</p>	319
<p>Figure 10-5. Thermal equilibrium band diagram of polysilicon, oxide and silicon after forming the p-MOSFET. <math>V_{gs} = 0V</math>. Note that the silicon conduction band is bend down by <math>qV_{fb}</math>.....</p>	320
<p>Figure 10-6. Flat band condition of the p-MOSFET with <math>V_{gs} = -V_{fb}</math>.....</p>	321
<p>Figure 10-7. Band diagram at <math>V_{gs} = V_T</math> for the p-MOSFET (top). Charge accumulation at <math>V_{gs} = V_T</math> (bottom) .....</p>	323

## LIST OF TABLES

Table 3-1. Fixed parameters and the corresponding potential variables used to restore the fixed parameters in Excel using 'goal seek' .....	88
Table 4-1. Comparison of some of the most important parameters of $\text{Si}_{1-x}\text{Ge}_x$ pseudomorphically grown over (100) Si substrate. [88], [111], [121] .....	145
Table 8-1. Some of the important material constants and other values used by the numeric simulator ISE-TCAD for the simulation of $\text{Si}_{1-x}\text{Ge}_x$ HMOSFET and conventional silicon MOSFET. $\text{Si}_{1-x}\text{Ge}_x$ has an x value of 0.24 and the n-type silicon substrate doping is $5 \times 10^{17}/\text{cm}^3$ .....	280

## ABSTRACT

Silicon Germanium ( $\text{Si}_{1-x}\text{Ge}_x$ ) is an alloy semiconductor that has caught considerable attention of the semiconductor industry in the past decade. Effects of strain in thin films are the reason for this. Strain leads to considerable deformation of bands providing enhanced mobility for both electrons and holes. Another important aspect of SiGe is the reduction of band gap. This makes band gap engineering feasible in all silicon technology. Yet another attractive point is the adaptability and compatibility of SiGe to silicon process technology.

In CMOS circuits the p-channel MOSFET needs more than double the area of the n-channel MOSFET due to the lower mobility of holes in silicon. Hence a p-channel hetero MOSFET (HMOSFET) is chosen as the object of this dissertation.

A simple general device structure that can provide considerable enhancement in performance, compared to a conventional MOSFET, is selected. A one dimensional Poisson equation is solved for this hetero junction device. Using these results an Excel spreadsheet is used as a tool to design a complete analytical program that can provide internal as well as terminal parameters of this device. The analytical program is tested by comparing the results with ISE-TCAD numerical device simulator results. The results were found to match very well. This analytical program yields results in a fraction of the time compared to numerical programs. For the device of choice variable parameters are identified. It is found that these parameters are interconnected in many ways and trade offs between them need to be applied.

From the front end of the spreadsheet input parameters can be varied and parameters like potentials, hole density and terminal characteristics can be plotted very easily while simultaneously computing other parameters like threshold voltage and saturation current.

The main contribution of this dissertation research is (1) Development of a very efficient and accurate analytical program to interactively design and optimize a p-channel HMOSFET (2) A detailed understanding and explanation of various design parameters, their implications, interdependency and trade offs (3) Study and explanation of certain special characteristics of p-HMOSFET like dual threshold voltage, low off-currents, structural limitations etc.

# CHAPTER 1

## INTRODUCTION

Silicon Germanium ( $\text{Si}_{1-x}\text{Ge}_x$ ) is an alloy semiconductor that shows great promise to integrate with silicon technology or even form a parallel fabrication line to silicon technology. There are several attractive features with Silicon Germanium that support the above statement. They are listed below:

1. First, band gap engineering is feasible in  $\text{Si}_{1-x}\text{Ge}_x$ . It was an exclusive realm of III-V compounds until recently. This is important because quantum well (QW) confinement of carriers and transport in these un-doped layers remote from the  $\text{SiO}_2$  interface is now possible. This can provide much faster devices than silicon complementary metal oxide semiconductor (CMOS).
2. Second, the mobility enhancements of both holes and electrons in compressively strained  $\text{Si}_{1-x}\text{Ge}_x$  or tensile strained silicon (Si) respectively can provide fast devices of both p-channel and n-channel type. This is very crucial for CMOS circuitry.
3. Third the processing of Silicon Germanium is very similar to that of silicon. It is to be noted here that the GaAs processing is not compatible with silicon and is more expensive to fabricate.
4. These Si/ $\text{Si}_{1-x}\text{Ge}_x$  based hetero junction transistors may be capable of operation at speeds close to the GaAs based modulation doped field effect transistor (MODFET). The GaAs based MODFET can provide high speed

devices of only n-channel type. The hole mobility in GaAs is inferior to that in silicon. On the contrary  $\text{Si}_{1-x}\text{Ge}_x$  based devices can provide fast devices of both n-channel and p-channel types. This is important in system on a chip (SOC) used in cell phones and global positioning system receivers (GPS) to name a few.

Design of Si/ $\text{Si}_{1-x}\text{Ge}_x$  based p-channel MOSFET devices with high operating speed, low operating power and low leakage currents is the objective of this research work. However, prior to embarking on this it is instructive to look back at the evolution, current status and future trends of silicon technology that has held a lion's share of the global semiconductor market since the inception of integrated circuits. In addition the exponential rise of battery operated systems like the cell phone, remote data logging systems and GPS, to name a few, have placed tremendous demand for low voltage and low power devices. These topics are briefly reviewed in sections 1.1 to 1.4 for background reasons. Limitations of Si CMOS are discussed in section 1.5. The objective of this research work, procedural details and general organization of the dissertation are given in section 1.6.

## **1.1 Short History of Silicon CMOS Devices**

Since the invention of the bipolar transistor in 1947 a new era of electronics began. The bipolar junction transistor (BJT) found its way into integrated circuits (IC) in 1960. The theory of inverting a silicon surface by application of an electric field to form a metal oxide semiconductor field effect transistor (MOSFET) was reported as early as 1925 [1]. But the device was not realized until 1960 due to fabrication problems related to the large

densities of Si/SiO<sub>2</sub> interface states. The ease of fabrication of the MOSFET, its planar technology and its tiny size were highly attractive in the fabrication of high-density integrated circuits (ICs) leading eventually to large-scale integration (LSI). However akin to bipolar transistors the integration density of single polarity MOSFETs (p-channel or n-channel alone) were limited due to their large standby power dissipation. The introduction of complementary MOSFET (CMOS) in 1963, with the formation of p-channel and n-channel MOSFETs on the same chip made it possible to have high integration density [3]. The CMOS technology made it possible to fabricate device pairs (n and p channel) with very low off currents on the same chip. In addition the MOSFET has very small feature size compared to the BJT, These two factors led to the evolution of VLSI (very large scale integration) around 1980 and ULSI (ultra large scale integration) by the year 2000. In addition the full rail-to-rail logic swing of CMOS ICs offer high noise margins [4] [ 5]. Starting with the introduction of CMOS ICs the silicon semiconductor industry was driven by the insatiable quest for higher speed of operation, lower operating voltages, lower standby power and smaller device sizes. The device speed kept on increasing and the feature size kept on decreasing making faster and faster high density ICs more economical to fabricate. Following Moore's law the density of devices on a chip doubled approximately every 18 months [ 7]. This evolution of high-speed high density ICs led to the rapid growth of information technology.

### **1.1.1 Technology node**

The progress in down scaling of CMOS devices was so aggressive that the predictions of the International Technology Roadmap for semiconductors (ITRS) for the year 2010

are almost here in 2008 [ 27]. The technology node in 2008 is at 45 nm. Today 45 nm devices are in full production and 33 nm devices are expected by the year 2009 [ 6]. The 33 nm technology node was a prediction for the year 2013 by the 2002 ITRS and the 2007 ITRS [ 27], [ 28]. The prediction by the ITRS only means that such devices will be in demand and are expected to be in production by the year 2013. It should be mentioned here that clever designs and introduction of new materials have taken the semiconductor technology beyond predictions of industry roadmaps.

Here a discussion of the technology node is appropriate. Technology nodes used to be historically linked to the introduction of new Dynamic Random Access Memory (DRAM) generations. Each generation has four times (4X) increase in bits/chip. The node number gives the lithography half-pitch of a DRAM. For as long as the above cycle followed Moore's law (1½ year cycle for 2X or 3 year cycle for 4X) the technology node and DRAM generations were synonymous. However, in the recent past faster introduction/optimization of product-specific technology and the increase in business and technology complexity tend to de-couple many of the technology parameters used to characterize 'advance to the next technology node'. For example the microprocessor unit (MPU) has driven the reduction of gate length or minimum feature size at a faster pace than the lithography half pitch. However, the lithography half pitch is the same for DRAM and MPU. For the 45 nm technology node mentioned before the physical gate length or minimum feature size is only 18 nm. In recent years the scaling trend does not seem to follow Moore's law. The technology generation seems to reduce by four times in

two years instead of three years. This shows that the device scaling is no longer directed by the ITRS predictions. Instead the scaling in the past few years is led by the industry.

### **1.1.2 Operating Speed**

In spite of the above advantages of CMOS devices bipolar devices still have higher driving capacity and hence operating speed. Due to this reason a high level of CMOS with a relatively low level of bipolar devices are integrated to form bipolar CMOS (BiCMOS) systems. The BiCMOS process allows the exploitation of the high level of integration of digital CMOS and the high speed of operation of analog bipolar transistors. CMOS and BiCMOS are two of the dominant semiconductor technologies today [8]. However, the much smaller feature size and extremely high packing density of CMOS, compared to bipolar transistors, makes them ideal for very large scale mixed signal systems incorporating RF, analog and digital designs to form 'system on a chip' (SOC) where high ON state drive currents are not necessary. Nowadays, at the 45 nm node, digital CMOS devices are used for analog designs like analog-to-digital converters (ADC), digital-to-analog converters (DAC), comparators and even for voltage references. Hence the operating speed and scalability of the CMOS is of vital interest. However, all Si devices can operate only up to a maximum frequency of around 5 GHz which is not fast enough for receiver front ends and amplifiers used in SOC.

### **1.1.3 Power Dissipation**

Low standby power dissipation is important for VLSI and ULSI circuits where the high packing density can lead to thermal degradation if the dissipated power is not systematically removed from the die. Hence any reduction in standby power dissipation

is highly desirable. At the same time higher drive currents are important for high speed of operation. In fact the speed performance and power intake are directly related. In the case of the MOSFET the optimization of speed performance comes with a penalty. If the on current,  $I_{on}$ , is increased to attain a higher speed of operation the off current,  $I_{off}$ , is simultaneously increased which in turn increases the standby power dissipation. These aspects will be discussed in chapter 8 with regard to the work presented here, vis-à-vis Si/Si<sub>1-x</sub>Ge<sub>x</sub> hetero junction MOSFET (HMOSFET). It will be seen in chapter 8 that the off current,  $I_{off}$ , in the Si/Si<sub>1-x</sub>Ge<sub>x</sub> based HMOSFETs are a few orders of magnitude less compared to those in MOSFETs. Low  $I_{off}$  along with higher drive currents and hence higher speed of operation makes the HMOSFETs highly desirable in the domain of ULSI as well as SOC designs. It may be surprising to note that present day digitally intensive microprocessor systems have heat dissipations as high as 100 W/cm<sup>2</sup> [ 39].

#### 1.1.4 Low Voltage Systems

About three decades ago low voltage systems were in demand only in a limited number of applications like watches, pace makers, integrated sensors and so on. But that scenario changed completely with the advent of battery operated systems such as portable computers, cell phones and remote data logging units. Low voltage, as low as single battery voltage, is in demand for portable communication devices, cell phones, remote control units and so on. Very low power dissipation is also crucial in these systems to reduce the drain on the battery. The power dissipation in a CMOS inverter is given by [9],

$$P_{total} = I_{off}V_{ds} + KC_LfV_{ds}^2 \quad (1.1)$$

where  $V_{ds}$  is the power supply voltage,  $I_{off}$  is the off-state current,  $K$  is the switching or duty factor,  $C_L$  is the load capacitance and  $f$  is the clock frequency. The first and second terms in equation (1.1) give the standby power and active power consumption respectively. As can be seen active power dissipation increases linearly with the clock frequency. In today's lap top computers with submicron CMOS devices operating at frequencies over a few hundred MHz the power dissipation can be very high making it difficult to cool the ICs hence leading to degradation of the IC in addition to the high battery drain.

Hence from the point of view of low voltage operation as well as to prevent degradation of the IC chip it is important to reduce the IC operating voltage, as can be seen in equation (1.1). In addition to packing more devices on a chip, scaling also demands lower voltages to avoid high fields that reduce device life time. This necessitates the use of devices capable of operating at very low supply voltages, as low as 0.8 to 1.5 Volts. The devices dealt with in this project are long channel devices and hence require higher operating voltage than the nanoscale devices. Considering this fact the p-HMOSFET designs and simulations carried out in this project are set at  $-1.5$  Volt.

## **1.2 Current Status of Silicon CMOS**

CMOS is the leading technology for VLSI. There are several reasons for this. Compared to all other device technologies CMOS has the lowest power-delay product, smallest size and highest packing density, CMOS circuits are easier to design and they are easier to fabricate than other technologies. Also device physics is easier to understand

compared to other devices [ 41]. However, it should be noted that in fast switching systems power dissipation in CMOS systems can be high and may even exceed bipolar devices [8]. This can be seen from equation (1.1).

Silicon CMOS and BiCMOS together has taken over 94 - 98% of the global semiconductor market share in the past couple of decades [23]. In order to improve the speed of operation and to increase the packing density silicon CMOS devices have been aggressively and continuously scaled down for past thirty years [ 7], [10] - [14]. The exponential increase in packing density as predicted by Moore's law and the associated exponential reduction in cost per function sustained the silicon industry unscathed so far [ 24]. The gate length reduced from 10 $\mu\text{m}$  in the year 1970 to 0.25 $\mu\text{m}$  and below by the year 2000. MOSFETs with channel lengths as low as 45 nm are in full production in the year 2007 [ 6]. The exponential reduction of feature size also led to an exponential increase in the cost of foundries [23]. The reduction of gate length demands scaling down of supply voltage to avoid subsurface punch through. Lower voltage operation in turn demands reduction of gate oxide thickness as well. This puts tremendous demands on the quality of the gate oxide. Lowering the trapped charges at the gate oxide interface as well as in the bulk oxide are important for MOSFET devices. But the scaling to the nanometer regime also increased the fields in the oxide as well as in the channel region [ 41]. Because of this it is more important to have better oxide quality in the nanometer devices. The scaling of CMOS by various methods such as constant voltage scaling, constant field scaling and generalized scaling were successful down to gate lengths of about 0.5  $\mu\text{m}$ . Further downscaling posed challenges due to short channel effects (SCE). To counter

these effects many design and process modifications were introduced from time to time, most of them being complex and expensive to implement [10], [17], [18].

### 1.3 Future of Silicon CMOS

In spite of the complexity and cost involved the short channel devices were expected to reach their limits around 45 nm channel length and about 3 nm gate oxide thickness by the year 2012 [10], [17], [22] Even though improvements in lithography has caught up with deep sub-micrometer feature sizes certain fundamental physics and process capabilities set limits to gate oxide thickness, effective channel length, junction depth and channel doping. Based on simulated as well as experimental studies the minimum values for gate oxide thickness  $t_{OX}$  were predicted to be around 2-3nm, the source and drain junction depth to be around 10 - 30nm, effective channel length  $L_g$  of about 60nm and the magnitude of threshold voltage of 0.2 – 0.3 Volt [10], [17]. The junction depth of the source and drain extension is limited by the increase in resistance. The gate oxide thickness is limited by the tunneling current. But contrary to expectations the CMOS devices underwent unimagined scaling to gate oxide thickness as low as 1.2 nm. Channel lengths as low as 32 nm are expected by 2008 – 2009 [29], [31]. The channel length and oxide thickness are related as shown in equation (1.2). For any oxide thickness there is a minimum channel length below which SCEs begin [10].

$$L_g = 45t_{OX} \quad (1.2)$$

Unique designs like double gated devices, FinFET, FlexFET and vertical MOSFET are being pursued with interest. New designs like FinFETs and FlexFETs fabricated in

CMOS process lines have potential to scale down to 10 nm [ 29] - [ 38] These devices are dual gate devices and are optimum structures to suppress short channel effects [ 44]. The 2003 Technology Roadmap for Semiconductors anticipates that double gate MOSFET and ultra thin SOI will reduce short channel effects [ 53]. SOI has the drawback of poor heat conduction and because of that the packing density is limited. Vertical structures such as FinFETs allows much higher packing density. Higher packing density will increase power density that will create more heat to be dissipated. At extremely scaled down devices the majority of the channel may get velocity saturated. Here the advantage of higher mobility becomes questionable since the velocity does not increase further. Saturation velocity for holes in silicon is  $1 \times 10^7$  cm/s. Based on simulations the saturation velocity for  $\text{Si}_{1-x}\text{Ge}_x$  is reported to vary approximately from  $1.6 \times 10^7$  to  $2.1 \times 10^7$  cm/s for a molecular Ge fraction of 0 – 0.5 [88]. Velocity overshoot in very short channel devices is also speculated. At extremely scaled down devices there are only few lattice atoms in the space from source to drain. Hence carriers may ballistically arrive at the drain end without having collisions or phonon interaction. For these nanoscale devices there is not an off state per se and an acceptable  $I_{\text{off}}$ . The  $I_{\text{ON}}/I_{\text{OFF}}$  ratios are as low as 100. So the leakage current poses a problem in terms of power and heat dissipation.

#### **1.4 Introduction to Silicon Germanium ( $\text{Si}_{1-x}\text{Ge}_x$ )**

Considering the scenario of silicon technology given in earlier sections and the increasing demand for low power and high frequency modern communication systems it would be desirable if high frequency silicon compatible devices could be integrated into

the silicon process. There seems to be a need for performance enhancement in the silicon ICs by some other alternate ways other than the scaling down approach. The scaling down path is becoming very difficult, expensive and complex to fabricate and to design with. The extremely long technological experience with silicon, its oxide of unsurpassed quality, abundant availability, extremely low cost of materials and production, the ultra large scale integration (ULSI) status of silicon CMOS and its unchallenged dominance in the semiconductor industry makes it very much undesirable, commercially and practically, to make a transition to a totally alternate technology.

Fortunately an improvement to silicon that stirred interest in the recent past has come to the rescue of the silicon industry. It may be noted that the earliest semiconductor material of choice was germanium (Ge) before silicon became the dominant material. In 1960s it was found that addition of germanium (Ge) to silicon makes a compound alloy semiconductor silicon-germanium ( $\text{Si}_{1-x}\text{Ge}_x$ ). Initial published works date back to 1955 [76], [77]. But only in the 1990s thin films and associated strain effects of this compound were studied. Superior properties of this compound semiconductor with associated strain revived the research interests in  $\text{Si}_{1-x}\text{Ge}_x$ . In applications like cell phones the high frequency front end used to be designed using III-V devices with the low frequency circuits designed using silicon devices. It is to be noted that III-V fabrication technology is not compatible with silicon fabrication technology and efforts to grow III-V devices on silicon substrates were not successful [75]. This has changed with the advent of silicon germanium. Now the front end is designed with  $\text{Si}_x\text{Ge}_{1-x}$  bipolar transistors and the back end is designed with Si CMOS using a BiCMOS technology.[ 100].

As mentioned in the beginning of the chapter silicon germanium ( $\text{Si}_{1-x}\text{Ge}_x$ ) is an alloy semiconductor formed by adding germanium to silicon. The diamond lattice and silicon type band gap are preserved for Ge fractions up to 0.85. The addition of germanium enhances the transport properties of this compound semiconductor. These enhanced properties can be utilized in the design of high-speed transistors. As mentioned earlier research on the compound semiconductor silicon-germanium ( $\text{Si}_{1-x}\text{Ge}_x$ ) is reported as early as 1955 [77] when the bulk properties were studied. The bulk  $\text{Si}_{1-x}\text{Ge}_x$  did not find much practical use as a compound semiconductor by itself. This led to disillusion of technical interest in  $\text{Si}_{1-x}\text{Ge}_x$  [80]. Revived interest in silicon germanium was reported in 1975 with the growth of super lattice structures of Si/SiGe [79]. Epitaxial growth of these hetero-junction films using ultra high vacuum chemical vapor deposition (UHVCVD) was undertaken by IBM in the year 1980 [46], [93].

Epitaxy is the preferred method to grow thin  $\text{Si}_{1-x}\text{Ge}_x$  layers on a (100) silicon substrate or another relaxed  $\text{Si}_{1-y}\text{Ge}_y$  bulk substrate of a different Ge fraction. This facilitates lattice matched or pseudomorphic growth. The molecular beam epitaxy (MBE) used in the fabrication of GaAs devices is too slow for silicon CMOS process lines in addition to being expensive. Silicon epitaxy traditionally used in CMOS and bipolar fabrication is chemical vapor deposition (CVD) at high temperatures, greater than  $1000^\circ\text{C}$ . If this high temperature CVD is used in the deposition of  $\text{Si}_{1-x}\text{Ge}_x$  the high temperature can lead to strain relaxation in the grown layer of  $\text{Si}_{1-x}\text{Ge}_x$ . This can lead to interfacial defects and degradation of carrier mobility. As will be seen in section 4.3 strain is of utmost importance to obtain high mobility in thin  $\text{Si}_{1-x}\text{Ge}_x$  layers. Low

temperature growth below about 350°C can lead to three dimensional growth and resulting in-plane defects. UHVCVD systems with a wide range of pressures as low as 10 Torr to atmospheric pressure and temperatures in the range of 500°C to 600°C are used in the growth of thin films of  $\text{Si}_{1-x}\text{Ge}_x$  [23], [111].

#### 1.4.1 Advantages of $\text{Si}_{1-x}\text{Ge}_x$

As mentioned before the enhancement in mobility and band gap engineering are two of the main attributes of  $\text{Si}_{1-x}\text{Ge}_x$  technology that makes this alloy semiconductor very much attractive in high speed transistor design. Band gap engineering gives the designer the ability to design a material of a desired band gap. This is achieved by exploiting the fact that increasing the Ge fraction reduces the band gap of silicon germanium. In an all silicon system the band gap is fixed at 1.12 eV. If the band gap can be changed junctions can be formed with different band gaps on either side or QW can be formed. These strain and Ge fraction dependent properties are discussed in detail in chapter 4.

MODFET in III-V system were the fastest transistor with cut off frequency  $f_T$  above 150 GHz with a transmission delay around few ps [78]. These devices exploit the band gap engineering and QW confinement to achieve this. MODFET inherits its high speed of operation by confining the carriers in an undoped QW channel away from the substrate surface.  $\text{Si}_{1-x}\text{Ge}_x$  provides these features in silicon technology. This opens up the possibility of the design of MODFETs capable of operating at very high frequencies, using Si and  $\text{Si}_{1-x}\text{Ge}_x$  systems.

Ultra high density Si ICs can be produced at a significantly low cost compared to other technologies. These are suited for high transistor count, high volume microprocessors and

memory applications. But the RF and microwave performance of Si is very poor. The maximum operating frequency of silicon devices is limited to about 5 GHz. At high electric fields the drift velocity reaches the saturation velocity ( $v_{sat}$ ) for holes which is equal to  $8.45 \times 10^6$  cm/sec. In the present day nanoscale devices a good portion of the channel is velocity saturated and the speed of the device is limited mostly by  $v_{sat}$  [ 31].

As will be seen in chapter 7 the off currents  $I_{off}$  for these  $Si_xGe_{1-x}$  devices are two to three orders of magnitude lower compared to the MOSFET. In the silicon ULSI chips where the heat dissipation is already a serious problem further scaling down is going to make the heat removal a challenge [ 39 ]. In that scenario the strained  $Si_{1-x}Ge_x$  hetero junction devices has a scope for replacement of silicon MOSFET as well as further down scaling.

Another aspect of these devices is the higher radiation tolerance compared to silicon. This will be desirable in space borne electronics as well as in radiation environments [ 111],[ 112] Third these devices may be faster either because of higher mobility, more velocity saturated portion of the channel compared to a Si MOSFET and due to  $v_{sat}$  increase with increased Ge fraction. This extra speed can be traded for lower power operation thus reducing the dissipation further [ 46]. Or the speed can be used to advantage by substituting smaller devices which can give the same on currents as the Si MOSFET thus increasing the packing density further.

#### **1.4.2 Silicon Germanium in all Silicon Technology**

If silicon MOSFET scaling did not come to a grinding halt at the 45 nm node and if scaling is proceeding down to expected 10 nm node why is an alternate technology

needed? To answer this question limitations of silicon CMOS need to be considered. First of all, the silicon band gap is fixed whereas band gap engineering is possible in Si/Si<sub>1-x</sub>Ge<sub>x</sub> system. Secondly, the high mobility of strained Si<sub>1-x</sub>Ge<sub>x</sub> provides much faster p-HMOSFETs than comparable all silicon p-MOSFET. It can be argued that in aggressively scaled devices the velocity saturation is going to occur anyway and hence the Si<sub>1-x</sub>Ge<sub>x</sub> may not give any advantage in speed. It has been speculated that the  $v_{sat}$  in Si<sub>1-x</sub>Ge<sub>x</sub> increases with the Ge fraction [88]. In addition due to the high mobility the Si<sub>1-x</sub>Ge<sub>x</sub> channels will be velocity saturated to a larger portion of the channel than the silicon MOSFET. Significant velocity overshoot also has been speculated for these devices [48], [81]. The third attractive feature of Si<sub>1-x</sub>Ge<sub>x</sub> HMOSFET is the fact that the off currents,  $I_{off}$ , are orders of magnitude less than those in MOSFETs. This is discussed in chapter 7. The low off current of Si<sub>1-x</sub>Ge<sub>x</sub> devices is hardly mentioned in many publications [45],[46], [50]. Conventional MOSFETs when scaled down from 250 nm to 65 nm,  $I_{off}$  increases from a few nA/ $\mu$  to hundreds of nA/ $\mu$ . The 2001 Technology Roadmap for Semiconductors points this out as a concern for scaled CMOS devices [26]. The high speed requirements demand higher  $I_{on}$  which simultaneously increases  $I_{off}$  in an all silicon CMOS. Thus, the passive power density is a significant problem in conventional CMOS [31].

Another advantage with the Si<sub>1-x</sub>Ge<sub>x</sub> system is the capability to integrate optoelectronic devices into CMOS systems [23]. Since Si<sub>1-x</sub>Ge<sub>x</sub> can be fabricated in Si CMOS process lines with minimal modifications high density Si CMOS as well as high speed Si<sub>1-x</sub>Ge<sub>x</sub> devices can co-exist on the same die.

Very high frequency devices can be derived from native silicon based devices with the introduction of  $\text{Si}_{1-x}\text{Ge}_x$ . If the  $\text{Si}_{1-x}\text{Ge}_x$  is introduced into the central region of a double gated device and if the gates are biased properly carriers will get confined to the central undoped  $\text{Si}_{1-x}\text{Ge}_x$  region [ 175]. The perpendicular electrical field is a minimum in this region. The mobility becomes high and carriers are away from the pseudomorphic  $\text{Si}/\text{Si}_{1-x}\text{Ge}_x$  interfaces. This can provide very high speed and HMOSFET devices with cut off frequencies as high as 100 GHz are reported. Fabricated devices with an all Ge channel having measured hole mobilities as high as  $9000 \text{ cm}^2/\text{Vs}$  at  $77^\circ \text{ K}$  are reported [170].

### **1.4.3 Current Status of $\text{Si}_{1-x}\text{Ge}_x$ Technology**

The 2003 Technology Roadmap for Semiconductors notes the importance of compound semiconductors [ 52]. It also notes that the challenge for compound semiconductors is to increase the number of devices per unit area and to integrate with CMOS technology [ 52].  $\text{Si}_{1-x}\text{Ge}_x$  with its associated process compatibility can easily integrate with the high density Si CMOS process. In addition strained  $\text{Si}_{1-x}\text{Ge}_x$  hetero bipolar transistors (HBT) can currently provide speeds close to III-V devices.  $\text{Si}_{1-x}\text{Ge}_x$  was introduced into the graded base hetero bipolar transistors (HBT) around 1987. Around the year 2000 these  $\text{Si}_{1-x}\text{Ge}_x$  HBTs have been introduced into the VLSI silicon technology, as a replacement to the high frequency GaAs bipolar transistors, in the BiCMOS circuits. Even though GaAs still has higher cut off frequency  $f_T$  the compatibility of Silicon Germanium HBTs with the silicon VLSI process is a definite advantage [98] - [104]. This also allows integration of both logic and high frequency

bipolar transistors on the same chip. These hetero bipolar transistors (HBTs) operate at very high frequencies and literally provide an equivalent substitute for GaAs devices on silicon wafer. The cut off frequencies for these devices range from 80 GHz to 350 GHz at room temperature [ 38], [ 95],[ 96], [ 93]. These are equally as fast as GaAs devices.  $\text{Si}_{1-x}\text{Ge}_x$  HBTs can be processed along with silicon CMOS without any difficulty in a BiCMOS process. Currently many foundries have BiCMOS fabrication lines using  $\text{Si}_{1-x}\text{Ge}_x$ . This is contrary to the prediction of 2001 Technology Roadmap for Semiconductors which stated that  $\text{Si}_{1-x}\text{Ge}_x$  presents severe challenges to cost effective CMOS integration [ 27]. It should be mentioned here that the introduction of  $\text{Si}_{1-x}\text{Ge}_x$  has changed the predictions of semiconductor technology road map of 2001 [ 100].

Silicon- $\text{Si}_{1-x}\text{Ge}_x$  BiCMOS is extensively used in wireless circuits operating in the range of 900 MHz to 2.4 GHz.  $\text{Si}_{1-x}\text{Ge}_x$  HBTs can operate above 65 GHz. Such high bandwidth is not needed in the above devices. But bandwidth can be traded for speed and power [ 46]. This reduces drain on the batteries.  $\text{Si}_{1-x}\text{Ge}_x$  HBTs can operate at very low power, have low noise, are more linear than Si BJTs and have more dynamic range [ 46]. So these devices are ideally suited for low noise high frequency applications such as low noise amplifiers (LNA) in the front end of cell phone receivers and similar front end units [ 100], [ 100].

Since the  $\text{Si}_{1-x}\text{Ge}_x$  based MOSFET has layers of  $\text{Si}_{1-x}\text{Ge}_x$  and Si it involves hetero junctions and is named HMOSFET to distinguish it from MOSFET. However, HMOSFETs are still in the developmental stage. There is particular interest in the p-channel HMOSFET (p-HMOSFET) since the hole mobility in CMOS circuits is

slightly less than half the value for electrons. Hence the design and analysis of a p-HMOSFET is chosen as the topic of this dissertation.

$\text{Si}_{1-x}\text{Ge}_x$  hetero CMOS (HCMOS) are also under study. One of the main advantages of these devices is the off current that is in the  $\text{pA}/\mu$  range. However there is very little mention of  $\text{Si}_{1-x}\text{Ge}_x$  CMOS for very low power applications in published works. The off currents in  $\text{Si}_{1-x}\text{Ge}_x$  HCMOS could be around  $5 \text{ pA}/\mu$ . Hence they are useful for very low power applications such as biomedical and wireless.  $\text{Si}_{1-x}\text{Ge}_x$  HCMOS compares favorably with aggressively scaled Si CMOS. It also has a better power-delay product than any leading analog technology [ 41]. Very low drain induced barrier lowering (DIBL) and very low  $1/f$  noise are also reported [ 160]. The ability to integrate many SiGe devices with CMOS is a great advantage. The low thermal budget of short channel CMOS is advantageous to  $\text{Si}_{1-x}\text{Ge}_x$  strained layers. Since the properties of  $\text{Si}_{1-x}\text{Ge}_x$  depends on the strain this factor is very important.  $\text{Si}_{1-x}\text{Ge}_x$  p-HMOSFETs have shown room temperature hole mobilities as high as  $1300 \text{ cm}^2/\text{sec}$  [ 161].

## 1.5 Limitations of Si CMOS

The exponential growth of silicon CMOS technology and its impressive dominance over other material systems in the global semiconductor market are due to three factors. (1) The  $\text{SiO}_2$  gate insulator is of unsurpassed quality and easily formed compared to other semiconductor systems. (2) The aggressive scaling of silicon CMOS devices brought around faster and smaller devices. (3) Silicon is an abundant raw material on earth's crust and is very economical for the semiconductor industry. Most of the cost comes from the device quality purification and processing.

Figure 1-1 shows a sketch of a p-channel MOSFET. Figure 1-2 shows its band diagram, at high gate voltage, along the mid section of the device starting at the poly gate interface. The proximity of the channel to gate provides good gate to channel coupling which translates into low threshold voltage and good transconductance  $g_m$ . But there are several drawbacks in this device structure limiting the overall performance. (1) Carriers are transported from source to drain in a region very close to the inferior Si/SiO<sub>2</sub> interface. The interface is inferior due to trapped charges in the SiO<sub>2</sub>, the interfacial defect states and surface roughness. (2) Carriers are confined to a region of very high perpendicular fields (perpendicular to the transport direction)  $\xi_{\perp}$  as indicated by the high slope of the band edge near the SiO<sub>2</sub> interface. (3) The carrier transport is in a silicon region that is moderately doped in long channel devices and highly doped in case of short

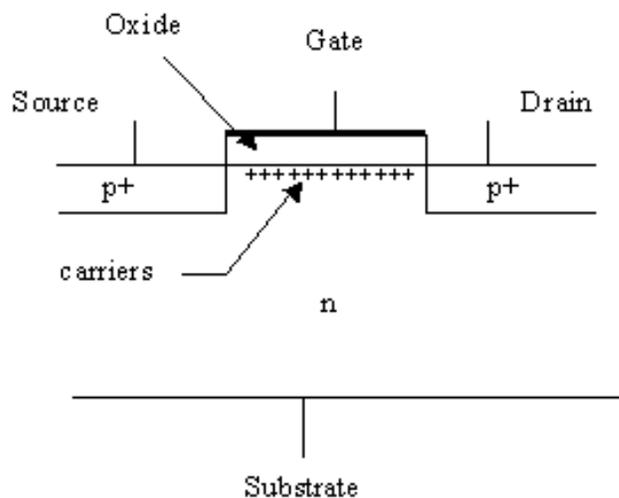


Figure 1-1. Sketch of a p-MOSFET.

channel devices. A high field is not desirable because dopant ions interact with the carriers through Coulomb forces and cause scattering. This reduces the effective mobility further by slowing down carriers. Intrinsic silicon has electron and hole mobilities of approximately 1450 and 450  $\text{cm}^2/\text{Vs}$  respectively. But in fabricated CMOS devices the above values drop down to as low as 250 and 100  $\text{cm}^2/\text{Vs}$  respectively [4]. These aspects will be discussed in detail in the following subsections, and ways to overcome these drawbacks will be considered

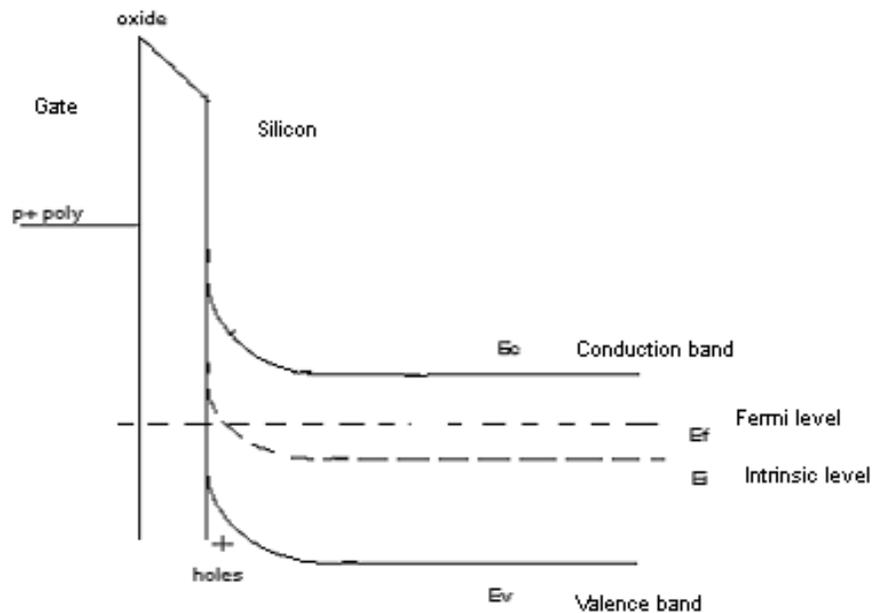


Figure 1-2. Band diagram of the p-MOSFET at high  $V_{gs}$ .

### 1.5.1 Si/SiO<sub>2</sub> Interface

Silicon has a crystalline lattice and SiO<sub>2</sub> is amorphous. The Si/SiO<sub>2</sub> interface is not lattice matched and has a large number of defect states. Even though these defects are considerably reduced through technological advancement defect densities of the order of  $10^{10}/\text{cm}^2$  are still present [4]. Electrons and holes can tunnel through the gate oxide and get trapped in the oxide. This is particularly true in scaled down nano-scale devices due to the high perpendicular fields. This trapping will modify the fields within oxide. A tunneling electron can gain high energy from the oxide field and cause impact ionization leading to the generation of holes that can get trapped in the oxide. Carriers (holes) arriving at the drain may possess very high kinetic energy and can cause impact ionization at the drain region leading to hot carrier generation. These hot carriers can get injected into the oxide as well. Apart from all the above ionizing radiation can produce carriers in the oxide and they may get trapped in it. Irrespective of the process how the charge is trapped in the oxide these charges modify the field near the silicon surface and can cause a rise in  $V_T$ . These degradations shorten device lifetime by causing threshold drift, oxide failure, drain erosion and gate leakage.

### 1.5.2. Perpendicular Fields

The silicon region near the oxide interface has very high perpendicular fields,  $E_{\text{perp}}$ . As shown in equation (1.3)  $E_{\text{perp}}$  is a weighted average value of perpendicular fields due to ions in the depletion region and holes in the inversion layer. In the case of p-channel MOSFET this field is given by [4],

$$E_{perp} = \frac{1}{\epsilon_{SI}} \left( |Q_D| + \frac{1}{3} Q_I \right) \quad (1.3)$$

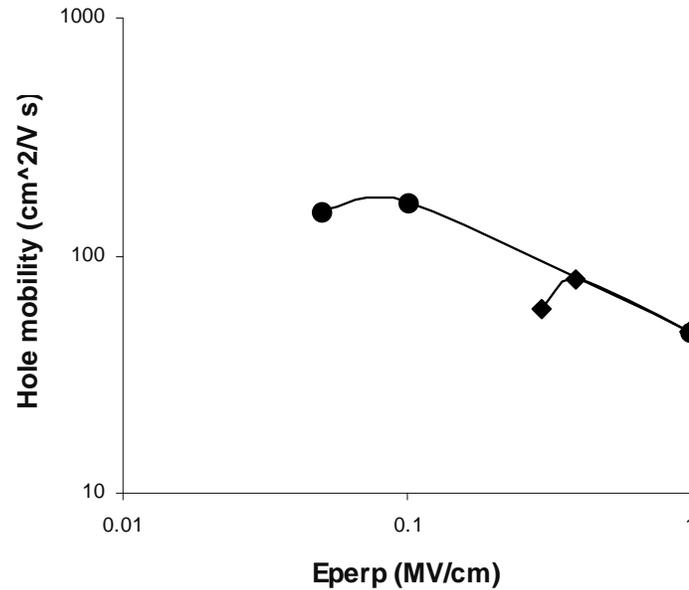


Figure 1-3. Hole mobility vs effective perpendicular field,  $E_{perp}$ . The upper and lower curves corresponds to doping densities of  $3.9 \times 10^{15}$  and  $3 \times 10^{17}$  respectively

where  $Q_D$  and  $Q_I$  are the depletion and inversion charges respectively. This field increases with the inversion charge. This field confines carriers close to the  $\text{SiO}_2$  interface. Since the  $\text{SiO}_2$  interface has traps and roughness, transport properties get degraded. Published curves showing the degradation in mobility with increasing  $E_{eff}$  are available [ 189].

Figure 1-3 shows such curves for holes corresponding to two doping concentrations of  $3.9 \times 10^{15}/\text{cm}^3$  and  $3 \times 10^{17}/\text{cm}^3$ . At low  $E_{eff}$  the corresponding mobilities are around  $170 \text{ cm}^2/\text{Vs}$  and  $80 \text{ cm}^2/\text{Vs}$ . For an  $E_{eff}$  of  $1.0 \text{ MV/cm}$  the mobility for both doping has reduced to about  $48 \text{ cm}^2/\text{Vs}$  [ 189]. If carriers are moved into a quantum well (QW) the

deleterious effect of the oxide interface can be avoided. This is what is done in the case of the  $\text{Si}_{1-x}\text{Ge}_x$  HMOSFET dealt with here.

### 1.5.3. Substrate Doping

As mentioned above the carrier transport in a MOSFET device is through a doped layer of silicon. Undoped or intrinsic silicon has electron and hole mobilities of approximately 1450 and 450  $\text{cm}^2/\text{Vs}$  respectively. As the doping levels are increased the mobility reduces due to the Coulomb scattering due to the ions present. Beyond a doping level of  $10^{16}/\text{cm}^2$  this drop is substantial. This drop can be represented by an empirical relationship as follows [105],

$$\mu = \mu_{\min} + \frac{\mu_o}{1 + (N / N_{\text{ref}})^2} \quad (1.4)$$

where  $\mu_{\min} = 54$ ,  $\mu_o = 407$  and  $N_{\text{ref}} = 2.35 \times 10^{17}/\text{cm}^3$  for holes. Figure 1-4 gives the above doping dependence on mobility. It is to be mentioned here that the doping density in the channel region normally ranges from  $10^{17} - 5 \times 10^{18}/\text{cm}^3$  for typical CMOS devices. As seen in the figure there is substantial reduction of mobility in this range due to increased interaction with charged ions.

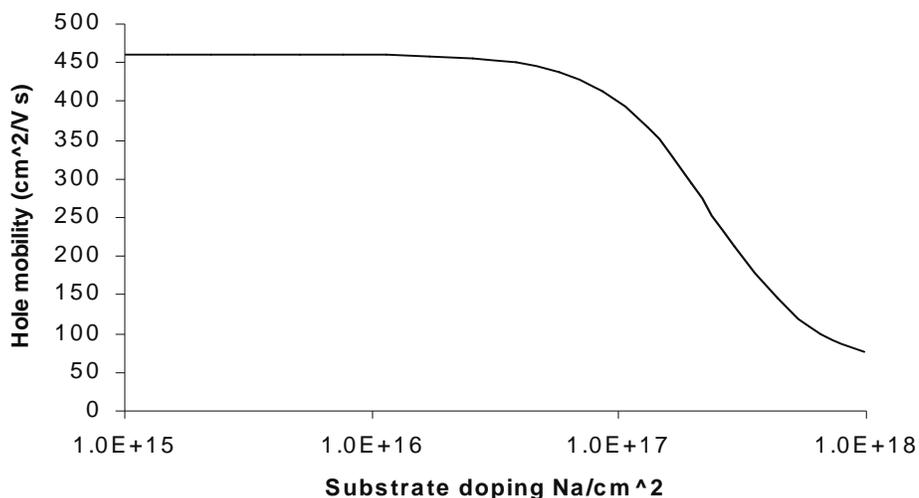


Figure 1-4. Hole mobility vs substrate doping density  $N_A$

#### 1.5.4. Improving Silicon CMOS Performance

In the previous subsections three prominent causes of degradation in silicon CMOS performance were discussed. In order to overcome or avoid them altogether the following need to be done.

1. Carriers should be moved away from the  $\text{SiO}_2$  interface. This could be done if a QW is formed below the  $\text{SiO}_2$  interface deep in the silicon substrate. Then carriers can be confined in the QW. Introduction of  $\text{Si}_{1-x}\text{Ge}_x$  can provide this QW. If the carriers are moved away from the silicon surface do we need a gate oxide? A metal-semiconductor Schottky gate can be used but the gate leakage currents will be higher than the oxide gate devices. This is due to the lower barrier at the metal-semiconductor contact.
2. It is also to be noted that the QW interface can be grown pseudomorphically. In contrast to the  $\text{SiO}_2$  interface this pseudomorphically grown interface is far

superior and is free of trapped charges. Interfacial defects in such pseudomorphically grown  $\text{Si}_{1-x}\text{Ge}_x$  interfaces are reported to be below  $10^9/\text{cm}^2$  [ 111 ].

3. The QW region need not be doped. Carriers can accumulate in the QW due to its proximity to the Quasi Fermi level compared to the rest of the valence band along the mid section of the device. So carrier transport can happen in the undoped QW region. This overcomes the mobility degradation due to stationary dopant ions and associated Coulomb scattering.

If all of the above three limitations of silicon could be overcome, with solutions considered here, much higher carrier mobilities could be achieved and higher operating speed devices can be realized. Unlike the conventional MOSFET case the carriers are away from the gate by the gate oxide and the cap layer (silicon thickness above the QW). In the conventional case the gate to channel capacitance is,

$$C_{OX} = \frac{\epsilon_{OX}}{t_{OX}} \quad (1.5)$$

whereas in this device the gate to channel capacitance is the series combination of gate oxide and the cap layer,

$$C_{GC} = \frac{\epsilon_{OX} \epsilon_{SI}}{t_{OX} \epsilon_{SI} + t_{CAP} \epsilon_{OX}} \quad (1.6)$$

where  $C_{OX}$  and  $C_{GC}$  are the gate-to-channel capacitance,  $\epsilon_{OX}$  and  $\epsilon_{Si}$  are dielectric constants of  $\text{SiO}_2$  and Si respectively,  $t_{OX}$  and  $t_{CAP}$  are thicknesses of gate oxide and cap layer respectively. Equation (1.6) will give a lower capacitance value than equation (1.5). For example, using 11.7 and 3.9 as  $\epsilon_r$  for Si and  $\text{SiO}_2$  respectively, equation (1.5) gives a value 33% more than equation (1.6). That means the gate control in the case of

this device is less than that in case of MOSFET. This smaller gate control tends to reduce the transconductance  $g_m$ . However the higher mobility in this device can yield higher currents that translates into higher  $g_m$ .

## **1.6 Objective and Organization of the Project**

In CMOS integrated circuits both n-channel and p-channel devices should operate at equal speed to enable optimization of logic circuits in terms of performance and integration density. As mentioned earlier room temperature electron and hole mobilities for minimally doped bulk silicon are around 1400 and 450  $\text{cm}^2/\text{V}\cdot\text{sec}$  respectively. In processed devices these further drop down to less than 250 and 100  $\text{cm}^2/\text{V}\cdot\text{sec}$  [106], [107]. Since the hole mobility is less than one half of the electron mobility p-channel devices have to be made more than double the size of the n-channel devices to have the same drive currents. This reduces the layout density on the die. In addition to consuming die area it also increases parasitic capacitances. Hence an improvement in p-channel devices can considerably improve the CMOS logic circuit layout density and performance. Hence the development of an analytical program to aid in the design of p-channel silicon/silicon-germanium based p-HMOSFET is the main goal of this research work. This analytical program gives results much faster than the time consuming numerical simulators. So this program can act as an interface between the designer and the numerical simulator.

### **1.6.1 Objective**

The objective of the work presented here is to develop an analytical program suitable for easy design of a p-HMOSFET by varying structural parameters in a generalized yet

simple device structure incorporating basic necessary layered structures. Such a generalized program is not found in literature. The published works mostly center around certain simulated or fabricated fixed structural designs [148 ], [172 ], [173 ]. These are not useful for designing optimized structural combinations. Some other published works are studies of variation of one of the parameters. For example cap layer optimization or channel optimization [ 158 ],[182 ]. Yet some others have studied certain particular characteristics like threshold voltage and hole densities [153 ], [168 ]. The generalized structure selected here is adapted from structures that seem to be the most common choices in published research works of the recent past. By choosing appropriate parameters many of these published structures can be duplicated in this program. However, it was found that many of the published works have structures too complex to allow testing with the program developed here. This is because these published devices have too many layers or complex growth patterns. Examples are fabricated structures with graded  $\text{Si}_{1-x}\text{Ge}_x$  substrate and graded channels [149], [151 ],[152 ]. Within its structural limitations the program developed here can assist in quickly adjusting the structural parameters to achieve the particular device performance sought. These parameters may be  $I_{\text{on}}$  and  $I_{\text{off}}$ . This has been successfully carried out in this work as discussed in chapter 7. This analytical program has been compared to numerical simulations and found to yield comparable results. Compared to the time consuming numerical computations with trial and error structural inputs this analytical program can provide results much faster even though the parameter stepping is still based on prudent

choice following the trends of the performance changes and guided by the design windows. So the objective of this project can be listed as follows,

1. A thorough literature survey has been carried out to select a suitable strained Si/Si<sub>1-x</sub>Ge<sub>x</sub> p- channel hetero junction MOSFETs (p-HMOSFET). Based on this a generalized structure is chosen encompassing the majority of published works of the past half a decade. This structure has the basic necessary layers and parameters to achieve a good deal of improvements over the conventional p-MOSFET. A consensus of a particular structure, like the MOSFET, has not yet been found in the literature for the p-HMOSFET.
2. The generalized p-channel hetero junction MOSFET, named p-HMOSFET, is analyzed using one dimensional Poisson equation. This gives the potential distribution along the central region of the device. This is a long channel device. The generalized device chosen here is very similar to few of the published works [168] and the analysis closely resembles one of the published works [184]. But the work presented here differs from the published works in three aspects. (1) The program developed here has an additional p-layer (2) the published works only presents studies of the dependence of some of the parameters on certain structural parameters, e.g. the threshold voltage dependence on cap layer thickness. (3) the program developed here can also study the parameter dependence. But more than that the program assists in the quick design of a p-HMOSFET. Thus the program developed here acts as an interface between the designer and the numerical simulator. The program gives a fairly close design that

need to be checked in the numerical simulator only a few times thus saving very much of the design time. Such a design program is not found in literature.

3. The result of the analysis is transformed into an analytical program that can take user inputs for device structure and other parameters and very quickly produce the terminal characteristics of the device without the need for the time consuming and tedious numerical algorithms. This is achieved through the use of Excel spread sheet.
4. Other than terminal characteristics the analytical program can also give the potential variation along the center of the device and hole density distribution among other parameters. These are compared to results of the numerical programs and the accuracy is found to be good.
5. Using this program the useful range of device parameters as well as design windows for p-HMOSFET devices have been produced.
6. After establishing the accuracy of the analytical program for long channel devices short channel comparisons are carried out and results compared to similar MOSFET devices.

### **1.6.2 Procedure**

The p-HMOSFET device layer sequence found in most of the literature can be divided into three categories. Starting from the substrate upwards,

1. SiGe channel/gate oxide (Figure 1-5)
2. Compressively strained Si<sub>1-x</sub>Ge<sub>x</sub> recessed channel/ silicon cap layer/gate oxide (Figure 1-6)

3. p-Si/ Si buffer/ compressively strained SiGe channel/ Si cap layer/gate oxide

(Figure 1-7)

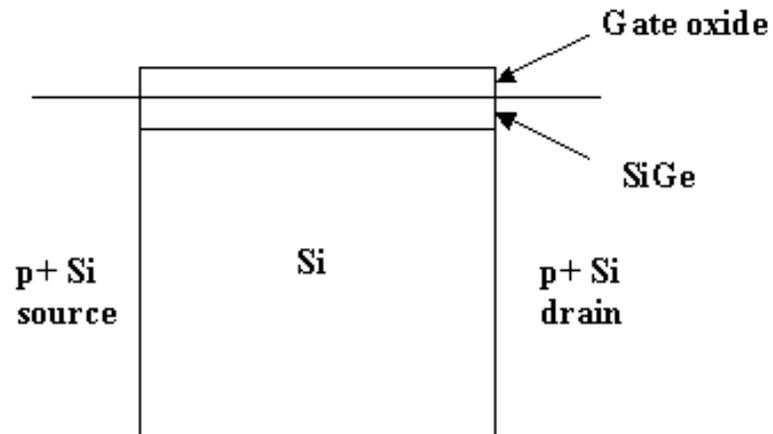


Figure 1-5. A simple SiGe MOSFET. It has a SiGe surface channel.

In this work a device as in item 3 in the list above has been adopted. The reason for this is that it is the simplest, with just the necessary layers to make most of the improvements compared to p-MOS. Further modifications of diverse individual nature are seen in literature. An exact direction for future mature devices or a consensus in the device structure, similar to MOSFET, has not been reached yet. But most of the recent

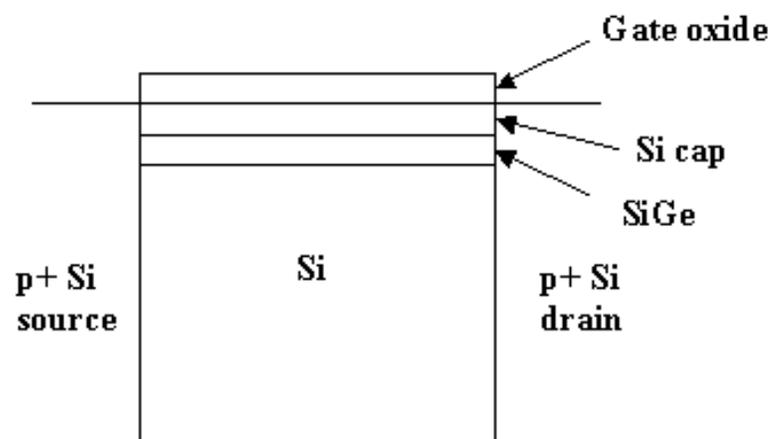


Figure 1-6. A recessed SiGe channel MOSFET. It has an undoped Si cap layer.

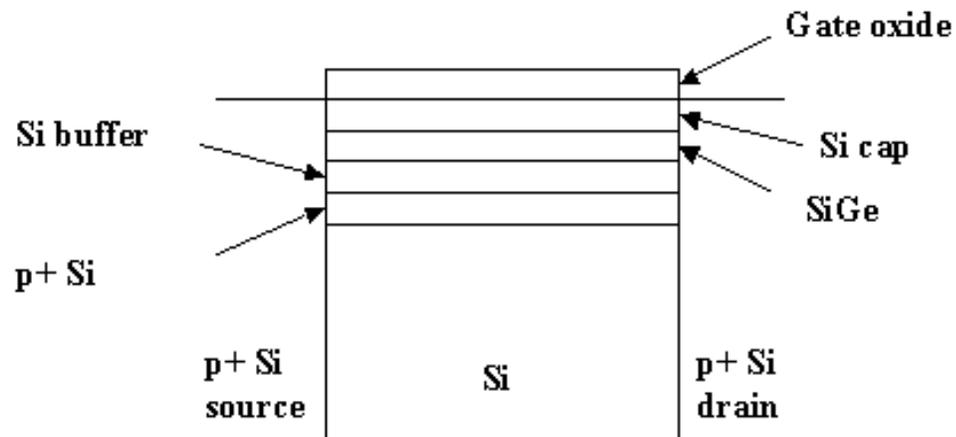


Figure 1-7. A recessed SiGe channel MOSFET with additional p+Si layer and an undoped Si buffer layer

published works seem to fall into the ones given in the above list. Taking the above structure the following are done to arrive at the objective.

1. A one dimensional Poisson equation is solved for the device structure
2. From the solution of the Poisson equation internal potentials and hole densities are derived
3. Making use of the solution of the Poisson equation an analytical program is written using the Excel spreadsheet with structural parameters left as variables of the design
4. Devices with varying structural values are simulated using the two-dimensional device simulator ISE-TCAD
5. The numerically simulated device structures are input into the analytical program and functions like potentials, hole density in QW and surface, hole sheet density and centroid of the hole charge are computed

6. The numerical results and analytical results are compared and modifications are introduced to the analytical program until the results agree. This results in an analytical program whose results closely agree with the numerical simulations.
7. Now the design can be done easily using the analytical program and the results need to be checked with numerical simulation only a few times with a few variations around the analytically obtained structural parameters.
8. Using the analytical program one parameter at a time was varied keeping the rest of the parameters fixed at chosen values. This is done for different sets of fixed parameters. These are also duplicated with the numerical simulator to make sure that the results agree
9. Using parametric variation of input data appropriate design windows are generated showing the suggested range of variation of each structural parameter
10. Using parametric variation of input data appropriate design windows are generated showing the suggested range of variation of each structural parameter
11. From the design window space optimal devices are formed from near the center of the window region and the terminal characteristics are computed using the analytical program. Once again the results are numerically simulated and the results are compared.

To study short channel device limitations the following is carried out. A p-MOSFET and a p-HMOSFET, both long channels, are designed for equal on currents  $I_{on}$ . The substrate doping densities are also chosen the same to simulate similar short channel behavior. Then their channel lengths are reduced systematically and the performance characteristics are compared.

### **1.6.3 Organization of chapters**

The overall organization of the dissertation along with a short summary of different chapters is given below.

Chapter 2: Gives a short overview of modern conventional CMOS devices. This serves as a basis for the derivation of performance parameters of the HMOSFET and also serves as a means to compare performance of the p-HMOSFET with p-MOSFET.

Chapter 3. Discusses the structure and simulation of an ideal simplistic QW device in silicon. This is a preliminary study to get an overall feeling of the performance and the feasibility of such a device. Instead of solving the full Poisson equation certain internal parameters are assumed. This makes the Poisson solution easier. Effects of variation of all structural parameters is given in graphical form.

Chapter 4: The effects of strain in  $\text{Si}_{1-x}\text{Ge}_x$  are discussed in this chapter. The concept and structure of a strained  $\text{Si}_{1-x}\text{Ge}_x$  p-HMOSFET device in general is discussed. Here advantages of  $\text{Si}_{1-x}\text{Ge}_x$  p-HMOSFETs are discussed along with a discussion of diverse structures found in published literature. A particular structure is chosen and the choice is justified. Along with the structural choice the potential and band gap arrangements are discussed to give a better understanding of the p-HMOSFET device operation.

Chapter 5: Discusses the design of a p-HMOSFET and the appropriate choice of parameters and their implications. The implications in the choice of different design parameters are discussed. The necessary design parameters are identified.

Chapter 6: The solution of the one dimensional Poisson equation and the development of the analytical HMOSFET design program are described in this chapter. The internal workings of the program are presented.

Chapter 7: This chapter discusses the numerical simulations of chosen device structures. These are long channel devices with a channel length of 0.5 microns. The numerical simulations are carried out using a two-dimensional numerical device simulator ISE-TCAD. The internal and external parameters from the analytical HMOSFET design program and numerical simulations are compared to determine the accuracy of the analytical program in Excel spreadsheet. A parametric study is carried out using the analytical HMOSFET design program and the results are compared to numerical simulation results. Based on these results design windows are created. A comparison of the proposed device and its performance parameters to the relevant published research from other sources is also given.

Chapter 8: This chapter discusses short channel devices and their simulations. The analytical program is modified for short channel p-HMOSFET and again chosen structures are simulated using the analytical program and are compared with the results of the numerical simulator. Mainly the focus is in scalability of p-HMOSFET and SCE.

Chapter 9: Here an attempt is made to optimize the p-HMOSFET structure. An Excel program is used to find the feasibility of the optimized structure. Then the numerical simulator is used to check the optimization and evaluate the performance.

Chapter 10: The summary and conclusions of this research work are given in this chapter. Advantages of the design program developed here and how it can be a useful interface between the designer and the time consuming numerical simulators is discussed here. Limitations of this work and suggestions for more advanced simulations are also given in this chapter.

## CHAPTER 2

### MODERN SHORT CHANNEL PMOS

CMOS has occupied an unscathed dominant share of the semiconductor industry for the past thirty years. One reason for this is CMOS has the least power dissipation of all competing technologies. Another reason is its planar process that is much easier to integrate into large systems. Yet another reason is that CMOS is economical to fabricate. Electronic evolution of CMOS was successful because it followed the three tenets:

- (1) the voltage or current gain available through these devices.
- (2) the fact that these devices amplify signals above the noise floor thus yielding high signal to noise ratio.
- (3) the scalability of these devices to smaller sizes enabling the integration of large number of devices [ 54 ].

Long channel MOSFET has a very simple structure as shown in Figure 2-1. This figure is reproduced here from Appendix A. The theory of operation of this device along with the necessary semiconductor physics are given in Appendix A. It is possible to find a closed form solution for its drain current and other parameters like transconductance. The Spice Level –1 simulation models used to simulate these devices were very simple with only less than a dozen model parameters. Hand calculations could be compared to simulation results to get an instant feed back. But the modern nanometer scale devices are very complex. The gradual channel approximation used to find the drain current in long channel devices are not valid with these modern devices. The electric field under the gate

is no longer one dimensional. On top of these most of the channel velocity is saturated. Spice simulation models like BSIM – 4 that is used to simulate these devices has nearly two hundred parameters [ 55 ], [ 56 ] Hand calculations are difficult or impossible. If hand calculations or Level –1 model simulation is applied to these devices the error could be over 100% [ 55 ]

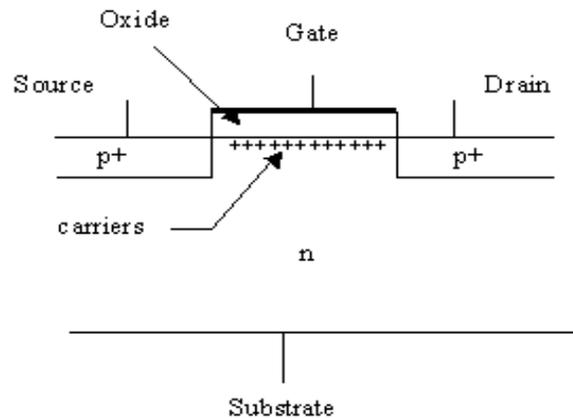


Figure 2-1. Cross sectional view of a basic long channel p-MOSFET

## 2.1 Scaling

The dominance of CMOS technology over the past three decades permitted the industry to focus its efforts on scaling. This led to remarkable performance, functionality and cost per function. Reduction in the channel length of a MOSFET can be described by scaling theory. A scaling parameter  $S$  is used to scale dimensions and operating voltages. The value of  $S$  is around 0.7 for one technology node to the next. Thus a 3.3 V device gets scaled down to 2.3 V, which is near the actual value used, which is 2.5 V. The main

benefits of scaling are (1) smaller device sizes and thus reduced chip area which is a premium and increased yield due to more dies per wafer, (2) lower gate delays which enables high speed of operation, (3) reduction in power dissipation [ 55 ].

The scaling down of MOSFET followed Moore's law that states that the transistor count will double in every  $1^{1/2}$  years. This rate has followed up to the nanometer technology node. Then the exponential rate seems to become more linear. Economical considerations are believed to be the reason for this [ 7 ], [ 57 ]. If a resistor is shunted by a capacitor the RMS thermal fluctuations (Johnson-Nyquist noise) is given by,

$$V_N = \sqrt{\frac{kT}{C}} \quad (2.1)$$

For allowable bit error in a digital system it was found that there is a limit to scaling due to this. This is given to be [ 57 ],

$$\frac{V_{TH}}{V_N} \geq 12 \quad (2.2)$$

Since scaling reduces the capacitances and since the thermal voltage  $V_{th}$  does not change there is a limit to scaling. This limit is reported to be around 40 nm. So it is speculated that scaling may not follow Moore's law beyond this limit. The technology node is currently at 45 nm. 33 nm node is already announced by Intel and IBM corporations and will be in production in the year 2008 - 2009. Research works have produced p-MOSFETs as low as 22 nm gate length [ 58 ] - [ 63 ]. The operating voltages of these devices are around 0.9 – 1.1 V. The threshold voltage  $V_T$  is as low as 0.17 V. Even zero  $V_T$  devices are announced [ 38 ]. The  $I_{ON}/I_{OFF}$  ratios of these devices are as low as 5000 which is very much inferior compared to  $10^6 - 10^7$  in long channel MOSFETs. Gate

oxide thickness of these devices is 1.2 – 2.2 nm. This is at the tunneling onset range and some devices have gate currents of around 0.3 nA/ $\mu$ .

## 2.2 Short Channel Effects

Other than the benefits of scaling down mentioned above there are also some unwanted side effects called short channel effects (SCE). This increases the sub-threshold leakage currents.

### 2.2.1 Velocity Saturation

One of the most important SCE in MOSFET stems from velocity saturation. When  $V_{ds}$  is small or the channel length  $L_g$  is large the horizontal electric field is low in the channel region. Then the carriers assume a linear low-field mobility relation between carrier velocity and field. But in modern nanometer scale transistors  $V_{ds}$  is relatively high and  $L_g$  is scaled down to be much smaller. So the lateral fields are very high. Then the carrier velocity saturates to a value,  $v_{sat}$ , in equilibrium with the thermodynamics of the lattice. Saturation velocity is a semiconductor material limitation. The hole saturation velocity,  $v_{sat}$ , is the same in silicon and in  $Si_{1-x}Ge_x$  and is equal to  $8.37 \times 10^6$  cm/s. Following the inception of  $v_{sat}$  the current saturates to a value of  $I_{Dsat}$  and will not increase any higher with increasing field. This in effect is a reduction of hole mobility and will increase the carrier sheet resistance. So in modern CMOS devices this has to be taken into account while calculating the drain current and other parameters.

### 2.2.2 Pinch off, Saturation and Channel Length Modulation

The above sub-section described velocity saturation of holes as a limiting mechanism to the increase in drain current. This, as discussed, is a material limitation. Drain current

can also get saturated by another mechanism called pinch off. Pinch off by itself need not be a short channel phenomena. It can happen in long channel devices also. It depends on the gate and drain voltage magnitudes. The drain current is given by the relation [4],

$$I_{ds} = \mu_{eff} \frac{W}{L} \int_0^{V_{ds}} [-Q_I(V)] dV \quad (2.3)$$

This states that the drain current is proportional to the area under the inversion charge density between  $V = 0$  and  $V = V_{ds}$ . Here  $V$  is the surface potential at any point along the channel and  $Q_I(V)$  is the inversion layer charge density at the point where the surface potential is  $V$ . At low  $V_{ds}$  the inversion charge density near the source and drain are nearly the same. As  $V_{ds}$  is increased  $I_{ds}$  increases but the inversion charge density near the drain end of the channel keeps reducing. This is because the potential available for inversion,  $V_{gs} - V_{ds}$  keeps reducing. Finally the charge density near the drain end falls to zero when  $V_{ds}$  reaches  $V_{Dsat}$ . This is called pinch off. Since  $V_{Dsat}$  is the drain voltage at which  $I_{ds}$  saturates the drain current becomes  $I_{Dsat}$ . Further increase in  $V_{ds}$  moves the pinch off point slowly towards the source by  $\Delta L$  thus shortening the channel to  $L - \Delta L$ . This is called channel length modulation. The potential at the pinch off point stays at  $V_{Dsat}$ . So the device looks like a reduced channel length device operating at  $V_{Dsat}$  instead of  $V_{ds}$ . Since there is an effective reduction in channel length there is a slight increase in  $I_{ds}$  in the region of  $V_{ds}$  above  $V_{Dsat}$ .

### 2.2.3 Drain Induced Barrier Lowering (DIBL)

DIBL happens in the sub-threshold region when a  $V_{ds}$  is applied. In modern nanometer devices  $V_{ds}$  is high and  $L_g$  is very short. Hence the drain voltage can reduce the source-

channel barrier causing excess leakage currents. This can be considered as a reduction in threshold voltage  $V_T$ . Without DIBL the sub-threshold slopes will be same for adjacent values of  $V_{ds}$  and the curves will be all parallel. When DIBL is present the slope of the sub-threshold region reduces indicating more leakage current. This is the reason for the high leakage current or low  $I_{ON}/I_{OFF}$  ratio noted earlier.

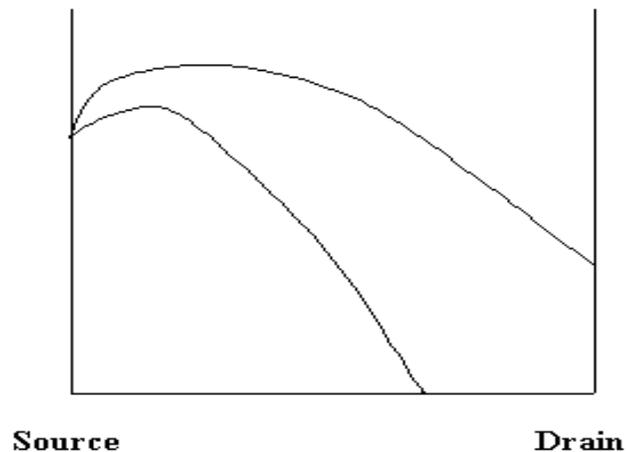


Figure 2-2. Potential plot from source to drain. As the gate length is reduced the potential barrier at the source junction reduces and carrier injection from source to channel occurs.

#### 2.2.4 Velocity Overshoot and Hot Carriers

Some carriers drifting near the drain end can attain energy much larger than the thermal energy of carriers under equilibrium. These carriers are called hot carriers. The velocity of these carriers can exceed the saturation velocity. This effect is called velocity overshoot. This enhances the speed of the MOSFET. But it also causes two bad effects. (1) these carriers can tunnel through the gate oxide causing gate current. (2) they can get trapped in the gate oxide and cause fluctuations in  $V_T$ . Hot carriers can also cause impact ionization at the drain junction and erode the drain increasing the channel resistance.

Lightly doped shallow drain implants are used to reduce this effect. The idea is to introduce a lightly doped short thin extension to the drain and source regions. This acts as a buffer between the channel and the drain. It is shown in Figure 2-3. The resistance of this region being higher it drops more voltage in a short distance. This reduces the field at the drain. This has the disadvantage of increasing the channel resistance and reducing the drain current.

### 2.2.5 Substrate Current Induced Body Effect

Hot carriers cause impact ionization and generate hole-electron pairs. In a p-MOSFET the hole goes to the gate oxide and the electron flows through the substrate resistance causing a substrate bias. This decreases the  $V_T$ . This increases the drain current but reduces the MOSFET output resistance causing a gain reduction.

### 2.2.6 Gate Tunnel Current

As gate oxide thickness is scaled down to 1.2 – 1.5 nm tunneling currents through the oxide can lead to gate leakage. The modern nano devices, as noted above, are in this range. High dielectric constant (high-K) materials are substituted for the gate oxide to reduce this problem. Hafnium nitride (HfSiON) is one of the materials for this. High K materials prevent the dielectric becoming too thin. The effective oxide thickness with this high K dielectric is given by,

$$EOT = \left( \frac{k_{SiO_2}}{k_{High-K}} \right) (thickness)_{High-K} \quad (2.4)$$

### 2.2.7 Gate Oxide Breakdown

For reliable operation the gate oxide electric field should be limited to 10 MV/cm or 1 V/nm. From the data noted above the modern nano devices are operating close to this value.

### 2.2.8 Sub-surface Punch Through

When the channel length  $L_g$  is very short and the  $V_{ds}$  is high the depletion regions of the source junction and the one from the drain junction can come into close proximity. This can reduce the barrier and cause a current flow from source to drain. If they overlap sufficient current can flow to cause a sub-surface break down.

### 2.2.9 Random Dopant Effects

In devices smaller than 100 nm the doping cannot be taken as uniform. The discrete random locations of them cause variations in threshold voltage and drain current. Statistical simulations of this random doping effect show 0.1 V variation in  $V_T$  at a channel length of 30 nm. Compared to the uniform doping case random doping simulations show lower threshold voltage [ 68 ]

### 2.2.10 Ballistic Transport

When gate length is made very small and gate oxide is made very thin ballistic transport from source to drain is possible. The velocity is limited by the oxide interface scattering only. Carriers can attain high velocities above saturation velocity since most of them pass from source to drain without scattering events with the lattice. For a fabricated 40 nm gate length p-MOSFET with oxide thickness of 1.5 nm operating at 1.5 V  $I_{DS}$  of 0.61 mA/ $\mu$  has been reported. This high current is attributed to ballistic transport [ 63 ].

### 2.2.11 Velocity Overshoot

When high lateral fields are present in nano scale devices the effective saturation velocity can exceed the bulk saturation value  $v_{sat}$ . This is termed as velocity overshoot. In a fabricated p-MOSFET device the hole velocity is high as  $5.5 \times 10^6$  cm/s and is still rising. This is speculated to be due to velocity overshoot [ 69 ] Saturation velocity for holes in silicon at 370 K is around  $7 \times 10^6$  cm.

### 2.3. Nano Scale Device Design Aspects

Many challenges have to be overcome while scaling down the MOSFET in the nanometer range. Some of them are technology node specific to enhance the performance. Some others are for overcoming SCE. Sketch of a modern nanometer scale device is shown in Figure 2-3. Scaling is done to increase the operating frequency. Simultaneously high  $I_{ON}$  is needed for fast operation. So different techniques are used to increase the drain current  $I_{ON}$  without affecting the  $I_{OFF}$ .

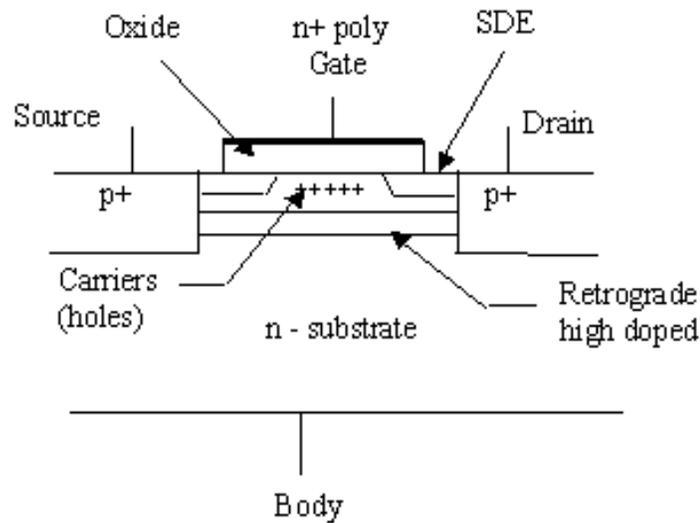


Figure 2-3. Sketch of a modern nanometer scale p-MOSFET. Drawing is not to scale. Source/Drain regions are enlarged to show details. The device has source/drain extensions (SDE) and a deep high doped Retrograde layer.

Source/Drain engineering is done to reduce the SCE and to increase  $I_{ON}$ . The deep source and drain junctions cause charge sharing. To reduce this effect short thin extensions of source and drain are used to connect the channel to the source and drain. These source/drain extensions (SDE) are 50 – 100 nm for 0.25  $\mu$  technology. For a 70 nm technology they are about 20 – 30 nm. They should also overlap the gate by 15 - 20 nm to prevent degradation of the drive current [10]. Even though thinner SDE is better for short channel suppression SDE thinner than 30 nm do not give improvement in drain current. This is due to the increased external resistance of the SDE itself.

Retrograde well engineering can be used to reduce SCE and improve drain current  $I_{DS}$ . When the channel length is very small a high substrate doping will be needed to reduce the source and drain junction depletion widths. This is needed to suppress sub-surface punch through by reducing depletion widths at source and drain junction regions. But a

high uniform doping can increase the  $V_T$  thus reducing the drain current. To avoid this a low doping is placed near the top oxide interface and a high doping is placed deeper down. This is named retrograde doping. A halo implant or a SDE is also used along with retrograde doping. Halo (pocket) is a high doped thin region placed near source and drain junctions by tilted implant. This halo or pocket implant gives additional control over  $V_T$ . This is because the pocket region has the lowest potential and it controls  $V_T$  by providing a barrier to the channel. The doping of this pocket can be controlled to control the pocket potential and hence  $V_T$ .

High dielectric constant materials are useful in nanometer devices to replace the gate dielectric. When the gate oxide reaches below about 1.5 nm tunneling can occur. Tunnelling will increase the gate leakage current. By substituting the gate oxide with a high dielectric material thicker insulators can be used as given in equation ( 2.4 ). This allows further scaling until it reaches the onset of tunneling [ 66 ]. These materials however need a thin oxide layer below it in order to reduce the interface states. It also need a metal gate electrode. This is due to the fact that these materials react with poly silicon to form  $\text{SiO}_2$ .

Channel engineering is yet another technique to improve the performance. If the silicon channel is strained the mobility increases due to the warping and separation of the valence bands. Epitaxially grown  $\text{Si}_x\text{Ge}_{1-x}$  source and drain can impart compressive strain to the silicon channel. This improves the drain current [ 67 ]. Additionally using  $\text{Si}_x\text{Ge}_{1-x}$  for the source and drain regions can reduce SCE. Since the band gap of  $\text{Si}_x\text{Ge}_{1-x}$  is less

than that of Si there is a band discontinuity at the source-channel junction. This barrier being a fixed barrier and not a depletable one the DIBL effects are suppressed [ 64 ].

## 2.4 Analysis of Short Channel MOSFET

The DC and small signal models derived for long channel MOSFET are not applicable for nano scale devices. The error could be over 100%. One of the important effects is velocity saturation over a good portion or the entire channel. This makes the square law model of  $I_{ds}$ , used in long channel devices, obsolete. In these nano scale devices  $I_{ds}$  varies linearly with overdrive voltage ( $V_{gs} - V_T$ ). This is contrary to the long channel devices where the  $I_{ds}$  has a square law dependence.

### 2.4.1 Drain Current

Carrier velocity with velocity saturation can be written as,

$$v_d = \frac{\mu_p \xi}{1 + \frac{\xi}{\xi_c}} \quad (2.5)$$

where  $\xi$  is the electric field at any point along the channel,  $\xi_c$  is a critical field and  $\mu_p$  is the low field mobility for holes. The drain current is given by,

$$I_D = WQ(y)v_d(y) \quad (2.6)$$

where  $W$  is the channel width,  $Q(y)$  is the charge density at point  $y$  in the channel and  $v_d(y)$  is the velocity at point  $y$ . The electric field is given by,

$$\xi(y) = \frac{dV}{dy} \quad (2.7)$$

Combining the above three equations, noting that  $Q(y) = C_{OX} (V_{gs} - V_T)$  and integrating from 0 to  $L$  for  $y$  and from 0 to  $V_{ds}$  for  $V$ ,

$$I_D = \frac{\mu_p C_{OX}}{2 \left( 1 + \frac{V_{DS}}{\xi_c L} \right)} \frac{W}{L} \left[ 2(V_{GS} - V_T)V_{DS} - V_{DS}^2 \right] \quad (2.8)$$

Let  $V_{Dsat}$  be the value of  $V_{DS}$  that marks the boundary of the triode linear region and the saturated region. Taking the derivative  $\delta I_D / \delta V_{DS}$ , knowing that the derivative is zero at this boundary, forming the differential equation and solving,

$$V_{Dsat} = V_{DS} = \xi_c L \left( \sqrt{1 + \frac{2(V_{GS} - V_T)}{\xi_c L}} - 1 \right) \quad (2.9)$$

The ratio under the square root can be expanded in Taylor series to get,

$$V_{Dsat} \approx (V_{GS} - V_T) \left( 1 - \frac{V_{GS} - V_T}{2\xi_c L} \right) \quad (2.10)$$

Substituting this  $V_{Dsat}$  in ( 2.9 ) for  $V_{ds}$  in ( 2.8 ),

$$I_{Dsat} = \frac{\mu_p C_{OX}}{2} \frac{W}{L} (V_{Dsat})^2 \quad (2.11)$$

Noting that  $v_{sat} = \mu_p \xi_c$  substitute ( 2.9 ) into ( 2.11 ) and taking the limit  $\xi_c \rightarrow 0$  for complete velocity saturation,

$$I_D = \mu_p C_{OX} (V_{GS} - V_T) \xi_c = W C_{OX} (V_{GS} - V_T) v_{sat} \quad (2.12)$$

This gives a linear relation of overdrive voltage to  $I_{DS}$ .

#### 2.4.2 Transconductance and Transition Frequency

As before substituting ( 2.9 ) into ( 2.11 ), finding  $\delta I_D / \delta V_{GS}$  and taking the limit

$\xi_c \rightarrow 0$ ,

$$g_m = k \frac{W}{L} (V_{GS} - V_T) \quad (2.13)$$

Transconductance increases when overdrive ( $V_{gs} - V_T$ ) increases or channel length decreases.

If  $C_{GS} \gg C_{GB} + C_{GD}$  transition frequency  $f_T$  can be written as,

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{GS}} \propto \frac{WC_{OX} v_{sat}}{WLC_{OX}} \propto \frac{v_{sat}}{L} \quad (2.14)$$

### 2.4.3 Scaling Limits and Reliability

Finally there are some additional issues with nanometer devices. The gate oxide thickness is aggressively scaled to the limit of breakdown in modern devices. 90 nm devices use oxide thickness of 2.2 nm or below with a supply voltage of 1.5. Oxide thickness in technology nodes like 45 nm may reach 1.2 – 1.5 nm which is the minimum to avoid excessive tunneling current. As discussed earlier the use of high-k dielectrics may push the problem further down the road. But as discussed earlier this brings additional problems with it too.

## CHAPTER 3

### PRELIMINARY STUDIES OF A QW MOSFET

A simplified conceptual structure of a modified MOSFET with the addition of a quantum well (QW) is studied in this chapter. One-dimensional Poisson equation is solved only in the QW region. By the assumption of some of the inner potentials, like the bulk potential, the solution of the Poisson equation is made simple. A simple program is written in Excel spread sheet incorporating the above Poisson solution, structural parameters and other variables. Parametric studies are carried out on the structural parameters and their effects on the device performance metrics are evaluated. A trade off study also has been carried out on the structural parameters. These are presented in this chapter.

#### 3.1 Introduction

Section 1.5.4 suggests some modifications that can be introduced to the Si CMOS structure so that the mobility of carriers (holes) can be improved. They are repeated here for convenience.

1. move carriers away from the inferior Si/SiO<sub>2</sub> interface
2. confine carriers in a QW
3. Grow the QW pseudomorphically so that the defect density at the QW walls is very low
4. Leave the QW undoped

To study the feasibility of such a structure and to make a rough estimate of its performance a simple analysis is performed on the structure shown in Figure 3-1. A QW is introduced deep in the Si substrate. A thin silicon cap layer is used to move the carrier transport layer (QW) away from the Si/SiO<sub>2</sub> interface. A p-delta layer is introduced below the QW. How a QW can be grown in Si MOSFET structure is not of concern here. The structure studied here is a mere conceptual one. The idea is to analyze the potential distribution in the structure. For this purpose a simple program is written using Excel spread sheet, incorporating the Poisson solution in the QW. A simple explanation of the operational steps of the program is given below. More elaborate details are given later in this chapter.

1.  $u(x)$  is the potential at the front (top) of the QW for any width of the QW  $x$ .  
This is an input parameter.  $u_{on}$  is the above potential in the ON state or with  $V_{gs}$  turned ON.. This is found by iteration.
2. The n-bulk potential  $u_{nB}$  at the back of the QW is an input parameter. From  $u_{nB}$  the thickness of the depletion layer and the number of positive ions in the bulk are calculated.
3. The number of negative ions in the p-delta layer,  $Q_d$ , is an input parameter
4. The field at the back of the QW,  $F_B$ , is calculated by taking the total ionic charge behind the QW. This includes the positive depletion charges in the bulk depletion region and  $Q_d$
5. Field at the front (top) of the QW,  $F$ , is calculated using the formula,

$$F = \text{Sign of } (u(x) - u_B) \sqrt{F_B^2 \frac{e^{u - u_B} - 1}{L_B^2}}$$

where  $L_B$  is the Debye length at the back of the QW and  $F_B$  is the field at the back of the QW.

6. The maximum thickness of the cap layer is calculated such that in the ON state the hole sheet density in the cap layer is a small fraction of the hole sheet density in the QW. Here this ratio is taken as a tenth. This cap layer thickness is a function of QW height,  $u_W$ , and the field at the front of the QW,  $F$ .
7. The gate voltage is an input parameter. Through iteration certain other parameters are changed to arrive at this input gate voltage. From the above the following performance metrics can be evaluated,
8. ON state hole sheet density in the QW is calculated
9. Hole sheet density ratio of ON and OFF states in the QW is calculated
10. Difference between QW highest potential and cap layer highest potential,  $u_{on} - u_{cap}$  is calculated

### 3.2 Solution of One Dimensional Poisson Equation

In order to arrive at the potential plots of the structure one-dimensional Poisson equation has been solved, starting from the back of the QW and proceeding towards the front (top) of the QW. It is to be noted here that this simple solution is possible because the bulk depletion potential drop,  $u_{nB}$  is assumed. But in the actual case this will not be available and the Poisson equation need to be solved starting from the bulk substrate and proceeding all the way to the gate. This detailed solution is given in Chapter 6 and

appendix C. The simplified solution of the Poisson equation in the QW is described below. The Poisson solution is carried out only in the QW. The rest of the potentials are either assumed or computed from the potentials at the boundary of the QW. The potentials are calculated only at the boundaries and they are joined by straight lines to get an approximate potential plot. So the quadratic variation of potential in the bulk and the exponential variation in the QW are not taken into account in the simple Excel program developed in this chapter. The program takes structural parameters, some potentials and other parameters. It can give instantaneous potential plots which allows visualization of the device functionality. It is also set up to run automatic 'goal seek' to find other parameters while keeping the desired value fixed. This is very useful in calculating variable parameters for a fixed gate voltage, for example.

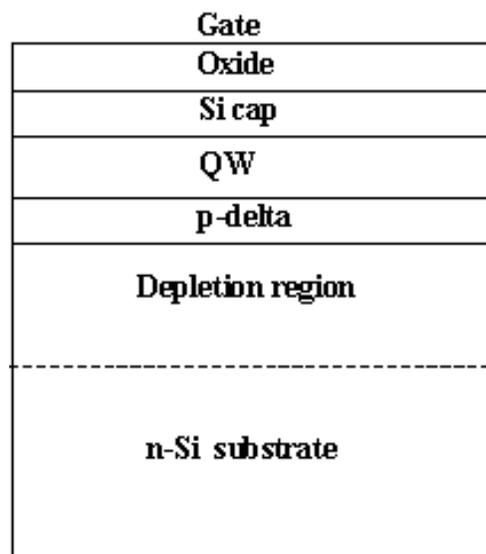


Figure 3-1. A modified structure of a MOSFET with the introduction of a QW. A p-delta layer is also introduced for additional reasons mentioned in the text.

The structure of the proposed QW device is sketched in Figure 3-1. Starting from the substrate the device consists of the substrate, p-delta layer, QW channel layer, Cap layer, Gate oxide and Gate. Valence band for this device is sketched in Figure 3-2 for the ON state. Poisson equation is solved starting from the bottom (left side) of the QW and moving towards the top (right side) of the QW. The work function of the gate is assumed to be the same as the substrate work function. First some parameters in Figure 3-2. need to be defined.

$n_B$  : Substrate doping density (/cm<sup>3</sup>) – an input quantity

$u_B$  : Potential at back of the QW (kT/q units)

$u_{nB}$ : potential drop in the bulk (kT/q units)

$u_W$  : QW height (kT/q units) – an input quantity

$F_B$  : Field at the back of the QW ((kT/q)/cm units)

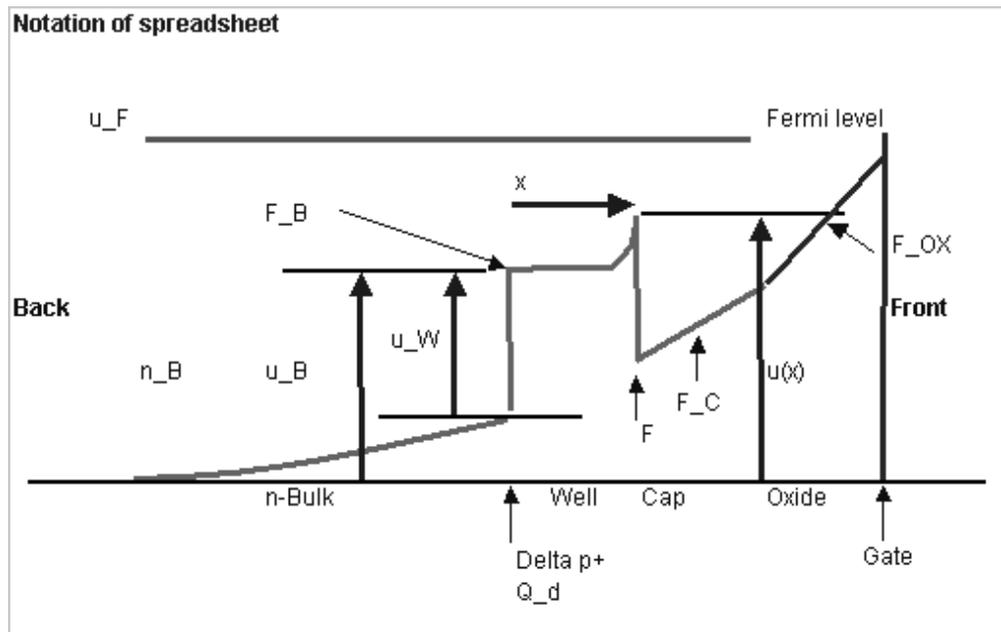


Figure 3-2. Valence band sketch of an ideal device in the ON state. The bulk potential,  $u_{nB}$ , ( $u_B - u_W$ ) is assumed.  $u_W$  and  $Q_d$  are input parameters.

$Q_d$  : p-delta layer doping density (/cm<sup>2</sup>) – an input quantity

$u(x)$  : QW potential at any distance  $x$  from the back of the QW (kT/q units)

$F$  : Field at the front of the QW ((kT/q)/cm units)

$F_C$  : Field in the cap layer ((kT/q)/cm units)

$F_{OX}$  : Field in the gate oxide

$u_F$  : Fermi level in the n-bulk (kT/q units)

$u_G$  : band gap of silicon (kT/q units)

$N_c$  : Conduction band density of states for silicon

The following are defined,

Fermi level in the n-bulk =  $u_G/2 - \ln(N_c/n_B)$

This is defined with respect to the n-bulk intrinsic potential away from the depletion region.

Debye length in the substrate  $L_{nB} = \sqrt{\frac{\epsilon_{Si} V_t}{qn_B}}$  cm

where  $V_t$  is kT/q,  $\epsilon_{Si}$  is dielectric constant of silicon,  $q$  is electron charge and  $n_B$  is the n-bulk doping density

The depletion width in the bulk is given by,

$$W_{nB} = L_{nB} \sqrt{2(u_{nB} - 1)} \text{ cm} \quad (3.1)$$

Now the two-dimensional charge number density in the bulk is given by,

$$Q_{nB} = n_B W_{nB} \text{ (/cm}^2\text{)} \quad (3.2)$$

Now the total number of charges in the bulk and p\_delta layer together is given by,

$$Q_I = Q_{nB} - Q_d \quad (3.3)$$

The field at the back of the QW,

$$F_B = \frac{qQ - I}{\epsilon_{SI} v_t} \quad (3.4)$$

The potential  $u_B$  and the field  $F_B$  at the back of the QW can be taken as boundary conditions to begin solving the Poisson equation.

The hole density in the quantum well is given by,

$$p(x) = N_V e^{-(u - F - u(x))} \quad (3.5)$$

within the QW the hole density at the back of the well becomes,

$$p_{-B} = N_V e^{-(u - F - u_{-B})} \quad (3.6)$$

Where  $u_{-B}$  is the value of  $u(x)$  inside the well at  $x=0+$ . The Poisson equation becomes,

$$\frac{d^2 u}{dx^2} = -\frac{Q}{\epsilon_{SI}} = -\frac{pq}{\epsilon_{SI}} = -C e^{u(x)} \quad (3.7)$$

where  $u$  and  $p$  are functions of  $x$ . and  $C$  is a constant.

$$C = \left( \frac{qN_V}{\epsilon_{SI}} \exp(-u - F) \right) \quad (3.8)$$

$$\frac{du}{dx} \frac{d^2 u}{dx^2} = \frac{1}{2} \frac{d}{dx} \left( \frac{du}{dx} \right)^2 \quad (3.9)$$

Integrating equation ( 3.7) using ( 3.9) we find,

$$\int_0^x dx \frac{1}{2} \frac{d}{dx} \left( \frac{du}{dx} \right)^2 = \int_0^x dx \frac{du}{dx} (-Ce^u) \quad (3.10)$$

$$\frac{1}{2} \left( \frac{du}{dx} \right)^2 \Big|_x - \frac{1}{2} \left( \frac{du}{dx} \right)^2 \Big|_0 = -C(e^u - e^{u-B}) \quad (3.11)$$

$$\frac{du}{dx} = \pm \sqrt{\left( \frac{du}{dx} \right)^2 \Big|_0 - 2C(e^u - e^{u-B})} \quad (3.12)$$

Taking dx to RHS and the root term to the left and integrating,

$$\int_0^x dx = \int_{u-B}^u \frac{du}{\sqrt{\left( \frac{du}{dx} \right)^2 \Big|_0 - 2C(e^u - e^{u-B})}} \quad (3.13)$$

where  $u_B$  is the value of  $u(x)$  on the well side of the n-bulk-to-well interface that is located at  $x=0+$ . The condition that  $x$  be real requires that the square root has a positive real argument. Otherwise the square root is imaginary. Also the plus sign in front of the square root is chosen to get a positive value for  $x$ .

If we set a variable  $y = ae^{-u/2}$  equation (3.14) can be re-written as,

$$\int_0^x dx = \frac{1}{\sqrt{2C}} \int_{u-B}^u \frac{du}{\sqrt{a^2 - e^u}} = -\frac{2}{a\sqrt{2C}} \int_{u-B}^u \frac{dy}{\sqrt{y^2 - 1}} \quad (3.14)$$

Here  $a$  is defined as  $a = \sqrt{(F - BL - B)^2 - 1}$ . The integral on the right side of the equation is tabulated [65]. Depending upon the value of  $a$  and  $C$  there are two solutions to  $x$ .

When  $a > 0$  and  $c > 0$  and applying the potential limits  $u_B$  to  $u$ ,

$$x = \frac{1}{a} (\ln(\sqrt{\exp(u - u_B) + a^2} - a) - \ln(\sqrt{\exp(u - u_B) + a^2} + a) + \ln(\sqrt{a^2 + 1} + a) - \ln(\sqrt{a^2 + 1} - a)) \quad (3.15)$$

When  $a > 0$  and  $c < 0$  and applying the limits,

$$x = \frac{2}{a} \left( a \tan \left( \frac{\sqrt{\exp(u - u_B) - a^2}}{a} \right) - a \tan \left( \frac{\sqrt{1 - a^2}}{a} \right) \right) \quad (3.16)$$

when  $u < u_B$  negative value of  $x$  is taken.

With  $u_B$  known above equations give  $x$  as a function of the potential at the front of the QW,  $u(x)$ . The above equations are programmed into an Excel spread sheet. If  $u(x)$  is assumed the width of the QW corresponding to the  $u(x)$  can be found. If on the other hand QW width,  $x$ , is given a 'goal seek' can be performed in Excel to find the potential at the front of the QW,  $u(x)$ . The Poisson solution has taken into account the hole density in the QW since hole density function  $p$ , as defined earlier, is used for charge density.

Next the field at the front of the QW is found from the field at the back of the QW, potential difference from front to back of the QW and Debye length.

$$\text{From equation (3.8)} \quad C = \left( \frac{qN_V}{\epsilon_{SI}} \exp(-u_F) \right) \quad (3.17)$$

$$u_F = u_G/2 - \ln \left( \frac{N_C}{n_B} \right)$$

$$e^{-u_F} = e^{-u_G/2 + \ln \left( \frac{N_C}{n_B} \right)} = e^{-u_G/2} e^{\ln \left( \frac{N_C}{n_B} \right)} = e^{-u_G/2} \frac{N_C}{n_B} \quad (3.18)$$

Debye length in the QW at its back in the  $p$ -delta is defined as,

$$L_B = \sqrt{\frac{\epsilon_{SI} V_t}{qp_B}} \quad (3.19)$$

$$C = \frac{qN_v \exp(-u_G/2)N_c}{\epsilon_{si} n_B} = \frac{qN_v N_c}{n_B \epsilon_{si}} \exp(-u_G/2) \quad (3.20)$$

$$= \frac{qn_i^2}{\epsilon_{si} n_B \exp(-u_G/2)} = \frac{kT}{q} \frac{1}{L_B^2} \frac{n_i^2}{p_B n_B \exp(-u_G/2)}$$

where  $u_G$  is the band gap and  $n_B$  is the n-bulk doping density.  $N_v$  and  $N_c$  are the density of states of valence and conduction bands respectively.

### 3.3 Field at the front of the QW

The field at the front of the QW can be calculated if the number of holes in the QW is known. The hole density at the front is given by,

$$p = N_v e^{-(u - F - u)} \quad (3.21)$$

The hole density at the back of the well at  $x=0+$  is given by equation ( 3.6). Dividing equation ( 3.21) by equation ( 3.6) and re-arranging terms and subtracting the hole density at the back ( $x = 0+$ ) the holes inside the QW is given by,

$$\Delta p = p_B (e^{u - u - B} - 1) \quad (3.22)$$

The field due to the holes in the well is given by,

$$F - \Delta p = \frac{Q}{\epsilon_{si}} = \frac{qp - B}{\epsilon_{si} V_t} (e^{u - u - B} - 1) = \frac{e^{u - u - B} - 1}{L_B^2} \quad (3.23)$$

in thermal units ( $kT/q$ ). The Debye length,  $L_B$ , inside and at the back of the well is given in equation ( 3.19). Now taking the vector sum of the field at the back and the field due to holes in the well given by equation ( 3.23) the field at the front of the well is given by,

$$F = \sqrt{F_B^2 + \frac{e^{u - u - B} - 1}{L_B^2}} \quad (3.24)$$

where  $F_B$  is the field at the back of the QW. The potentials used in the above equations are at the top of the well, at back and front. So the well height does not figure out in the field  $F$ .

Knowing the field at the front and the back of the QW carrier sheet density in the QW can be found using Coulomb's law and from this the carrier hole sheet density can be found.

$$Q = \epsilon_{SI} (F - F_B) \quad (3.25)$$

QW carrier sheet density is given by,

$$p_{sh} = \frac{Q}{q} \quad (3.26)$$

### 3.4 The Gate Voltage

The field at the top (right) of the QW is the same as the field in the cap layer since there are no charges in the cap layer. So  $F = F_C$  in Figure 3-2. The potential in front of the QW,  $u(x)$ , drops down by the well height,  $u_W$ , before entering the cap layer. This dropped down potential is continuous across the QW - cap layer interface. So the potential at the top (right) of the cap layer is,

$u_C = u(x) - u_W - F x_C$ , where  $x_C$  is the cap layer thickness. and  $u(x)$  is the potential at the front of the QW, which is a function of QW width,  $x_W$

The field in the oxide,

$$F_{OX} = \frac{\epsilon_{SI}}{\epsilon_{OX}} F \quad (3.27)$$

Again the potential is continuous across the cap layer - gate oxide interface.

So gate voltage  $V_{gs} = u_C - F_{OX} x_{OX}$ , where  $x_{OX}$  is gate oxide thickness. Using the QW Poisson solution and the above expressions a simple program is written in Excel work sheet to plot the potential (valence band) along the device from left to the right..

### 3.5 Simulation Results

With the above set up for a quick analysis some of the parameters like hole sheet density in the QW, the hole sheet density ON/OFF ratio in the QW and the sub-threshold slope are calculated. The input parameters of interest are listed below for clarity.

1.  $u_W$ : QW height in  $kT/q$
2.  $x_W$  : QW width in nano meters
3.  $n_B$  : substrate doping density/ $cm^3$
4.  $Q_d$  : p-delta doping density/ $cm^2$

The above four parameters are fixed for a particular fabricated device. They are however varied one at a time in the simulation, keeping the other three fixed, to study the parametric influence on device characteristics. The four variables used to re-establish the other three fixed parameters are given below. This is needed to get the original values for the three parameters when the fourth one is varied. So the adjusting potentials are,

1.  $u_{on}$ : The potential at the top (right) of the QW in the ON state
2.  $u_{off}$ : The potential at the top (right) of the QW in the OFF state
3.  $u_{nB_{on}}$  : The bulk potential in the ON state or high gate voltage.
4.  $u_{nB_{off}}$  : The bulk potential in the off state or zero gate voltage.

$u_{nB\_off}$  is not zero at zero gate voltage because ions in the p-delta layer will cause some depletion in the bulk. This is shown in Figure 3-5.

### 3.5.1 Goal Seek to Set QW Width and Gate Voltage

For any step in any of the first four input parameters some of the four potential variables have to be adjusted to keep the other three input parameters fixed. For example assume that the effect of QW height  $u\_W$  is to be studied.  $u\_W$  is varied by a step. This will change the QW width  $x\_W$  and the gate voltage  $V_{gs}$ . So a ‘goal seek’ is used in the Excel spread sheet to restore  $x\_W$  to its original fixed value by changing the potential drop in the substrate for the ON state,  $u_{nB\_on}$ . This restores the QW width  $x\_W$  for the ON state. The ON state  $V_{gs}$  is restored to the given value (1.5 V) through ‘goal seek’ using the potential at the front of the QW for the ON state,  $u_{on}$ , as a parameter.

In order to restore  $x\_W$  for the OFF state a ‘goal seek’ is again applied using the potential at the front of the QW for the OFF state,  $u_{off}$ , as a parameter. To restore  $V_{gs}$  to zero for the OFF state ‘goal seek’ is used with the substrate potential drop,  $u_{nB\_off}$  for the OFF state as a parameter. If  $V_{gs}$  is restored through ‘goal seek’ the QW width  $x\_W$  will change. This does not make sense since it is a fixed value for a device and cannot be changed either in the ON or OFF case. So the goal seek has to be applied again and again until  $V_{gs}$  and  $x\_W$  settle down to their fixed values.

Even though any of the parameters can be used in ‘goal seek’ the above selection of potentials were found to be fast in converging without problems. The fixed parameters and the ‘goal seek’ variables are given below in Table 3.1.

### 3.5.2 Excel Program and Device Details

The reference device structure details are given below. Following structural parameters are kept fixed since these cannot be varied in an actual device. But, in simulations, individually they can be used as a running parameter for parametric study while keeping the other three fixed.

Table 3-1. Fixed parameters and the corresponding potential variables used to restore the fixed parameters in Excel using ‘goal seek’

Fixed parameter	Goal seek variables
QW width, $x_W$ , for ON state	Substrate potential, $u_{nB}$ , for ON state
$V_{gs}$ (1.5 V) for ON state	QW potential at front, $u_{on}$ , for ON state
QW width, $x_W$ , for OFF state	QW potential at front, $u_{off}$ , for OFF state
$V_{gs}$ (0 V) for OFF state	Substrate potential, $u_{nB\_off}$ for OFF state

The input device parameters for the simulations are chosen to be,

1. QW height  $u_W = 10 \text{ kT}/q$
2. QW width  $x_W = 5 \text{ nm}$
3. p-delta doping,  $Q_d = 1.2 \times 10^{12}/\text{cm}^2$
4. Substrate doping  $n_B = 5 \times 10^{18}/\text{cm}^3$

Using ‘goal seek’ approach discussed in section 3.4.1 any value of QW width,  $x_W$ , and any value of  $V_{gs}$  can be realized. If  $V_{gs}$  is given and want to find any of the four structural parameters listed in 1 to 4 under section 3.4 it can be done through ‘goal seek’. In this case  $V_{gs}$  can be taken as the target cell and any one of the four parameters (1 to 4 in the above list) of interest can be varied to arrive at a structural value of that parameter for that particular  $V_{gs}$ .

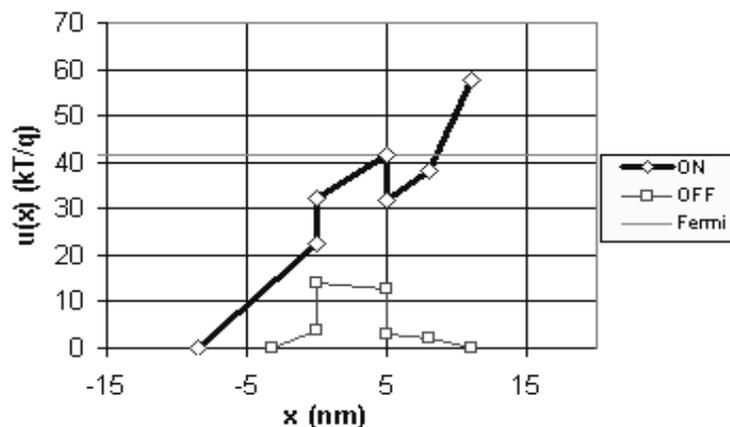


Figure 3-3. ON and OFF potentials with a substrate doping of  $5 \times 10^{18}/\text{cm}^3$ .  
 $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 1.2 \times 10^{12}/\text{cm}^2$

Figure 3-3 shows the potential plots for the ON and the OFF states for a device with structural parameters given in 1 to 4 at the start of this sub-section. The display panel for this graph is given in Figure 3-4. It should be noted that the potentials are approximate since the quadratic behavior in the bulk and the exponential behavior in the QW are not taken into account. The potentials are calculated only at the boundaries and these values

Bulk doping	<b>n_B3</b>	<b>5.00E+18</b> ← Input		
Number of ions in delta layer/cm <sup>2</sup>	Q_d_set3	1.20E+12		
Potential drop in bulk n-layer (kT/q)	u_nB_on3	22.43341778		
Potential drop in bulk n-layer (kT/q)	<b>u_nB_off3</b>	<b>4.021332116</b> ← Input		
Well height (kT/q)	<b>u_W_set3</b>	<b>10</b> ← Input		
			well width (nm)	gate V
potential at front of well (on)	u_on3	41.68944781	4.99982	1.50E+00
potential at front of well (off)	u_off3	12.84000576	5.060863853	1.01E-03

S,on/off

p/cm<sup>2</sup>

ON 3.11E+11      S 1.33E+02mV/dec

OFF 1.68E+00ON/OFF 1.85E+11

Figure 3-4. Display panel for Figure 3-3.

are simply joined by straight lines. This is true for all simulation results presented. It can be seen that the QW width is iterated to 5 nm for the ON and the OFF state. The gate voltage is iterated to 1.5 V for the ON state and a near zero value for the OFF state. In Figure 3-3 the upward tilt of the QW shows that the positive ions in the bulk depletion region are stronger than the negative ions in the p-delta layer for the ON state. This causes a positive field in the QW. In the OFF state the downward bending of the QW indicates that the negative ions in the p-delta layer are stronger than the positive ions in the bulk depletion region. This is because the bulk depletion width is reduced which reduces the bulk positive ions whereas the p-delta ions stay the same since they cannot change. The device has a ON/OFF hole density ratio of  $1.85 \times 10^{11}$  and a  $S$  of 133. In comparison a long channel MOSFET has a ON/OFF ratio of around  $10^7$ .

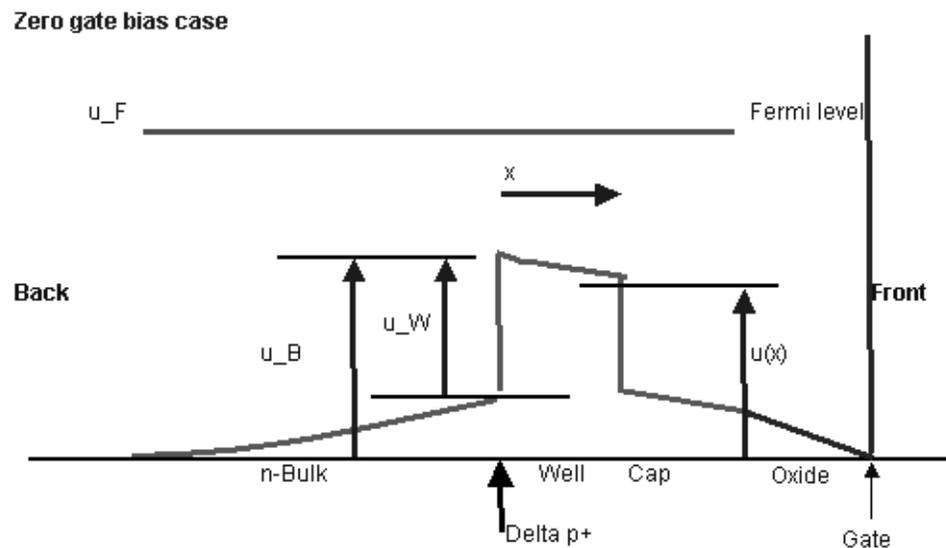


Figure 3-5. Valence band diagram at zero gate bias. Notice that the bulk potential drop is not zero due to the negative ions in the p-delta layer. The QW tilts downwards indicating that the ions in the p-delta layer is stronger than the positive ions in the substrate.

Before starting with parametric simulation it is instructive to study the test structure given at the beginning of sub-section 3.4.2. For this structure the dependence of hole sheet density, ON/OFF hole sheet density ratio and the sub-threshold parameter  $S$  upon  $x_W$ ,  $u_W$  and  $Q_d$  are simulated.

### 3.5.3 ON State Hole Sheet Density in the QW

The QW hole sheet density dependence on QW width,  $x_W$ , is given in Figure 3-6. There is an increase in hole density with increase in  $x_W$ . This is possible since there is more region or more states of the QW

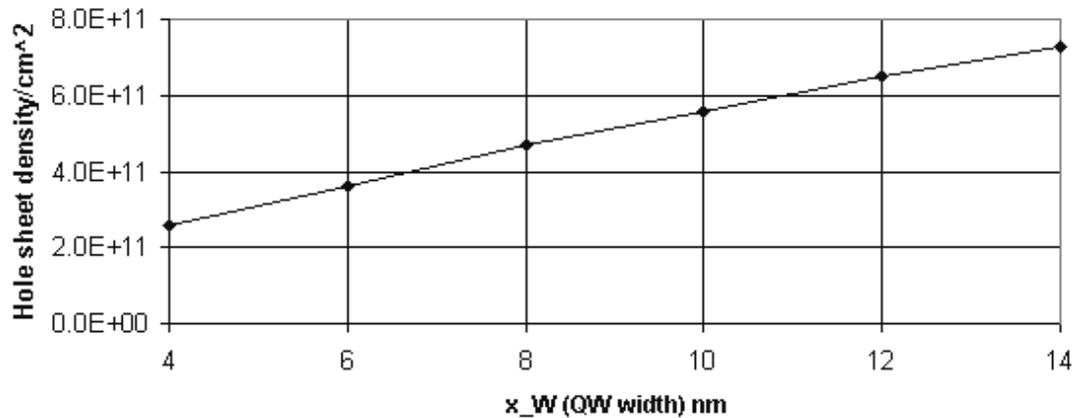


Figure 3-6. Hole sheet density dependence on QW width.  $n_B = 5\text{e}18/\text{cm}^3$ ,  $u_W = 10 \text{ kT}/q$ ,  $Q_d = 1.2\text{e}12/\text{cm}^2$ ,  $V_{gs} = -1.5 \text{ V}$

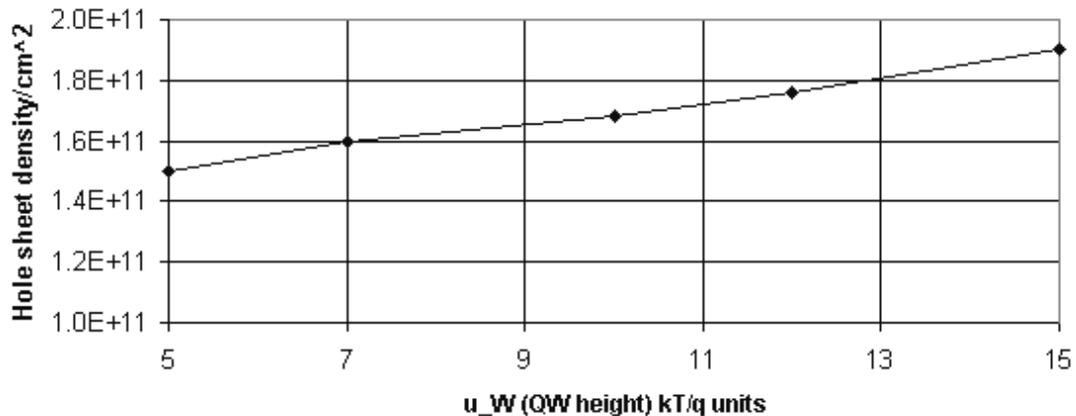


Figure 3-7. Hole sheet density dependence on QW height.  $n_B = 5e18/cm^3$ ,  $Q_d = 1.2e12/cm^2$ ,  $x_W = 5$  nm,  $V_{gs} = -1.5$  V

available for population in the ON state. Figure 3-7 gives the hole density dependence on QW height. The rate of increase of hole density is higher with increase in QW width  $x_W$  than with QW height. This may be due to more regions of the top of the QW being closer to the Fermi level in the ON state in the case of a wider QW. Hence more states may be available for populating.

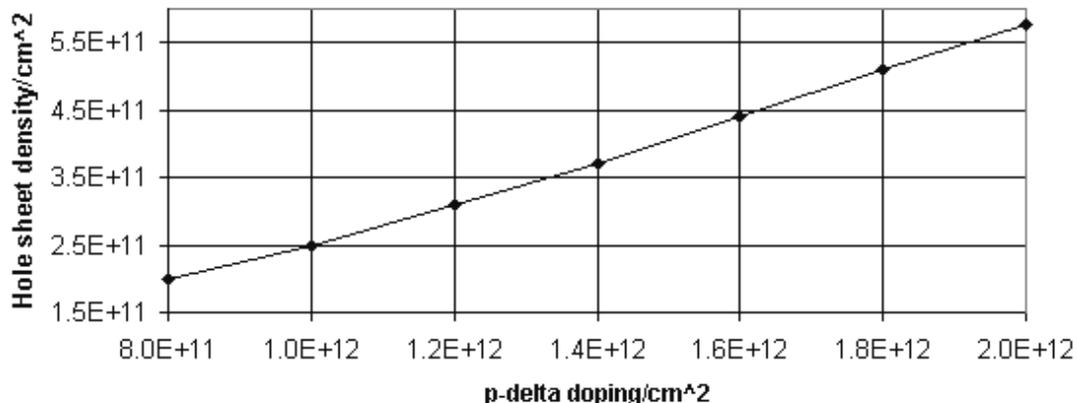


Figure 3-8. Hole sheet density dependence on p-delta doping.  $n_B = 5e18/cm^3$ ,  $x_W = 5$  nm,  $u_W = 10$  kT/q,  $V_{gs} = -1.5$  V

Dependence of hole sheet density on p-delta doping is given in Figure 3-8. There is a steady increase in hole sheet density with increase in p-delta doping. A higher p-delta doping will move more of the QW towards the Fermi level or try to make the QW more or less parallel to the Fermi level, thus populating with more holes.

### 3.5.4 OFF State Hole Sheet Density in the QW

It is desirable to keep the top of the QW as low as possible in the OFF state in order to reduce the QW carrier population in the OFF state and hence  $I_{\text{OFF}}$ . Figure 3-9 and Figure 3-10 gives the ON and the OFF state potential plots for a QW height of  $4 \text{ kT}/q$  and  $10 \text{ kT}/q$  respectively. In both figures the lower curve shows the valence band in the OFF state. It has to be mentioned here that the potential plots shown here are approximate since the potentials at the boundaries only are calculated. These boundary potentials are

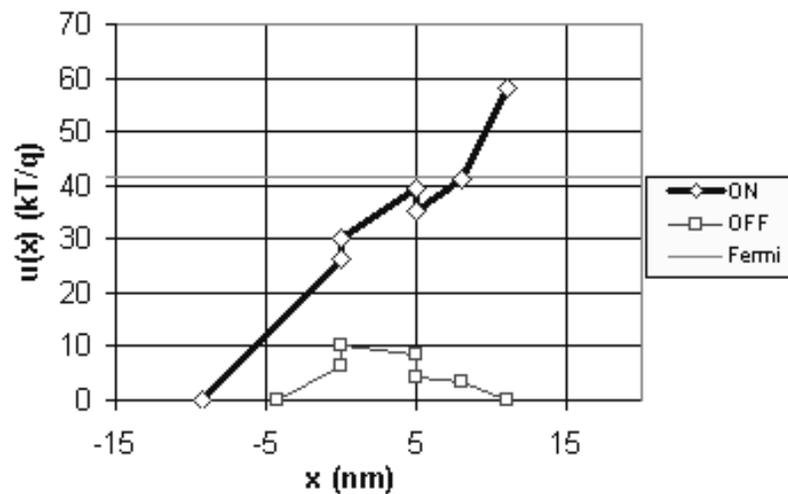


Figure 3-9. Valence band plots for the ON and OFF states with a QW height of  $4 \text{ kT}/q$ . The marker just to the right of the QW is the top of the cap layer. This point is lying above the QW and forms a parasitic surface channel.

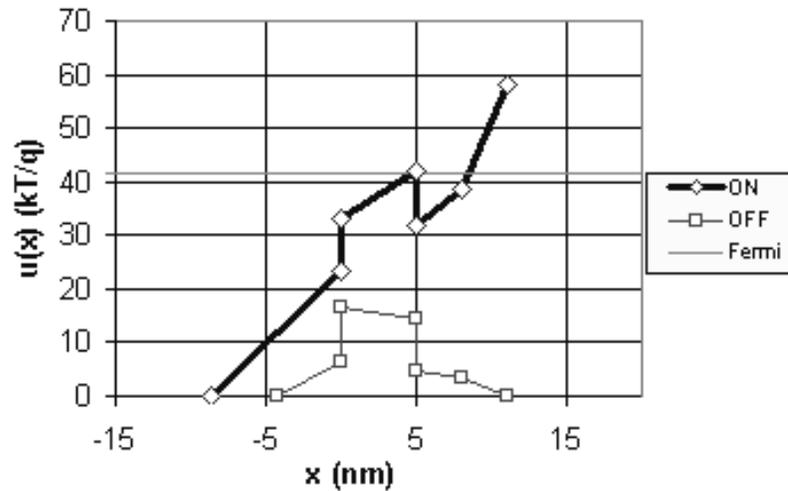


Figure 3-10. Valence band plots for ON and OFF states with a QW height of  $10 \text{ kT}/q$ . It can be seen that the top of cap layer (marker to the right of the QW) is lying below the top of the QW. Now the QW has much higher carrier population compared to the cap layer. This figure is the same as Figure 3-3.

then joined by straight lines. So the quadratic behavior of potential in the bulk and the exponential behavior of potential in the QW are missing. It can be seen that the top of the QW in the OFF state lies at a higher value in Figure 3-10 compared to Figure 3-9. This means that a higher QW height is not desirable since  $I_{\text{off}}$  increases with increase in QW height.

### 3.5.5 Hole Density (ON/OFF) Ratio in the QW

The hole sheet density ratios between ON and OFF states are studied in this section. The hole sheet density ratio is defined as the ratio of QW hole sheet density in the ON state to the QW hole sheet density in the OFF state. Figure 3-11 shows the sketch of the valence band in the ON state. Figure 3-12 shows the effect of QW width  $x_W$  on the ON/OFF ratio of the QW hole sheet density. The dependence of the ON/OFF ratio on

$x_W$  seems weak. This may not be the case for other combination of parameters. If the bulk and p-delta ions are equal leading to a flat band QW most of the QW will stay close to the Fermi level in the ON state. This can cause a linear increase of hole sheet density, with increasing  $x_W$ , in the ON state. But in the OFF case p-delta ions are stronger and the QW tilts downwards very much reducing the OFF state holes in the QW. The OFF state holes in the QW becomes independent of  $x_W$  because only few holes get populated at the back end of the QW. In such a case there will be an almost linear rise in ON/OFF ratio. This seems to be the case in

Figure 3-12. There is a linear rise in the ratio even though the rise with increasing  $x_W$  is weak.

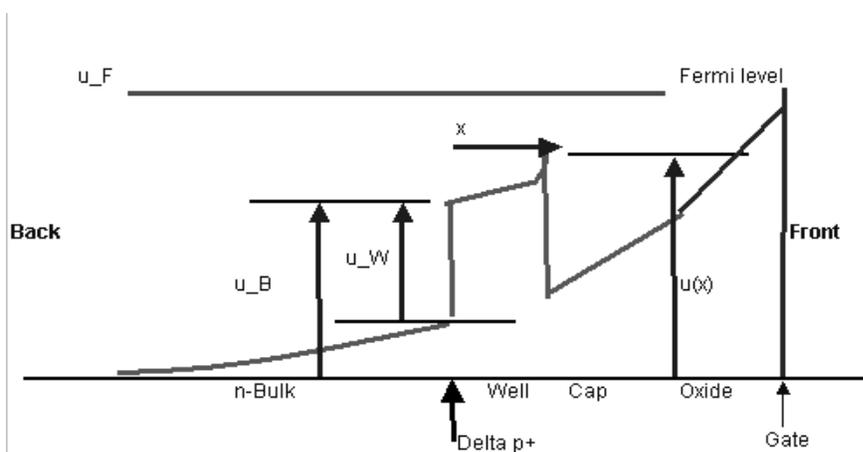


Figure 3-11. Valence band sketch of the device in the ON state.

Figure 3-13 shows the effect of increasing the QW height  $u_W$  upon the ON/OFF hole sheet density ratio. Increase in  $u_W$  will increase the carrier density in the ON state. But in the OFF state this will place the top of the QW closer to the Fermi level thus increasing

the OFF state carrier density. Because of this there is a considerable drop in the hole sheet density ratio. This drop is seen in the figure. The effect seems to be strong giving nearly three orders of magnitude drop for the range of  $x_W$  considered.

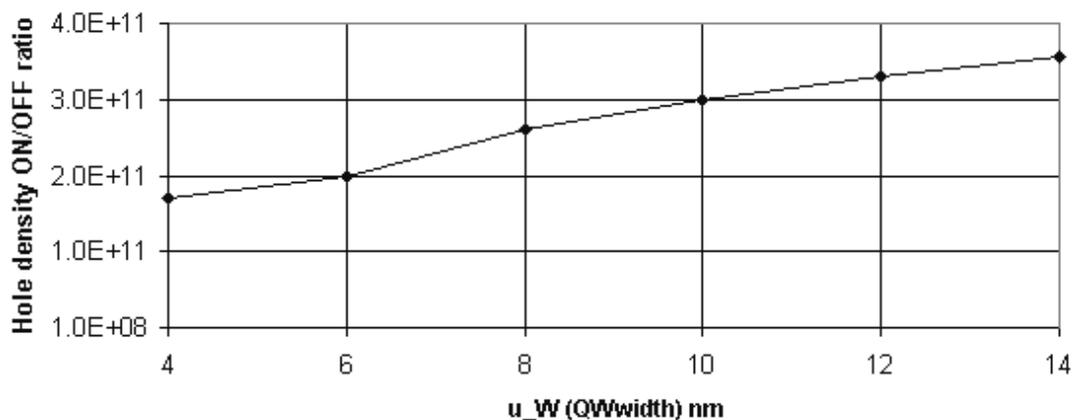


Figure 3-12. Hole sheet density ON/OFF ratio dependence on QW width.  $n_B = 5e18/cm^3$ ,  $Q_d = 1.2e12/cm^2$ ,  $u_W = 10$  kT/q

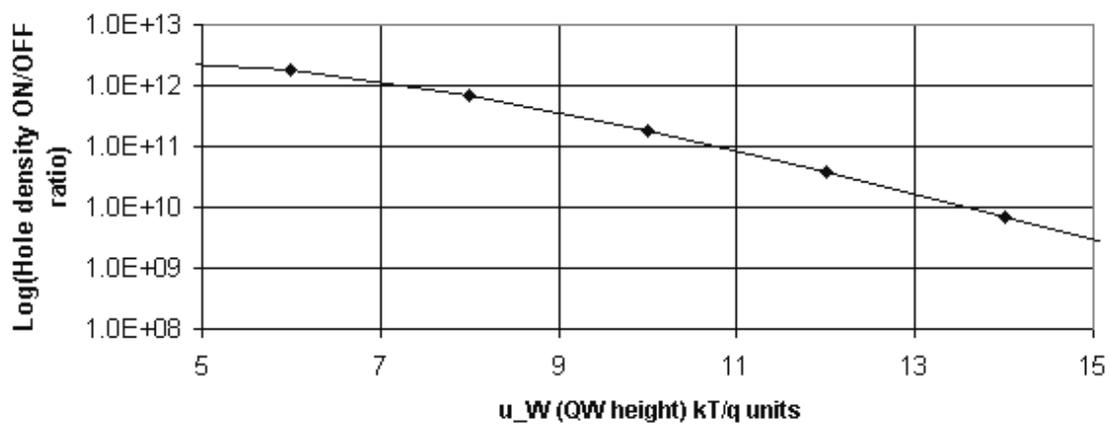


Figure 3-13. Hole density ON/OFF ratio vs QW height.  $n_B = 5e18/cm^3$ ,  $u_W = 10$  kT/q,  $Q_d = 1.2e12/cm^2$

Figure 3-14 gives the dependence of ON/OFF ratio on the p-delta doping density. As the p-delta doping is increased hole sheet density increases slightly as seen in Figure 3-8.

But the sub-threshold parameter,  $S$ , also increases very much as seen in Figure 3-18 This indicates more OFF state hole density. Hence the ON/OFF ratio decreases very much. In fact the curve in Figure 3-14 has nearly the same shape as the  $S$ . Another explanation could be that the substrate positive ion density is much higher in the OFF

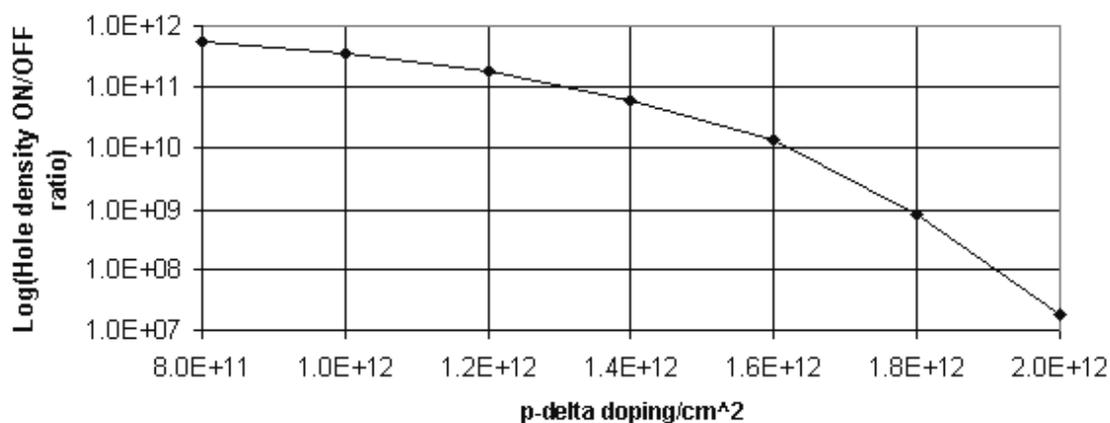


Figure 3-14. Hole sheet density ON/OFF ratio vs p-delta doping.  
 $n_B = 5e18/cm^3$ ,  $u_W = 10 \text{ kT}/q$ ,  $x_W = 5 \text{ nm}$   
 state due to the higher p-delta doping. This will raise the back of the QW upwards

increasing the OFF state hole population.

### 3.5.6 Parasitic Channel

In this sub-section the formation of a parasitic channel in the cap layer is studied.

Figure 3-9 shows the valence band plots for the ON and OFF states for a small QW height of  $4 \text{ kT}/q$ . The marker to the right of the QW is the top of the cap layer. It can be seen clearly that the top of the cap layer is lying above the top of the QW. So the top of the cap layer will confine much more carriers (holes) than the QW. Since these carriers at the top of the cap layer are lying close to the Si/SiO<sub>2</sub> interface the mobility degrades to the value of a conventional MOSFET. So the advantage of higher mobility in the

undoped QW is lost. The device behaves as a conventional MOSFET. To alleviate this degradation the QW height has to be kept higher. As can be guessed from Figure 3-9 a thicker cap layer can also worsen the situation even for higher QW heights. So the cap layer thickness should be kept low.

Figure 3-10 shows the valence bands for the ON and the OFF states with a QW height of  $10 \text{ kT/q}$ . Here it can be noted that the top of the QW is lying above the top of the cap layer. Hence the QW will confine much more carriers (holes) than the cap layer. Since the holes are mostly in the QW they experience the higher mobility of the undoped QW channel. From the above discussion it can be seen that a higher QW and a thin cap layer are desirable for high performance devices.

### 3.5.7 Sub-threshold Slope

In this sub-section parametric dependence of sub-threshold slope has been studied by changing the parameters one at a time. Sub-threshold parameter,  $S$ , is a measure of the ON/OFF ratio of hole sheet density in the QW.  $S$  is the inverse of the sub-threshold slope. So a smaller  $S$  is desirable. In the case of well designed MOSFET  $S$  ranges from 80 to 100.

The sub-threshold parameter is calculated as follows. The sub-threshold parameter  $S$  for a MOSFET is defined as,

$$S = \left( \frac{d(\text{Log}(I_{DS}))}{dV_{GS}} \right)^{-1} \quad (3.28)$$

$I_{DS}$  is proportional to hole sheet density. Hence from Figure 3-15 the ratio in the parenthesis, which is the slope of the curve, can be approximately written as,

$$Slope = \frac{Log(p_{ON}) - Log(p_{OFF})}{V_{GS}} = \frac{Log(p_{ON} / p_{OFF})}{V_{GS}} \quad (3.29)$$

In the simulation  $I_{ON}$  and  $I_{OFF}$  are not computed. Instead hole sheet densities are calculated. Since  $I_{ON}$  and  $I_{OFF}$  are proportional to hole sheet densities in the ON and OFF states respectively  $S$  can be written as,

$$S = \frac{1.5 \times 1000 \text{ mV}}{Log(p_{ON} / p_{OFF})} \text{ mV / dec} \quad (3.30)$$

where  $V_{gs}$  is assumed to be equal to 1.5 V. This approximate equation for  $S$  is used to calculate the sub-threshold parameter  $S$ .

Dependence of sub-threshold parameter  $S$  on QW width is given in Figure 3-16. There is a very slight decrease in  $S$  with increase in  $x_{W}$ . As the QW width increases more of the QW region will be closer to the Fermi level and this increases the hole density, in the ON state. But in the OFF state due to the downward tilt of the QW the increase in hole density will be less. So a slight reduction in  $S$  is possible. Sub-threshold parameter dependence on QW height is given in Figure 3-17. There is a large increase in sub-threshold parameter with increase in QW height. This can again be inferred from Figure 3-5.

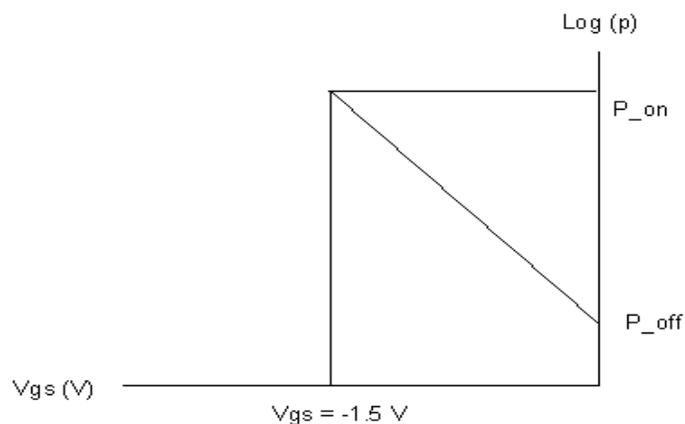


Figure 3-15. Log(hole sheet density) vs  $V_{gs}$ , sub-threshold plot for the p-MOSFET device. Since hole sheet density is proportional to the drain current the ratio of hole sheet densities in the ON and OFF states can be used instead of  $I_{ON}/I_{OFF}$

A higher QW height will move the top of the well towards the Fermi level thus populating the QW with more holes in the OFF state. Even though a higher QW will accumulate more holes in the ON state the same will lead to more OFF state carriers.

This will be discussed further towards the end of the section and trade offs will be discussed. Figure 3-18 shows the sub-threshold parameter dependence on p-delta doping.

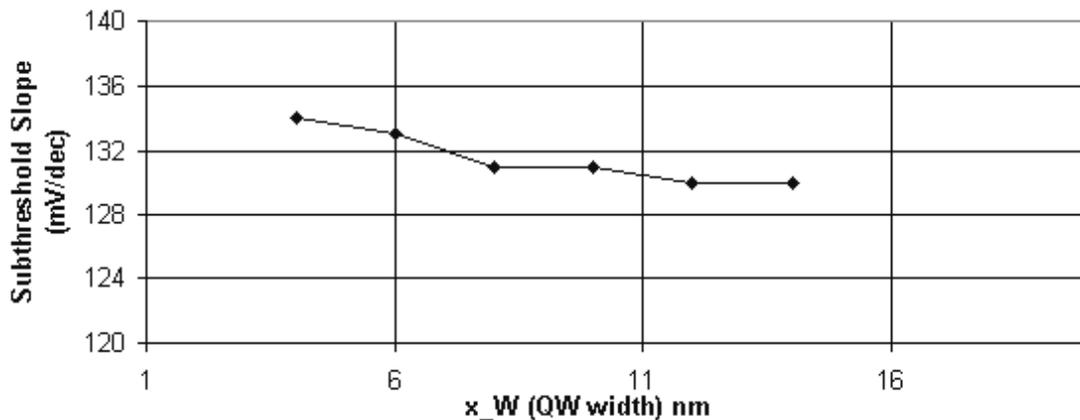


Figure 3-16. Sub-threshold slope dependence on QW width,  $x_W.u_W = 10 \text{ kT}/q$ ,  $n_B = 5e18/\text{cm}^3$ ,  $Q_d = 1.2e12/\text{cm}^2$

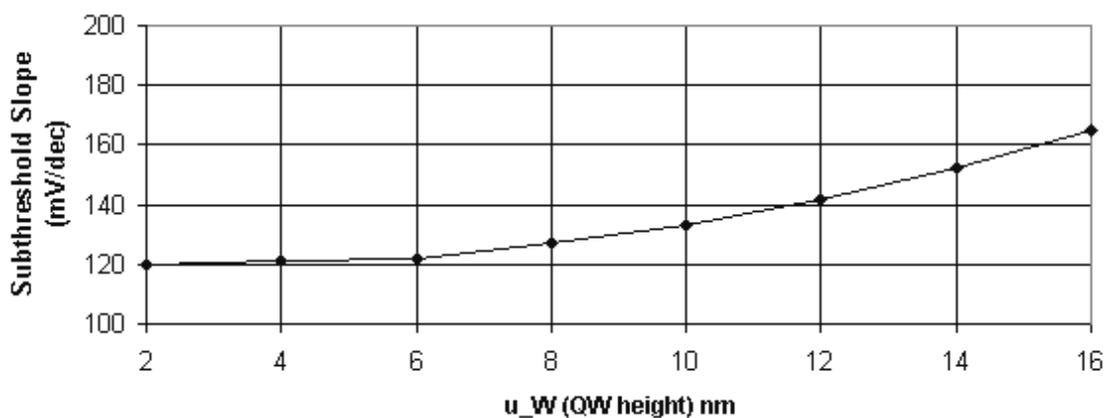


Figure 3-17. Sub-threshold parameter,  $S$ , dependence on QW height,  $u_W$ .  
 $n_B = 5e18/cm^3$ ,  $Q_d = 1.2e12/cm^2$ ,  $x_W = 5$  nm

There is a large increase in the value of  $S$  with p-delta doping. In the OFF state there could be more positive ions in the substrate to compensate the large number of negative ions in the strong p-delta layer. This will increase the substrate potential drop raising the QW towards the Fermi level in the OFF state increasing the carrier density.

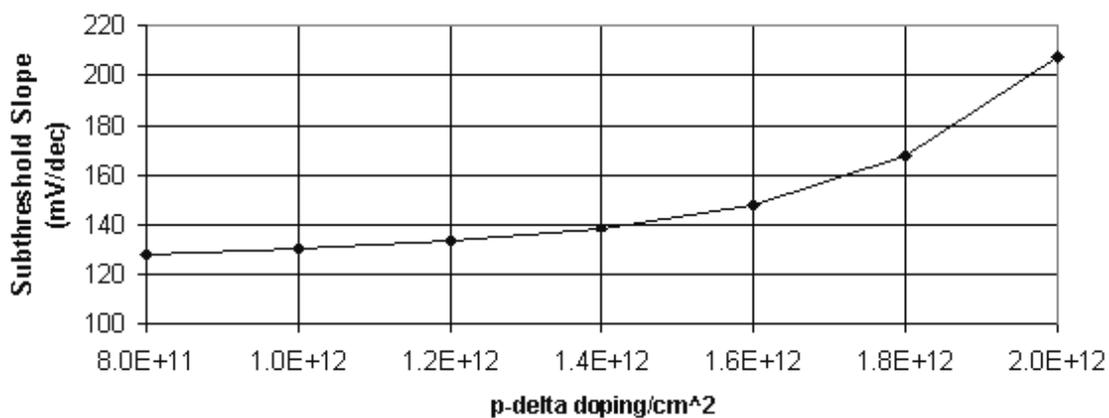


Figure 3-18. Sub-threshold slope dependence on p-delta doping.  $n_B = 5e18/cm^3$ ,  
 $u_W = 10$  kT/q,  $x_W = 5$  nm

### 3.6 Trade Off Studies

In this section effects of structural parameters upon performance metrics and behavior of the device are studied. The limits or ranges of these parameters, trends on device performance, trade offs and their influence on the device performance like the QW hole sheet density, the QW hole sheet density ON/OFF ratio, the sub-threshold parameter and potentials are studied in particular. The structural parameters studied are listed below. Unlike normal simulations where a standard procedure is applied to all simulations varied and diverse techniques are applied here. Since simulations, observations and varied methods are used to evaluate the performance of the device the metrics are not listed. In the order presented here the four structural parameters studied are,

1. p-delta doping density,  $Q_d$ , /cm<sup>2</sup>
2. QW height,  $u_W$ , kT/q
3. QW width,  $x_W$ , nm
4. Substrate doping,  $n_B$ , /cm<sup>3</sup>

#### 3.6.1 Trade off: p-delta Doping Density

##### 3.6.1.1 The ON State

In this sub-section a comparative study is carried out for three different ON states. The substrate doping density, QW width and QW height are kept fixed in all three ON states. For the three ON states the p-delta negative ions is chosen to be (1) less than the bulk depletion positive ions, (2) equal to the bulk positive ions and (3) greater than the bulk depletion ions. The potentials for these three choices are conceptually sketched in

Figure 3-19, Figure 3-20 and Figure 3-21 respectively. The gate voltage is assumed to be the same in all three ON states. The discussion given here is based on sketches given in this section. They are not from simulations. However a few simulation plots are included to make a reality check and to support the arguments made using sketches. Figure 3-22 is one such sets of plots generated using the solution of the Poisson equation and Excel spread sheet simulation program discussed in sub-section 3.2.

The field at the back of the QW is given by  $\epsilon\xi = Q$ .  $\epsilon$  is the dielectric constant of silicon,  $\xi$  is the field at the back of the QW and  $Q$  is the net charge to the left of the QW. So the field is positive if  $Q$  is positive and negative if  $Q$  is negative.

Figure 3-19 shows the valence band edge when the p-delta negative ions are lower than the bulk depletion positive ions. So the field at the back of the QW is positive. So the front side of the QW is near the Fermi level and gets highly populated. Since the p-delta ions are not very strong bulk depletion due to p-delta ions will be less and  $u_B$  will reduce moving the QW downwards.

Figure 3-20 has the same number of ions in the p-delta layer and the bulk depletion. So the QW is flat near the back end of the well. In the ON state there are large number of holes in the well. This lifts the valence band higher and higher towards the right edge of the QW. This increases the hole density towards the right edge of the QW. This appears to give a higher hole sheet density. But note that the bulk depletion will be less since the p-delta ions are not very strong. This will reduce  $u_B$  and move the QW downwards. The bulk potential may change a little compared to Figure 3-19 This is shown in Figure 3-22. This may modify the hole density in the QW a little.

Figure 3-21 has more p-delta negative ions than the bulk depletion positive ions. So the field at the back of the QW is negative and the band bends downwards. Now the backside of the QW gets populated more. This backside population will be higher than the backside population in Figure 3-19. Since p-delta ions are stronger the bulk depletion due to p-delta ions will be more,  $u_B$  will be more and the QW will be moved upwards. This will increase the holes at the left edge of the QW further. This is because the bulk potential drop and the bulk positive ions will be higher than the same in Figure 3-19 for the same gate. So this configuration may give the highest hole sheet density.

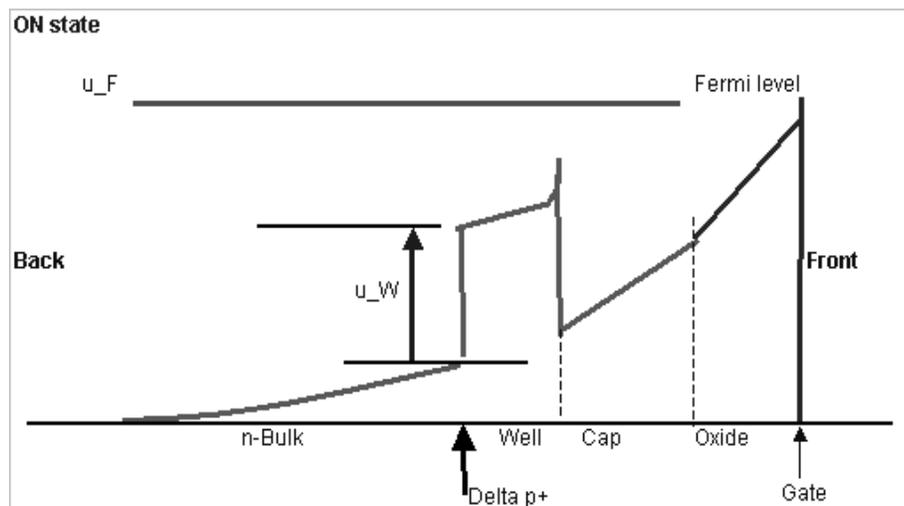


Figure 3-19. ON state sketch. Negative p-delta ions are less than the positive bulk depletion ions. The field in the QW is positive since the positive ions is stronger.

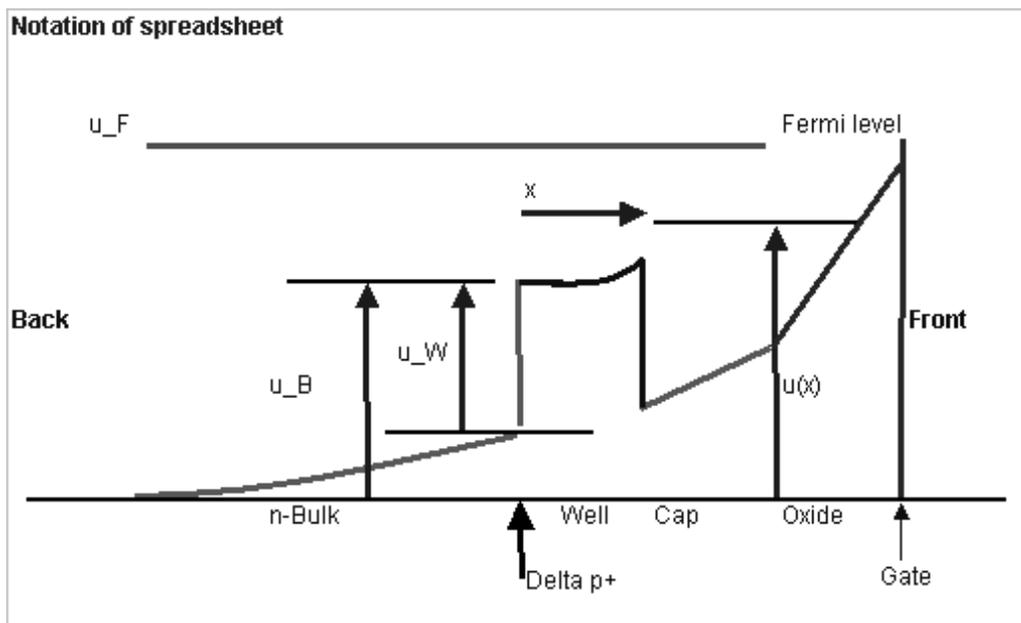


Figure 3-20. ON state sketch .  $p$ -delta negative ions is equal to the bulk depletion positive ions. Since the ions cancel each other the field at the back of the QW is zero and the QW is flat at the left side.

voltage. So this case may have the maximum hole sheet density. The higher bulk depletion happens to cancel the high  $p$ -delta ions. This is confirmed by simulations using the Excel program described under sub-section 3.2. Figure 3-22 shows the simulation results. The parasitic surface channel was thought to get stronger with increase in  $p$ -delta doping. Figure 3-22 show that the difference between the QW potential and the cap potential ( $u_o - u_{cap}$ ) remains the same. But both these potentials are increasing with increasing  $p$ -delta doping. Since the carrier to potential relation is exponential this could cause a parasitic channel degradation.

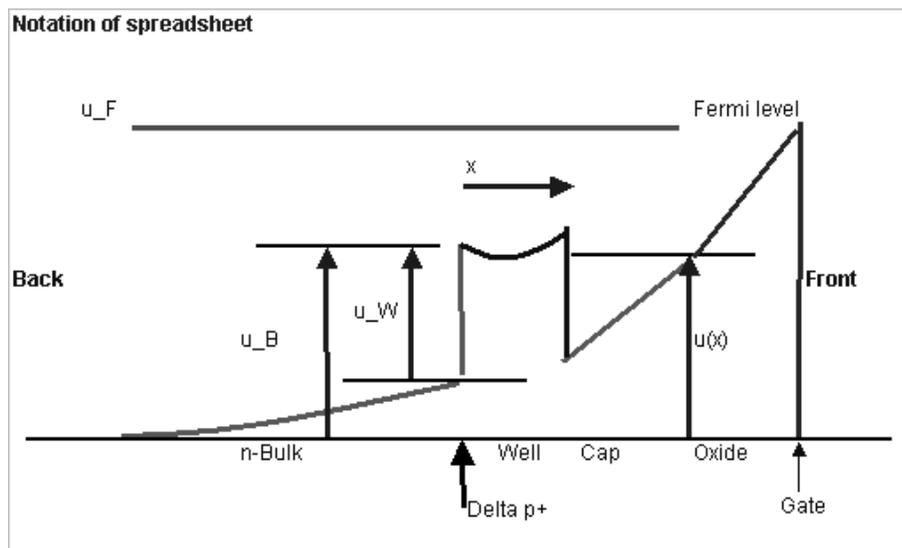


Figure 3-21. ON state sketch. p-delta negative ions are more than the bulk depletion positive ions. So the field in the QW is negative and the band bends downwards at the back of the QW.

### 3.6.1.2 Conclusion for the ON state

From the above discussions Figure 3-21 seem to give the maximum hole population in the ON state. So it is better to have higher number of ions in the p-delta than the bulk. This lifts the whole QW towards the Fermi level providing more states for holes to occupy. This increases the hole sheet density. Figure 3-22 shows the bulk potential,  $u_{nB\_on}$ , cap potential,  $u_{cap}$ , and the QW potential in the ON state,  $u_{on}$ . There is an increase in bulk potential,  $u_{nB\_on}$ , with an increase in p-delta doping. This will lift the QW higher towards the Fermi level. This shows that higher p-delta doping is better to have higher hole density. In this figure the rise in cap layer potential with increased p-delta doping is not significant. Also the difference between QW potential at the front and cap layer potential is nearly the same for the p-delta doping range simulated. This may suggest that the p-delta doping hardly affects the ratio of hole sheet density in the QW to

the hole sheet density in the cap layer. But there is slight steady increase in both potentials. The carrier and potential relation being exponential there may be some degradation.

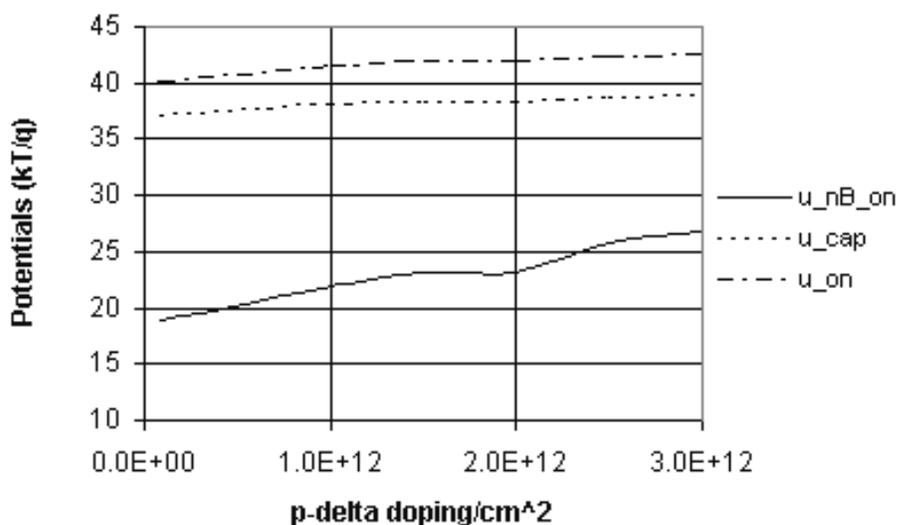


Figure 3-22. Simulated potentials, using the Excel program, in the ON state for a device with  $u_W = 10$  kT/q,  $x_W = 5$  nm,  $n_B = 5e18/cm^3$ .  $u_{nB}$  is the bulk depletion potential,  $u_{cap}$  is the potential at the top of the cap layer,  $u_{on}$  is the potential at the top of the QW in the ON state or  $V_{gs} = 1.5$  V

### 3.6.1.3 Off state

Next, we consider the effect of the above three choices of the ON state upon the OFF state. Figure 3-23 has the band structure when the p-delta ions are less than the depletion ions for the ON state. In the OFF state  $V_{gs} = 0$  V and the bulk potential and the bulk positive ions will be less than the ON state. But the p-delta ions cannot change. So the QW will tilt more downwards than the Figure 3-19.

Figure 3-23 had less p-delta ions than the bulk depletion in the ON state. Now in the OFF state with  $V_{gs} = 0$  V the bulk depletion will reduce which can cause p-delta ions to



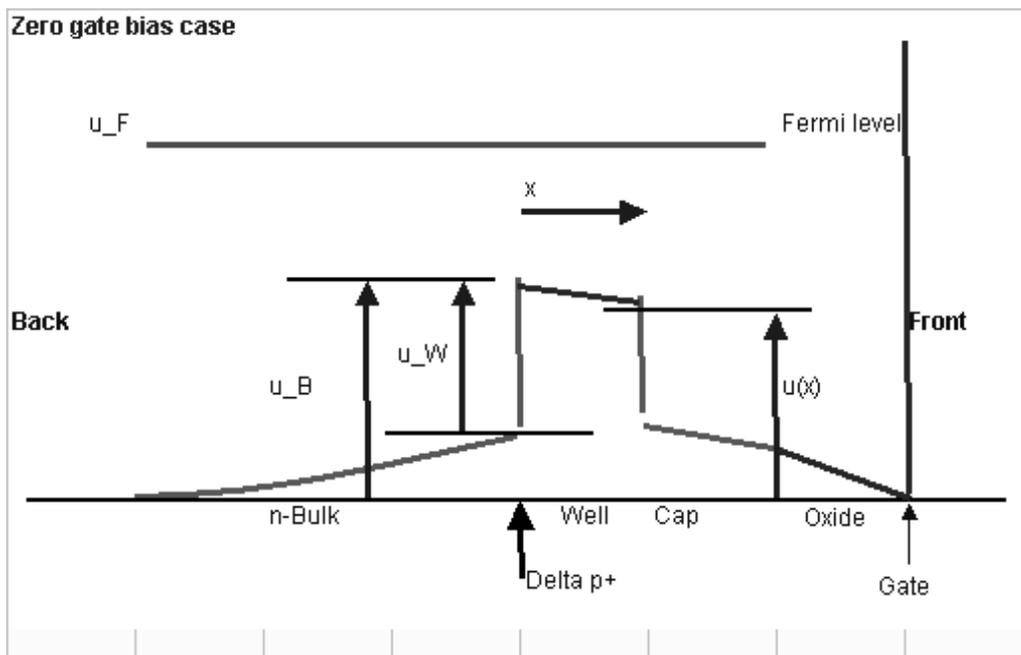


Figure 3-24. OFF state sketch. p-delta ions were equal to the depletion ions in the ON state. But now the p-delta ions are stronger. This can increase  $u_B$  and lift the QW

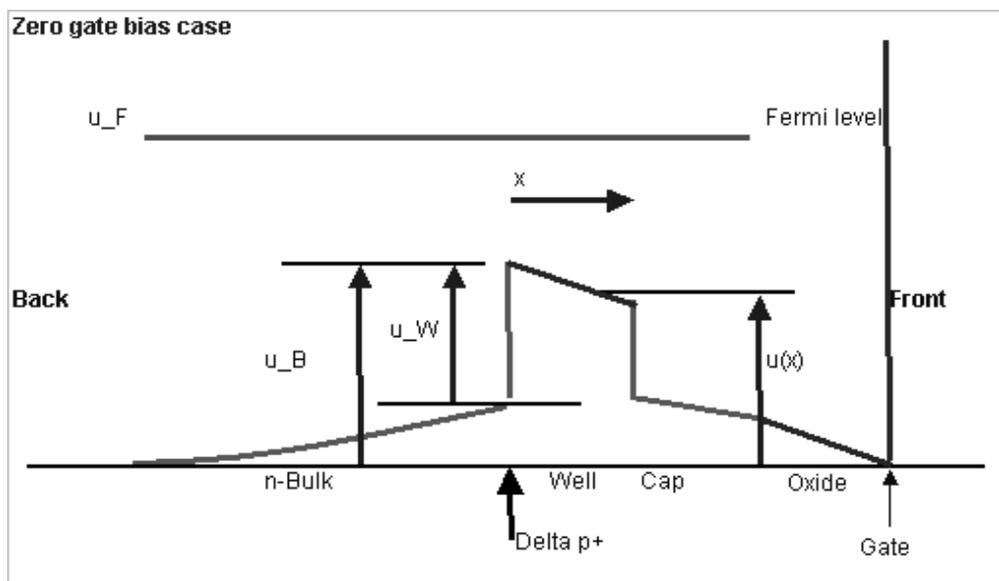


Figure 3-25. OFF state sketch. p-delta ions were more than bulk depletion ions in the ON state. Now p-delta ions are very high leading to high  $u_B$  and lifts the QW. This case may have the highest OFF state hole density

In simulations it was found that the bulk potential drop is adjusting very much so that the discussion for the OFF state above may not be fully correct. First the OFF state QW looks like in Figure 3-24 for low p-delta doping and like in Figure 3-25 for high p-delta doping. The simulated results of these conditions are shown in Figure 3-26 and Figure 3-27 respectively. Second the bulk potential drop increases very much with increased p-delta doping as seen in Figure 3-28. When the bulk potential drop is high the QW will be raised towards the Fermi level in the OFF state and the carrier population will increase.

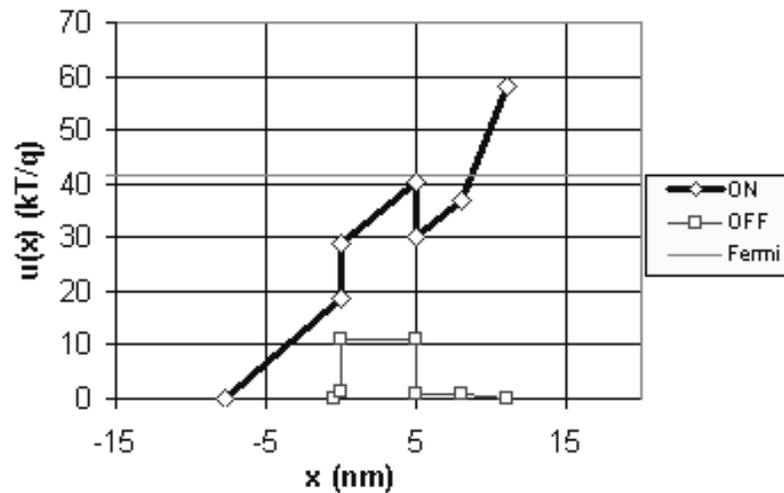


Figure 3-26. ON and OFF state potentials for a device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $n_B = 5 \times 10^{18} / \text{cm}^3$ . The p-delta doping is at a low value of  $9 \times 10^{10} / \text{cm}^2$

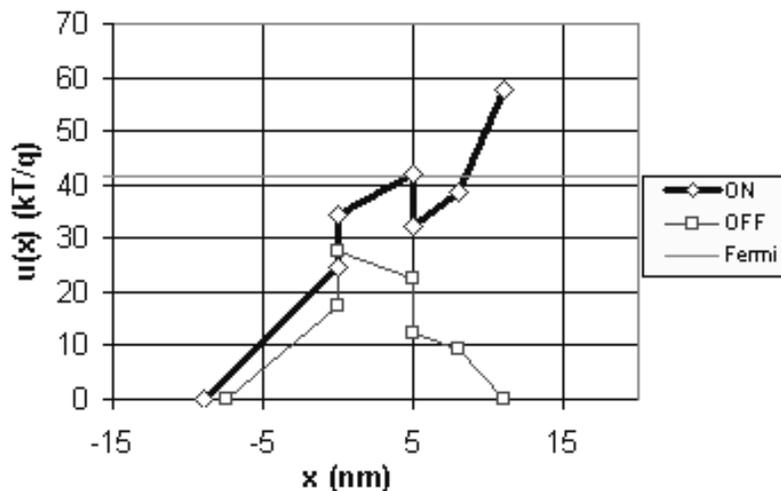


Figure 3-27. ON and OFF state potentials for a device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $n_B = 5e18/\text{cm}^3$ . The p-delta doping is at a high value of  $2e12/\text{cm}^2$ . Notice the high OFF state potential.

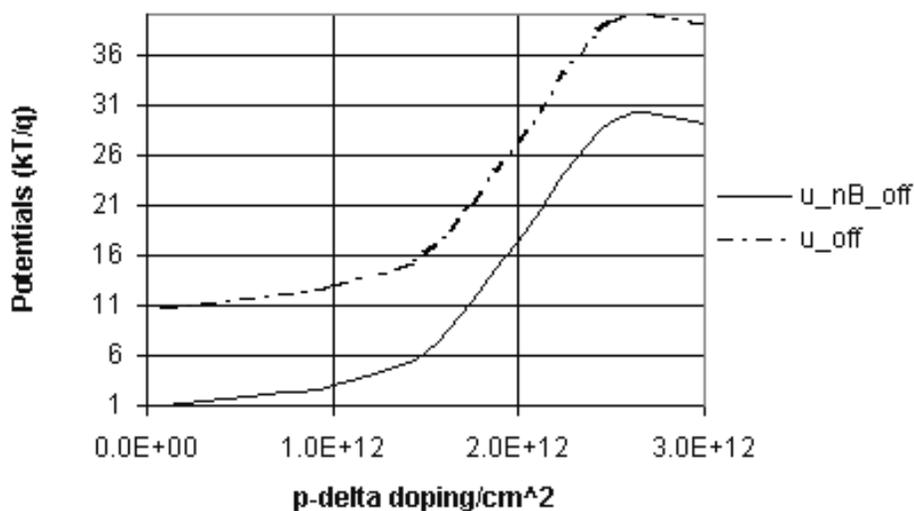


Figure 3-28. Simulated potentials in the OFF state for a device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $n_B = 5e18/\text{cm}^3$ .  $u_{nB\_off}$  is the bulk depletion potential,  $u_{off}$  is the potential at the top of the QW. This simulation used the Excel program discussed earlier.

#### 3.6.1.4 Conclusion for the OFF state

Band structures as shown in Figure 3-24 or Figure 3-25 reduces the carrier population in the OFF state. In simulations shown in Figure 3-28 the OFF state bulk depletion potential increases very much with increase in p-delta doping. This will increase the carrier population in the OFF state since the QW is raised. To get the best band structure, shown in Figure 3-25, a high p-delta doping is desired, as seen in Figure 3-27. But a higher doping also increases the hole density in the OFF state as shown in Figure 3-27.

It can be seen in Figure 3-28 that the difference between the OFF state bulk potential,  $u_{nB\_off}$ , and the QW potential,  $u_{off}$ , is almost the same regardless of p-delta doping. This is around a value of 10 which is the QW height. This is due to the fact that the contribution of QW holes to the potential in the OFF state is very low. But it can be noted from Figure 3-28 that a p-delta doping less than  $1 \times 10^{12}$  has the least value for the bulk depletion potential,  $u_{nB\_off}$ .

#### 3.6.1.5 Conclusion for p-delta Doping Simulations

As can be seen from the simulated potential plots of Figure 3-22 and Figure 3-26 to Figure 3-28 the p-delta layer has a strong influence on the device characteristics. In the ON case it appears that a flat band QW operation is the best choice. This yields the maximum QW population. From the simulation results for ON state given in Figure 3-22 the bulk depletion potential,  $u_{nB\_on}$ , increases with increasing p-delta doping. So a higher p-delta doping will increase the QW hole population. Figure 3-22 also shows that the difference between the QW potential,  $u_W$ , and the cap layer potential,  $u_{cap}$ , remains nearly the same irrespective of p-delta doping. So the p-delta doping density do not seem to worsen the parasitic channel effects in the cap layer.

In the OFF case a p-delta doping density stronger than the bulk depletion ionic sheet density seems to be best in order to reduce OFF currents. Figure 3-24 or Figure 3-25 seem to be good choice for low hole density in the OFF state. Comparing the simulation results of Figure 3-26 and Figure 3-27 it can be found that a low p-delta doping is better for lower OFF state hole density. In Figure 3-28 the substrate potential increases very much above a p-delta doping of  $1 \times 10^{12}/\text{cm}^2$ . Earlier it was noted that a higher p-delta doping gives a higher hole density. But the p-delta doping has to be limited to  $1 \times 10^{12}/\text{cm}^2$  in order to reduce the hole density in the OFF state. So a p-delta doping below  $1 \times 10^{12}/\text{cm}^2$  is an optimum choice for the device structure considered.

### **3.6.2 Tradeoff: QW Height**

In this section the effects of the QW height  $u_W$  upon the ON and OFF state behavior are studied. It is found that there are two conflicting issues. A high QW height is desirable in the ON state to avoid the surface parasitic channel in the cap layer. A low QW height is desirable in the OFF state to reduce  $I_{\text{OFF}}$ . Hence there is a trade off.

Previously it was shown that a low QW height leads to the formation of a degraded

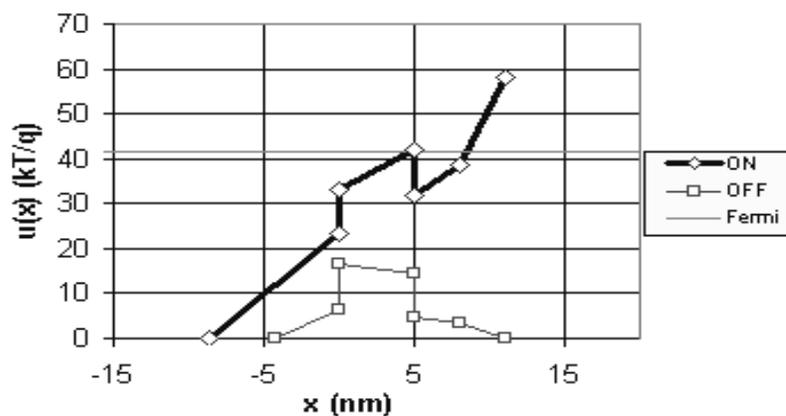


Figure 3-29. Valence band plots for ON and OFF states with a QW height of 10 kT/q. It can be seen that the top of cap layer (marker to the right of the QW) is lying below the top of the QW. Now the QW has much higher carrier population compared to the cap layer.

parasitic cap layer channel. From Figure 3-9 a shallow QW causes the formation of the parasitic channel. From Figure 3-29 higher QW height causes more OFF state hole density in the QW. So there is a trade off between these two quantities, parasitic channel and  $I_{OFF}$ . To study this further the QW height was increased from a low value towards a high value in increments. The ratio of the QW channel to the parasitic channel is studied at each step. Instead of assessing the OFF state hole sheet density the sub-threshold parameter  $S$  and the hole sheet density ON/OFF ratio were studied.

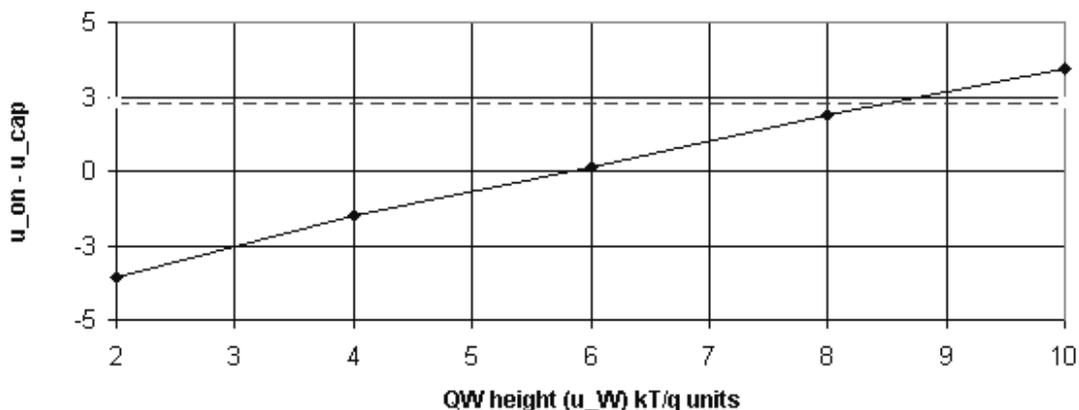


Figure 3-30. ( $u_{on} - u_{cap}$ ) vs QW height. A difference of 2.3 corresponds to the QW hole sheet density being ten times the hole sheet density in the cap layer. The value 2.3 is plotted with the dashed line. Substrate doping  $n_B = 5e18/cm^3$ , QW width  $x_W = 5$  nm and p-delta doping density  $Q_d = 1.25e12/cm^2$ .

Figure 3-30 plots the difference between ON state potential at the top of the QW,  $u_{on}$ , and the potential at the top of the cap layer,  $u_{cap}$ . Or the plot is ( $u_{on} - u_{cap}$ ) vs QW height. The curve meets the demarcation line (2.3) near a QW height of 9 kT/q. So a minimum QW height of 9 kT/q is needed to keep the parasitic channel negligible compared to the QW channel. However, this value for QW height depends on the selection of other parameters like QW width, substrate doping density, cap layer thickness and p-delta doping density.

### 3.6.2.1 Optimization of QW Height

Figure 3-31 plots the sub-threshold parameter  $S$  vs QW height. At a QW height of 9 kT/q the parameter  $S$  has degraded from 125 to about 140. So increasing QW height much higher is going to increase the device OFF current Figure 3-32 shows the hole density ON/OFF ratio vs QW height. The ON/OFF ratio has degraded from 1012 to  $5.95 \times 10^{10}$ . So from both the sub-threshold parameter (or IOFF) and the ON/OFF ratio a

higher QW is not desirable. So for the structural parameters of choice a QW height of 9  $kT/q$  is optimal. The parameters corresponding to this value are,

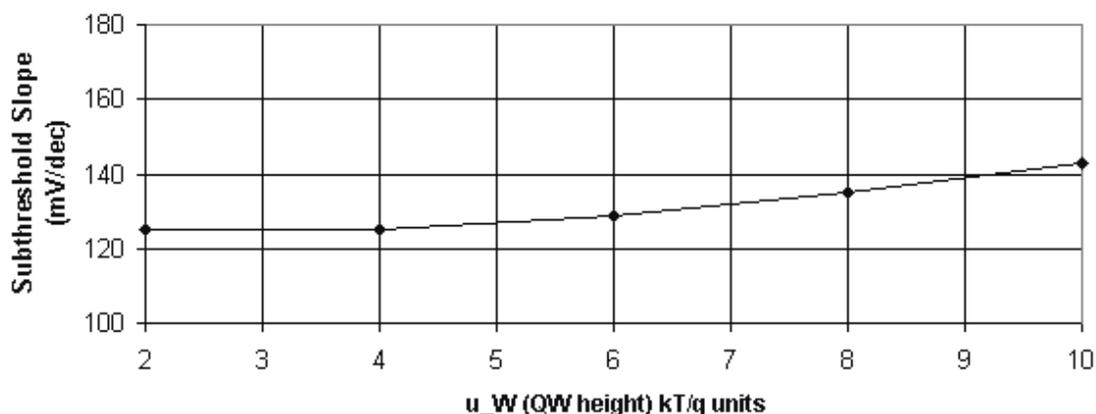


Figure 3-31. Sub-threshold slope parameter  $S_t$  vs QW height.  $n_B = 5e18/cm^3$ ,  $x_W = 5$  nm,  $Q_d = 1.25e12/cm^2$ .

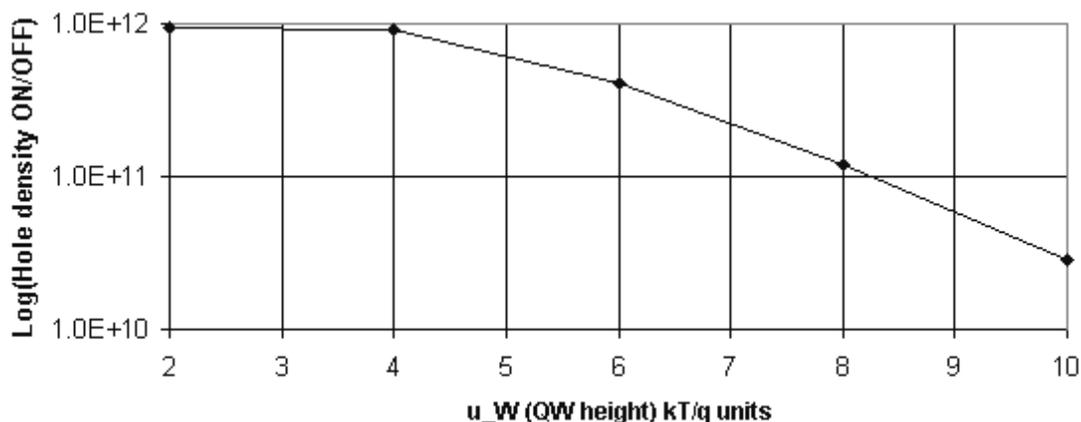


Figure 3-32. Log(hole sheet density ON/OFF) vs QW height.  $n_B = 5e18/cm^3$ ,  $x_W = 5$  nm,  $Q_d = 1.25e12/cm^2$ .

Optimum QW height  $u_W = 9$  (to keep the cap hole density  $< 1/10$  of QW hole density)

Sub-threshold parameter  $S = 140$  (which is much higher than a desired value of 100)

Hole sheet density ON/OFF =  $5.95e10$  (which is better than the same for MOSFET,  $10^7$ )

### 3.6.2.2 Conclusion for QW height Simulations

. In this sub-section the effect of QW height is discussed. QW height is found to be a very influential parameter. A low height can lead to surface channel formation. But a large height can lead to more off currents. In Figure 3-30 it was found that a height above  $9 \text{ kT/q}$  is needed to keep the surface channel hole density one tenth of the QW hole density. But in Figure 3-32 a height above  $4 \text{ kT/q}$  leads to a reduction in ON/OFF ratio of holes in the QW. At  $9 \text{ kT/q}$  the ratio degrades nearly two orders of magnitude. So to make a better device some other parameter need to be adjusted to get out of this quagmire. Reducing the QW width  $x_W$  can improve the hole sheet density ON/OFF ratio in the QW. But  $x_W$  chosen in the simulation is already a low value of 5 nm. So the substrate doping density,  $n_B$ , and the p-delta doping density,  $Q_d$ , could be varied to arrive at a reasonable set of parameters. This will require a lot of simulations repeating the results given here for a whole set of choices.

### 3.6.3 Trade off: QW width

Out of the four structural parameters p-delta doping density and QW height were discussed in previous sub-sections. In this sub-section trade off studies on the fourth parameter, QW width, is discussed. By keeping the other three parameters at their fixed values the QW width alone is increased step by step, from a starting low value. Same as in previous sub-sections performance metrics (1) QW ON state hole sheet density, (2) QW hole density ON/OFF ratio and (3) the difference between the QW right side potential and the cap layer right end potential,  $u_{on} - u_{cap}$ , is studied. The three fixed parameters are assigned the following values,

Substrate doping density,  $n_B = 5 \times 10^{18}/\text{cm}^3$

p-delta doping,  $Q_d = 2 \times 10^{12}/\text{cm}^2$

QW height,  $u_W = 10 \text{ kT}/q$

### 3.6.3.1 Simulation Results

Figure 3-33 shows the variation of sub-threshold parameter,  $S$ , with variation of QW width  $x_W$ . For a variation of  $x_W$  from 3 nm to 15 nm  $S$  varies from nearly 220 mV/dec to below 160 mV/dec. This shows that a wider QW is desirable to reduce the sub-threshold parameter,  $S$ . But even then the values of  $S$  are undesirably higher. This can be improved by selecting a low p-delta doping density.

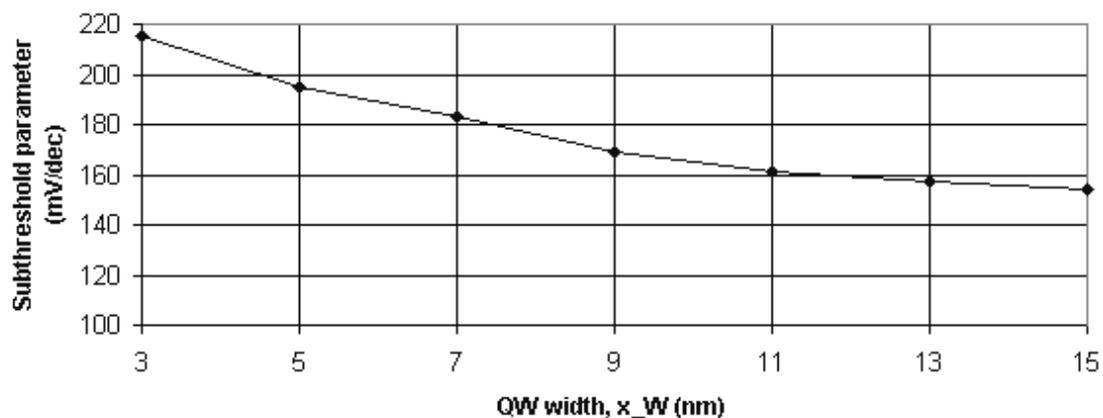


Figure 3-33. Variation of sub-threshold parameter,  $S$ , with variation of QW width,  $x_W$ .  $u_W = 10 \text{ kT}/q$ ,  $Q_d = 2 \times 10^{12}/\text{cm}^2$ ,  $n_B = 5 \times 10^{18}/\text{cm}^3$

The variation of the QW hole sheet density with variation of QW width is plotted in Figure 3-34. There is a slight increase of hole sheet density with increasing QW width. The increase is only less than a factor of 2 and is not significant in comparison to the

effect of other three parameters. As seen in previous sub-sections the QW height has a strong effect on the hole sheet density.

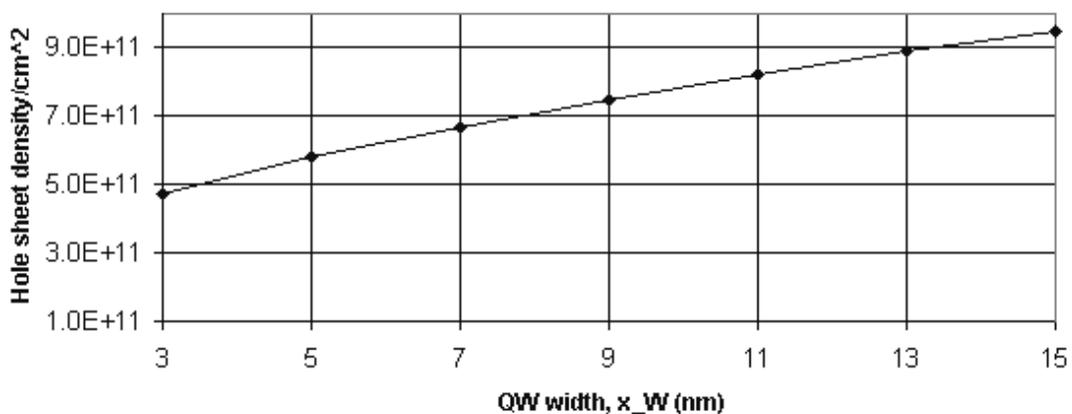


Figure 3-34. Variation of QW hole sheet density with variation of QW width,  $x_W$ .  $u_W = 10 \text{ kT}/q$ ,  $Q_d = 2 \times 10^{12}/\text{cm}^2$ ,  $n_B = 5 \times 10^{18}/\text{cm}^3$

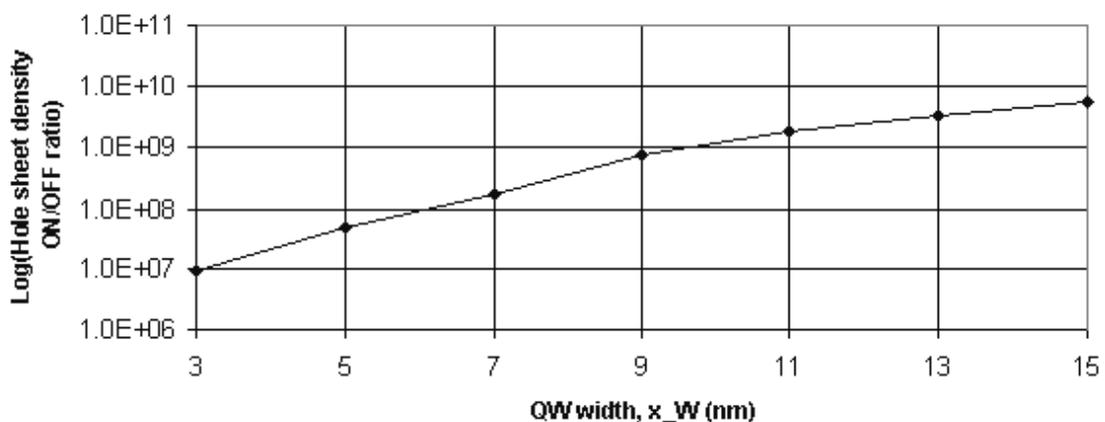


Figure 3-35. Hole sheet density ON/OFF ratio vs QW width.  $u_W$  is  $10 \text{ kT}/q$ ,  $Q_d = 2 \times 10^{12}/\text{cm}^2$ ,  $n_B = 5 \times 10^{18}/\text{cm}^3$

Hole sheet density ON/OFF ratio is plotted in Figure 3-35. There is large rise in ON/OFF ratio with increasing QW width,  $x_W$ . The increase is about three orders of magnitude for a variation of  $x_W$  from 3 nm to 15 nm. So a wider QW will give very

good OFF state behavior. The rise in the ON/OFF ratio also tends to saturate for a QW width of around 15 nm. So a QW width of 15 nm can be considered the optimum value.

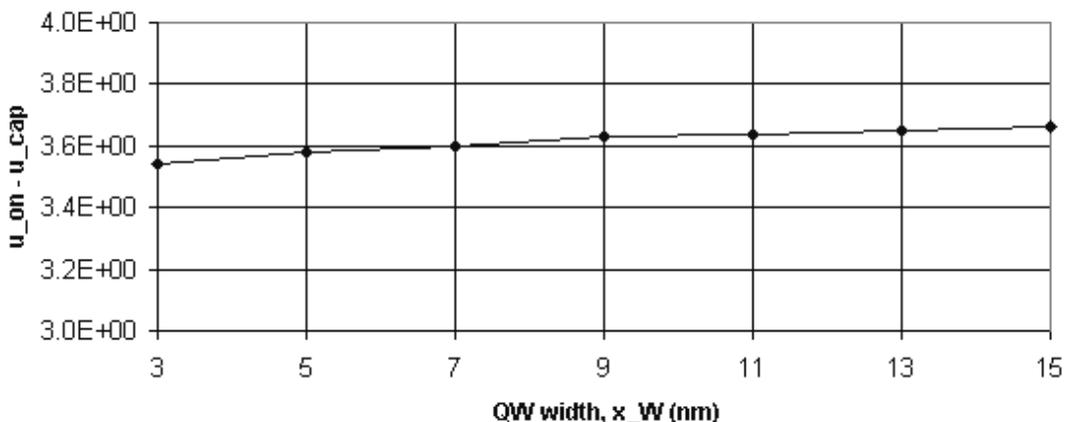


Figure 3-36. Difference between QW potential and cap layer potential. As discussed in previous sub-sections a minimum difference of 2.3 is needed to keep the hole sheet density in the cap layer less than a tenth of the QW hole sheet density.

The difference between the front end potential of the QW and the front end potential of the cap layer is plotted in Figure 3-36. The variation of this difference with QW width is minimal. In previous sub-sections it was mentioned that a minimum difference of 2.3 is needed to keep the cap layer hole sheet density a tenth below the QW hole sheet density. Figure 3-36 shows that the difference is above 2.3 for a variation of QW width from 3 to 15 nm.

### 3.6.3.2 Conclusion

Contrary to the other three structural parameters (QW height, p-delta doping and substrate doping) a wider QW seems to improve all metrics considered here. The improvement is very pronounced in the case of sub-threshold parameter,  $S$ , and the hole density ON/OFF ratio in the QW. This situation can be improved by selecting a low p-delta doping density. The hole density ON/OFF ratio tends to saturate near a QW width

of around 15 nm. So a QW width of 15 nm can be considered an optimum value. The improvement in the case of hole sheet density and cap layer hole density is minimal. But the values in Figure 3-36 shows that the cap layer density is less than a tenth of the QW hole sheet density. From the above it can be concluded that QW width is a free parameter. So  $x_W$  can be arbitrarily chosen to adjust the effects of changing other three parameters, viz. QW height, p-delta doping and substrate doping. A wider QW seems to be a better choice.

### 3.6.4 Trade off: Substrate Doping Density

Effects on device performance and trade offs of p-delta doping density, the QW height and QW width were discussed in previous sub-sections. Next the trade off of substrate doping density will be studied. From the simulations carried out, using the simple structure and program discussed in the beginning of this chapter, the substrate doping density does not seem to have much influence in device metrics such as hole sheet density, sub-threshold parameter etc.

#### 3.6.4.1 Simulation set up

The simulation results with discussions are presented below. Already it was discussed before that there are four structural parameters of interest in this device. The values of these parameters are given in parenthesis.

1. QW height,  $u_W$  (10 kT/q)
2. QW width,  $x_W$  (5 nm)
3. p-delta doping density/cm<sup>2</sup>,  $Q_d$  ( $8 \times 10^{10}$ /cm<sup>2</sup>)
4. Substrate doping density/cm<sup>3</sup>,  $n_B$  ( $1 \times 10^{17} - 1 \times 10^{19}$ /cm<sup>3</sup>)

In this section the substrate doping density is varied with the other three parameters kept fixed at the above values.

### 3.6.4.2 Potential plots

Figure 3-37 - Figure 3-41 shows the potential plots for the ON and the OFF states with a substrate doping variation from  $1 \times 10^{17}$  to  $1 \times 10^{19}/\text{cm}^3$ . In the ON case the substrate potential drop,  $u_{nB\_on}$ , reduces only very little up to a doping density of  $1 \times 10^{18}/\text{cm}^3$ . For higher doping densities there is a large reduction. In the OFF state the QW looks almost the same throughout the range of doping density considered here.

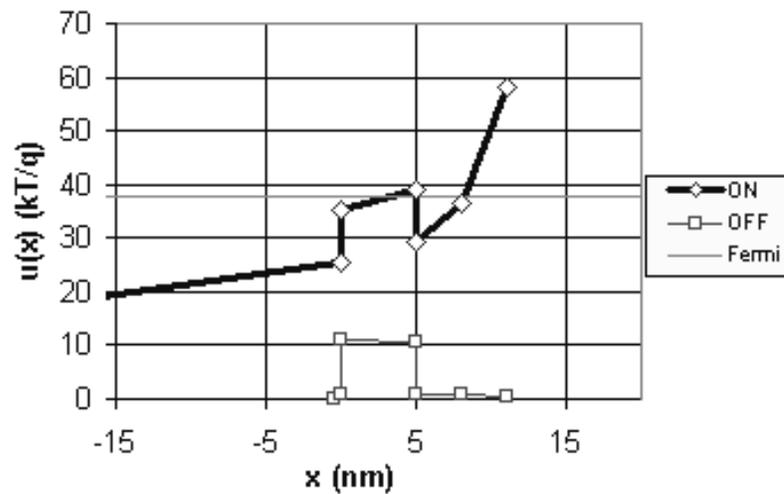


Figure 3-37. Potential plots for the ON and the OFF state for a device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 8 \times 10^{10}$ ,  $n_B = 1 \times 10^{17}$ .

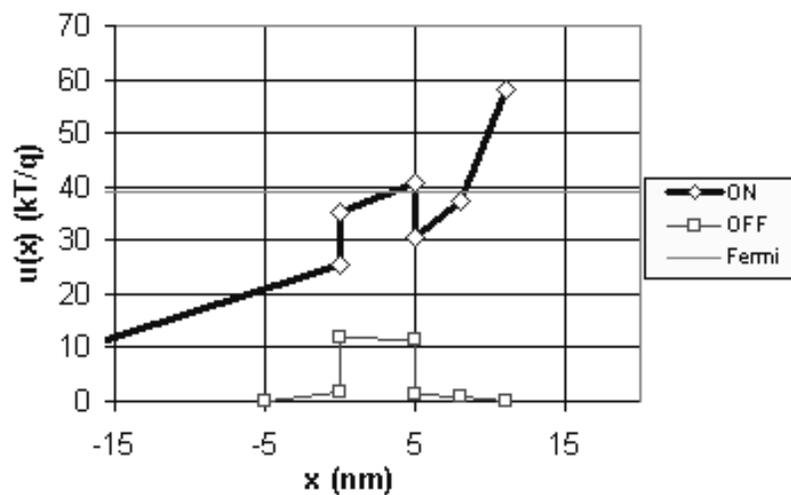


Figure 3-38. Potential plots for the ON and the OFF states for a device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 8 \times 10^{10}$ ,  $n_B = 5 \times 10^{17}$ .

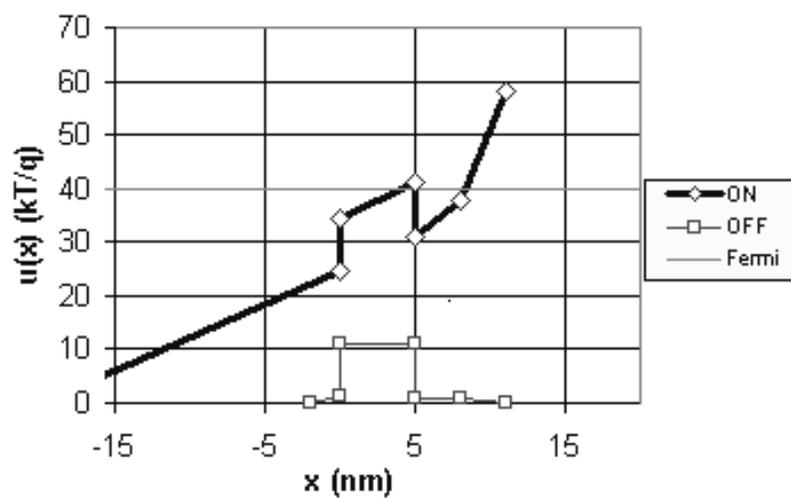


Figure 3-39. Potential plots for the ON and the OFF state for a device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 8 \times 10^{10}$ ,  $V_{gs} = 1.5 \text{ V}$ ,  $n_B = 1 \times 10^{18}$ .

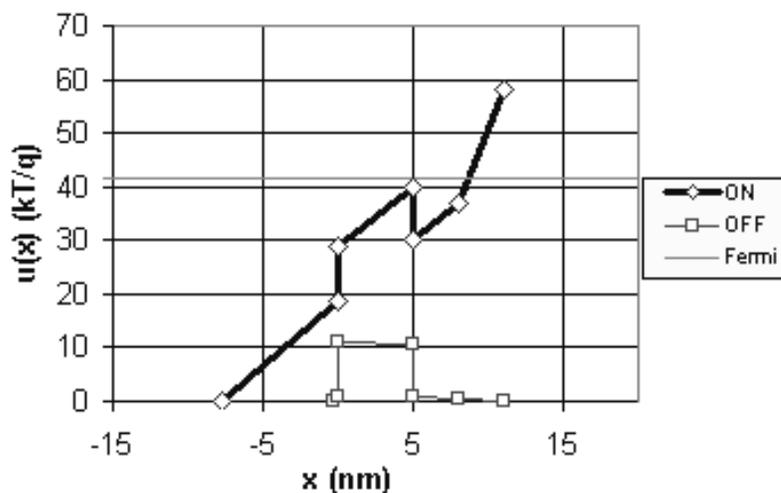


Figure 3-40. Potential plots for the ON and the OFF state for a device with  $u_W = 10$   $kT/q$ ,  $x_W = 5$  nm,  $Q_d = 8 \times 10^{10}$ ,  $V_{gs} = 1.5$  V,  $n_B = 5 \times 10^{18}$

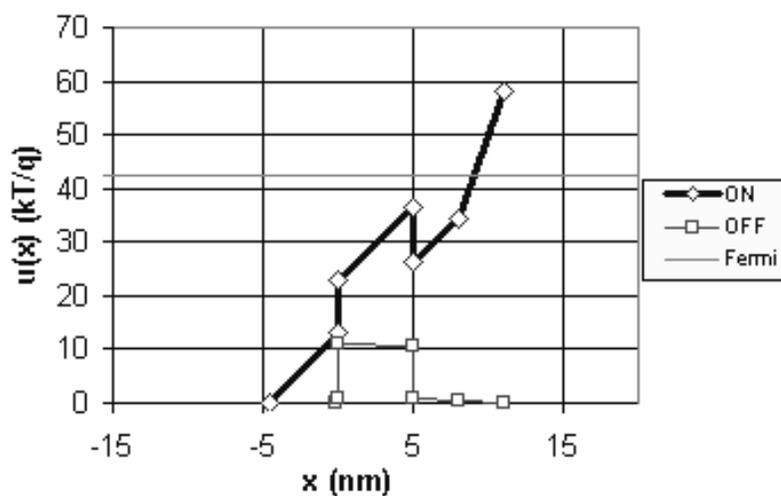


Figure 3-41. Potential plots for the ON and the OFF state for a device with  $u_W = 10$   $kT/q$ ,  $x_W = 5$  nm,  $Q_d = 8 \times 10^{10}$ ,  $V_{gs} = 1.5$  V,  $n_B = 1 \times 10^{19}$

Accordingly the peak of the QW top also moves down very little up to a doping density of  $1 \times 10^{18}/\text{cm}^3$  and moves down very fast afterwards. At a high doping of

$1 \times 10^{19}/\text{cm}^3$  the cap layer potential is near to the QW potential. So the surface parasitic channel gets stronger.

#### 3.6.4.3 Simulation results

Next quantified plots of simulation results are given. Figure 3-42 shows the QW hole sheet density variation with  $n_B$ . Up to a doping density of  $1 \times 10^{18}/\text{cm}^3$  the hole sheet density remains nearly the same. As seen in Figure 3-43 the substrate potential drop remains nearly the same up to  $1 \times 10^{18}/\text{cm}^3$ . So the top of the QW is nearly at the same level compared to the Fermi level below a doping of  $1 \times 10^{18}/\text{cm}^3$ . So the hole sheet density remains nearly the same.

Figure 3-44. shows the ON/OFF hole sheet density ratio vs. substrate doping. Up to a doping density of  $1 \times 10^{18}/\text{cm}^3$  there is a slight improvement in the ratio. Above a doping density of  $5 \times 10^{18}/\text{cm}^3$  the ratio drops steeply. Variation of sub-threshold parameter,  $S$ , is plotted in Figure 3-43.  $S$  shows a slight improvement up to a doping density of  $1 \times 10^{18}/\text{cm}^3$  and above  $5 \times 10^{18}/\text{cm}^3$  deteriorates fast. The difference between the potential at the front of the QW and the top of the cap layer is plotted in Figure 3-43. As discussed in earlier sections a difference of 2.3 is needed to keep the inferior cap layer

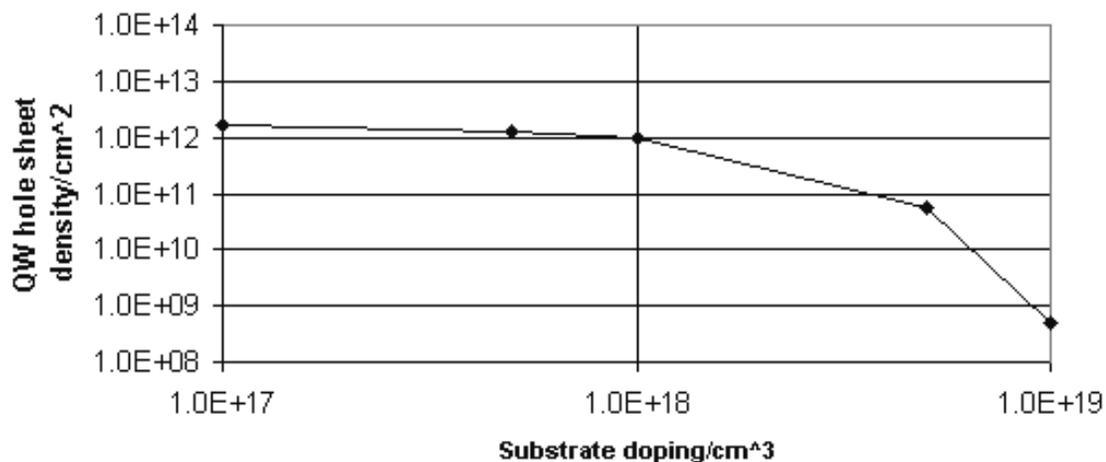


Figure 3-42. QW hole sheet density vs substrate doping,  $n_B$ .  $u_W = 10 \text{ kT}/q$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 8 \times 10^{10}$ ,  $V_{gs} = -1.5 \text{ V}$ .

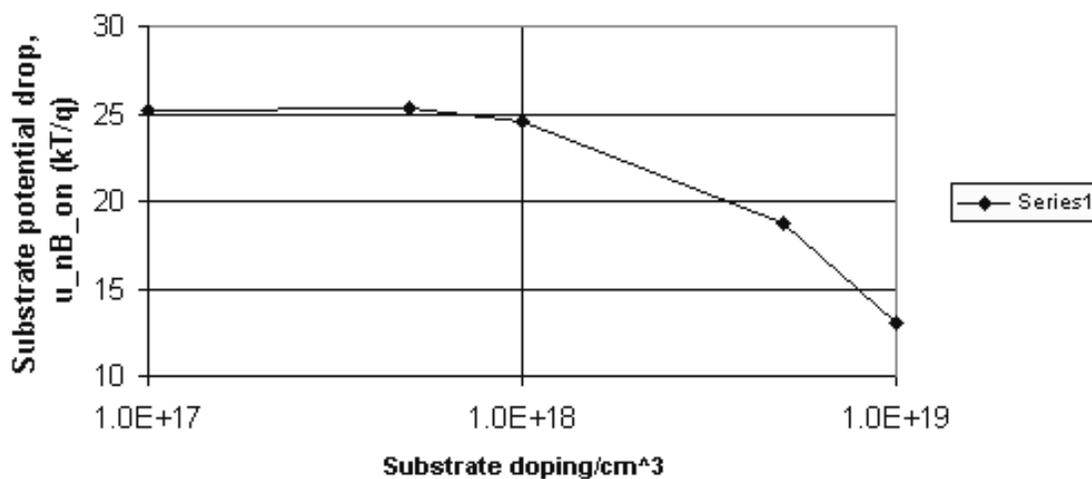


Figure 3-43. Substrate potential drop vs substrate doping for the device with  $u_W = 10 \text{ kT}/q$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 8 \times 10^{10}$ ,  $V_{gs} = 1.5 \text{ V}$ .

hole sheet density below a tenth of the QW hole sheet density. Up to a doping of  $5 \times 10^{18}/\text{cm}^3$  this is true. Finally Figure 3-47 shows the peak QW potential in the OFF state. The potential is nearly the same for the range of substrate doping density variation considered here. So the OFF state leakage is independent of substrate doping.

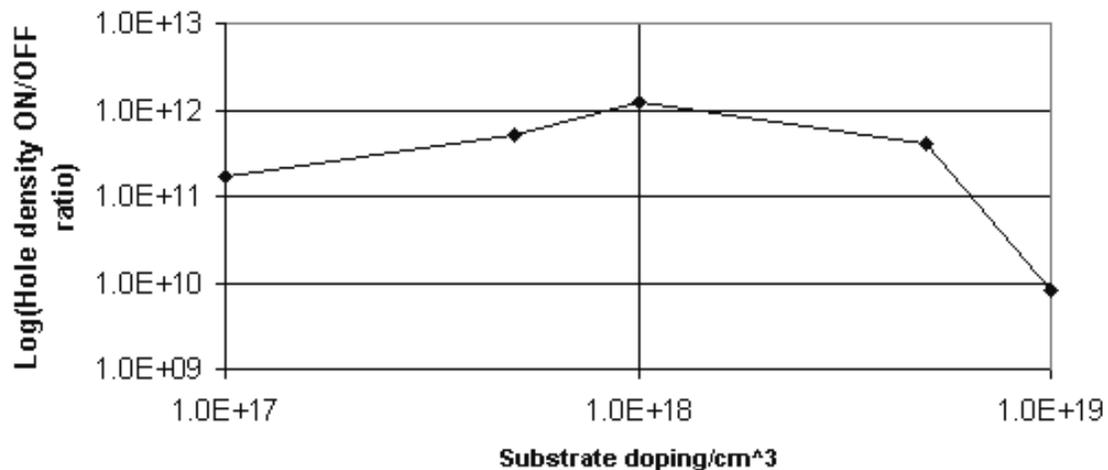


Figure 3-44. Log(hole density ON/OFF ratio) vs substrate doping for the device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 8 \times 10^{10}$ ,  $V_{gs} = 1.5 \text{ V}$ .

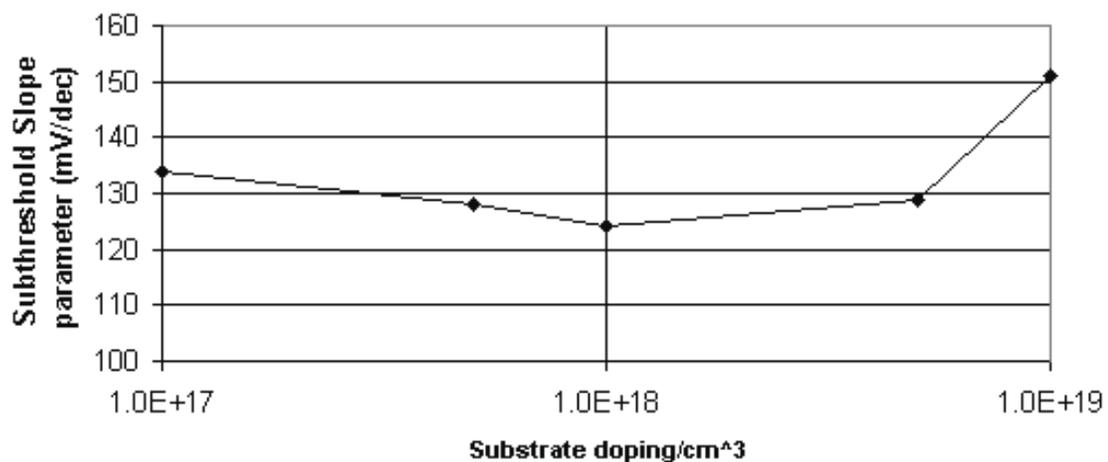


Figure 3-45. Sub-threshold slope parameter,  $S$ , vs. substrate doping density for the device with  $u_W = 10 \text{ kT/q}$ ,  $x_W = 5 \text{ nm}$ ,  $Q_d = 8 \times 10^{10}$ .

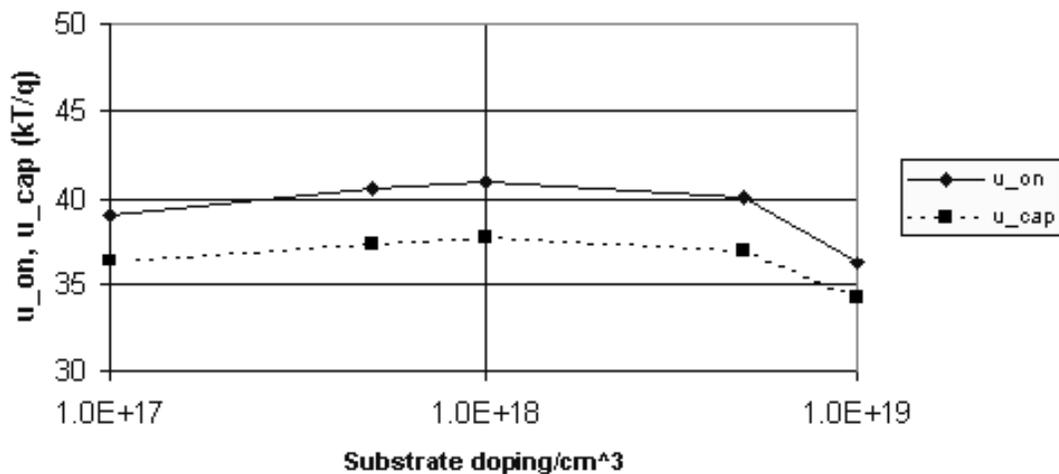


Figure 3-46. The difference between the front side QW potential and the potential at the top of the cap layer vs substrate doping density.  $V_{gs} = -1.5$  V

#### 3.6.4.4 Conclusions

In this section the effect of substrate doping density upon various device metrics are presented. A low p-delta doping density of  $8 \times 10^{10}/\text{cm}^2$  was used since higher densities were giving convergence problems. The substrate doping density considered is from  $1 \times 10^{17}$  to  $1 \times 10^{19}/\text{cm}^3$ . It has been found there is no dependence of OFF state QW hole sheet densities over the range of the above doping densities. Even though the hole

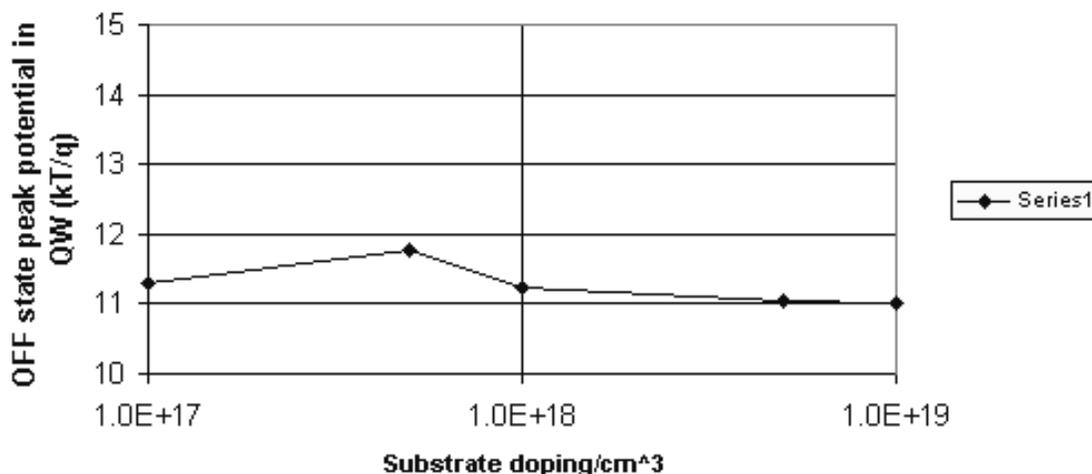


Figure 3-47. OFF state peak QW potential vs. substrate doping. This potential is nearly the same throughout the doping density range considered.

density ON/OFF ratio and the sub-threshold parameter,  $S$ , do not show considerable variation up to a doping density of  $5 \times 10^{19}/\text{cm}^3$  the hole sheet density in the ON state drops by more than an order of magnitude above  $1 \times 10^{18}/\text{cm}^3$ . So the substrate doping density should be limited to  $1 \times 10^{18}/\text{cm}^3$ . It is also seen that the difference between the peak QW potential and the peak cap layer potential is sufficient to keep the hole sheet density in the cap layer less than a tenth of the hole sheet density in the QW up to a doping density of  $5 \times 10^{18}/\text{cm}^3$ . Considering Figure 3-42, Figure 3-44 and Figure 3-45 the effect of substrate doping density on device performance is not much and hence it is not a good variable to adjust the device metrics. It is also to be noted from Figure 3-44 that the ON/OFF hole density ratio (and hence ON/OFF current ratio) in this device is few orders of magnitude higher than the ON/OFF current ratio in conventional MOSFET.

#### 3.6.4.5 Overall Discussion

Figure 3-48 and Figure 3-49 show the potential distribution without and with the p-delta layer respectively. These plots are generated using the program described in

appendix B. These potential plots give insight into the shape of the QW as well as the position of the top of the cap layer in the ON state. It can be seen that with the inclusion of the p-delta layer the buffer potential drop can be avoided. This, as seen in Figure 3-48 and Figure 3-49 moves more area of the QW closer to the Fermi level thus increasing the carrier density. At the same time it moves down the cap layer considerably down from the Fermi level thus reducing the parasitic channel effects. So in spite of the simple nature of this analysis it gives important information of the conceptual device structure of a p-channel hetero junction MOSFET (p\_HMOSFET) which is described in more detail in chapter 5.

The objectives of this simplified analysis program are listed in the introduction of this chapter. In the previous sections the following simulated results are given,

1. QW ON state hole sheet density dependence on structural parameters
2. QW ON/OFF hole sheet density ratio dependence on structural parameters
3. Sub-threshold parameter  $S_t$  dependence on structural parameters

The electric field in front of the QW is not explicitly given but is used in the calculation of  $V_{gs}$ . The maximum cap layer thickness to keep the percentage of surface channel hole density to any given minimum value is available but they are not brought out in the graphical form. Through 'goal seek' both the QW width,  $x_W$ , and the gate voltage,  $V_{gs}$ , can be set to any pre-defined values. 'Goal seek' can be performed with  $V_{gs}$  as the target cell and the front side or the back side potential of the QW as the variable. In general this simple program helped to identify the behavior and structural dependence of

vital parameters of this QWMOS. In general the feasibility of the device discussed in this chapter looks promising in the design of a p-HMOSFET.

It is to be noted here that depending on the ionic strengths of the bulk depletion region and the p-delta layer the QW band edge may tilt up or down respectively and this can make considerable difference in the results reported in this section

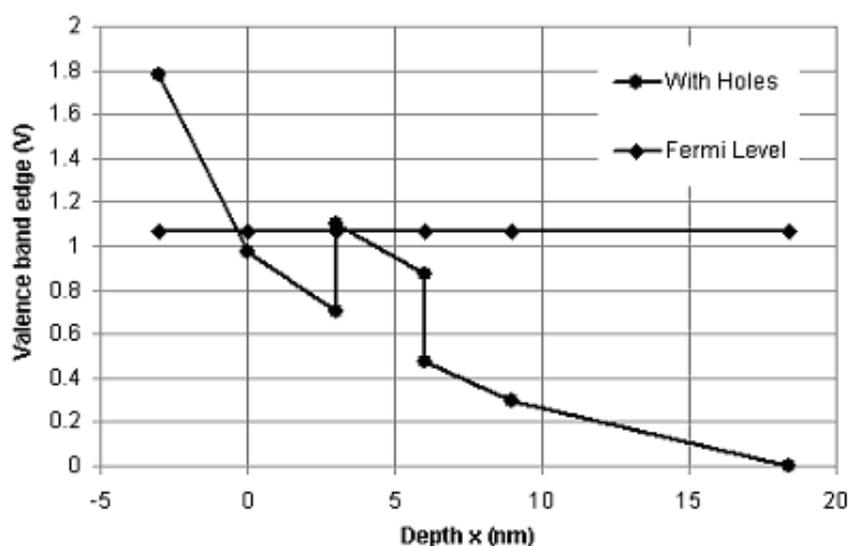


Figure 3-48. Potential plot of an ideal device.  $t_{ox}$ ,  $X_{cap}$ ,  $X_{QW}$  and  $X_{buf}$  are 3 nm and the substrate n-type doping is  $4 \times 10^{18}/\text{cm}^3$ . There is no p-layer. So the potential keeps rising. The diamond marks are on Fermi level and dots represent potentials including the effects of holes in the QW.

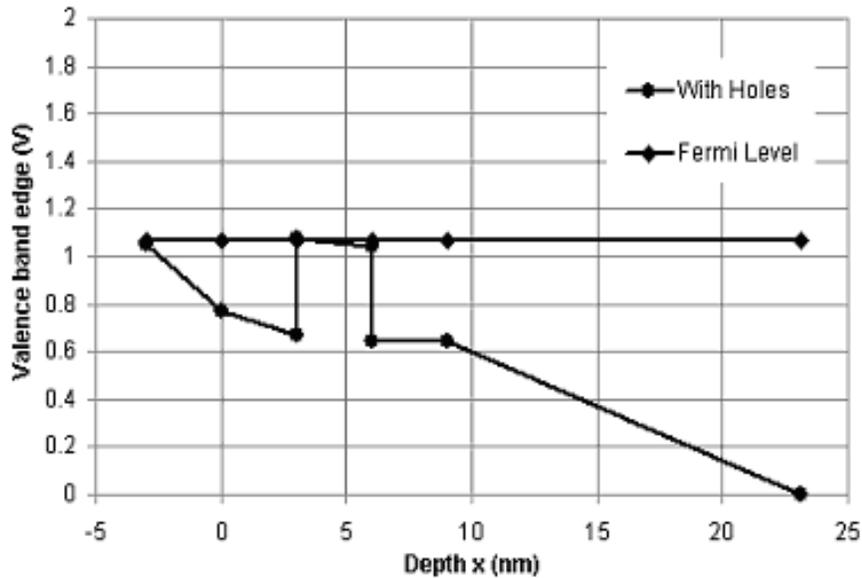


Figure 3-49. Potential plot of an ideal device with the inclusion of a p-doped  $\delta$ -layer with doping density of  $5.64 \times 10^{12}/\text{cm}^2$ . The layer dimensions and doping are the same as before.  $t_{\text{OX}}$ ,  $X_{\text{cap}}$ ,  $X_{\text{QW}}$ ,  $X_{\text{buf}}$  are 3 nm and the n-substrate doping is  $4 \times 10^{18}/\text{cm}^3$ . Notice that the p-layer doping is selected such that the field in the buffer layer is zero.

### 3.7 Lessons Learned from the Above Simulations

So far in this chapter a simple QW device is considered in silicon MOSFET, the Poisson equation is solved in the QW, a simple Excel spreadsheet program is written to find the potentials for a given gate voltage and a complete parametric study is carried out to get a rough idea of the influence of parametric variations on the device characteristics. Only the potentials at the boundaries of the QW are calculated using Poisson solution. Other potentials at other boundaries are calculated using the input potentials and the QW boundary potentials and fields. These boundary potentials are joined by straight lines to get an approximate potential distribution in the device. So the quadratic variation of

potential in the bulk or the exponential variation of the potential in the QW are not taken into account.

The fixed parameters for the device are,

1. QW width,  $x_W$ , in nm
2. QW height,  $u_W$ , in  $kT/q$
3. Substrate doping,  $n_B/cm^3$
4. p-delta ion density,  $Q_d/cm^2$

These parameters are fixed for a device and do not vary. But for parametric simulation studies one parameter is varied at a time keeping the other parameters at fixed values. But stepping of one parameter will change the other three parameter. To bring them back to their fixed values the following potentials are used in iteration mode,

1. Potential at right (top) of QW in ON state,  $u_{on}$ , in  $kT/q$  units
2. Potential at right (top) of QW in OFF state,  $u_{off}$ , in  $kT/q$  units
3. Potential in the bulk in ON state,  $u_{nB_{on}}$ , in  $kT/q$  units
4. Potential in the bulk in OFF state,  $u_{nB_{off}}$ , in  $kT/q$  units

The simulations using the Excel spread sheet program, developed using the QW Poisson equation solution of sub-section 3.2, are studied. The hole sheet density in the QW in the ON state, ON/OFF hole sheet density in the QW and the sub-threshold parameter,  $S_t$ , are the important metrics. In fact ON/OFF ratio and  $S_t$  are related. But the ON/OFF ratio is a vital metric that can quickly describe the device more effectively than  $S_t$  or the OFF state hole sheet density.

It was found that QW width,  $x_W$ , has only a weak effect on ON state hole sheet density in the QW. But both QW height,  $u_W$ , and p-delta doping density,  $Q_d$ , have strong effects on ON state hole sheet density in the QW. A higher  $u_W$  is desirable for higher hole density ratio in the QW. But this also increases the OFF state QW hole sheet density. A lower  $u_W$  and wider cap layer will create an inferior parasitic channel in the front (top) of the cap layer and this will, in addition to degrading mobility, shield the QW charges. Equal number of ions in the bulk and the p-delta, or flat band QW, seems to yield more ON state hole sheet density in the QW.

A higher  $u_W$  is desirable to keep the cap layer hole density low. But it was found that this reduces the ON/OFF ratio of hole sheet density in the QW. A value of  $9 \text{ kT/q}$  was found to be optimum to keep the cap layer hole sheet density a tenth of that in the QW. But a  $u_W$  above  $4 \text{ kT/q}$  seems to degrade the ON/OFF ratio. But even at  $9 \text{ kT/q}$  the ON/OFF ratio is three orders of magnitude greater than the ratio in conventional MOSFET. So the  $u_W$  of  $10 \text{ kT/q}$  used in the simulations is justifiable.

The QW width,  $x_W$ , has only a weak effect on sub-threshold parameter  $S_t$ . But  $u_W$  and p-delta doping density,  $Q_d$ , have strong influence. Wider  $u_W$  was also found to degrade  $S_t$ . For both  $S_t$  and hole sheet density ON/OFF ratio a wide  $u_W$  is not preferred.

$Q_d$  seems to have no effect on the ratio of QW hole sheet density/Cap layer hole sheet density. A higher  $Q_d$  above  $10^{18}/\text{cm}^2$  reduces the hole sheet density in the QW. A higher  $Q_d$  also increases  $u_{nB\_off}$  which will increase the hole sheet density in the QW in the OFF state. A  $Q_d$  less than  $10^{12}/\text{cm}^2$  has the least effect in the OFF state. The ON state hole sheet density was also found to drop for a  $Q_d$  above  $10^{18}/\text{cm}^2$ . A wider QW

improves all metrics slightly. But above  $x_W$  of 15 nm ON/OFF hole sheet density ratio saturates. So  $x_W$  can be used as a free parameter to adjust other parameters upto a value of 15 nm. So 15 nm can be taken as the upper limit for  $x_W$ . The substrate doping density,  $n_B$ , is also found to be a weak parameter. The ratio of bulk ions and the p-delta ions have very strong influence on device characteristics in the ON and OFF states.

### 3.8 Summary

In this chapter preliminary studies of a modified MOSFET (QWMOS) is explored. A recessed undoped QW channel region is introduced below the channel region of a conventional MOSFET. This can be realized by having a thin cap layer that bridges between the QW and the gate oxide. A p-delta layer is also introduced below the QW. The idea is to confine holes in the QW and achieve improved transport properties.

The aim is to find the feasibility and a rough estimate of the performance advantage of such a device. The Poisson equation is solved only in the QW region. A simple program is written in Excel spread sheet taking the Poisson solution and the structural inputs. Poisson solution is used to find the potential and field at the front (top) of the QW. Potentials at other boundaries are found from the QW boundary potentials or by iteration keeping fixed inputs like gate voltage and width of the QW as target cells. Then these boundary point potentials are joined with straight lines to get an approximate potential distribution. Even though the program is not very accurate it can give potential distribution instantaneously which helps to make a guess of the functionality check.

Four of the input structural parameters are treated as fixed parameters. These are QW height, QW width, p-delta doping density and substrate doping density. In a realistic

device these cannot be changed. Here the parametric study of the influence of these structural parameters on device performance requires the change of these parameters. So in the simulations one of these structural variables is changed in steps while keeping the other three structural variables fixed at their input values. Some performance metric has to be observed at each step while the above structural variable is stepped. Four monitoring parameters of vital interest in this QWMOS are hole sheet density in the QW in the ON and OFF states, the ratio of the QW hole sheet density in the ON state to the QW hole sheet density in the OFF state and the sub-threshold parameter  $S_t$ . In addition to these metrics some potentials also need to be monitored in the QWMOS.

Each of the observing metrics is plotted against the stepping structural parameter to find out the trends of variation of each of these metrics and also to evaluate if the variation is in the desired or undesired direction. This will allow to choose a range of values for that particular structural parameter. In some cases the variation of the structural parameter may be desirable for one metric but undesirable for another. Then an optimum value or range of values has to be fixed for that structural variable. In some cases an optimum value of the stepping structural variable can be found that improves one or more of the important performance metrics can be improved to the desired values at the expense of some other not so important performance metrics.

After completing the above parametric simulation studies it is felt that this is a tangible device and holds high promise in performance enhancement of MOSFET. Now the question is how to realize such a QWMOS. This will be discussed in detail in chapter 5.

## CHAPTER 4

### EFFECTS OF STRAIN ON $\text{Si}_{1-x}\text{Ge}_x$ AND HMOSFET

In this chapter characteristics of thin layers of compressively strained  $\text{Si}_{1-x}\text{Ge}_x$  is discussed. Further the exploitation of these characteristics in the design of high-speed p-channel hetero junction MOSFET (p-HMOSFET) is explored. In particular strain-induced effects of thin layers of  $\text{Si}_{1-x}\text{Ge}_x$  are of particular interest from the standpoint of device design. Ge fraction  $x$  and the associated compressive strain is the vital mechanism by which novel high speed  $\text{Si}_{1-x}\text{Ge}_x$  based hetero junction devices are designed. A detailed introduction to  $\text{Si}_{1-x}\text{Ge}_x$  and a discussion of its current technological status are given in chapter 1. So no introductory material will be presented here. Instead the strain dependent effects and their usefulness in the design of hetero junction p-HMOSFET will be the main topic of this chapter. The following are covered in this chapter,

First sub-section will go through some of the general properties of  $\text{Si}_{1-x}\text{Ge}_x$ . The second section has some strained thin film growth related details. The third section will elaborate the effect of compressive strain on  $\text{Si}_{1-x}\text{Ge}_x$ . In the fourth section a general introduction to p-HMOSFET is given. Prior to choosing an appropriate structure of the p-HMOSFET for this dissertation a detailed survey of the published research work has been carried out. The structures of these published research works are presented in the fifth section. Then the sixth section will detail the particular structure chosen for this research work and discuss the details of the structure.

## 4.1 Properties of $\text{Si}_{1-x}\text{Ge}_x$

Ge is miscible in Si for the complete range from zero to 100% to form a solid state alloy semiconductor. However, for Ge fraction  $x < 0.85$  the compound semiconductor preserves the silicon like diamond lattice and the silicon like band structure. The site occupancies of Ge in  $\text{Si}_{1-x}\text{Ge}_x$  are not completely random since the alloy solidifies into a few long range ordered distributions of Ge in the diamond lattice [103]. This is important to get consistent characteristics in different samples. Lattice constants for Si and Ge are 0.5431nm and 0.5657nm respectively. Because of this difference the lattice constant of the compound semiconductor increases steadily with increasing Ge fraction, starting with the lattice constant of silicon. The lattice constant  $a(x)$  of unstrained  $\text{Si}_{1-x}\text{Ge}_x$  can be expressed empirically as [80],

$$a(x) = 0.5431 + 0.002733x^2 + 0.01992x \text{ (nm)} \quad (4.1)$$

The properties of bulk or unstrained  $\text{Si}_{1-x}\text{Ge}_x$  are not discussed here in detail since the strained  $\text{Si}_{1-x}\text{Ge}_x$  is of interest in the design of p-HMOSFET. A brief mention of the band gap and transport properties for the bulk or relaxed  $\text{Si}_{1-x}\text{Ge}_x$  are given in section 1.4 of chapter 1. Addition of Ge to Si leads to a reduction of the indirect band gap as shown in Figure 4-3. The upper curve refers to thick layers of bulk unstrained or relaxed

$\text{Si}_{1-x}\text{Ge}_x$ . In bulk  $\text{Si}_{1-x}\text{Ge}_x$  the initial reduction of band gap is nearly linear up to a Ge fraction of about 0.4. Between a fraction of 0.4 to 0.85 the band gap nearly saturates. Above a Ge fraction of 0.85 the band gap reduction is much faster, finally falling to the

Ge band gap of 0.66 eV. Thick layers of bulk unstrained or relaxed  $\text{Si}_{1-x}\text{Ge}_x$  finds use as a substrate to provide tensile strain to overgrown layers of Si. This is used in the case of n-channel devices. The lower curve in Figure 4-3 refers to compressively strained thin layers of  $\text{Si}_{1-x}\text{Ge}_x$ . The fall in band gap is nearly linear up to a Ge fraction,  $x$ , of 0.4. At a value of  $x = 0.5$  the band gap has fallen from 1.12 eV of Si to about 0.7 eV. The reduction in band gap is much more pronounced in strained thin layers compared to bulk unstrained  $\text{Si}_{1-x}\text{Ge}_x$ . This drop in band gap can be effectively used for band gap engineered devices as described in sub-section 4.3.1.

Another important concern in the case of  $\text{Si}_{1-x}\text{Ge}_x$  is the poor thermal conductivity. This is true with alloy semiconductors and is true with GaAs, for another example. As can be seen in Table 4.1 the thermal conductivity is roughly 18 times lower compared to Si. Addition of Ge to Si increases its resistivity and this increases thermal resistivity, following Matthiesen's rule. This demands proper heat removal schemes. However as will be seen in chapter 8 the off currents in these p-HMOSFETs are two to three orders of magnitude lower than that of MOSFETs for the same ON currents. Hence the standby heat generation will be much lower in these p-HMOSFET devices. In the case of present day nano meter silicon MOSFETs the OFF state currents are orders of magnitude higher compared to long channel devices. As pointed out in the 2003 Technology Road Map for Semiconductors this is a serious concern for MOSFETs since the packing density is rising with nano meter devices and the leakage currents are rising with down scaling [ 52].

## 4.2 Strained $\text{Si}_{1-x}\text{Ge}_x$

As indicated in chapter 1 the introduction of Ge into Si lattice and the associated compressive strain dramatically changes the properties of silicon. The Ge fraction alone does not produce much of these effects. But the addition of Ge gives a compound semiconductor whose band gap can be designed at will by the designer. When thin layers of this compound semiconductor are overgrown on silicon substrate a compressive strain is produced in these  $\text{Si}_{1-x}\text{Ge}_x$  layers because of the larger lattice constant of  $\text{Si}_{1-x}\text{Ge}_x$ . This strain magnifies the effects to a higher degree. Some of these effects are band gap reduction, band offset at interfaces, increase in effective hole mobility and increase in saturation velocity  $v_{\text{sat}}$ . These effects are discussed in the following sub-section 4.3.

### 4.2.1 Pseudomorphic Growth of $\text{Si}_{1-x}\text{Ge}_x$

As noted before  $\text{Si}_{1-x}\text{Ge}_x$  has a slightly larger lattice constant than that of silicon. Hence the thin layer of  $\text{Si}_{1-x}\text{Ge}_x$  grown over a (100 oriented) silicon substrate will experience a lattice deformation leading to an in plane compressive strain. This is schematically shown in Figure 4-1. The growth is pseudomorphic in the sense that the growing  $\text{Si}_{1-x}\text{Ge}_x$  bonds are attached to the underlying Si atom bonds. That means there are hardly any un-terminated dangling bonds. This gives a very low defect interface. At the same time it causes the overgrown  $\text{Si}_{1-x}\text{Ge}_x$  bonds to compress and lead to a compressive in plane strain in the  $\text{Si}_{1-x}\text{Ge}_x$ . As shown in the figure the over grown

$\text{Si}_{1-x}\text{Ge}_x$  layer is tetragonally distorted with a compressive in plane strain and a tensile strain along the growth direction. This leads to changes in the band structure of  $\text{Si}_{1-x}\text{Ge}_x$ . In case of thick or relaxed  $\text{Si}_{1-x}\text{Ge}_x$  layers this distortion soon vanishes over a small

growth thickness and physical properties of bulk  $\text{Si}_{1-x}\text{Ge}_x$  are maintained afterwards. Since many important characteristic changes are associated with the strain it is important to keep the layer thin enough so as to preserve the strain. In the UHV-CVD epitaxy the growth is pseudomorphic in spite of slight lattice mismatch between the silicon and the overgrown  $\text{Si}_{1-x}\text{Ge}_x$  layer. Silicon has a lattice constant of 5.43 Å whereas the lattice constant of  $\text{Si}_{1-x}\text{Ge}_x$  is given by the equation. ( 4.1). In the extreme case of pure Ge grown over Si, this mismatch may be as high as 4.2 % . But for good quality device applications the Ge content is found to be between 0.3 and 0.5, later on in chapter 7. For this range of  $x$  the lattice mismatch is much less than 4.2 %.

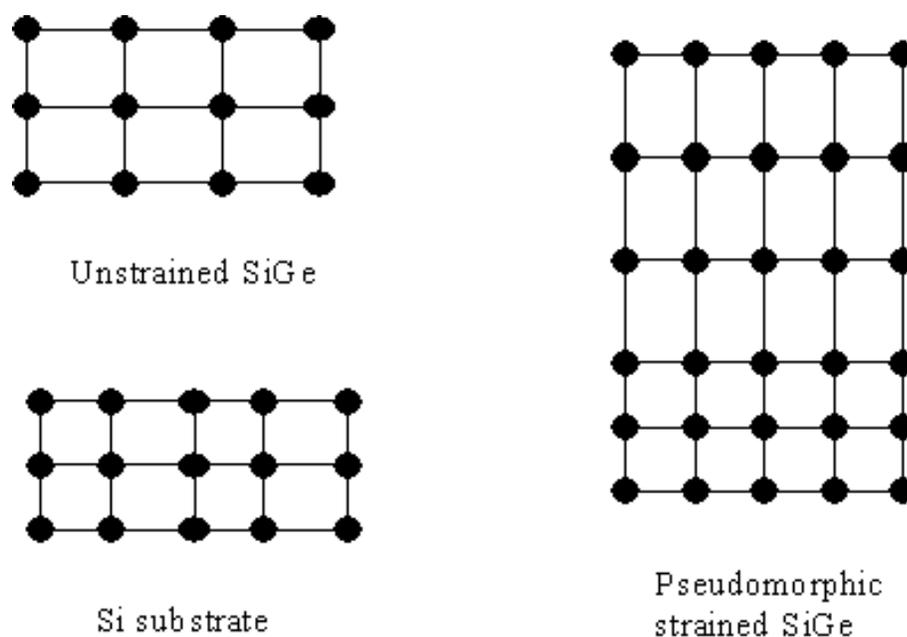


Figure 4-1. Lattice matched growth of  $\text{Si}_{1-x}\text{Ge}_x$  over silicon substrate. The overgrown  $\text{Si}_{1-x}\text{Ge}_x$  lattice is compressed in the plane of growth leading to a tetragonal distortion

### 4.2.2 Critical Thickness

When the  $\text{Si}_{1-x}\text{Ge}_x$  layer growth thickness is increased there exists a critical thickness ( $h_c$ ) beyond which the pseudomorphic growth is hampered and associated strain effects vanish. Defects begin to appear due to broken bonds and relaxation of the tetragonal deformation occurs. Above  $h_c$  the layer keeps losing the built in strain. This degrades the electronic properties of the grown  $\text{Si}_{1-x}\text{Ge}_x$  layer to its bulk values [80]. For thickness above the critical value the grown  $\text{Si}_{1-x}\text{Ge}_x$  film relaxes to its bulk lattice constant. The band gap changes to that of the bulk  $\text{Si}_{1-x}\text{Ge}_x$  shown by the upper curve of Figure 4-3. It can be seen in Figure 4-3 that the band gap reduction in relaxed layer is much less compared to the strained layer. This will have serious consequences as will be discussed in sub-section 4.3.2. The critical thickness is related to the Ge fraction as well as growth conditions like temperature and pressure..

The critical thickness calculations most commonly quoted in literature are based on an equilibrium model of Mathews and Blakeslee (MB theory) [129]. Critical thickness  $h_c$  falls off with increasing strain (or increasing  $x$ ), increasing CVD growth temperature and increased rate of growth. Out of the orientations of silicon the (100) plane gives the maximum critical thickness for an overgrown  $\text{Si}_{1-x}\text{Ge}_x$  film [130], [131]. Figure 4-2 shows a sketch of this critical thickness for  $\text{Si}_{1-x}\text{Ge}_x$  grown over (100) silicon substrate. Solid line is from Mathews-Blakeslee theory. The dashed line is from equation. ( 4.2) given below. Defect free films of higher thickness than predicted by MB theory has been reported [ 114], [130].. These films are experimentally grown at 550°C. For  $\text{Si}_{1-x}\text{Ge}_x$

grown over (100) silicon substrate at low temperatures the critical thickness can be expressed by a semi-empirical formulation as a function of Ge fraction  $x$  [23], [132],

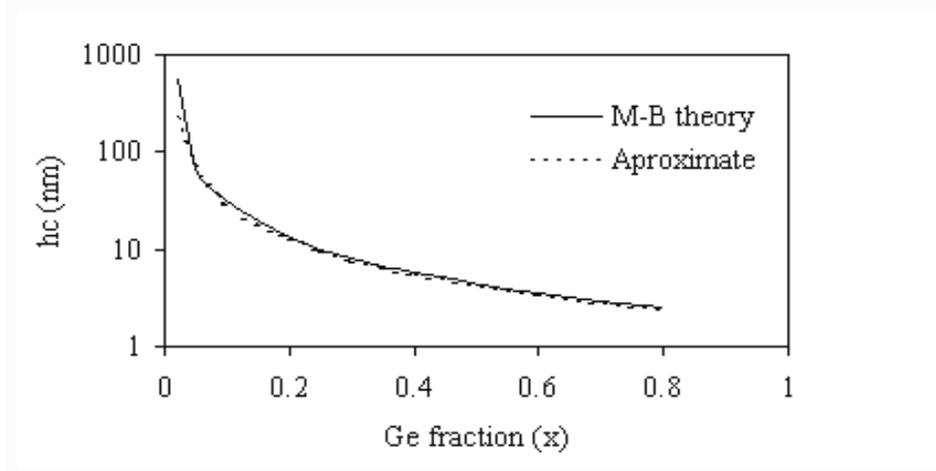


Figure 4-2. Critical thickness  $h_c$  as a function of Ge fraction  $x$ . Solid line is from Mathews - Blakeslee theory. The dashed line is from equation ( 4.2)

$$h_c \approx 1.7793 x^{-1.2371} \text{ nm} \quad (4.2)$$

It has been found that low temperature growth at 550°C yields higher  $h_c$  than high temperature growth at 900°C [129], [131]. If the Ge content in the film is graded, from zero to the desired value, during growth, the resulting dislocation density at the surface can be reduced to as low as  $10^6/\text{cm}^2$  [111]. This can also increase the critical thickness [135]. Compared to the above defect density the best  $\text{SiO}_2$  interface defect density is around  $10^9/\text{cm}^2$ . Surface dislocation densities as low as  $2000/\text{cm}^2$  have been experimentally demonstrated in SiGe [137]. Grading the Ge fraction can be used to advantage in locating the conduction channel away from interfaces as will be discussed in the sub-section 4.5.2.

### 4.3 Compressive Strain Induced Effects in $\text{Si}_{1-x}\text{Ge}_x$

As discussed before if a  $\text{Si}_{1-x}\text{Ge}_x$  layer less than the critical thickness  $h_C$  is grown over a Si substrate there will be compressive in plane strain. A few important changes occur if a thin  $\text{Si}_{1-x}\text{Ge}_x$  film is pseudomorphically grown along the [100] direction over a silicon substrate. (1) there is a band gap reduction, (2) there is band splitting in the valence band, (3) there is a lowering of the effective mass (4) there is a rise in hole mobility. The valence band characteristics and the hole mobility are considered here in detail since these are important in the design of a p-HMOSFET. Direct computation of the compressive strain and its inclusion in the numerical simulator or analytical program is not easy since it involves energy band calculations. The numerical simulator ISE-TCAD [109] that is used to verify the analytical design program uses published values for the parameters of strained  $\text{Si}_{1-x}\text{Ge}_x$ . Since the experimental values are only available for some discrete Ge fraction  $x$  the values in between are interpolated. The analytical analysis and design program, of chapter 6, uses equations (4.4), (4.5) and (4.9) to compute valence band offset, band gap and mobility respectively. Table 4.1 gives some of the relevant properties of  $\text{Si}_{1-x}\text{Ge}_x$  pseudomorphically grown on (100) Si substrates.

#### 4.3.1 Band Gap Engineering Using $\text{Si}_{1-x}\text{Ge}_x$

A thin film of  $\text{Si}_{1-x}\text{Ge}_x$  grown over silicon (100 plane) experiences compressive strain due to the larger lattice constant of  $\text{Si}_{1-x}\text{Ge}_x$ . This compressive strain causes the band gap to drop monotonically as shown by the lower curve in Figure 4-3. The drop in the band gap for this thin film has a much steeper slope compared to the bulk  $\text{Si}_{1-x}\text{Ge}_x$  discussed earlier. This drop in band gap is nearly linear up to an  $x$  of about 0.5.

The band gap drops from 1.12 eV for pure silicon to nearly 0.7 eV for a x about 0.5 [ 108]. This gives a wide range of energy for band gap engineered devices such as p-HMOSFET. The key to band gap engineered devices is the ability to alter or

Table 4-1. Comparison of some of the most important parameters of  $\text{Si}_{1-x}\text{Ge}_x$  pseudomorphically grown over (100) Si substrate. [88], [ 111], [121]

Parameter	Si	Ge	$\text{Si}_{0.7}\text{Ge}_{0.3}$ relaxed	$\text{Si}_{0.7}\text{Ge}_{0.3}$ strained	GaAs	Units
Lattice Constant (A)	5.431	5.658	5.493	5.431	5.653	A
Crystal structure	Diamond	Diamond	Diamond	Diamond	Zinc blende	
Energy gap	1.12	0.66	1.04	0.9	1.42	eV
Valence band offset				0.22		eV
Dielectric constant	11.7	16.2		13.05	13.2	
Intrinsic carrier density	$1 \times 10^{10}$	$2 \times 10^{13}$	$1 \times 10^{12}$	$1 \times 10^{12}$		
Hole mobility	450	1900	350	885	400	$\text{cm}^2/\text{Vs}$
$N_V$	$1.8 \times 10^{19}$	$5 \times 10^{18}$	$1.2 \times 10^{19}$	$6.2 \times 10^{18}$		
$m_{hh}$	0.54	0.33	0.4	0.257	0.38	$m_0$
$m_{lh}$	0.15	0.043	0.2		0.38	$m_0$
Electron affinity	4.05	4.0				eV
vsat	$1 \times 10^7$			$1.13 \times 10^7$		$\text{cm/s}$
Thermal conductivity	148	60		8		$\text{W/cm K}$

design the desired band gap and then pseudomorphically grow these materials of different band gap energies. This causes valence band discontinuities. This band discontinuities can be used to form quantum wells. So far band gap engineering was not

feasible in silicon and remained a sole design tool for the III-V based devices.  $\text{Si}_{1-x}\text{Ge}_x$  thus brings into the all silicon semiconductor industry the ability to design devices using band gap engineering. Figure 4-3 shows the energy gap  $E_g(\text{Si}_{1-x}\text{Ge}_x)$  as a function of the Ge fraction  $x$  in the film. The bottom two curves are for strained  $\text{Si}_{1-x}\text{Ge}_x$  thin films. For strained  $\text{Si}_{1-x}\text{Ge}_x$  the heavy hole (hh) band moves up thereby reducing the band gap as shown in Figure 4-6. There are many empirical equations in literature for the computation of the band gap of strained  $\text{Si}_{1-x}\text{Ge}_x$  as a function of the Ge fraction in the alloy. The band gap reduction is mostly in the valence band for compressively strained  $\text{Si}_{1-x}\text{Ge}_x$  and is negligible in the conduction band (less than 20 meV). For strained  $\text{Si}_{1-x}\text{Ge}_x$  grown over (100) silicon the band gap reduction can be computed by the following empirical formula [80],

$$\Delta E_g = 0.96x - 0.43x^2 + 0.17x^3 \text{ eV} \quad (4.3)$$

The band gap can also be computed by knowing the valence band offset as described in section 4.3.2 below. There are many published formulas for band gap that differ from each other. But the value of the upward shift of the valence band seems to be in agreement in many publications [ 86], [ 110]. Figure 4-3 shows the effect of increasing the fraction of Ge on the band gap of bulk and coherently strained  $\text{Si}_{1-x}\text{Ge}_x$  thin film grown on Si (100) substrate. Figure 4-4 shows plots of the strained  $\text{Si}_{1-x}\text{Ge}_x$  band gap as computed using equation (4.5) and the piecewise linear interpolation used in the numerical simulator ISE-TCAD [109]. ISE-TCAD uses published experimental values of the band gap for strained  $\text{Si}_{1-x}\text{Ge}_x$  films grown over (100) silicon. The direct inclusion of

strain effects is a formidable task and there is no publication found in literature to do that. The closest one gives some band gap coefficient computations. The approximate computation using the experimentally fitted empirical equation (4.5) seems to match the simulator values fairly close up to a x of 0.5.

Hence the equation (4.5) assuming a linear reduction of valence band is used in the analytical program developed in this work. Since the linear equation gives results close to the numerical simulator values analytical HMOFET design program results can be compared to the numerical simulation results. It may also be seen that the curves in Figure 4-3 and Figure 4-4 are in fairly close agreement for the strained  $\text{Si}_{1-x}\text{Ge}_x$  films.

#### 4.3.2 Band Edge Alignment in Si/Si<sub>1-x</sub>Ge<sub>x</sub>

When  $\text{Si}_{1-x}\text{Ge}_x$  is grown over (100) silicon the band gap reduction of the  $\text{Si}_{1-x}\text{Ge}_x$  leads to a discontinuity in the band gap at the interface. The band alignment is Type I as shown in

Figure 4-5 [80], [110], [111]. The band edge discontinuity in conduction and valence bands are represented by  $\Delta E_c$  and  $\Delta E_v$  respectively. For all Ge fractions of  $\Delta E_c$  is less than 20 meV which is less than the thermal voltage  $kT/q$  and can be neglected [80], [110]. For  $\text{Si}_{1-x}\text{Ge}_x$  with Ge fractions less than 0.85 grown over a (100) silicon substrate  $\Delta E_v$  can be expressed as [15], [79], [110], [111], [111],

$$\Delta E_v = 0.74x \quad (4.4)$$

where x is the fraction of Ge content.

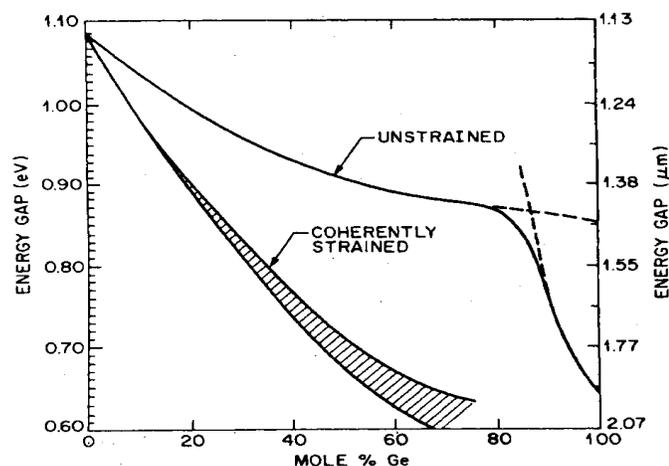


Figure 4-3. Band gap of  $\text{Si}_{1-x}\text{Ge}_x$  as a function of the Ge fraction  $x$ . The top line is for bulk or relaxed  $\text{Si}_{1-x}\text{Ge}_x$ . For thin strained  $\text{Si}_{1-x}\text{Ge}_x$  films (bottom two curves) the band gap reduction is more pronounced [114].

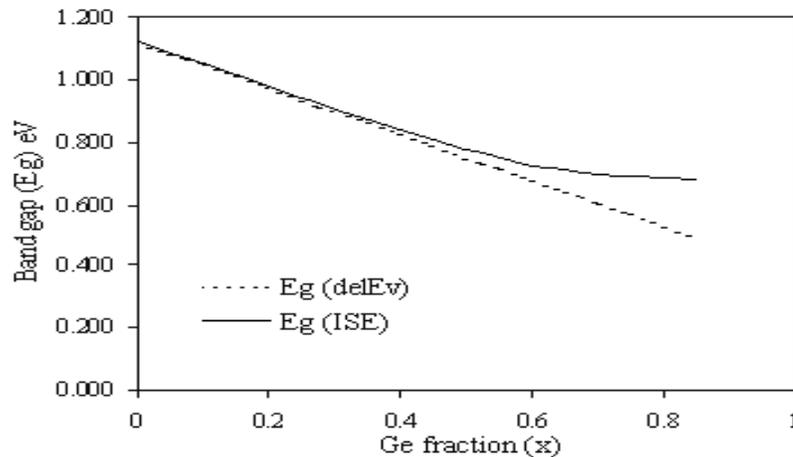


Figure 4-4. Band gap  $E_g(\text{Si}_{1-x}\text{Ge}_x)$  as a function of Ge fraction  $x$ . Using the linear interpolation used in ISE-TCAD (solid), Using the linear approximation of equation (4.5) (dashed)

Using the equation (4.4) the band gap of strained  $\text{Si}_{1-x}\text{Ge}_x$  can be approximated as,

$$E_g(\text{SiGe}) \approx E_g(\text{Si}) - 0.74x = 1.124 - 0.74x \quad (4.5)$$

where  $E_g(\text{Si})$  is the band gap of Si which is 1.124 eV. It is evident that a higher Ge fraction leads to a higher valence band discontinuity. It is to be noted that the band offsets are crystallographic orientation dependent [110]. If a quantum well is formed by sandwiching the  $\text{Si}_{1-x}\text{Ge}_x$  layer between two silicon layers the band edge discontinuity allows confinement of holes in the  $\text{Si}_{1-x}\text{Ge}_x$  layer. This confinement is used in the design of p-HMOSFET. The higher the Ge content the higher the valence band discontinuity and higher the confinement.

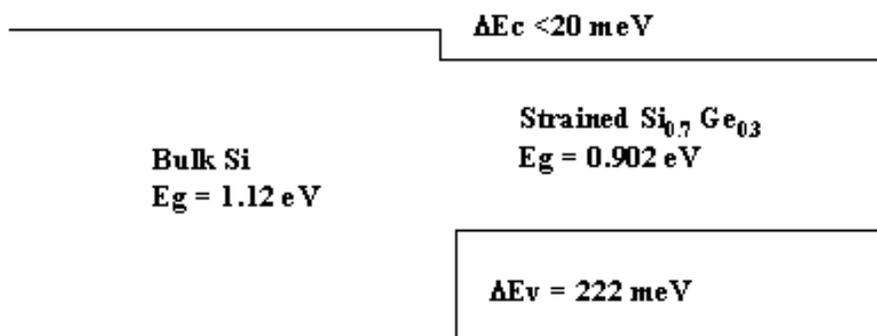


Figure 4-5. Band edge alignment diagram for a strained  $\text{Si}_{1-x}\text{Ge}_x$  film grown over (100) silicon for an  $x$  value of 0.3. The alignment is of type I with a valence band discontinuity of approximately 222 meV and a negligible conduction band discontinuity which is less than 20 meV

### 4.3.3 Transport Properties

The transport properties of compressively strained  $\text{Si}_{1-x}\text{Ge}_x$  thin films are different from that of silicon. This includes effective hole mass, hole mobility and density of states.

#### 4.3.3.1 Effective Hole Mass

Effective hole mass depends on the shape of the valence band. Figure 4-6 shows the comparison of valence bands in unstrained and strained  $\text{Si}_{1-x}\text{Ge}_x$  grown over (100)

silicon substrate. In strained  $\text{Si}_{1-x}\text{Ge}_x$  the light hole band moves down and the heavy hole band moves up. The upward shift of heavy hole band may suggest that the effective hole mass may increase with strain. But a careful observation shows that strain also causes a warping in the heavy hole band. Many publications have reported the heavy hole band near the center of the Brillouin zone to be highly non-parabolic [ 90], [ 91]. The effective mass of the hole in the heavy hole (hh) band is given by,

$$m_{hh}^* = \left( \frac{\hbar}{2\pi} \right)^2 \frac{1}{d^2 E / dk^2} \quad (4.6)$$

The warping of the heavy hole band causes an increase in the  $d^2 E / dk^2$  term where E is the energy and k is the momentum vector. Since the effective hole mass is inversely proportional to the above term this warping in turn reduces the heavy hole mass. Significant reduction in heavy hole effective masses with increasing Ge fraction have been reported [ 85], [ 110]. It is reported that for a x value of 0.3 the effective mass reduces by a factor of 3.4 [ 83]. Effective masses for the hh (raised band) and lh (lowered band) are given by the equations ( 4.7) and ( 4.8) respectively [ 92],

$$\frac{m^*}{m} = -2.84x^3 + 4.68x^2 - 2.87x + 0.89 \quad (4.7)$$

$$\frac{m^*}{m} = -0.14x^3 + 0.36x^2 - 0.37x + 0.25 \quad (4.8)$$

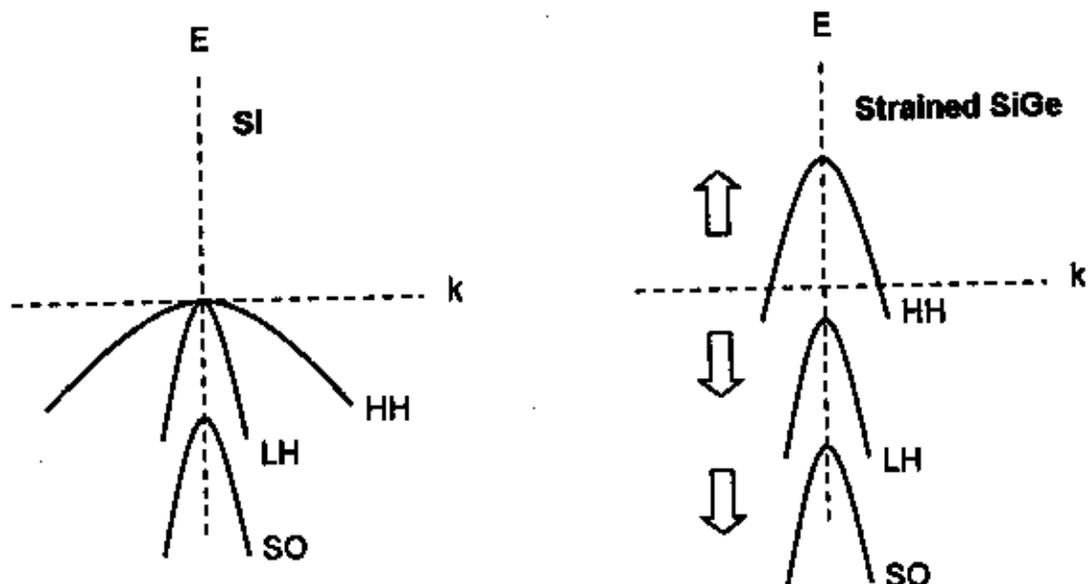


Figure 4-6. The valence band diagrams for unstrained (left) and strained (right)  $\text{Si}_{1-x}\text{Ge}_x$  grown over (100) silicon substrate. Notice the warping of the heavy hole band and the lowering of the light hole band with strain [ 111].

#### 4.3.3.2 Hole Mobility

Mobility is limited by scattering mechanisms. The main scattering mechanisms in silicon are ionized impurity scattering, acoustic phonon, optical phonon, interface impurities and interface roughness. The ionized impurity scattering is unimportant because undoped  $\text{Si}_{1-x}\text{Ge}_x$  is used in the p-HMOSFET for the quantum well hole transport channel. In  $\text{Si}_{1-x}\text{Ge}_x$  alloy scattering due to the difference in potentials of Si and Ge atoms also should be taken into account. The reported studies still are not in agreement on the importance of alloy scattering amidst other scattering mechanisms, particularly phonon scattering. Some theoretical studies argue that alloy scattering is not important compared to the acoustic phonon at room temperature [ 82], [ 83]. But in some other studies the inclusion of alloy scattering seems to reduce the mobility by a factor of nearly four [ 84],

[ 85]. The difference of these results stems from the fact that the alloy potential used in the calculation varies widely from 0.2 to 0.8 [ 86]. One of the above theoretical studies also suggests that alloy scattering is only important below 100° K where the phonon scattering becomes comparable to it [ 82]. The study also finds that surface roughness and surface impurities do not contribute much towards hole mobility reduction. The mobility also reduces with increase in carrier (hole) concentration in the QW. For a x value of 0.2 the hole mobility seems to reduce from approximately 1000 to 600 for a typical carrier density variation from 0 to  $10^{12}/\text{cm}^2$  [ 82]. So the acoustic phonon, optical phonon and carrier density seem to have adverse effects on mobility at room temperature.

At the center of the Brillouin zone the heavy hole and light hole bands are degenerate in silicon, with a degeneracy of four. The lifting of the heavy hole band reduces the degeneracy to two. From published simulations it has been predicted that over 75% of holes will occupy the raised hh band for any appreciable Ge fraction in  $\text{Si}_{1-x}\text{Ge}_x$  [ 90]. Due to these factors inter valley scatter for holes in the raised heavy hole band is reduced. Reduction of the hole mass and reduction of the inter valley scattering lead to an effective increase in hole mobility in the strained  $\text{Si}_{1-x}\text{Ge}_x$  film. Intrinsic undoped silicon has a hole mobility of around  $470 \text{ cm}^2/\text{V}\cdot\text{sec}$ . Intrinsic Ge has a hole mobility of around  $1900 \text{ cm}^2/\text{V}\cdot\text{sec}$ . The hole mobility in intrinsic strained  $\text{Si}_{1-x}\text{Ge}_x$  can be approximated by a linear dependence to the Ge fraction as [ 83], [109],

$$\mu_h \approx 470 + (1900 - 470)x \quad (4.9)$$

The above is low field mobility. The numerical program ISE-TCAD also uses the same linear model. As the field along the transport direction is increased the mobility reduces. This field dependence can be computed using Caughey-Thomas formula.

$$\mu_h(E) = \frac{\mu_{low}}{\left[1 + \left\{\frac{\mu_{low} E}{v_{sat}}\right\}^\beta\right]^{1/\beta}} \quad (4.10)$$

where  $\mu_h(E)$  is the effective mobility.  $\mu_{Low}$  is given by equation (4.9), the saturation velocity  $v_{sat}$  for holes in  $Si_{1-x}Ge_x$  is  $8.37 \times 10^6$  cm/s and  $\beta$  is 1.2.  $E$  is the electric field along the transport direction or the parallel field in V/cm. Hole mobility in bulk silicon and strained silicon germanium are plotted against  $E_{par}$  in Figure 4-7. At low fields the mobilities in  $Si_{1-x}Ge_x$  are far higher than the mobility in silicon. But in short channel devices where the field values along transport direction may exceed  $1 \times 10^5$  V/cm mobilities in strained  $Si_{1-x}Ge_x$  and silicon fall off to nearly equal values. This is because the saturation velocity for holes in  $Si_{1-x}Ge_x$  is taken to be the same as in silicon, equal to  $1 \times 10^7$  cm/s [105].

Quite curiously the above fact is not seen discussed in literature. This is an important point to note that the advantages of low field higher mobility may be lost in scaled down velocity saturated devices. Since no published experimental data is available this result cannot be taken for granted but should be kept in mind. There are published results that indicate a rise in hole saturation velocity  $v_{sat}$  in  $Si_{1-x}Ge_x$  with increase in the value of  $x$  [88].  $v_{sat}$  value of  $2.1 \times 10^7$  cm<sup>2</sup>/V-s has been reported for Ge fraction  $x$  of 0.5 [88].  $Si_{1-x}Ge_x$  devices have the advantage that a much more portion of the channel will be velocity

saturated compared to silicon short channel devices. This is due to its higher low field mobility. As an example for a Ge fraction  $x$  of 0.3 an increase in mobility by a factor of 2.4 compared to silicon has been reported [ 83]

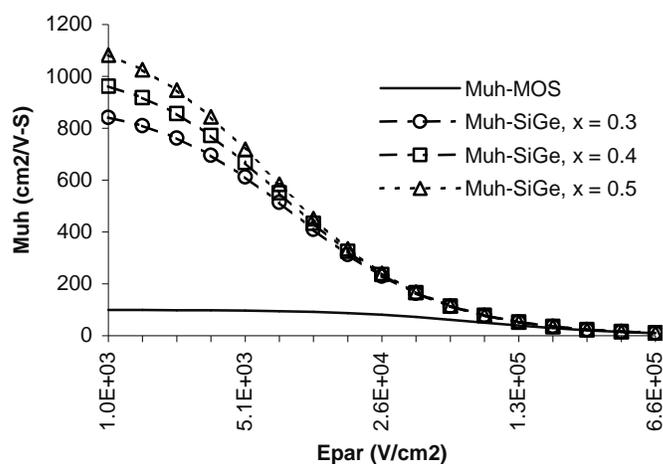


Figure 4-7. Hole mobility as a function of the electric field along the direction of transport. The field dependence is computed using Caughey-Thomas formula. The mobilities are for low doped silicon MOS and strained  $\text{Si}_{1-x}\text{Ge}_x$  with  $x$  values of 0.3, 0.4 and 0.5.

#### 4.3.3.3 Velocity Saturation

There are expectations that the saturation velocity,  $v_{\text{sat}}$ , may increase with the Ge fraction  $x$  [88], [ 89]. There is a reduction of the effective hole mass with the addition of Ge to Si. In the bulk  $\text{Si}_{1-x}\text{Ge}_x$  the effective hole mass changes only slightly with  $x$  but in strained layers the drop is significant [ 85]. This is due to the warping of the hh band as discussed in section 4.3.3.1. This allows carriers to be accelerated to velocities higher than the saturation velocity ( $v_{\text{sat}}$ ). This is explained as follows. Since the addition of Ge retains the Si like band structure the optical phonon retains a high energy Si like spectra.

But the addition of Ge reduces the effective mass. This means the velocity of holes has to increase before carriers can interact with the optical phonon since momentum of holes has to be increased to the same level as in silicon. In terms of physics this can be written as,

$$\frac{1}{2}mv^2 = n\frac{h}{2\pi}\omega_p \quad (4.11)$$

Where  $m$  is the effective mass,  $v$  is the velocity,  $h$  is the Plank's constant,  $\omega_p$  is the phonon frequency and  $n = 1, 2, \dots$ . Since  $\omega_p$  is unchanged and  $m$  has reduced  $v$  has to increase. It is to be noted here that strain may also change the phonon spectra, most probably to higher values.

Fabricated p-HMOSFET have been reported to show 50% increase in  $v_{\text{sat}}$  at a Ge concentration  $x$  of 0.5 compared to Si control p-MOSFET [88]. This was found to be particularly true in short channel (200 nm) devices where most of the channel gets velocity saturated. In compound semiconductors the maximum velocity to which carriers can be accelerated is determined by the phonon emission or collision with the lattice. As discussed above the phonon scattering occurs at a higher velocity in  $\text{Si}_{1-x}\text{Ge}_x$ . Thus it is speculated that  $v_{\text{sat}}$  could possibly increase with  $x$  in  $\text{Si}_{1-x}\text{Ge}_x$ . [88], [ 89]. Since the cut off frequency of a MOSFET is directly proportional to the carrier velocity  $\text{Si}_{1-x}\text{Ge}_x$  HMOSFETs can operate at higher frequencies than Si MOSFETs. Since mobility is inversely proportional to effective mass the mobility increases with increasing values of  $x$ . The variation of mobility with  $x$  seems to be near linear from that of silicon to that of germanium [ 83].

Another observation is that for fabricated p-HMOSFET an increase in  $g_m$  of about 50% has been measured over Si control devices. This is also found to increase with Ge fraction  $x$  and with shorter channel length [ 81], [88]. In Figure 4-7  $v_{sat}$  for  $Si_{1-x}Ge_x$  and silicon are assumed to be the same. According to Figure 4-7 both control and  $Si_{1-x}Ge_x$  devices should have given same  $g_m$  in short channel devices since they both would have had the mobility degraded to the same value. Hence this experimental observation supports the fact that the value of  $v_{sat}$  increases with Ge fraction  $x$ . It is also speculated that the increase in mobility saturates around a Ge fraction of 0.4 – 0.5 due to alloy scattering.

#### 4.3.3.4 Valence Band Density of States

The hole density of states is altered due to the valence band changes with the addition of Ge. As noted earlier the more the Ge content the more the heavy hole band moves up along with more warping of the band. Due to this a reduction in density of states (DOS) for holes can be expected with increasing  $x$  values [ 91]. At the center of the Brillouin zone DOS drops sharply and almost linearly to about half the value in silicon for an  $x$  value of 0.05. Thereafter the drop is slow and near linear. At a  $x$  value of 0.3 DOS falls to a fifth of the value in silicon [ 83]. This reduction in DOS reduces the density of hole confinement in the quantum well channel. However this reduction does not look realistic particularly the large reduction at very low value of  $x$ . A theoretical simulation gives the following formula for DOS [ 91],

$$N_V = 2[m_U^{3/2} + m_M^{3/2} \exp\left(-\frac{\Delta E_V}{kT}\right)\left(\frac{kT}{h^2}\right)^{3/2}] \quad (4.12)$$

$m_U$  and  $m_M$  are the upper (hh) and lower (lh) hole effective masses respectively. For a  $x$  value of 0.1 to 0.3 normalized  $N_V$  drops from 0.3 to 0.2 [ 83] and an average value of 0.25 could be used in calculations.

#### 4.3.4 Summary of Strain Induced Effects

In summary increasing Ge content improves mobility but at higher values the alloy scattering is speculated to limit the hole mobility. For intrinsic  $\text{Si}_{1-x}\text{Ge}_x$  this limiting seems to happen around an  $x$  value of 0.5. The effective mass for holes reduces with increasing  $x$ . This is speculated to be due to the warping of the hh band. With increasing  $x$  values the hh band moves upwards and the warping of the band increases. The reduced effective mass brings up the speculation that holes can be accelerated above the  $v_{\text{sat}}$  of silicon since the holes have to catch up with the momentum before interacting with phonons which retain the silicon like high energy spectra. There seems to have evidence in the rise of  $v_{\text{sat}}$  with Ge fraction  $x$ . Since the hh band moves up and lh band moves down the degeneracy is reduced to two. This will reduce inter-valley scattering. The density of states for holes also decreases with increasing fraction of Ge. This is due to the warping of the raised hh band. But the reduction seems to be slow above a fraction of 0.05 or for any practical value of  $x$ . Both these factors lead to an observation that a high percentage of Ge may not be desirable. In chapter 7 the maximum useful range of  $x$  is found to be around 0.5.

The compressive strain reduces the band gap of  $\text{Si}_{1-x}\text{Ge}_x$ . The reduction in the band gap is nearly linear with increasing Ge fraction,  $x$ . This reduction in band gap can be used in the fabrication of QWs. Since the band gap keeps reducing with increasing Ge fraction,

$x$ , the band gap can be engineered to the desired value. It is to be noted that the critical thickness reduces with increasing  $x$ . Films grown above the critical thickness will lose the built in strain and the advantages of strained  $\text{Si}_{1-x}\text{Ge}_x$  will be lost.

#### **4.4 Introduction to p-HMOSFET**

The device characteristics of interest depend on the application. The following are important in digital systems. (1) small size, (2) high ON current, (3) low switching power loss, (4) low OFF currents and (5) small gate capacitance. Small size allows very large density of integration that improves the yield from each die. High  $I_{\text{on}}$  is required to charge the gates at the output node faster in order to increase the switching speed. It is to be noted that high  $I_{\text{on}}$  can be obtained by increasing the width  $W$  of the device but doing so also increases the gate capacitance and hence the gate delay. In CMOS systems the power loss  $P_{\text{D}}$  happens only at switch charge time. In order to reduce the quiescent standby power loss and battery drain  $I_{\text{off}}$  must be as low as possible. The gate capacitance of the device should be small so that they can be charged faster improving the gate delay. Instead if this device is to be used in microwave systems like amplifiers, oscillators and so on the linearity of the gate characteristic ( $I_{\text{ds}}$  vs  $V_{\text{gs}}$ ), high  $I_{\text{on}}$  and low noise or low noise figure (NF) are important. Linearity is important to reduce distortion of the signal. Linearity of the circuit could be achieved by feedback also. High  $I_{\text{on}}$  is required to drive the stages to follow. The noise figure (NF) is the ratio of signal to noise ratio (SNR) at the input to SNR at the output. Low NF is important while processing low level signals which are common in RF and analog systems like receivers.

In section 4.3.1 it was mentioned that the band gap of a thin strained  $\text{Si}_{1-x}\text{Ge}_x$  layer grown over (100) silicon substrate experiences a reduction in band gap compared to the band gap of silicon. Simultaneously the band edges of silicon substrate and  $\text{Si}_{1-x}\text{Ge}_x$  form a type I band edge alignment as shown in

Figure 4-5. The band discontinuity at the conduction band is negligible whereas the same at the valence band can be many times the thermal voltage  $kT/q$  depending on the fraction of Ge,  $x$ . This suggests that a quantum well for holes can be formed if a thin strained  $\text{Si}_{1-x}\text{Ge}_x$  layer is sandwiched between a silicon substrate and an overgrown silicon cap layer. If this quantum well is positioned inside the depletion region of a silicon MOSFET holes can be confined into it and transported through this quantum well layer instead of the inferior surface inversion layer of a MOSFET. This can dramatically increase the hole mobility due to three aspects. (1) ionic Coulomb scatter and associated reduction in mobility can be reduced since the QW need not be doped. (2) the pseudomorphically grown  $\text{Si}/\text{Si}_{1-x}\text{Ge}_x$  interface has far less defect densities and hardly any trapped charges compared to the inferior  $\text{Si}/\text{SiO}_2$  interface [183]. (3) the hole mobility in the  $\text{Si}_{(1-x)}\text{Ge}_x$  is higher than the hole mobility.

However the idea of confining carriers into an undoped quantum well layer and transporting through this undoped layer is not new. In 1990s research on III-V superlattices led to a new device named MODFET (modulation doped field effect transistor) where an undoped GaAs quantum well layer has been used for the above purpose. n-channel devices with very high transconductance, lower switching voltages, lower power dissipation and switching speeds in picosecond range could be produced

[ 119] [ 116]. These MODFETs are being used as microwave amplifiers, voltage controlled oscillators and high speed digital switching devices in communication systems. A sketch of a basic device of this sort and its band diagram are given in Figure 4-8 and Figure 4-9. Since III-V compounds do not have a good quality insulator equivalent to the superior  $\text{SiO}_2$  in silicon systems the MODFET evolved from MESFET (metal semiconductor field effect transistor) where a Schottky metal gate is used instead of the  $\text{SiO}_2$  insulating gate of the MOSFET (metal oxide semiconductor field effect transistor). This has the disadvantages of lower gate breakdown voltage and higher gate leakage current than MOSFETs. The application of a gate voltage depletes the surface AlGaAs doped layer. These carriers, electrons, diffuse into the undoped GaAs quantum well layer where they get confined in a thin layer adjacent to the AlGaAs-GaAs interface. Until recently these devices coexisted with Si CMOS integrated circuits (ICs), in discrete form, in communication systems since the silicon could not provide the necessary microwave performance. The feasibility to alter the band gap of silicon along with the simultaneous improvement in mobility, by the introduction of germanium, changed this scenario. Instead of GaAs devices high speed bipolar transistors using  $\text{Si}_{1-x}\text{Ge}_x$  base are being used in front end RF systems.

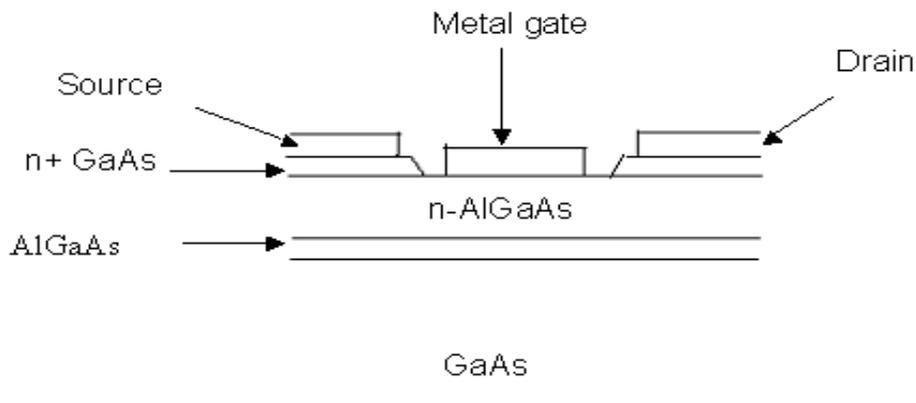


Figure 4-8. Sketch of a GaAs MODFET

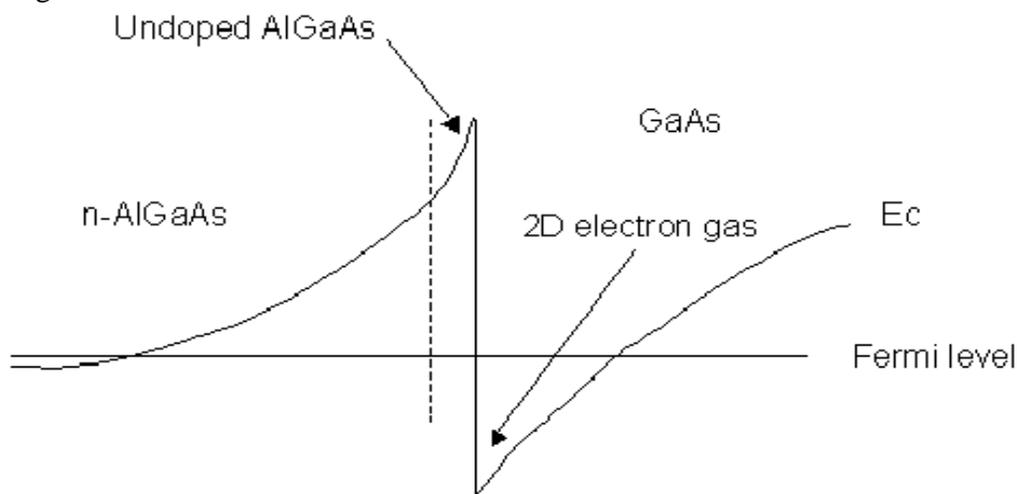


Figure 4-9. Band diagram for the MODFET

The impressive advances and dominance of silicon technology owes a good share to the silicon planar MOSFET process. So in spite of the impressive performance of HBTs mentioned in chapter 1 there is a need for MOSFET devices using  $\text{Si}_{1-x}\text{Ge}_x$ . MOSFETs offer devices with low power, low noise, high frequency operation and very high density of integration. These are desirable for portable battery operated equipments and cellular devices. Silicon grown over  $\text{Si}_{1-x}\text{Ge}_x$  experiences tensile strain. Tensile strain in silicon lowers the in plane electron mass. In addition the band splitting reduces the inter valley

scattering. These two effects together gives enhanced electron mobility in strained silicon [144], [147]. Room temperature low field electron mobility as high as  $2600 \text{ cm}^2/\text{Vs}$  are predicted [122]. As described in section 3.2  $\text{Si}_{1-x}\text{Ge}_x$  grown on (100) silicon gets compressively strained and provides high hole mobility. These two processes together are suitable for hetero complementary MOSFET (HCMOS). The main objective of this research, as stated in chapter 1, is to study, evaluate and select a generalized structural scheme of a p-HMOSFET which is deemed to be simple enough to fabricate and provide a device with superior characteristics compared to Si MOSFET. Hence a p-channel HMOSFET will be discussed further. Some of the important structures found in literature are reviewed below. From these a generalized structure incorporating most of the desirable features is arrived at for study and design in this project.

#### **4.5 Survey of p channel HMOSFETs**

A literature survey of p-HMOSFETs over the past two decades is presented here. This gives an idea of different structural designs. A discussion of these designs also is given. The discussion points out the advantages and disadvantages of certain structures. . Many different structures are discussed. Out of these structures a few seem to appear in recent publications. Here the idea is to come up with a simple structure that can provide a good deal of improvement in performance over the MOSFET. Then the selection of a structure for this work and the justification for the choice is discussed.

##### **4.5.1. Main p-HMOSFET Structures**

Initial interest in the use of  $\text{Si}_{1-x}\text{Ge}_x$  in MOSFET structures started around 1985 [114]. The earlier designs of HMOSFET copied the GaAs MODFET instead of evolving from

MOSFET. So similar to the MODFET the earlier HMOSFET also had a recessed channel layer. But doped layers were introduced on top of the channel layer, top and bottom of the channel layer or had a thick doped surface layer. The doped layer at top needs a spacer layer below it to reduce channel layer Coulomb scatter due to ions in it. Apart from this the doped layer has to be sufficiently thick to contain the full depletion width, to avoid forming an intermediate QW, when full gate voltage is applied. These structures had their disadvantages. The doped layer and the spacer layer above the channel layer in fact degrades the performance of a MOSFET quite a bit. First of all the channel is pushed down deeper and this reduces the gate-channel coupling and yields poor transconductance. It also necessitates high voltage for its operation. The spacer layer has to be very thin to avoid degradation of transconductance further. So the Coulomb coupling of the negative Boron ions in the p-layer degrades the mobility of holes in the

$\text{Si}_{1-x}\text{Ge}_x$  QW channel [72], [122], [148], [148], [194]. Further a undoped cap layer is needed at the top to form the gate oxide. All the above together degrades the performance very much. A sketch of the layering sequence for such a general device is given in

Figure 4-10. This type of device structure however do not seem to appear in publications in the past decade.

A simple modification to the silicon CMOS that can yield an improvement in performance is to strain the silicon channel region. This does not require any new designs. The process flow follows exactly same CMOS fabrication sequence. In the end the source and drain diffusions are etched away and  $\text{Si}_{1-x}\text{Ge}_x$  is deposited using chemical vapor deposition.  $\text{Si}_{1-x}\text{Ge}_x$  has a higher lattice constant than silicon. Due to this reason the

silicon in the channel region gets unilaterally compressively strained along the channel direction. This leads to a type II quantum well at the surface channel region. Such a structure is shown in Figure 4-11. In addition to the ability to trap both electrons and holes the strain also increases the mobility of both carriers. These types of devices are already being manufactured as a precursor to  $\text{Si}_{1-x}\text{Ge}_x$  CMOS [ 49] - [ 71 ]. Straining the silicon channel using  $\text{Si}_{1-x}\text{Ge}_x$  source and drain is reported to give an improvement in mobility of about 20%. Presently this type of devices are in active research.

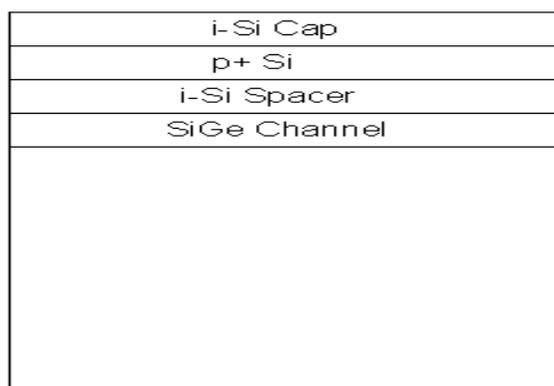


Figure 4-10. Layer sequence for some earlier SiGe p-HMOSFET structures. It has a p+ layer on top of the SiGe QW channel layer. This is an adaptation from the III-V MODFET.

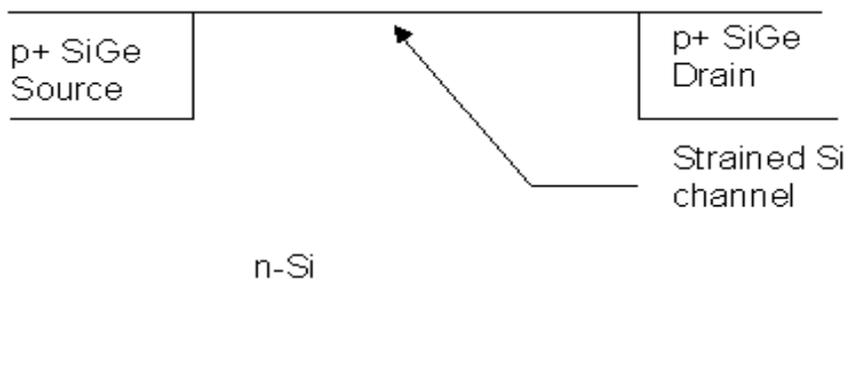


Figure 4-11. Strained silicon channel p-MOS

The increased mobility of a thin layer of strained  $\text{Si}_{1-x}\text{Ge}_x$  can be easily taken to advantage in a simple MOSFET by simply replacing the conducting channel region with  $\text{Si}_{1-x}\text{Ge}_x$ . Even though some improvement in mobility and saturation current  $I_{\text{Dsat}}$  can be achieved the inferior  $\text{SiO}_2$  boundary will degrade the performance very much just like in the case of a MOSFET. In scaled down sub-micron devices the perpendicular field will be very high and this will degrade the performance still further by forcing carriers towards the  $\text{SiO}_2$  interface. However it is easy to manufacture such a device by a simple near surface Ge ion implantation followed by a low temperature anneal. 18% to 80% increase in  $I_{\text{Dsat}}$  for Ge fraction of 0.18 to 0.4 is claimed for such devices [153], [154].

In order to be able to derive the full mobility improvement the strained  $\text{Si}_{1-x}\text{Ge}_x$  channel should be recessed into the substrate and away from the  $\text{SiO}_2$  interface. To do this a thin layer of undoped silicon cap layer can be formed over the recessed  $\text{Si}_{1-x}\text{Ge}_x$  channel layer. This cap layer will also provide the silicon to form the gate oxide. Since the band gap of  $\text{Si}_{1-x}\text{Ge}_x$  is less and because most of the band offset occurs in the valence band a quantum well for holes will form and the collected holes in this undoped QW layer will drift with improved high mobility. If pseudomorphically grown the interface roughness and interface defects will be very less and scattering from the quantum well walls will be much lower. In commonly used UHV-CVD, particularly with hydrogen passivation, the interface impurities are reported to be as low as  $10^6/\text{cm}^2$  [113]. This is lower than the best  $\text{Si}/\text{SiO}_2$  interface in MOSFET. Such low interface impurities are of no concern in devices. Starting with a highly planar silicon surface and using low temperature in the range of 400 – 500 °C very smooth interfaces can be achieved [111].

The layering sequence below the gate oxide, for such a device, will be undoped Si cap/undoped  $\text{Si}_{1-x}\text{Ge}_x$  channel/n-Si substrate [ 89]. [155] - [ 160].

There are references to structures with an additional p+ doped layer below the  $\text{Si}_{1-x}\text{Ge}_x$  channel. This changes the internal potentials and thus can be used to adjust  $V_T$ . A buffer or spacer layer will be needed between the channel and the p+ layer to reduce the Coulomb scattering due to the negative ions in this p+ layer. So the total layering sequence for such a device will be undoped Si cap/undoped  $\text{Si}_{1-x}\text{Ge}_x$  channel/undoped Si buffer/p+ doped Si/n-Si substrate [174]. This structure and the structure in the previous paragraph (without p+ layer) seem to lead the recent publications of Si/SiGe hetero MOSFET. Sketches of such structures are given in Figure 4-12 and Figure 4-13.

There are few other structures reported. But they have complex fabrication sequence like having tapered or graded  $\text{Si}_{1-x}\text{Ge}_x$  substrate or other complications. These structures do not seem to appear in publications of the past decade. Hence they are not discussed here.

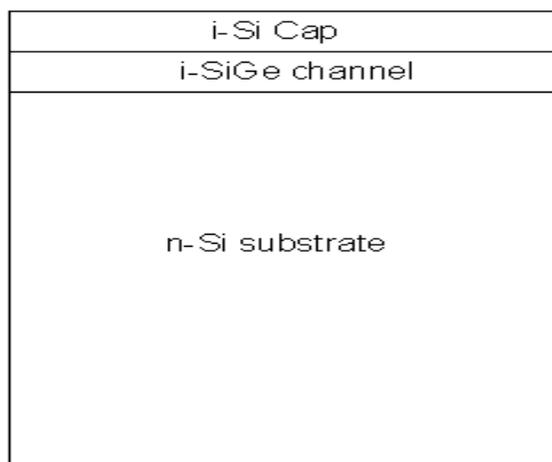


Figure 4-12. A simple structure of p-HMOSFET. It has a recessed SiGe channel

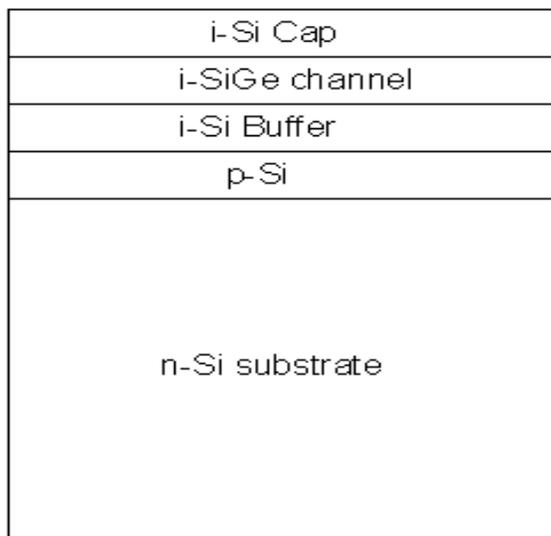


Figure 4-13. A p-HMOSFET structure that seems to have promise for general acceptance. It has an additional p-Si layer that can be used to adjust  $V_T$  and thus increase the hole density in the QW.

#### 4.5.2. Improved p-channel HMOSFET Structures

An improvement to the above recessed channel p-HMOSFET is suggested to keep most of the holes in the  $\text{Si}_{1-x}\text{Ge}_x$  QW layer and reduce leakage to the surface channel. This is done by grading the Ge fraction from a low value at bottom to the desired value at top of the QW channel. This makes valence band of the QW channel triangular in nature with the peak at the top. Though slightly difficult to fabricate this has certain advantages. The grading also increases the critical thickness and thicker QW layer can be grown. The carriers are pushed to the top of the QW where they are closest to the gate. This will give some improvement in the  $I_{D\text{sat}}$ . But it has to be kept in mind that the QW is narrow and making it triangular may lead to quantization and reduce the available levels at the peak of the Ev band [149], [162], [182]. A sketch of such a device is shown in Figure 4-14.

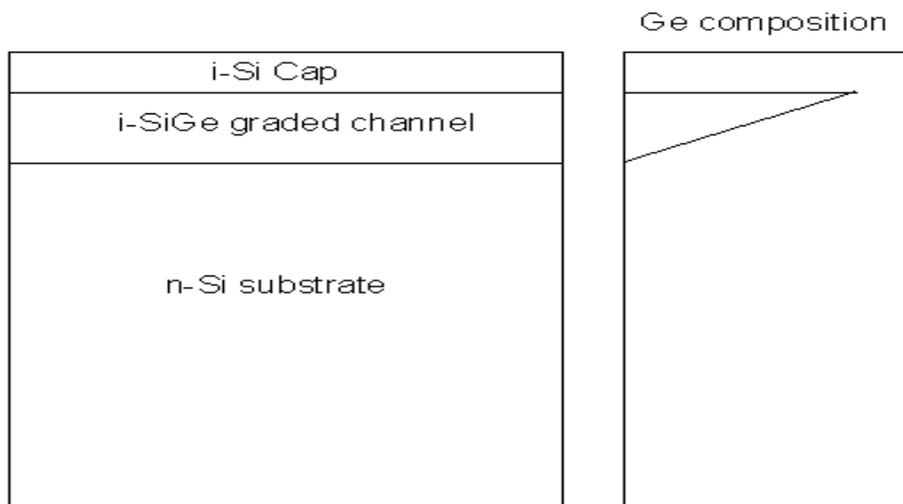


Figure 4-14. A p-HMOSFET structure with graded SiGe channel

In the recent past a merger of SOI (silicon on insulator) and  $\text{Si}_{1-x}\text{Ge}_x$  seem to have caught interest. SOI being a thin body it is much more radiation tolerant than bulk Si substrate. It has also less parasitic capacitances. Since the devices are well isolated and there are no n or p wells latch up is not a problem. This is a very good radiation hardened device which will find applications in space electronics and high radiation environments. The layer sequence for such a device is as follows. Below the gate oxide the structural sequence is Si cap/ $\text{Si}_{1-x}\text{Ge}_x$ /oxide (BOX) [ 162] - [ 165]. A sketch of the device is in Figure 4-15.

Another new design which is getting attention recently is a double gate structure, both in Si and  $\text{Si}_{1-x}\text{Ge}_x$  based devices. In the  $\text{Si}_{1-x}\text{Ge}_x$  structure the undoped strained

$\text{Si}_{1-x}\text{Ge}_x$  layer is sandwiched between two p-Si layers. There is a gate each on either side, top and bottom. On application of gate voltage symmetrically to both gates the

valence band can be made to peak at the center of the  $\text{Si}_{1-x}\text{Ge}_x$  layer. Thus carriers are totally

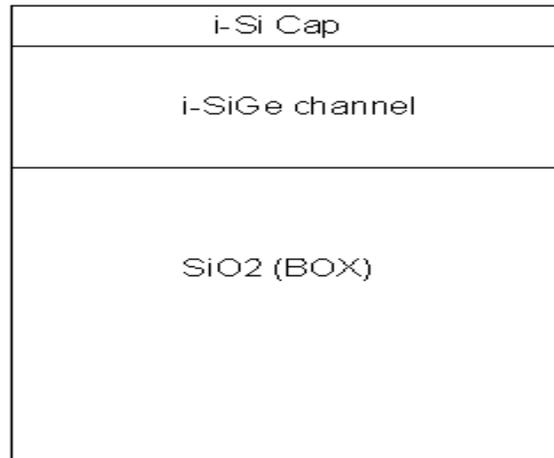


Figure 4-15. Recessed SiGe channel HMOSFET in SOI process.

moved away from interfaces, they are in a low field region, the channel is undoped and the channel is strained. All these together yields very high mobilities in spite of the fact that the gate channel capacitance is lower than that in a similar all Si structure. Due to the high mobility and lower gate capacitance it achieves very high speed of operation [ 33.], [ 34]. Cut off frequency  $f_T$  of 1000 GHz and  $I_{ON}$  of 3.6 mA/ $\mu$  are reported. This is the highest frequency of operation and on currents for  $\text{Si}_{1-x}\text{Ge}_x$  based devices reported in literature. A sketch of the device and band diagram are given in Figure 4-16.

#### 4.6 Generalized Structure for this Project

Some of the most relevant p-HMOSFET structures found in literature are discussed above. Most of the publications center around a particular structure either fabricated or simulated. Due to this fact it is difficult to make a good comparison between them in terms of structural effects on performance. In the analytical analysis and design program

developed in this dissertation is a structure that can accommodate the necessary layers and other parameters discussed in previous subsections. The layering sequence given in

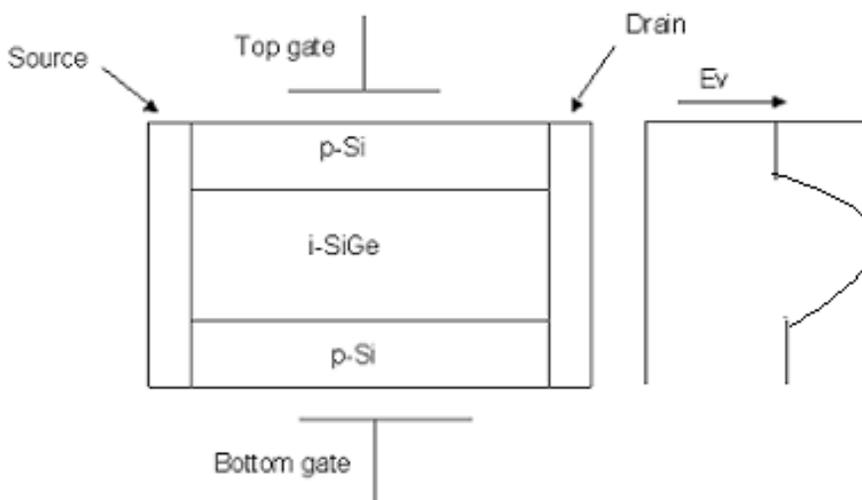


Figure 4-16. Double gated p-HMOSFET. The channel is at the center, away from interfaces.

Figure 4-13 is all inclusive of all previous devices. This however does not include the earliest structures given in the first paragraph of subsection 4.5.1. These are the ones with p+ doped layer above the QW. Some of these are also grown on a relaxed  $\text{Si}_{1-x}\text{Ge}_x$  starting layer instead of silicon. These are more complex to fabricate. These are not considered good structures since they have a doped thick layer on top of the channel layer. In case of relaxed  $\text{Si}_{1-x}\text{Ge}_x$  starting layer the starting surface is highly defective and good interfaces cannot be obtained [80]. So except those earliest structures others can be simulated in the analytical simulation and design program developed in chapter 6 by adjusting the input parameters fed into the program. A consensus of the future device structure is not very clear now. But the published works seem to lead certain structures in fabrication or analytical modeling. Figure 4-13 gives one such structure. These structures

are all merged in the analytical analysis and design program developed so that they can be simulated. For example structure in Figure 4-12 can be simulated by adjusting the other layers and keeping only the cap layer and the QW. This will be shown in chapter 6.

The device proposed in this work has a thin intrinsic silicon cap layer, a thin

$\text{Si}_{1-x}\text{Ge}_x$  quantum well (QW) channel layer, an undoped silicon buffer layer, a p-doped layer and then the n-doped silicon substrate, in that order. A variable thickness p-doped layer is selected instead of p-delta doped layer used in certain other designs. This has the following advantages. The delta-doped layer is most commonly formed by high energy ion implantation which will damage the quantum well layer and cause strain

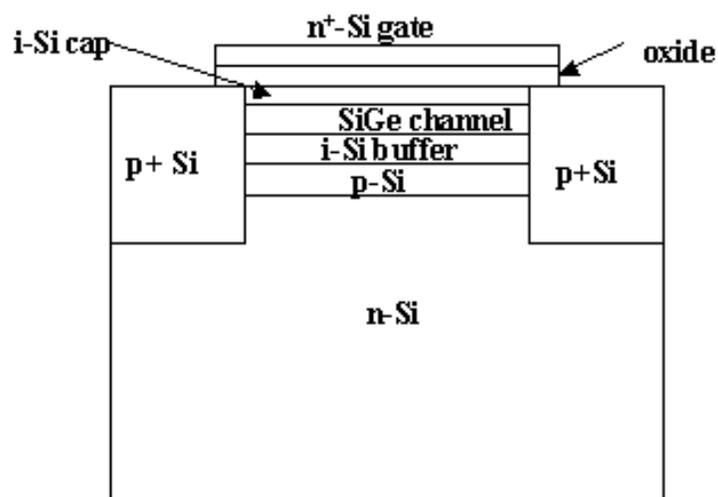


Figure 4-17. Structure of the proposed p-HMOS. It has an undoped silicon cap layer, undoped  $\text{Si}_{1-x}\text{Ge}_x$  quantum well layer, undoped silicon buffer layer and a variable thickness p-doped layer followed by an  $n$ -Si substrate.

relaxation by breakage of interface bonds. This will deteriorate the interface by introducing dangling bonds at the interface and degrade the performance. A variable thickness p-layer can be used as an independent means of threshold voltage adjustment without undesirable doping of the n-substrate, as is done in short channel Si MOSFET. A highly doped p-delta layer is going to be bridging the source to drain. A high doping in this delta layer will reduce the junction barrier at the source. The device structure is shown in Figure 4-17. As indicated above this is a merger of structures found in literature. Even though n+ gate p-MOSFET is more prone to SCE the introduction of  $\text{Si}_{1-x}\text{Ge}_x$  channel can suppress this to some extent. The p-layer can enhance carrier confinement in the QW channel and give the correct threshold voltages along with the n+ gate [174]. Curiously a similar structure is reported in an SOI device without the .

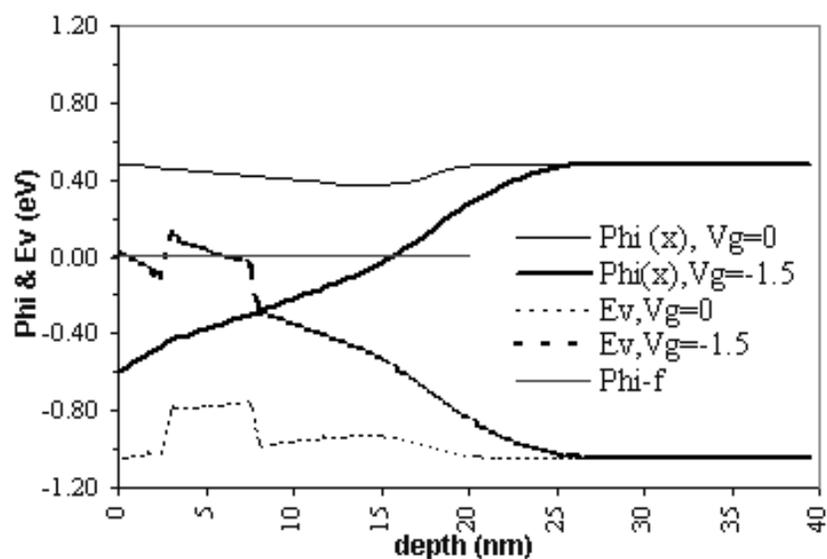


Figure 4-18. Potential (top two curves) and valence band edge (bottom two curves) of a p-HMOSFET for  $V_{gs} = 0$  (outer curves) and  $V_{gs} = -1.5$  V (inner curves). The straight line is the Fermi level.

p-layer. This has the same layering sequence up to the p-layer followed by silicon and bulk oxide [ 179]. Figure 4-18 shows the potential and valence band plots from the surface of the cap layer all the way into the bulk. The smooth curves are potentials. The valence band curves have a visible quantum well between 3 to 8 nm. These two are shown for low and high gate voltages. The straight line is the Quasi Fermi level for holes. This is just repeated here from chapter 7 to give an idea of the band structure of the general device selected for this work.

#### 4.7 Summary

Effects of strain on thin films of  $\text{Si}_{1-x}\text{Ge}_x$  and their use in the design of p-HMOSFET is the main topic of this chapter. The lattice constant of Ge is slightly more than the lattice constant of silicon. So the lattice constant of  $\text{Si}_{1-x}\text{Ge}_x$  steadily increases when the Ge fraction in the alloy is increased. When this alloy semiconductor  $\text{Si}_{1-x}\text{Ge}_x$  is pseudomorphically (lattice matched) grown over a silicon substrate there will be compressive strain in  $\text{Si}_{1-x}\text{Ge}_x$  film in the growth plane.

The compressive strain causes a reduction in band gap. The reduction is mostly in the valence band. This leads to a discontinuity in the valence band. If the  $\text{Si}_{1-x}\text{Ge}_x$  film is grown in between silicon substrate and a top silicon cap layer type I QWs can be formed. Only very thin films of  $\text{Si}_{1-x}\text{Ge}_x$  can be grown since there is a critical thickness beyond which the strain relaxes. The critical thickness for  $\text{Si}_{1-x}\text{Ge}_x$  for a x of 0.5 is only about 7 nm.

The compressive strain also causes a reduction in the effective hole mass, increases the hole mobility and reduces the valence band density of states. There is speculation that

the reduction of effective hole mass will cause an increase in the saturation velocity beyond that of silicon.

Further an introduction to p-channel hetero MOSFET or p-HMOSFET is given. The basic idea is to confine holes in a recessed undoped QW and transport through the QW. The strain induced mobility enhancement and the lack of dopant ions can accelerate holes to higher velocity compared to a MOSFET. If the  $v_{\text{sat}}$  for  $\text{Si}_{1-x}\text{Ge}_x$  is higher the holes can attain velocities higher than the  $v_{\text{sat}}$  in silicon, even in the velocity saturated regime of transport. A literature survey of the recent most popular p-HMOSFET structures is presented. Using this as a basis a simple p-HMOSFET structure with just the necessary structural components is selected for design and analysis in this dissertation. The selection is further justified.

## CHAPTER 5

### DESIGN CONSIDERATIONS OF A P-CHANNEL HMOSFET

Structural design considerations of a p-HMOSFET is discussed in this chapter. The structure of the device is presented and each of the structural components are discussed in detail. The limitations to be imposed on the structural components as well as the range of values for the structural parameters are described in detail. Now a p-HMOSFET can be designed by simply choosing appropriate structural parameter values from within the range for each structural component. The design will be carried out in chapter 6 and chapter 7 where the designed device is simulated.

#### 5.1. Introduction

In chapter 1 basic ideas of hetero junction MOSFET (HMOSFET) were discussed .In chapter 3 and chapter 4 different HMOSFET structures and their simulation results were reviewed. In conventional silicon MOSFET the effective hole mobility is two to three times lower than that of electrons. In order to get the same current as its counter part n-MOSFET the p-MOSFET has to be made two to three times larger in area. This increases the capacitances and slows down the circuit. So, the p-HMOSFET designs are pursued with interest in order to overcome the above drawback by exploiting the higher hole mobility achievable in strained  $\text{Si}_{1-x}\text{Ge}_x$ . But as discussed in chapter 4 most of the designs were not properly focused on their integration into silicon technology and are too complex and not very effective. This is because these designs basically followed the abundant experience available in AlGaAs/GaAs MODFETs using Schottky gate. The

same structure was initially copied into  $\text{Si}_{1-x}\text{Ge}_x$  devices with the exception of an insulated gate. In addition complex substrates were used in many of these devices. As mentioned in sub-section 4.2.2 there is mention of a critical thickness that depends on the Ge content,  $x$ . In order to control this critical thickness as well as to control the strain in the  $\text{Si}_{1-x}\text{Ge}_x$  layer substrates with a different or graded Ge fractions were used. All these together made these designs ineffective, complex and costly to fabricate. Remember that one of the salient strengths of Si-CMOS is its simplicity in fabrication. Later on in the past decade new designs eliminating the above drawbacks began to emerge. Designs started to evolve from MOSFET features instead of the MODFET. These designs are simple enough and are suitable for low voltage operation at or below  $-1.5$  V. In addition, as will be seen in chapter 8, the low leakage current of p-HMOSFET is very attractive for battery operation.

Curiously, a very elaborate and exhaustive discussion of the p-HMOSFET structure with varying layer sequences and properties is reported as early as 1994 [174]. Reference [174] explains in quite detail the significance and choice of gate material, effect of Ge profile in the SiGe channel, threshold adjustment, effect of silicon cap layer and gate oxide. One of the designs mentioned uses a p+ delta layer just below the SiGe channel to adjust potential to increase hole density. A very thin buffer layer is present between channel and the p+ delta layer. The channel profiling used in this case may effectively move carriers to the top of the channel thus reducing the ionic scatter from the p+ delta layer as well as improving gate-channel coupling [182]. Contrary to the majority of earlier publications where p+ delta layer was placed above the channel layer here

placement of the p+ delta layer is recommended below the channel layer. The doping of this layer is used for  $V_T$  adjustment. A qualitatively optimized design for  $0.2\mu\text{m}$  gate length is simulated and a  $V_T$  value of  $-0.35\text{ V}$  is reported. The same device was also fabricated. With grading of the Ge profile during growth thicker channel layers can be grown. Analytical derivation of a similar structure as above, without the p+ delta layer, is available with derivation for the vital parameters [184]. The analysis gives a closed form equation for the threshold voltage. 2D numerical simulation of sub-nanometer structures with Ge fractional grading of the QW channel are also reported [149], [182].

However almost all above designs are either fabricated with fixed parameters, simulations of only some chosen structures or optimization of a device for a single parameter, eg: the cap layer. These simulations do not give insight into the physics happening within the device since only the output parameters are plotted. Hence they cannot be used for diverse designs or optimization. A general structure that seems to catch up momentum in recent publications is as follows. The basic structure consists of gate oxide, silicon cap, SiGe channel, an optional buffer layer, optional p+ delta layer and n-substrate. The dimensions and properties of various layers are not analyzed or studied in detail, in these publications, other than in a heuristic and intuitive fashion.

In Section 5.3 the structure of the device chosen for this project along with the significance of each layer and their influence on the device behavior are discussed in detail. Design variables are identified. Since the HMOSFET has many parameters a notation for the device is suggested to explain the device in a condensed manner.

## 5.2. Material Considerations

The theory of operation of HMOSFET is similar to that of the n-channel GaAs MODFET [ 161]. In the case of p-HMOSFET the carrier transport occurs in the undoped QW region yielding very high mobilities. Since the carrier transport occurs in a thin QW layer the material considerations of this layer are paramount. Factors like interface impurities, interface roughness, interface morphology and layer thickness play a crucial role in the transport. UHVCVD is used to grow the thin  $\text{Si}_{1-x}\text{Ge}_x$  layer [23], [ 161]. It is important that the starting silicon surface is atomically smooth in order to start a lattice matched or pseudomorphic growth of the  $\text{Si}_{1-x}\text{Ge}_x$  layer. The growth is pseudomorphic since the silicon growth surface and the grown  $\text{Si}_{1-x}\text{Ge}_x$  have slightly different lattice constants but have the same crystal structure allowing lattice matched growth with associated strain [ 111]. It can be noted here that a higher Ge fraction can produce more defects like dangling bonds at the interface, which will impede transport of carriers. Higher Ge fraction also puts a limit on the layer thickness due to the reduced critical thickness as discussed in chapter 4. The ultra high vacuum as well as the low temperature growth improves the quality of the interfaces. Hydrogen passivation of the silicon starting surface reduces surface reactivity thus keeping the surface devoid of impurities. Carefully grown QW layers with all of the above growth conditions yield very low defect interfaces with impurities as low as  $10^6/\text{cm}^2$  [ 111 ]. This is better than the best Si/SiO<sub>2</sub> interface which has impurity concentrations of  $10^9/\text{cm}^2$ .

## 5.3. p-channel HMOSFET Structural Details

A sketch of the proposed p-HMOSFET is given in

Figure 5-1. It consists of a n+ poly silicon gate, thin gate oxide layer, thin undoped silicon cap layer, strained undoped SiGe QW channel layer, undoped silicon buffer layer, p-doped silicon layer and a n-doped substrate. Each of these layers has its own significance and has to be designed appropriately. In the following sections each of the above is discussed separately starting from the substrate upwards, which make explanations easier.

### **5.3.1. The Substrate**

The substrate doping is similar to that in conventional MOSFET. The HMOSFET structure is shown in

Figure 5-1. All layers above the p-Si are undoped and can be considered to be nearly insulators. Depletion due to gate voltage occurs, under the p-Si layer, in the n-Si substrate. Mostly this doping controls the threshold voltage. However some control is possible by the p-Si layer doping since this doping can alter the potential distributions within the device. But this p-layer cannot be doped heavily since it will constitute a leakage path from source to drain. This can be seen with the help of the top part of the band diagram given in Figure 5-2. The bottom part of this figure shows the barrier at the source-substrate junction. The barrier at the source junction through the p-layer is less than the same through the substrate. It should also be noted that these top layers are all thin constituting only 18 – 20 nm including the p-Si layer. So for a shallow 30 nm diffusion the bottom half of the source and drain is in contact with the n-Si substrate. This will be a route for sub-surface punch through. Raised source and drain is a solution to this problem. In order to find a minimum value for the n-Si substrate doping, for a particular

channel length, an approximate formula suggested for short channel MOSFET may be used [185] This doping may be sufficient to stop subsurface punch through from source to drain under the gate length of the device. The HMOSFET structure is quite different from the MOSFET structure, so it is not expected that equation ( 5.1) will be very accurate for the HMOSFET. However, this formula is used only as a guideline to estimate the doping where short channel effects begin to be a problem.

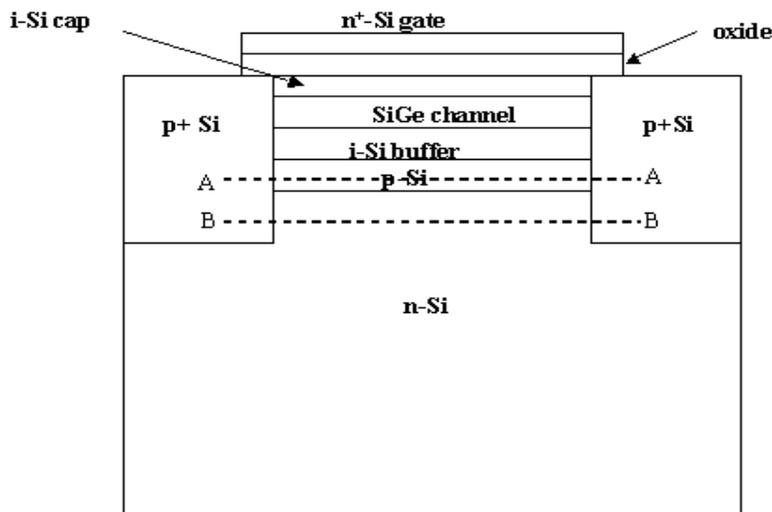


Figure 5-1. Sketch of the generalized p-HMOSFET used in this study. It consists of a n+ gate, gate oxide, undoped Si cap layer, undoped strained SiGe channel layer, undoped Si buffer layer, p-doped Si layer and n-doped substrate.

This estimate is used only as a starting point to discuss short channel effects using the numerical simulator later on.

$$L_{\min} = A(x_j t_{ox} (W_s + W_d)^2)^{1/3} \quad (5.1)$$

where A is a proportionality factor,  $x_j$  is the source and drain junction depth,  $t_{ox}$  is the gate oxide thickness and  $W_s + W_d$  is the sum of the source and drain depletion widths with  $V_{ds} = -1.5$  V.

### 5.3.2 The p-doped Si Layer

Instead of the high dose delta layer used in many publications [148], [168] a moderately doped p-Si layer is substituted. Both the ion beam and the high temperature used in delta implant will degrade the channel layer interfaces and relax the built in strain. This will degrade the mobility considerably. Mostly the threshold voltage is controlled by the depletion in the n-Si substrate below the p-Si layer in

Figure 5-1. This p-Si layer gives

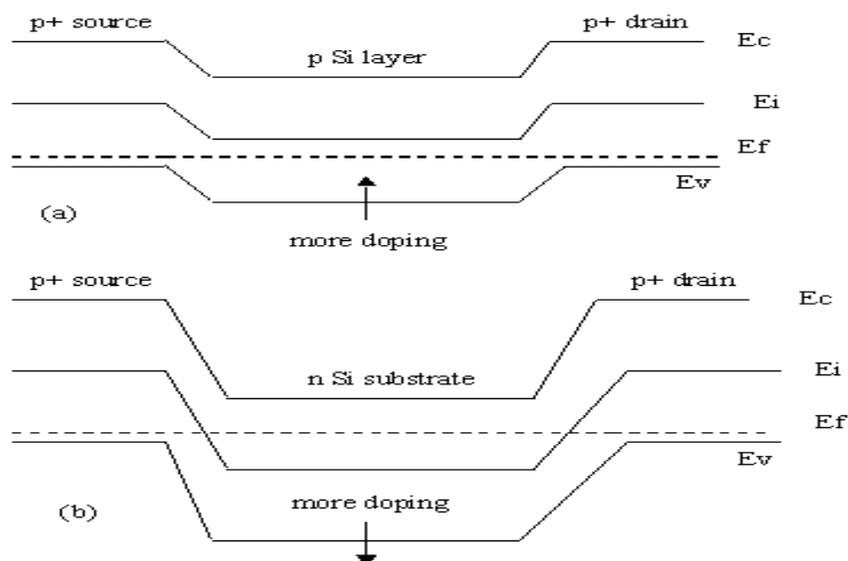


Figure 5-2. (a) Band diagrams from source to drain through p-Si layer through the section A-A (b) through Si substrate through the section BB of Figure 5-1. The barrier at the source through the p-layer is less than that through the substrate.

only a slight increase in the threshold voltage. This is because the downward band bending due to the p-Si layer has to be compensated by an additional small gate voltage increase. Since this p-layer is of finite size it can be grown at low temperature CVD thus avoiding the damage caused by ion bombardment. This layer has a fixed doping of

$5e^{18}/\text{cm}^3$  in the simulations presented here. This value is chosen since lower values were not very effective in the adjustment of  $V_T$ . By varying the thickness of this p-Si layer the effective number of holes in this layer can be varied. As shown in Figure 5-3 a lower doping of this layer gives more barrier at the source – p-Si junction reducing the off state leakage currents caused by thermionic and diffusion processes. To keep these currents low at least a barrier of  $3kT/q$  must be kept. So the doping parameter in the p-Si layer should be  $(E_g / 2) - (3kT / q)$  or lower.

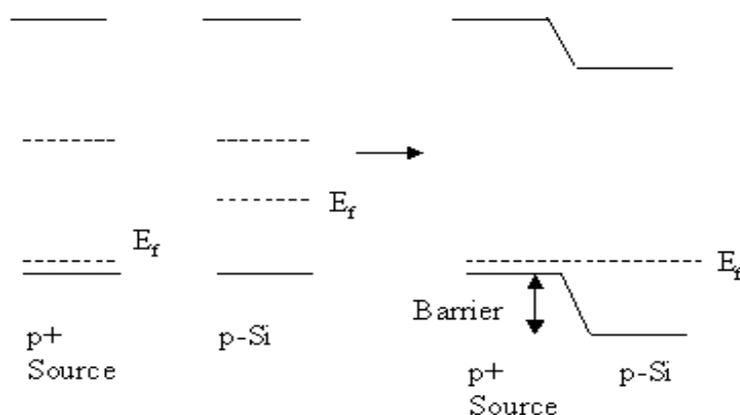


Figure 5-3. Source to p-layer barrier at the source junction. Higher p-layer doping reduces this barrier

### 5.3.3 The Undoped Si Buffer Layer

A thin undoped Si buffer layer is placed above the p-doped silicon layer. This reduces the Coulomb scattering of carriers in the QW channel layer due to the negative ions in the p-doped layer. If this undoped layer is too thick, it can cause a hump in the valence band and collect holes leading to a parasitic QW which will reduce the hole density in the QW well channel. This is because the available carriers are shared by the quantum well and

the hump. Considering that the hole wave functions from the  $\text{Si}_{1-x}\text{Ge}_x$  channel can penetrate roughly 2nm to this layer [ 158] a minimum of 3nm thickness is desirable. In addition, as seen from simulations, more than an order of difference seem to exist in the hole population between the top and bottom of the  $\text{Si}_{1-x}\text{Ge}_x$  quantum well. This is because the top of the QW is closer to the quasi Fermi level when a gate voltage is applied. As seen in chapter 3 the shape of the top of the QW is under the influence of the number of p-layer dopant atoms. So most of the carriers are up and away from this buffer layer. This will further reduce the Coulomb scatter due to negative ions in the p-doped layer. This buffer layer also affects the threshold voltages but its effect was found to be very low, in simulations of chapter 7, compared to  $N_{\text{sub}}$  and p-layer thickness.

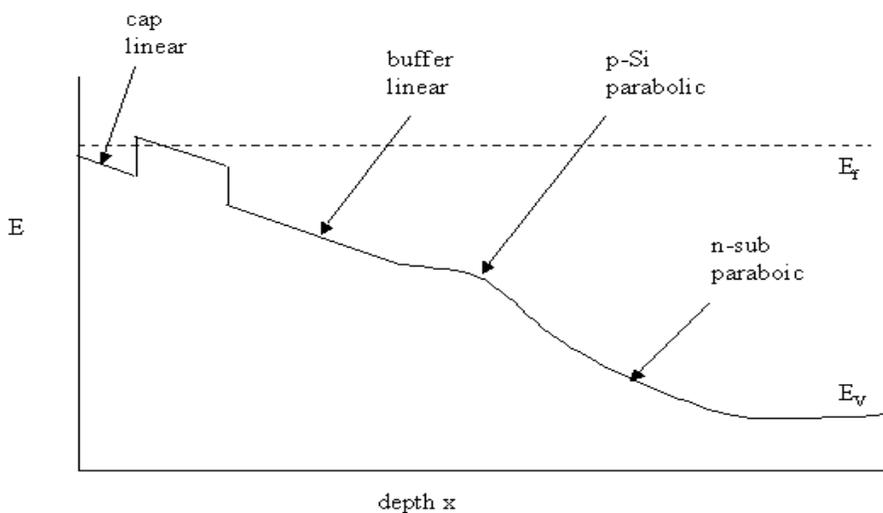


Figure 5-4. Valence band diagram of p-HMOSFET at high  $V_{\text{gs}}$ . The n-Si substrate parabolically bends the band upwards and the p-Si layer within its width bends it downwards. Both the buffer layer and the cap layer have no charges. Hence the transition through them is linear. The QW is nonlinear since the hole distribution within it and the potential are coupled.

### 5.3.4 Strained Si<sub>1-x</sub>Ge<sub>x</sub> Quantum Well Channel Layer

Carriers, in this case holes, are confined to this QW layer. Conduction from source to drain occurs through this layer. Hence this layer is most important. At the same time design of this layer is more complex since it depends on many factors. Maintaining the compressive strain in this layer and keeping the interfaces defect free are important to attain high mobility. If this layer is too thin quantization will occur and the hole density will be reduced. A thin layer will also degrade mobility due to higher interaction with the interfaces. If the Ge fraction is high the abrupt transition from the undoped Si buffer layer discussed above to Si<sub>1-x</sub>Ge<sub>x</sub> layer, and vice versa at the top of the Si<sub>1-x</sub>Ge<sub>x</sub> layer, will introduce more dangling bonds or defects at the interfaces leading to degradation in mobility. Hence the choice of Ge fraction and the thickness of this channel layer are very important. However, this layer thickness should be less than the critical thickness mentioned in chapter 4. As described in chapter 4 the critical thickness is only 7 nm for a Ge fraction, x, of 0.5.

Compressive strain of this layer along the channel is necessary to gain the higher hole mobility and hence higher ON currents. The strain increases with the Ge fraction x. The valence band discontinuity increases with Ge fraction. The band gap reduction in the Si<sub>1-x</sub>Ge<sub>x</sub> channel is mostly in the valence band, almost 97%, and can be approximated as below [80]. This is copied here from chapter 4.

$$\Delta E_v = 0.74x \quad (5.2)$$

In order to avoid spillage of holes into the cap layer lying above the QW channel layer sufficient band edge discontinuity should be provided. A minimum fraction of 0.3 is

found to be necessary, in simulations of chapter 7, to keep holes confined to the QW channel layer even with a thin cap layer as thin as 3 nm. The linear equation for valence band discontinuity is not accurate above a Ge fraction of 0.6. Moreover as mentioned above higher values of  $x$  will degrade interfaces due to the abrupt transition in lattice constant. Hence a range of values of  $x$  between 0.3 to 0.6 is chosen for simulations.

In order to preserve the strain in the  $\text{Si}_{1-x}\text{Ge}_x$  layer the thickness has to be below the critical thickness as discussed in Chapter 4. This depends on Ge fraction and growth conditions. For increasing Ge fractions the critical thickness reduces. As an example the critical thickness is only around 9 nm for a Ge fraction of 0.3. This falls to about 7 nm for a Ge fraction of 0.5. Other than this there will be quantization effects and band splitting in the quantum well if it is too thin. This will lead to reduced hole density confinement due to reduction in density of states. But the lowermost energy wave function is shown to spread only to about 5 nm [ 87 ], [ 176 ]. Considering the above the  $\text{Si}_{1-x}\text{Ge}_x$  quantum well channel thickness is kept fixed at 5 nm in simulations.

### **5.3.5 The Undoped Si Cap Layer**

The undoped cap layer and the recessed undoped  $\text{Si}_{1-x}\text{Ge}_x$  quantum well channel are main reasons for the superior performance of HMOSFETs. The cap layer separates carriers physically from the inferior Si/SiO<sub>2</sub> interface. The carriers confined in the

$\text{Si}_{1-x}\text{Ge}_x$  channel layer lie close to the interface formed by the QW channel layer and the cap layer. This interface is pseudomorphically grown and have very low defects and impurities as discussed in chapter 4. So the carrier mobility is not very much degraded. The choice of this cap layer thickness is important. The hole wave functions from the

$\text{Si}_{1-x}\text{Ge}_x$  channel layer are found to spill over by about 2nm [ 158]. So a thickness more than 2 nm will avoid any interaction of holes in the quantum well with the Si/SiO<sub>2</sub> interface.

If the cap layer is too thick surface inversion will take place before the QW inversion as shown in Figure 5-5. For the HMOSFET there are two threshold voltages, one for the QW channel and the other for the surface inversion. They must be well separated to avoid parallel channel at the surface in the cap layer. From this viewpoint it is better to keep the

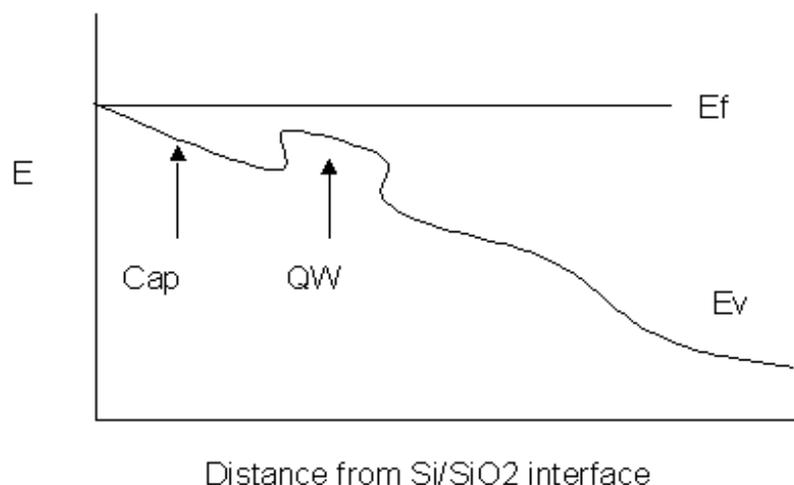


Figure 5-5. Valence band plot at high  $V_{gs}$  value. The thick cap layer leads to inversion near Si/SiO<sub>2</sub> interface.

cap layer thickness a minimum. From this and the wave function containment in the QW channel layer a thickness of 3 nm is chosen. Also this is a thickness that can be grown with minimum defects. This value is kept constant in the simulations. The cap layer thickness also affects the threshold voltage. The depletion region is below the p-Si layer. A thicker cap layer will reduce the gate coupling or capacitance to this depletion region. So the threshold voltage  $V_T$  will increase. It is to be clarified here that  $V_T$  is the

threshold voltage for the QW channel and  $V_{TS}$  is the threshold voltage for the surface channel formation.  $V_{TS}$  is higher than  $V_T$ , in magnitude, in a well designed device. This device can only be operated between these two threshold levels.

### **5.3.6 Gate Oxide**

The lower the gate oxide thickness the larger the gate to channel capacitance and larger the transconductance. However to keep the gate leakage currents due to tunneling low at least a thickness of 3 to 4nm is needed [168]. Also at 3nm good quality oxides can be grown. Hence, gate oxide thickness is kept fixed at 3nm.

### **5.3.7 Gate Material**

The gate material work function selection is important since it determines threshold voltage. The gate material can be chosen as n+ or p+ polysilicon. Selection of p+ gate will give near zero  $V_T$ . This is not desirable since the OFF currents will be high. n+ gate gives a  $V_T$  near 0.5 - 1.0 V. Hence n+ polysilicon gate is chosen. The value of  $V_T$  mentioned here is high compared to MOSFET devices but simulations in chapter 8 shows that HMOSFET can give currents comparable to MOSFET even with this high  $V_T$ .

### **5.3.8 Fixed Parameters and Design Variables**

From the discussion of structural parameters above a list of parameters can be made.. As discussed some of the parameters can be set at their optimum values. Changing these will only make the device inferior. So they are kept fixed. The rest of the parameters are varied in the simulation to arrive at appropriate ranges and design windows. So the parameters are grouped into fixed parameters and design variables. The following parameters are at their optimum feasible values and are treated as fixed parameters.

1. n+ poly silicon gate
2. Gate oxide thickness = 3 nm
3. Undoped Si cap layer thickness = 3 nm
4. Undoped Si<sub>1-x</sub>Ge<sub>x</sub> QW channel layer thickness = 5 nm
5. p-doped layer doping density =  $10^{18}/\text{cm}^3$

This leaves the following parameters as design variables,

1. Substrate doping density  $N_{\text{sub}}/\text{cm}^3$
2. p-doped layer thickness  $X_p$
3. Undoped buffer layer thickness  $X_{\text{buf}}$
4. Undoped SiGe quantum well Ge fraction  $x$

All of the above parameters were discussed in the previous sub-sections.

### 5.3.9 Device Notation

The layers and doping concentrations are discussed above. Since many parameters are involved it is difficult to explain the device. Hence it is good to have a notational scheme that can be easily understood. The gate oxide thickness is assumed to be fixed at 3 nm. So this can be omitted from the notation. Starting from the cap layer up to the substrate the following have to be specified. The layer thicknesses, the Ge fraction of the QW, p-layer doping density and substrate doping density. So the following notation includes all the above in a condensed manner.

$$X_{\text{cap}}, X_{\text{SiGe}}, X_{\text{buf}}, X_p/x/N_p, N_{\text{sub}}$$

The above are thickness of the cap layer, thickness of the SiGe QW layer, thickness of the buffer layer, thickness of the p-layer/ Ge fraction in the QW/ p-layer doping density, substrate doping density.

## **5.4 Summary**

Design considerations related to structural parameters are discussed in this chapter. First the necessity of the pseudomorphic growth of the QW layer and the cap layer was stressed. This is to make sure that the QW walls have the least defect density. The significance of the Ge fraction,  $x$ , in the QW was discussed. The significance of the width of the QW was discussed. The importance of the proper choice of  $x$  was stressed.

Then the structural parameters of the p-HMOSFET were identified. The limitations, ranges and optimum values of these parameters were discussed. Some of the parameters were identified to be fixed parameters since it is better to leave them at their optimum values. Rest of the parameters were identified as design variables.

Unlike conventional MOSFET the p-HMOSFET requires several parameters to be specified to explain the device. This is difficult to convey without being much verbose. So a compact notational definition of the p-HMOSFET is proposed.

## CHAPTER 6

### ANALYTICAL ANALYSIS AND DESIGN PROGRAM

The main goal of this project is to develop an easy to use analytical program that can quickly analyze a device structure that has the basic structural parameters to design a strained  $\text{Si}_x\text{Ge}_{1-x}$  p-channel HMOSFET (hetero junction MOSFET). A device structure of this sort was introduced in chapter 4. Unlike MOSFET which has adopted a finite structure more than three decades back a common-stay general device structure is yet to be adopted for HMOSFET. The structure discussed in chapter 4, Figure 4.17 has the necessary basic structural parameters and looks like a structure that has the potential to becoming a generally acceptable one. Hence that structure is chosen for the design program. This program should provide internal parameters like potentials and hole densities as well as terminal characteristics. This enables the designer to quickly evaluate effects of any design changes and enables optimization of the device by knowing certain parameters like hole density ratio  $p_h/p_0$  where  $p_h$  is the hole sheet density in the quantum well and  $p_0$  is the hole sheet density at the top of the cap layer (oxide interface), both at ON state. In this chapter the development and validation of such an analysis program is presented.

First of all the designer should arrive at a near optimal design by adjusting the input parameters. Once a near optimal design is reached the design can be fine tuned by running a few simulations in the numerical simulator. First of all this saves a whole lot of time. Secondly it is very easy to visualize the impact of parameter changes instantly and

make corrections quickly and get the results for evaluation. It is not possible in the numerical simulator. The numerical simulator takes too much time to draw the device and input parameters, eg. doping profiles. The simulator takes too much time. It is difficult to do the book keeping of all the inputs and outputs. The simulator do not give an idea of the Physics involved. On the contrary plots like potentials and hole density distribution can be instantly viewed in the analytical program which gives a quick feed back for further parameter changes thus quickly progressing towards an almost mature design. So the analytical analysis and design program developed here can be looked upon as an user interface between a start up design and the numerical simulator. That means most of the design is carried out quickly in the analytical program and only a few numerical simulations are required.

## **6.1 Solution of One Dimensional Poisson Equation**

In order to make an analysis of the device internal potentials are of vital importance. All other properties of the device depend on them. Since a simplistic implementation of the analysis is desired it is assumed that the device is a long channel device. This avoids the two dimensional nature of potential contours of short channel devices. What this means is that a one dimensional Poisson solution of the potentials could be derived. Later in chapter 8 a correction is added to this program to approximately solve short channel p-HMOSFET. So in this section the solution of the Poisson equation will be discussed. The structure of the device is reproduced here since references are made to different labels in the figure. Figure 6-1 shows a sketch of the proposed p-HMOSFET that is used for analysis in this project. It has the dimensions and co-ordinates which appear in the

Poisson solutions. It is the same structure described in chapter 4, Figure 4.17. It is reproduced here for easy reference and appropriate labels are added to it.

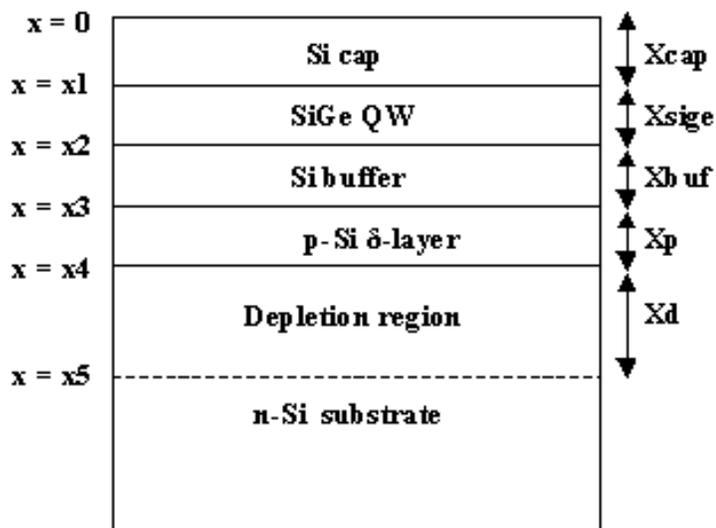


Figure 6-1. Sketch of a p-HMOSFET showing the different regions, dimensions and co-ordinates.  $\text{Si}_{1-x}\text{Ge}_x$  QW is the conduction channel layer.

The solution details of the Poisson equation are given in Appendix C. The Poisson solution closely follows one of the published works [184]. But there are differences in the structure and what is being accomplished. In the device used in this project there is an additional p-layer to provide some independent control of the threshold voltage. The published work is about few parametric studies whereas the work presented here is aimed at providing an easy to use full analysis and design program. The analytical program to follow the Poisson solution provides most of the necessary parameters, internal and external. This will be discussed in the program section below. The necessary solutions and the derivation of equations to be used in the program are discussed here. A list of few

variable names is given below for easy reference before discussing the Poisson solutions.

Rest of the symbols found in equations can be found at the beginning of Appendix C.

1.  $K_{buf}$  – A constant parameter defined in the buffer layer
2.  $K_{sige}$  – A constant parameter defined in the SiGe layer
3.  $K_{cap}$  – A constant parameter defined in the cap layer
4.  $\xi_4$  – Electric field at  $x = x_4$
5.  $\phi_4$  – Potential at  $x = x_4$
6.  $N_p$  – p-layer doping density per  $cm^3$
7.  $N_D$  – substrate doping density per  $cm^3$

The Poisson solution for the structure of Figure 6-1 is given in Appendix – C. The potential solutions within each region are first used for plotting overall potentials, hole densities, integrated hole sheet density, centroid of the hole density distribution and so on. The centroid is used for the correction of gate-QW hole channel capacitance correction along with  $C_{OX}$ . The potential at any point  $\phi(x)$  along the center of the device can be found by adding up potentials of each region and the potential in the region where the point is located. So first a potential plot is generated,  $\phi_x$  vs  $x$ . Then the interest is to find the hole density distribution. Using equation ( 6.9) and neglecting the density of states term the hole density can be written as,

$$p(x) = N_D \exp\left(\frac{q}{kT}\right)(\phi_{TH} - \phi_x) \quad (6.1)$$

As shown in Appendix C section C6 the above equation can be written as,

$$p(x) = N_D \exp\left(-\frac{q}{kT} \Phi_x\right) \exp\left(\frac{\Delta E_v}{kT}\right) \quad (6.2)$$

If the second exponent due to the valence band discontinuity is removed this is the same as the MOSFET case. Since  $\phi_x$  is known the hole density at any location  $p(x)$  vs  $x$  can be plotted.

The aim is to find the hole densities in the quantum well and at the surface channel near the oxide interface for any applied gate voltage. Next the threshold voltages need to be found. This device has two threshold voltages as mentioned in chapter 5. One is when the quantum well channel layer begins to get populated with holes. This is the desired one and is represented by  $V_T$ . Further increase in gate voltage will begin to populate the cap layer at the oxide interface, at some higher value, just like a MOSFET. This is an undesirable channel as described in chapter 5 and is represented by  $V_{TS}$ . Knowing these values the drain saturation voltage  $V_{Dsat}$  and current  $I_{Dsat}$  can be calculated. In order to do this the potential at the top of the quantum well  $\phi_H$  and the surface potential in the cap layer  $\phi_0$  are needed. If, at  $x = x_4$ ,  $\xi_4$  and  $\phi_4$  are known the rest of the potentials can be found. In the QW layer and the cap layer a correction term  $H(\phi)$  need to be taken into account in order to account for the holes in the QW. For  $V_{gs} > V_T$   $\xi_4$  can be found if  $X_{dmax}$  is known. The relationships between them are given below,

$$\xi_4 = \frac{\sqrt{2}}{\beta L_D} \sqrt{(-\beta \phi_4 - 1)} \quad (6.3)$$

$$X_d = \sqrt{2}L_D\sqrt{-\beta\phi_4 - 1} \quad (6.4)$$

$$\xi_4 = \frac{X_{d \max}}{\beta L_D^2} \quad (6.5)$$

$$\phi_4 = \frac{-1}{\beta} \left( \frac{\xi_4^2 \beta^2 L_D^2}{2} + 1 \right) \quad (6.6)$$

$$\beta = \frac{q}{kT} \quad (6.7)$$

$$L_D = \sqrt{\frac{\epsilon_{SI} kT}{q^2 N_D}} \quad (6.8)$$

$X_{d\max}$  is the depletion depth in the substrate when the quantum well begins to fill with holes or at  $V_{gs} = V_T$ ,  $k$  is the Boltzman constant and  $L_D$  is the Debye length. Another parameter to be introduced is the threshold potential  $\phi_{TH}$ . This is the value of the potential  $\phi_H$  at the top of the QW, at  $x = x_1$ , when holes begin to fill in the quantum well [184]. Here the threshold is defined similar to the MOSFET case. Since there is no inversion in the QW it is defined as the potential twice negative than that of the substrate. Unlike in the case of MOSFET it may not mean that there are as many holes as there are electron in the substrate. Since the valence band is lifted up in the QW that discontinuity is directly added to the above. The third term can be found by equating the carriers in Si and SiGe for the same  $\phi_B$  [184]. Another definition could be the potential at which the hole density in the QW becomes equal to electron density in the substrate. At this condition

the depletion layer charge no longer depends on gate voltage. This is confirmed from simulations with the numerical simulator ISE-TCAD [109].

$$\phi_{TH} = 2\phi_B + \frac{\Delta E_V}{q} - \frac{kT}{q} \ln\left(\frac{N_V(Si)}{N_V(sige)}\right) \quad (6.9)$$

$$\phi_B = -\frac{kT}{q} \ln\left(\frac{N_D}{n_i}\right) \quad (6.10)$$

$N_V(Si)$  and  $N_V(Si_xGe_{1-x})$  are the valence band density of states for silicon and  $Si_xGe_{1-x}$  respectively. Referring to Appendix – B section B4 the form of the electric field inside the SiGe QW complicates integration to find the potential in the QW. The electric field equation is copied here for reference since a condensed term  $H(\phi_H)$  appears in the equations to follow.

$$\xi_x = \frac{(qNdXd - qNaXp)}{\epsilon_{sige}} [\sqrt{(1+H(\phi_x))}] \quad (6.11)$$

$$H(\phi_x) = \frac{2kTN\epsilon_{sige}}{(qNdXd - qNaXp)^2} \{ \exp[(\Phi_{TH} - \Phi_x)(q/kT)] - \exp[(\Phi_{TH} - \Phi_{x2})(q/kT)] \}$$

To proceed with the integration in the cap layer an assumption is made that the contribution of holes in the QW to the potential is small and hence the  $H(\phi_x)$  term is negligible [184]. Referring to Appendix – C, equation A7 and noting the fact that the potentials are negative the second exponent in the  $H(\phi_x)$  can be neglected. Above the QW in the cap layer the  $H(\phi_x)$  term becomes a constant  $H(\phi_H)$  since the contribution of holes stops at  $x = x1$ . When the number of holes in the QW is small ( $V_{gs} < V_T$ ) the first exponent is also small and the  $H(\phi_H)$  term can be neglected. When there are many holes

in the QW ( $V_{gs} > V_T$ )  $H(\phi_H) \gg 1$  and only this term need to be retained under the square root.

Knowing the gate voltage  $V_{gs}$  and the gate to semiconductor work function difference the following equation can be written by inspection. A quantum well potential  $\phi_H$  is present at the top of the QW . Then the field at the back of the QW ( $x = x_2$ ), modified by the  $H(\phi_H)$  function to include the field due to holes, is used to compute the field at the gate,

$$V_{gs} - \phi_{ms} = \phi_H - \left( \xi_4 - \frac{qN_p X_p}{\epsilon_{Si}} \right) \sqrt{1 + H(\phi_H)} \left( X_{CAP} + \frac{\epsilon_{Si} X_{OX}}{\epsilon_{OX}} \right) \quad (6.12)$$

In the above it can be assumed that at the threshold point there are not much holes in the quantum well and hence the  $H(\phi)$  function can neglected. Knowing that  $\phi_H = \phi_{TH}$  when  $V_{gs} = V_T$  the threshold voltage can be written as,

$$V_T = \Phi_{ms} + \Phi_{TH} - \left( \xi_4 - \frac{qN_p X_p}{\epsilon_{Si}} \right) \left( X_{cap} + \frac{\epsilon_{Si}}{\epsilon_{OX}} X_{OX} \right) \quad (6.13)$$

The potential at the surface  $x = 0$  can be written as below. This is different from the equation (C.12) given in appendix C It assumes that  $\phi_4$  is known. This is what the Excel spread sheet program is going to do.

$$\Phi_0 = \Phi_4 + K_{cap} - \xi_4 (X_p + X_{buf} + RX_{sige}) - \left( \xi_4 - \frac{qN_p X_p}{\epsilon_{Si}} \right) \sqrt{1 + H(\Phi_H)} X_{cap} \quad (6.14)$$

$$K_{CAP} = \frac{qN_p X_p^2}{2\epsilon_{Si}} + \frac{qN_p X_p}{\epsilon_{Si}} X_{BUF} + \frac{qN_p X_p}{\epsilon_{SiGe}} X_{SiGe} \quad (6.15)$$

Again the potential at the surface can be equated,

$$V_g - \phi_{MS} = \phi_0 - \epsilon_{OX} X_{OX} \quad (6.16)$$

At  $V_{gs} = V_{TS}$   $\phi_0 = \phi_{TS}$  where  $\phi_{TS} = 2\phi_B$ . Again neglecting the  $H(\phi_H)$  term the surface threshold  $V_{TS}$  can be written as,

$$V_{TS} = \phi_{TS} + \phi_{MS} - \left( \xi_4 - \frac{qN_P X_P X_{OX}}{\epsilon_{Si}} \right) X_{OX} \quad (6.17)$$

In order to find the hole densities for any gate voltage  $\phi_4$  must to be found. This is done by finding  $\xi_4$  first as shown below and then using equation (6.6) to find  $\phi_4$ . The surface potential  $\phi_0$  can be related to  $V_{gs}$  and the field at interface  $x_4$  can be solved as shown below.

$$\begin{aligned} Vg - \phi_{MS} &= \phi_0 - \epsilon_{OX} E_{OX} \\ &= \phi_4 + K_{cap} - \xi_4 (X_P + X_{buf} = RX_{sige} + X_{cap}) + \frac{qN_P X_P}{\epsilon_{Si}} X_{cap} \\ &\quad - \left( \frac{\epsilon_{Si}}{\epsilon_{OX}} \xi_4 - \frac{qN_P X_P}{\epsilon_{OX}} \right) X_{OX} \end{aligned} \quad (6.18)$$

Solving,

$$\xi_4 = \frac{-C_{13} + \sqrt{C_{13}^2 - 4C_{14}}}{2} \quad (6.19)$$

$$C_{13} = \frac{2}{\beta L_D^2} \left( X_P + X_{BUF} + RX_{SIGe} + X_{CAP} + \frac{\epsilon_{Si}}{\epsilon_{OX}} X_{OX} \right) \quad (6.20)$$

$$C_{14} = \frac{2}{\beta L_D^2} \left( V_g - \phi_{MS} - K_{CAP} - qN_P X_P \left( \frac{X_{CAP}}{\epsilon_{Si}} + \frac{X_{OX}}{\epsilon_{OX}} \right) + \frac{1}{\beta} \right) \quad (6.21)$$

$C_{13}$  contains constant terms and  $C_{14}$  contains  $V_{gs}$  and some constant terms. To keep the focus clear they are not listed here. What is important is that the gate voltage  $V_{gs}$  is part

of this equation. So electric field at  $x = x_4$  ( $\xi_4$ ) can be found for any  $V_{gs}$  input using this equation. Note that this procedure is for the sub-threshold region. For the above threshold region  $\xi_4$  is found from  $X_{dmax}$  using ( 6.5).  $X_{dmax}$  is found as follows. The potential at the top of the quantum well can be written as,

$$\phi_H = \phi_4 - \xi_4 C_9 + K_{cap} \quad (6.22)$$

$C_9$  is a structural constant and  $K_{cap}$  is an integration constant from the solution of the Poisson equation Substituting for  $\phi_4$  and  $\xi_4$  from equations ( 6.5) and ( 6.6) and noting that  $\phi_H = \phi_{TH}$  at threshold voltage  $X_d$  can be solved for at the threshold voltage

$$X_{dmax} = -C_9 + \sqrt{C_9^2 - C} \quad (6.23)$$

$$C_9 = X_p + X_{BUF} + RX_{SiGe} \quad (6.24)$$

$$C = 2\beta L_D^2 \left( \phi_{TH} + \frac{1}{\beta} - C_{CAP} \right) \quad (6.25)$$

$C_9$  contains some structural constants and  $C$  contains some constants and  $\phi_{TH}$ .  $C$  also contains a  $L_D$  term. So  $X_{dmax}$  is both a function of structure and substrate doping. In order to find the hole density for  $V_{gs} > V_T$  (magnitude) the potential  $\phi_H$  is needed. How to find  $\phi_H$  for an applied  $V_{gs}$  is explained next. The equation below is written by inspection. The field at the bottom of the QW passes through the QW modified by the field contributed by the holes present. Apart from that top of the QW is at a potential  $\phi_H$ . Extending this potential and the field to the gate the following equation can be written.

$$V_g - \phi_{MS} = \phi_H - \left( \xi_4 - \frac{qN_p X_p}{\epsilon_{Si}} \right) \left( X_{cap} + \frac{\epsilon_{Si}}{\epsilon_{OX}} X_{OX} \right) \quad (6.26)$$

Taking  $\phi_H$  to the left, taking log and after some manipulation,

$$\phi_H = \phi_{TH} - 2v_{th} \ln \left( \frac{\phi_H - V_g + \phi_{MS}}{C_1 \sqrt{C_H}} \right) \quad (6.27)$$

$$C_1 = \left( \xi_4 - \frac{qN_P X_P}{\epsilon_{Si}} \right) \left( X_{cap} + \frac{\epsilon_{Si}}{\epsilon_{OX}} X_{OX} \right) \quad (6.28)$$

$$C_H = \frac{2qN_D \epsilon_{SiGe}}{\beta(\epsilon_{Si} \xi_4 - qN_P X_P^2)} \quad (6.29)$$

The potential at the top of the QW  $\phi_H$  is needed to find the hole density. Note that  $\phi_H$  in equation ( 6.27) has a circular relationship and an iteration is required to find  $\phi_H$ .

Now all necessary values are available and  $\xi_4$  can be found for any gate voltage. From  $\xi_4$   $\phi_H$  can be found and then hole density in the QW can be found.  $\phi_0$  can be found from equation ( 6.14). That can be used to find the surface hole density. The analytical program uses few of the above equations in its computations. So the analytical program will be discussed next.

## 6.2 Analytical Analysis and Design Program

In the previous section the solutions of the one dimensional Poisson equation was discussed The solutions were also manipulated and set up to compute the parameters of design interest. One of the vital parameters is difference between the threshold voltages  $V_T$  and  $V_{TS}$ . This tells how much margin is available for operating voltage. It also tells how inferior or superior the mobility is going to be. A high  $V_{TS}$  is desirable to reduce the inferior surface channel conduction. Another parameter of interest is the ratio of the

quantum well hole density to the cap layer surface hole density,  $p_H/p_0$ . This in fact is dependent on  $V_T$  and  $V_{TS}$  but gives a quantitative assessment rather than a qualitative assessment. A higher hole density in the QW is desirable. Knowledge of hole sheet density could be used in current calculations. From the terminal characteristic point of view the saturation current  $I_{Dsat}$ , drain saturation voltage  $V_{Dsat}$ , sub-threshold parameter  $S$ ,  $V_T$  and  $V_{TS}$  are of interest.

The analytical program developed here is written with the above requirements, which a device designer is interested in. Such a program is not found in the literature. Programs found in the literature mainly focus on couple of parameter studies like how the cap layer thickness affects performance, for an example. There are publications on fabricated devices. But that is not a way to make an optimum device. So the type of design program developed here is of much value from the design point of view. The overall structure choice is fixed. That means additional layers cannot be added. But layers can be removed by making the dimension and doping densities negligible. However the overall structure chosen for this program seems to have enough layers to make an optimum device and similar structure appears in many publications in the recent past.

### **6.2.1 Program Inputs and Outputs**

The device was explained in chapter 4 and the layer details are given in Figure 6-1. Using this program effects of structural and dopant changes can be found almost instantaneously. This is not possible with numerical simulators like the one used here for comparing the analytical program results. For each change one has to wait for many minutes or hours to get a response, has to process the data and make plots to visualize the

results and it is difficult to keep track of the direction the designer is headed for.

Referring to Figure 6-1 the following parameters can be changed in the analytical program and the outputs can be evaluated until desired design performance is achieved,

1. Cap layer thickness  $X_{\text{cap}}$
2.  $\text{Si}_x\text{Ge}_{1-x}$  QW thickness  $X_{\text{SiGe}}$
3. Buffer layer thickness  $X_{\text{buf}}$
4. p-layer thickness  $X_p$
5. Substrate doping density  $N_{\text{sub}}$
6. p-layer doping density  $N_p$
7. Device channel length  $L_g$
8. Drain voltage  $V_{\text{ds}}$
9. Gate voltage  $V_{\text{gs}}$

The program can provide the following output plots for changes in the structure.

Unlike the numerical simulator these plots need not be individually selected and plotted.

All of these plots are available immediately after completing the analysis, which is only

minutes at the maximum. To sort out all of the following from a numerical simulator

output will take hours.

1. Plot of potential  $\phi_x$  vs depth  $x$
2. Plot of valence band edge  $E_V$  vs depth  $x$
3. Plot of QW potential  $\phi_H$  vs gate voltage  $V_{\text{gs}}$
4. Plot of QW hole density  $p_H$  vs gate voltage  $V_{\text{gs}}$
5. Plot of surface channel density  $p_0$  vs gate voltage  $V_{\text{gs}}$

6. Plot of ratio of  $p_H/p_0$  vs gate voltage  $V_{gs}$
7. Drain characteristic of the device ( $I_{ds}$  vs  $V_{ds}$ )
8. Gate characteristic of the device ( $I_{ds}$  vs  $V_{gs}$ )

In addition to the above one can also have the following output values.

1. The QW threshold voltage  $V_T$
2. The surface threshold  $V_{TS}$
3. Quantum well hole density  $p_h$  for a given  $V_{gs}$
4. Surface hole density  $p_0$  for a given  $V_{gs}$
5. Ratio of  $p_h/p_0$  for a given  $V_{gs}$
6. Hole sheet density  $p_s$  for a given  $V_{gs}$
7. Drain saturation voltage  $V_{Dsat}$
8. Drain saturation current  $I_{Dsat}$
9. Sub-threshold parameter S

Apart from the above it is also possible to perform a 'goal seek' in Excel to get an exact drain current  $I_{Dsat}$  by changing any one of the parameters

### 6.2.2 Dissection of the Program

The program is implemented in Excel spread sheet. The parameters set up in section 6.1 are computed by the program and they are then used to find the parameters listed in previous sub-section. The steps of the program are as follows. These can be seen on the left side of the Design Program page of Appendix D.

1. Calculate  $\phi_{TH}$  using equation ( 6.9)

2. Calculates  $X_{dmax}$  using equation ( 6.23)
3. Calculates  $V_T$  using equation ( 6.13)
4. Calculates  $\phi_{TS}$  using equation ( 6.10) (which is double the value)
5. Calculates  $V_{TS}$  using equation ( 6.17)
6. Calculates  $\phi_H$  (PHIH!) using equation ( 6.27) and the gate voltage
7. Using  $\phi_{TH}$  and  $\phi_H$  compute the  $H(\phi_H)$  function
8. Computes  $p_h$  using  $\phi_H$
9. Computes surface potential  $\phi_0$
10. Computes surface hole density  $p_0$
11. Computes  $\phi_x$
12. Computes hole density  $p(x)$
13. Scans  $V_{gs}$  and computes  $\xi_4$  , potential at  $x_4$
14. Computes  $\phi_H$  from  $\xi_4$  . This equation is a circular one and need iteration
15. Computes hole density in QW  $p_H$
16. Computes  $\phi_0$
17. Computes  $p_0$

The above gives all the necessary internal parameters to plot. Figure 6-2 shows a flow chart of the program to find the hole density  $p_H$  in the QW. When  $V_{gs} < V_T$  (magnitude) the contribution of holes in the quantum well to the field can be neglected and the  $H(\phi_x)$  term in equation ( 6.11) can be neglected. The field  $\xi_4$  at  $x = x_4$  can be computed using

equation ( 6.19). The potential  $\phi_4$  at  $x = x_4$  can be computed using equation ( 6.6). Using  $\phi_4$  and knowing that the field is only  $\xi_4$  the potential  $\phi_H$  can be found. From this the hole density in the QW  $p_H$  can be calculated. If  $V_{gs} > V_T$  then the  $H(\phi_x) \gg 1$  in equation ( 6.11) Now it is more difficult to find the quantum well potential  $\phi_H$ . Here  $X_{dmax}$  is found using equation ( 6.23). Then  $\xi_4$  is found using equation ( 6.5). Now  $\phi_H$  can be found using equation ( 6.27). This equation has a circular relationship to  $\phi_H$ . Hence iteration is done to find the value of  $\phi_H$ . These details are shown in Figure 6-2. A starting seed value must be given. Normally the value from previous run will be used. But if iteration fails a seed value will be needed. A value of 1 or above is enough as learned from simulations. The neglect of the second exponent term in equation ) causes a small error in surface potential  $\phi_0$  compared to the numerical simulations. However this is less than 5 % of numerically simulated  $\phi_0$ .

The surface potential is directly found from equation ( 6.14).  $H(\phi_H)$  can now be calculated since  $\phi_H$  is computed in the flow chart. In equation ( 6.11) the second exponent is neglected and  $\phi_x$  is substituted with  $\phi_H$ . The main design page of the program is shown in page - 1 of Appendix C under parameter design. Top left of the panel has the input variables. Bottom left has the output variables which can be monitored almost instantaneously. Right side shows three sets of plots of interest. Page – 2 is the computational page.  $\Phi_{TH}$  is  $\phi_{TH}$ ,  $VT_{-}$ ,  $VTS_{-}$ ,  $Xdmax_{-}$  are  $V_T$ ,  $V_{TS}$  and  $X_{dmax}$  respectively.  $H\_PhiH$  is  $H(\phi_H)$ .  $\phi_{H-}$  is  $p_h$ .

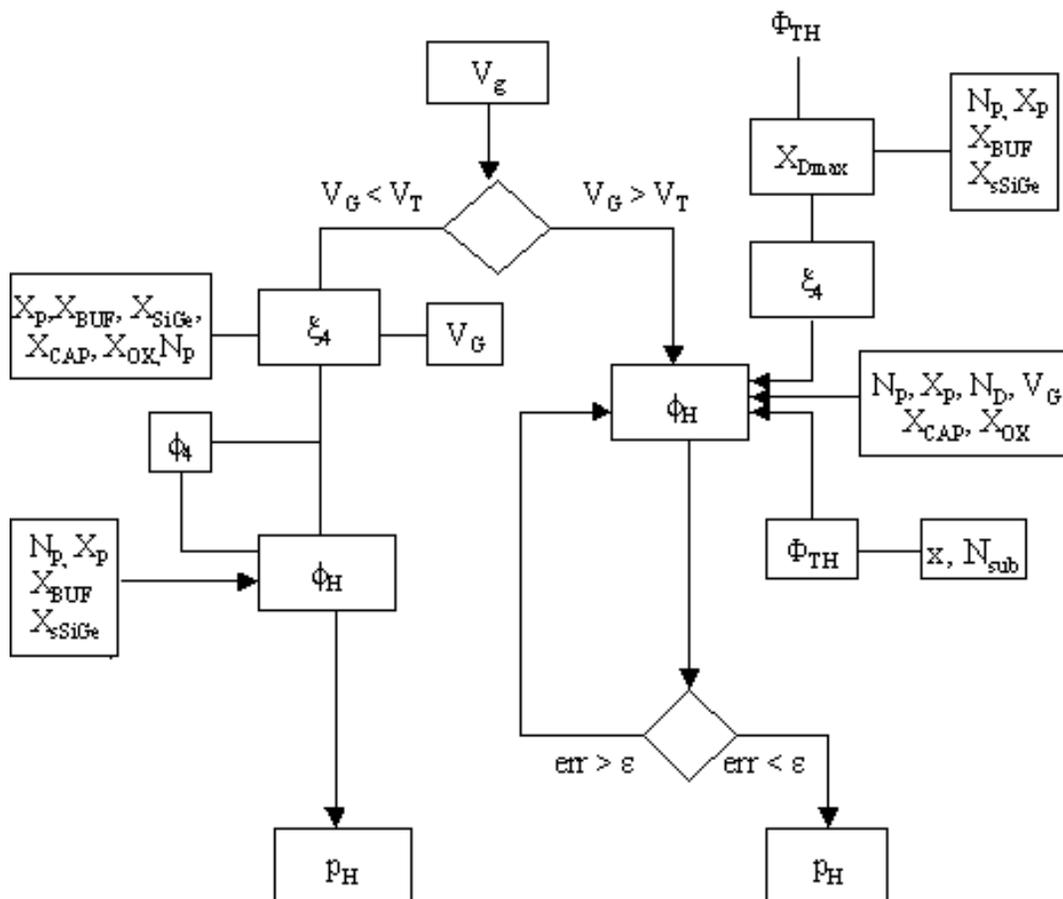


Figure 6-2. Flow chart of the analytical program showing the computation of hole density  $p_H$  in the QW.

Now the program computes the terminal characteristics using the following equations.

Critical field,

$$E_C = \frac{2vsat}{\mu_0} \quad (6.30)$$

$$V_{dsat} = \frac{V_g - V_T}{1 + \frac{V_g - V_T}{L_g E_C}} \quad (6.31)$$

$$I_{dsat} = C_{OX} (V_g - V_T - V_{dsat}) vsat \quad (6.32)$$

The following formula which takes into account velocity saturation is used for plotting drain characteristic [118], The main interest is in the  $I_{Dsat}$  but the linear region is used here only to generate  $I_{ds}$  vs  $V_{ds}$  plots.

$$I_D = A \left( V_g - V_T - \frac{V_D}{2} \right) \frac{V_D}{\left( 1 + \frac{V_D}{\beta} \right)} \quad (6.33)$$

$$A = \frac{W\mu q\beta}{L_g} \quad (6.34)$$

$$\beta = L_g E_C \quad (6.35)$$

### 6.3 Verification of the Analytical Program

After the analytical program is completed and found to be functional it is necessary to evaluate its accuracy by comparing the results with another simulator, preferably a numerical based one. In this case the ISE-TCAD numeric simulator was used [109]. A device with structural parametr values similar to published works mentioned in chapter 4 was set up both in the analytical program and the numeric simulator. The device details are as follows,

1. Undoped cap layer thickness  $X_{cap}$  of 3 nm
2. SiGe QW channel layer thickness  $X_{sige}$  of 5 nm
3. Undoped buffer layer thickness  $X_{buf}$  of 5 nm
4. Doped p-layer thickness  $X_p$  of 5 nm
5. SiGe QW Ge fraction of 0.3
6. p-layer acceptor doping density of  $5 \times 10^{18}/\text{cm}^3$
7. Substrate donor doping density of  $5 \times 10^{18}/\text{cm}^3$

Based on the notation given in chapter 5 this device can be represented as 3,5,5,5/0.3/5e18,5e18. This device is simulated both by the analytical program and the numeric simulator. Since the internal functions like potentials give more accurate picture of the device function they are compared. The potential along the depth of the device starting at the oxide interface, potential at the top of the SiGe QW, hole density at the top of the QW, hole density at the top of the cap layer and integrated hole sheet density in the QW are compared. These are given in the following sub-sections.

### 6.3.1 Potential functions

Since the internal potential distribution determines characteristics of the device and carrier density distributions the same has to be accurately evaluated in the analytical program. The device described in section 6.3 was simulated using both the analytical program and the numerical simulator from ISE-TCAD [109]. Figure 6-3 shows the results. The match between the two results is extremely good for low and high  $V_{gs}$  and are indistinguishable. For  $V_{gs}$  value near  $V_T$  (-1.2 V) a slight split is visible, above the SiGe QW, in the cap layer. This is because the  $H(\phi_H)$  function used to incorporate the effect of holes in the QW. This function has a discontinuity in the analytical model. At low  $V_{gs}$  contribution of  $H(\phi_H)$  function is very low and is neglected compared to one in equations like ( 6.12). For high  $V_{gs}$   $H(\phi_H)$  is very large compared to unity. So near  $V_T$  the approximations do not work well and iteration using the equation ( 6.27) do not produce correct results. Another point to observe is that the potential, particularly at the top of the SiGe QW located at 3 nm, varies only slightly for  $V_{gs}$  above  $V_T$ . An enlarged view of the top region is given in Figure 6-4. The potential for  $V_{gs}$  equal to -1.5 is not very different

from the  $V_{gs}$  equal to  $-1.2$  case. This is because the QW is nearly at inversion for the latter case and holes accumulated in the QW shields the depletion layer in the substrate.

A sharp bend is noticeable at the top of the QW. This is due to the large number of holes accumulated at the top of the QW whose effect is represented by the function  $H(\phi_H)$ . The iteration overestimates  $H(\phi_H)$  near  $V_{gs} = V_T$  region.

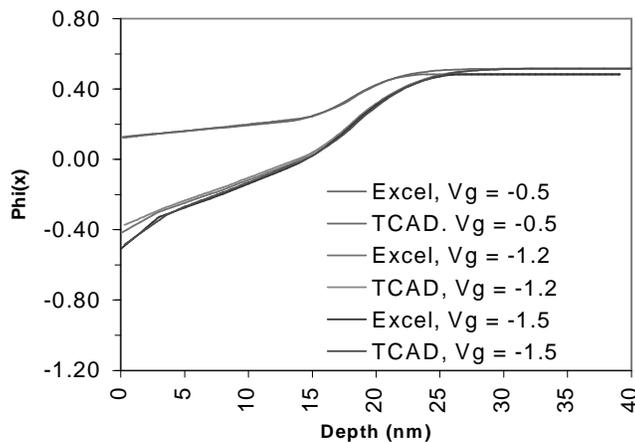


Figure 6-3. Potential variation with depth for the device 3,5,5,5/0.3/5e18,5e18 for  $V_{gs} = -0.5, -1.2, -1.5$ . The device has a  $V_T$  of  $-1.25$ . The match is very good for  $-0.5$  V and  $-1.5$  V. Near  $-1.2$  V (near  $V_T$ ) there is a discontinuity in the potential function approximation.

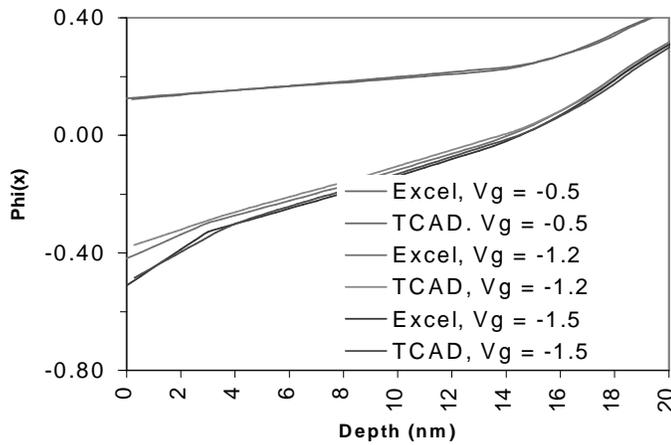


Figure 6-4. An enlarged view of Figure 6-3. Both low and high  $V_{gs}$  simulations match extremely well but a slight split is obvious for  $V_{gs}$  near  $V_{gs} = -1.2$  V. This is explained in Figure 6-3.

Potential plots comparison between analytical program and numerical simulator along depth for a device 3,5,5,5/0.3/5e18,5e18 is shown in Figure 6-5 and Figure 6-6.

Figure 6-5 is for a  $V_{gs} = -1.2$  V and Figure 6-6 is for a  $V_{gs} = -1.5$  V. The plots are displayed separately to show the near perfect match of the result from the analytical program to the result from the numeric simulator.

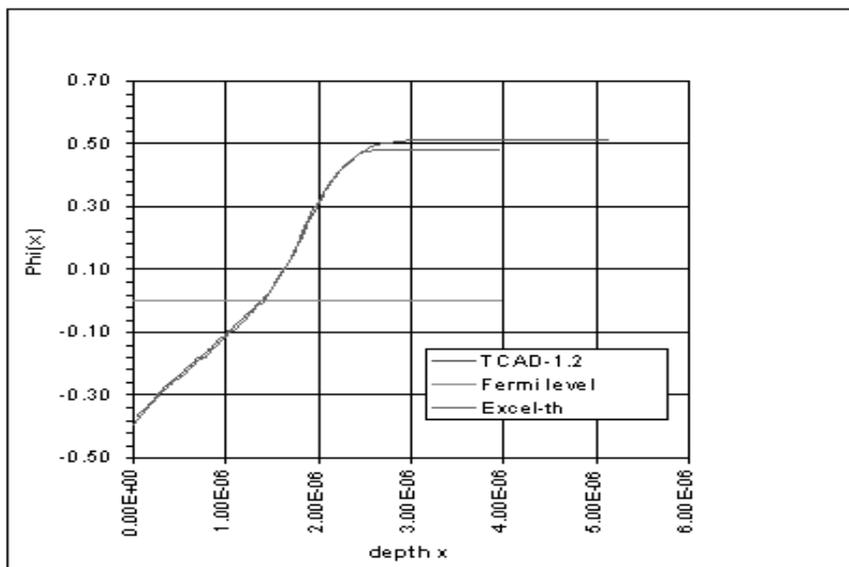


Figure 6-5. Comparison of potential plots from analytical program and numerical simulator.  $V_{gs} = -1.2$  V. Device is 3,5,5,5/0.3/5e18,5e18

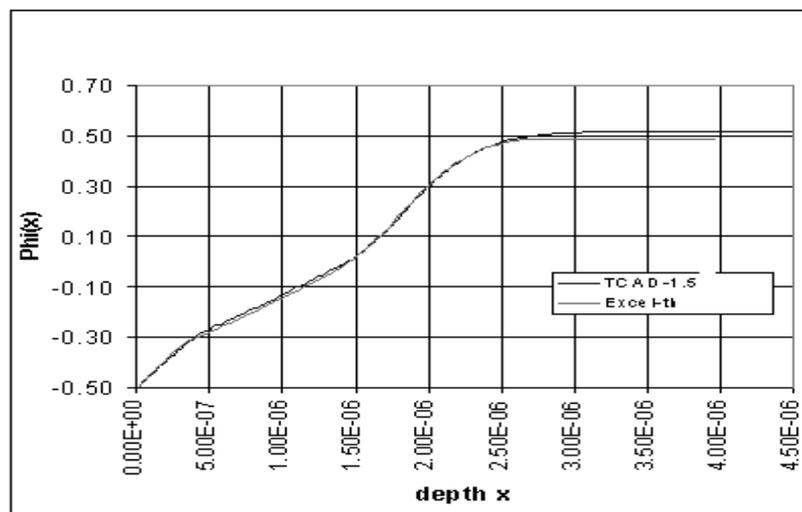


Figure 6-6. Comparison of potential plots from analytical program and numerical simulator.  $V_{gs} = -1.5$  V. Device is 3,5,5,5/0.3/5e18,5e18

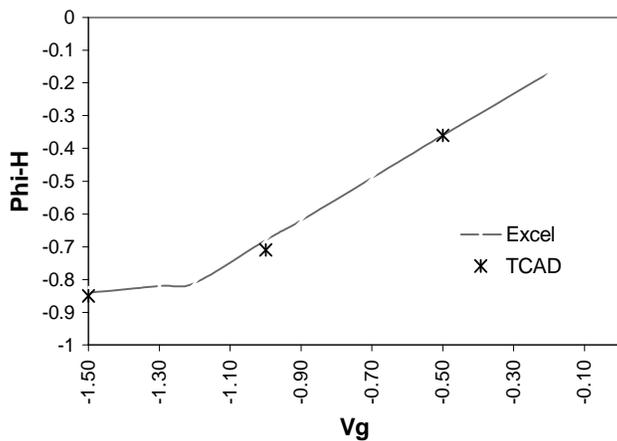


Figure 6-7. Potential at the top of the QW for the device 3,5,5,5/0.3/5e18,5e18 as a function of  $V_{gs}$ . Three numerical simulated values are shown by asterisks

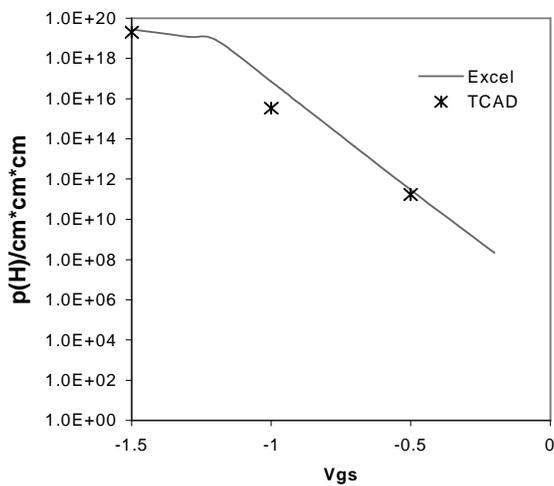


Figure 6-8. Hole density at the top of the QW  $p_H$  for the device 3,5,5,5/0.3/5e18,5e18 as a function of  $V_{gs}$ . Numerical simulation values are shown by asterisks for  $V_{gs} = -0.5$ ,  $-1.0$  and  $-1.5$  V

The potential at the top of the QW is shown in Figure 6-7. Numerical simulation values are shown for  $V_{gs}$  of  $-0.5$ ,  $-1.0$  and  $-1.5$ . These points lie very close to the curve

thus showing the accuracy of the analytical values. It can be seen that for  $V_{gs}$  values above  $V_T$  of  $-1.2$  V, to the left of  $-1.2$  V, the potential is nearly saturated and do not vary much. This is similar to the case of a conventional MOSFET where the surface potential is pinned at  $V_T$ .

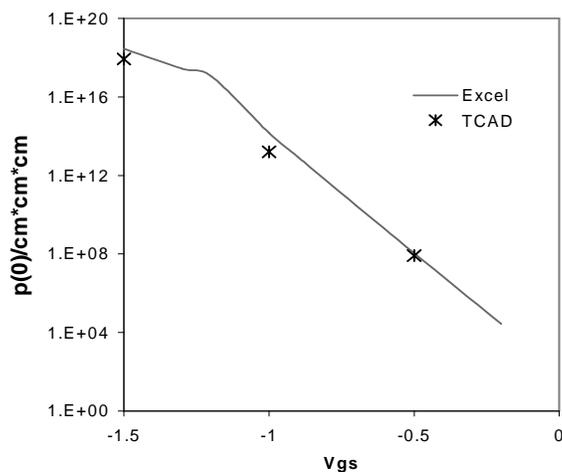


Figure 6-9. Hole density at the top of the cap layer for the device 3,5,5,5/0.3/5e18,5e18 as a function of  $V_{gs}$ . Three numerical simulation values are shown at three points by asteriks

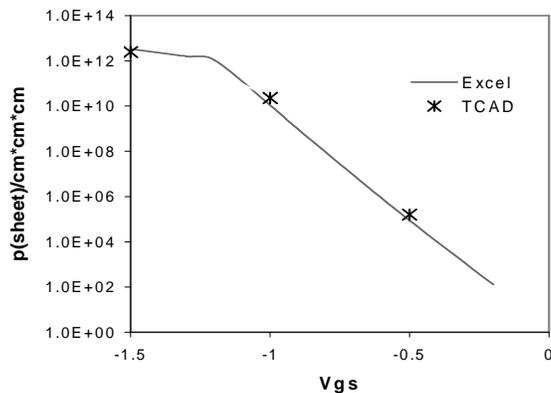


Figure 6-10. Integrated hole sheet density in the QW for the device 3,5,5,5/0.3/5e18,5e18 as a function of  $V_{gs}$ . Numerical simulator outputs are shown by asteriks

### 6.3.2 Hole Densities

Once the potentials are compared with numerical simulations and found to match well the hole densities are computed using the analytical program and compared to the results from the numerical simulator. The hole density at the top of the QW is given in Figure 6-8. Numerical simulation values of the same are also shown for three  $V_{gs}$  of  $-0.5$ ,  $-1.0$  and  $-1.5$  V. In the low and high  $V_{gs}$  the match is very good. The middle point is overestimated in the analytical program. This is due to the approximation of the  $H(\phi_H)$  function above and below  $V_T$ .

Comparing Figure 6-7 and Figure 6-8 the latter shows a higher percentage difference. This is due to the fact that the hole densities are exponential functions of potentials. However the match is very good at low and high  $V_{gs}$ . Figure 6-9 shows the hole density at the cap layer along with three numerical simulation values. The match between analytical and numerical simulations is reasonably good. It can be seen that these densities are an order of magnitude lower than the QW hole density given in Figure 6-7. Finally Figure 6-10 gives the integrated hole density in the QW or hole sheet density. The numerical simulation points again lie very close to the analytical curve. This sheet density could be used in computing on currents using the charge sheet approximation model [188].

## 6.4 Summary and Conclusions

In this chapter one-dimensional Poisson equation is solved along the mid-section of the p-HMOSFET. The results are manipulated to get the necessary parameters to evaluate the device. An analytical program is set up in Excel spread sheet and these equations are

programmed into it. A list of the variable input parameters and the working of the program are discussed. This program is easy to use, much less time consuming compared to the numerical simulators and presents many desirable internal and external parameters of the device. This makes the analytical program very useful for design of the p-HMOSFET.

## CHAPTER 7

### LONG CHANNEL DEVICE SIMULATIONS

Results of potentials and hole densities are compared to two dimensional numerical simulations using ISE-TCAD [109] in chapter 6. The accuracy of the analytical model is found to be good. In this chapter the dependencies of device characteristics on design variables are studied for the case of a long channel device. Based on these results appropriate design windows are created. From the design window an optimized device is formed and its terminal characteristics are computed using the analytical program. The same device is simulated numerically and the terminal characteristics are compared with the analytical results.

In section 7.1 design variables are varied one at a time and their effects on the internal and terminal characteristics of the device are studied using the analytical program. In section 7.2 design windows for appropriate device characteristics are constructed based on results from section 7.1. In section 7.3 an optimal device is constructed based on the design windows and its analytical terminal characteristics are compared to the numerical simulation results. Section 7.4 gives a summary of the results of the long channel device simulations and comparison with published results discussed in chapter 4.

#### **7.1 Parametric Simulations of Design Variables**

The structure of the device used is the same as the one described in chapter 4. This device is a sort of optimal device obtained through trial and error simulations using the numerical simulator. Maxima of  $I_{on}$  was the selection criteria. As mentioned in Chapter 5,

sub-section 5.3.9 the device is written in notational form as 3,5,5,5/0.3/5e18,5e18<sup>1</sup>. It is a long channel device with a channel length of 500 nm . The source and drain regions are p+ doped silicon. The design variables, selected in section 5.3, are given below in order of their relative influence on device characteristics.

1. Ge fraction of the SiGe QW channel, x
2. Substrate doping density, Nsub
3. p-doped layer thickness, Xp
4. Buffer layer thickness, Xbuf
5. p-layer doping density, Np

Total number of dopant atoms in the p-layer is given by  $N_p X_p$ . Hence the parameter p-layer doping density  $N_p$  variation is not simulated. Instead the doping density  $N_p$  is kept fixed at a moderate value of  $5e18/cm^3$  and the thickness  $X_p$  is varied. The p-layer doping density used here is lower than the delta doping density used in publications that are over  $10^{19}/cm^3$ . The lower doping density provides more built in voltage at the source junction comprising of p+ doped source and the p-layer. This helps to reduce off state thermionic and diffusion currents through the p-layer.

---

<sup>1</sup> Xcap = 3 nm/Xsige = 5 nm/Xbuf = 5 nm, Xp=5nmGe fraction in QW = 0.3,  
p-layer doping density =  $5 \times 10^{18}/cm^3$ /Nsub doping density =  $5 \times 10^{18}/cm^3$

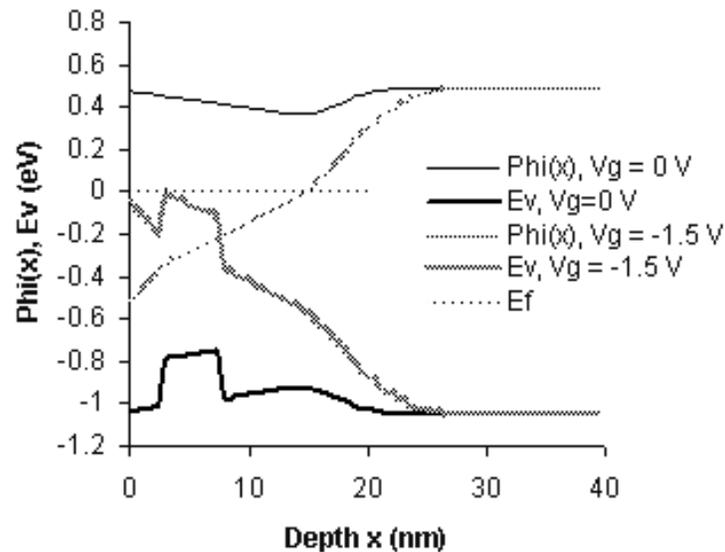


Figure 7-1. Potential and valence band plots for  $V_{gs} = 0$  V and  $V_{gs} = -1.5$  V. The device is 3,5,5,5,0.3/5e18,5e18. The lowermost curve and the uppermost curve corresponds to valence band edge  $E_v$  and potential  $\Phi(x)$  respectively for a  $V_{gs} = 0$  V

### 7.1.1 Ge fraction, $x$ , of the SiGe QW Channel

Germanium fraction<sup>2</sup>  $x$  of the QW is found to have the highest influence on the hole density at the top of the QW as well as the hole density ratio<sup>3</sup> compared to other variables. In order to get higher  $I_{ON}$  the hole density in the QW must be high. The hole density ratio<sup>3</sup> is important and should be high to ensure that conduction is confined to the QW channel with the contribution of the surface channel being negligible. The surface channel has highly degraded mobility, compared to the QW channel, due to the proximity

<sup>2</sup>  $x$  is the Ge composition given in  $\text{Si}_{1-x}\text{Ge}_x$

<sup>3</sup> Hole density ratio = Hole density at the top of the QW/hole density at the top of the cap layer

to the SiO<sub>2</sub> interface. For lower hole density ratios substantial conduction occurs in the surface channel thereby reducing the overall effective mobility.

In this subsection the value of Ge fraction  $x$  alone is varied from 0 to 0.6 keeping other structural variables fixed. The structure of the device is the same as the one given in section 5.3.9 and is given by the notation 3,5,5,5/ $x$ /5e18,5e18 (see footnote 1). It is a long channel device with a channel length of 500 nm. As mentioned in section 5.3.4 the useful range of  $x$  lies between 0.3 and 0.6. Below 0.3 the hole population at the top of the cap layer (surface channel) increases since the discontinuity of the valence band edge ( $\Delta E_V$ ) at the QW is lower. This can be seen in Figure 7-2 where the valence band edge and the Fermi level are plotted for a  $V_{gs}$  of  $-1.5$  V and  $x = 0.1$ . The surface channel is lying closer to the Fermi level than the QW top. Hence most of the conduction occurs in the surface channel and the mobility degrades to values found in conventional MOSFETs. In Figure 7-4 the situation is reversed. Now the top of the QW lies far higher than the surface band edge for an  $x$  value of 0.5. Hence most of the conduction happens in the QW layer thus yielding enhanced mobility. Above 0.6 the strain at the top of the quantum well might reduce, due to lower critical thickness. This will reduce the strain induced enhancement in mobility. For  $x > 0.6$  the mobility will tend to drop towards the bulk SiGe mobility and the device will tend to function like a recessed channel MOSFET, except mobility is better because the QW is undoped. Figure 7-5 shows the variation of hole density at the top of the QW with variation of  $x$  for a  $V_{gs}$  of  $-1.5$  V and a  $V_{ds}$  of  $-0.05$  V. With increasing value of  $x$  the hole density at the top of the QW increases. The density increases from  $1.8 \times 10^{18}/\text{cm}^3$  to  $4.5 \times 10^{19}/\text{cm}^3$  for  $x = 0.3$  to  $x=0.6$ . As expected Ge

fraction has the largest influence, than other structural parameters, upon hole concentration in the channel since the valence band edge ( $E_V$ ) in the QW is lifted towards the hole quasi Fermi level in a linear manner.

Figure 7-6 shows that there is a huge variation in the hole density ratio with increasing Ge fraction  $x$ . It varies from 0.01 to over 6100, being 7 for an  $x$  of 0.3. It can be seen that the increase in the ratio is slow up to a Ge fraction of 0.1. This is due to the fact that valence band edge ( $E_V$ ) at the top of the cap layer lies above that in the QW, as seen in Figure 7-2 and the surface hole density is more than the hole density in the QW. This region is not useful since the mobility degrades due to the surface channel adjacent the  $\text{SiO}_2$  interface. To get a ratio of unity a value of  $x$  about 0.22 is required for this structure. A hole density ratio above ten is preferable to ensure that only insignificant conduction happens through the surface layer. At around an  $x$  value of 0.3 this ratio is approximately ten. So the lower limit of  $x$  can be fixed at a value of 0.3. The large increase in the ratio, with increasing  $x$ , is due to the fact that the top of the QW is raised towards the quasi Fermi level and this leads to an exponential rise in hole density in the QW.

Threshold voltage  $V_T$  variation with Ge fraction  $x$  is shown in Figure 7-7.  $V_T$ , the QW channel threshold, varies approximately from  $-1.5$  to  $-1$  V, with a value of  $-1.25$  for an  $x$  of 0.3.  $V_{TS}$ , the surface threshold<sup>4</sup>, varies approximately from  $-1.5$  to  $-2$  V, being  $-1.7$  for an  $x$  of 0.3.  $V_{TS}$  sets the limit for the highest operating voltage since the surface channel conduction dominates for  $V_{gs}$  above  $V_{TS}$ . The build up of holes at the surface further

---

<sup>4</sup>  $V_T$  is the threshold voltage for the QW top which is defined as the gate voltage needed to get a hole density equal to the  $N_{sub}$  doping density of the substrate below. Similarly  $V_{TS}$  is the threshold voltage for the surface channel.

shields the QW and the enhanced mobility advantage is lost very fast. In order to operate the device in a high  $I_{ON}$  mode a higher  $V_{gs}$  is desired. This is possible only if the difference between  $V_T$  and  $V_{TS}$  is large with  $V_{TS}$  being more negative than  $-1.5$  V. So a high Ge content is preferable. The difference between  $V_T$  and  $V_{TS}$  increases as  $x$  is increased. At a value of  $x$  equal to zero there is no valence band discontinuity and the device has only a surface channel and acts merely like a MOSFET. As  $x$  is increased the QW height increases and  $E_V$  at the surface is lowered since the top of the QW almost enters the Fermi level. This causes a reduction in magnitude of  $V_T$  and an increase in the magnitude of  $V_{TS}$  as seen in the Figure 7-7. which is desirable for higher voltage operation, upto  $-1.5$  V. It can be seen that an  $x$  value of about 0.5 is needed for a  $V_T$  of  $-1.0$  V for this structure. At this value of  $x$   $V_{TS}$  is around  $-2.0$  V. This gives sufficient margin for  $-1.5$  V operation. So this value of  $x$  can be taken as an upper limit. For  $x$  greater than 0.5 the reduction of  $V_T$  will be countered by the strain relaxation imposed by the critical thickness. So an  $x$  value of 0.3 to 0.5 can be chosen as a useful range of Ge fraction in SiGe QW.

In summary the Ge fraction has a strong effect on QW hole density and a huge effect on hole density ratio (footnote 3). Increasing  $x$  causes  $V_T$  to reduce and  $V_{TS}$  to increase which is exactly what is needed to make a device with high  $I_{ON}$  and capable of operating at a higher voltage. Unlike MOSFET this device cannot be operated at  $V_{gs}$  much higher than the threshold voltage  $V_T$  to get higher currents. The maximum  $V_{gs}$ , equal to  $V_{dd}$ , is limited to  $V_{TS}$ .

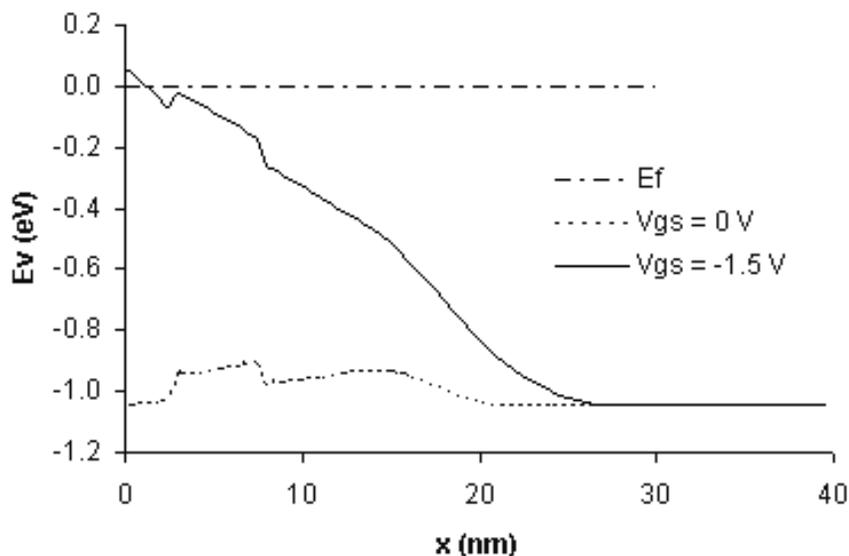


Figure 7-2. Valence band plots starting from the top of the cap layer ( $x = 0$ ) for a  $x = 0.1$ , for  $V_{gs} = 0$  V and  $V_{gs} = -1.5$  V,  $V_{ds} = -0.05$  V. Device is 3,5,5,5/0.1/5e18,5e18.  $L_g = 500$  nm. Note that the surface channel is predominant.

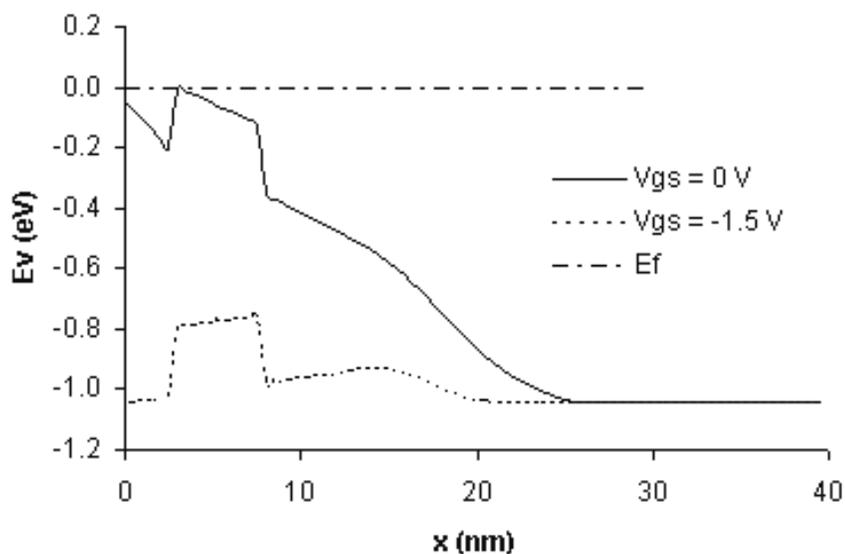


Figure 7-3. Valence band plots starting from the top of the cap layer ( $x = 0$ ) for a  $x = 0.3$ , for  $V_{gs} = 0$  V and  $V_{gs} = -1.5$  V,  $V_{ds} = -0.05$  V. Device is 3,5,5,5/0.3/5e18,5e18.  $L_g = 500$  nm.

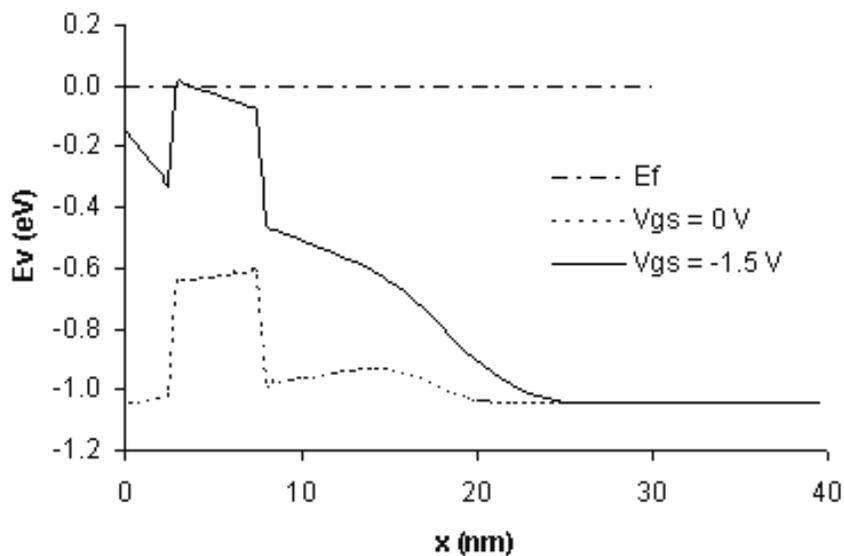


Figure 7-4. Valence band plots starting from the top of the cap layer ( $x = 0$ ) for a Ge fraction  $x = 0.5$ , for  $V_{gs} = 0$  V and  $V_{gs} = -1.5$  V,  $V_{ds} = -0.05$  V. Device is 3,5,5,5/0.5/5e18,5e18.  $L_g = 500$  nm.

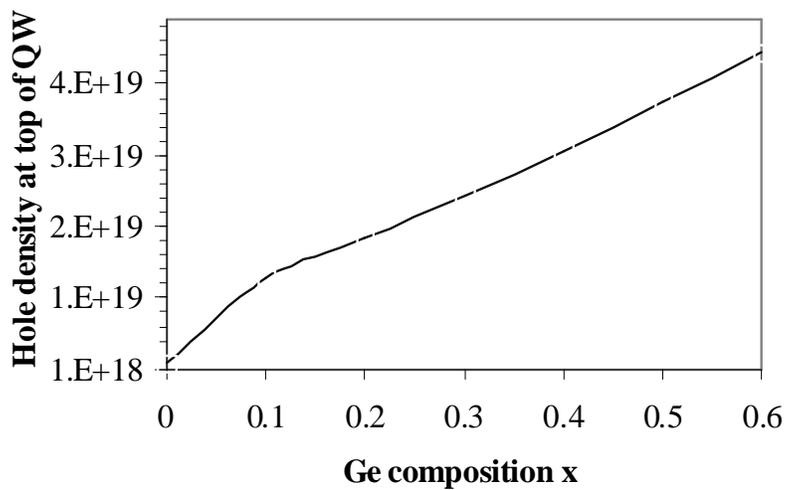


Figure 7-5. Hole density variation at the top of the QW with variation of Ge fraction  $x$  in the QW. The device is 3,5,5,5/ $x$ /5e18,5e18 (see footnote 1),  $L_g = 500$  nm,  $V_{gs} = -1.5$  V,  $V_{ds} = -0.05$  V.

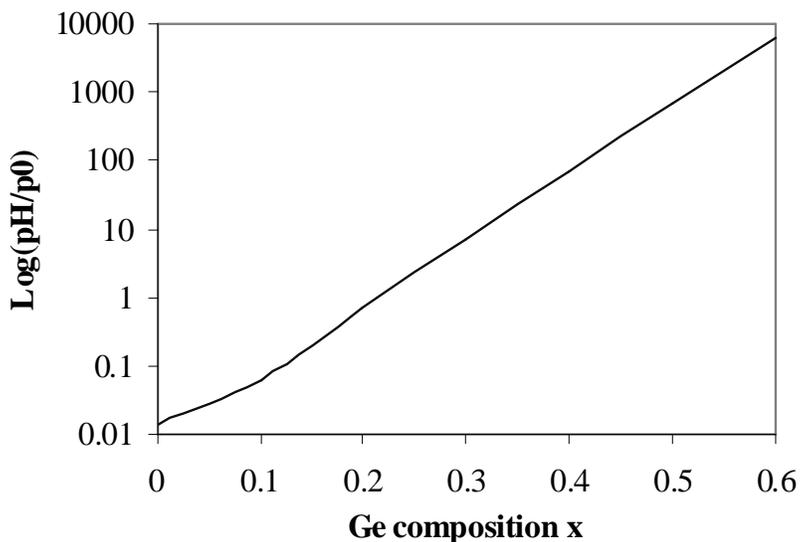


Figure 7-6. Variation of the ratio of hole density (see footnote 3) at top of QW to that at the top of the cap layer with variation in Ge fraction  $x$ .  $V_{gs} = -1.5$  V,  $V_{ds} = -0.05$  V. Device is  $3,5,5,5/x/5e18,5e18$ ,  $L_g = 500$  nm

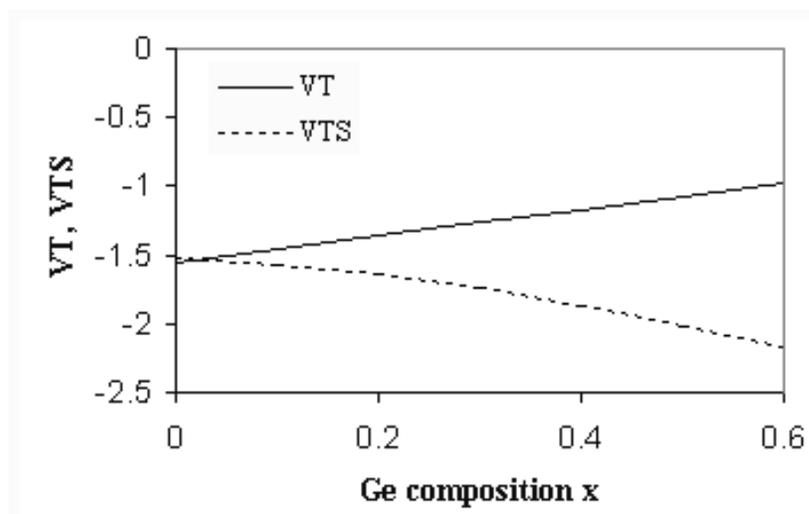


Figure 7-7. Threshold voltage variations with variation of Ge fraction  $x$  of the QW channel layer. Device is  $3,5,5,5/x/5e18,5e18$ ,  $L_g = 500$  nm

### 7.1.2 Substrate Doping Density, $N_{sub}$

The minimum doping density of the substrate has to be selected such that subsurface punch through does not happen. So  $N_{sub}$  cannot be varied arbitrarily. Minimum value of

$N_{\text{sub}}$  can be found by adding the source and drain depletion widths and equating to channel length  $L_g$ .  $V_T$  is a function of the substrate doping. In order to have high  $I_{\text{ON}}$   $V_T$  has to be low.

In the following  $N_{\text{sub}}$  is varied and its effect on threshold voltage and QW hole density are studied. Figure 7-8 shows the effect on hole density at the top of the QW for variation in  $N_{\text{sub}}$  for a  $V_{\text{gs}}$  of  $-1.5$  V and a  $V_{\text{ds}}$  of  $-0.05$  V.  $V_T$  is calculated from structural parameters. Potential and valence band edge  $E_v$  along the depth is plotted in Figure 7-9 for a low and high  $N_{\text{sub}}$ . The potentials for both plots seem to overlap in the cap layer. As can be seen in Figure 7-8 the substrate doping has only a small effect on the hole density. For both cases the top of the QW is located nearly at the same level. So both cases have the same hole density. This makes the field in the cap layer same for both cases. So the potential distribution in the cap layer for both cases are the same.

Figure 7-10 shows the ratio of the hole density at the top of the QW to the hole density at the top of the cap layer as a function of  $N_{\text{sub}}$ . In Figure 7-8 for a variation of doping from  $1 \times 10^{18}/\text{cm}^3$  to  $1 \times 10^{19}/\text{cm}^3$  the hole density varies only from  $2.9 \times 10^{19}/\text{cm}^3$  to  $2.3 \times 10^{19}/\text{cm}^3$  or by 21%. But the ratio of hole densities varies from 13 to 5.4. The ratio falls as  $N_{\text{sub}}$  is increased. A higher hole density ratio is desirable to have a  $V_{\text{TS}}$  much higher than  $V_T$  so that the surface channel can be made insignificant at the maximum gate voltage of  $-1.5$  V.

Figure 7-11 shows the dependence of  $V_T$  and  $V_{\text{TS}}$  on substrate doping. For a doping variation from  $1 \times 10^{18}/\text{cm}^3$  to  $1 \times 10^{19}/\text{cm}^3$   $V_T$  changes from  $-0.92$  to  $-1.4$  V and  $V_{\text{TS}}$  changes from  $-1.4$  to  $-1.9$  V. From the figure it could be noted that the difference

between  $V_T$  and  $V_{TS}$  remains unchanged. The minimum value of  $N_{\text{sub}}$  is  $2 \times 10^{18}/\text{cm}^3$  to have a  $V_{TS}$  of  $-1.5$  V. Whilst the variation of  $V_T$  and  $V_{TS}$  are controlled by the doping  $N_{\text{sub}}$  their difference is controlled by the QW Ge fraction  $x$  and the cap layer thickness. Since these two parameters are not changed the difference between the threshold voltages remains unchanged. It can also be seen in Figure 7-11 that a low  $N_{\text{sub}}$  is needed to get  $V_T$  in the range of  $-1$  V.  $V_T$  adjustment over a range of  $0.5$  V can be done by changing  $N_{\text{sub}}$  for the structure considered here.

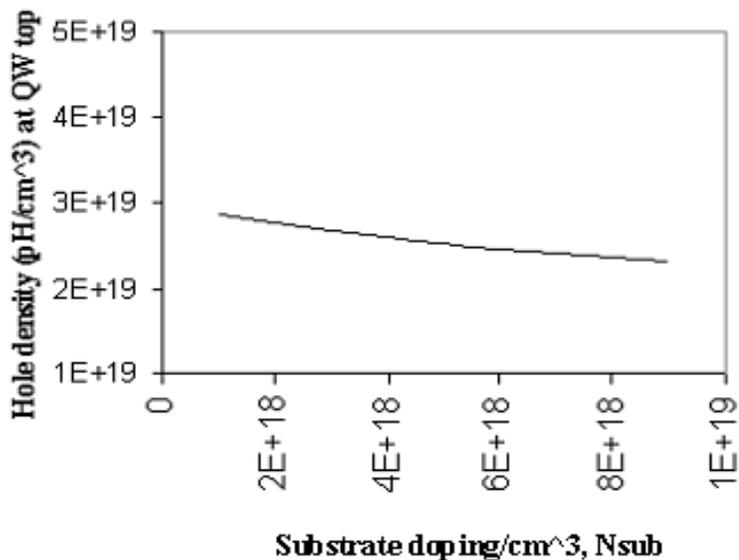


Figure 7-8. Hole density at the top of the QW as a function of substrate doping for the device 3,5,5,5/0.3/5e18,Nsub.  $V_{\text{gs}} = -1.5$  V,  $V_{\text{ds}} = -0.05$  V

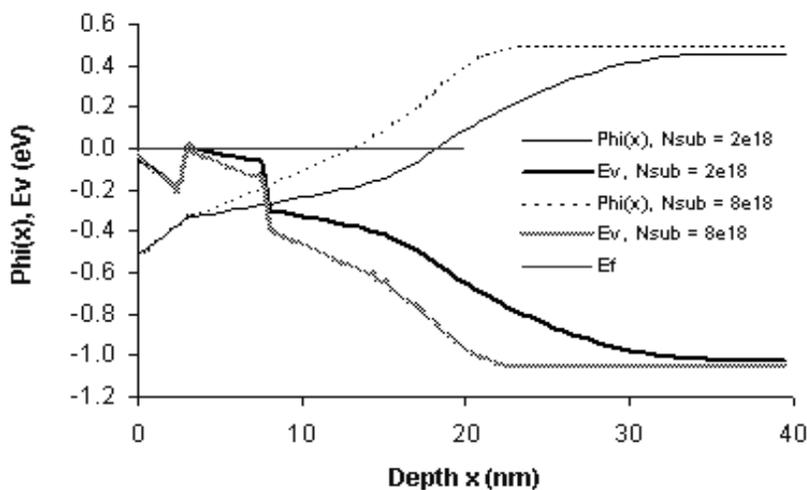


Figure 7-9. Potential and valence band edge plots along the depth starting from the top of the cap layer ( $x = 0$ ).  $V_{gs} = -1.5$  V,  $V_{ds} = -0.05$  V. The device is 3,5,5,5/0.3/5e18, $N_{sub}$ . Two sets of curves are for  $N_{sub}$  values of  $2 \times 10^{18}/\text{cm}^3$  and  $8 \times 10^{18}/\text{cm}^3$ . Solid curves are Ev and  $\Phi(x)$  for  $2 \times 10^{18}/\text{cm}^3$

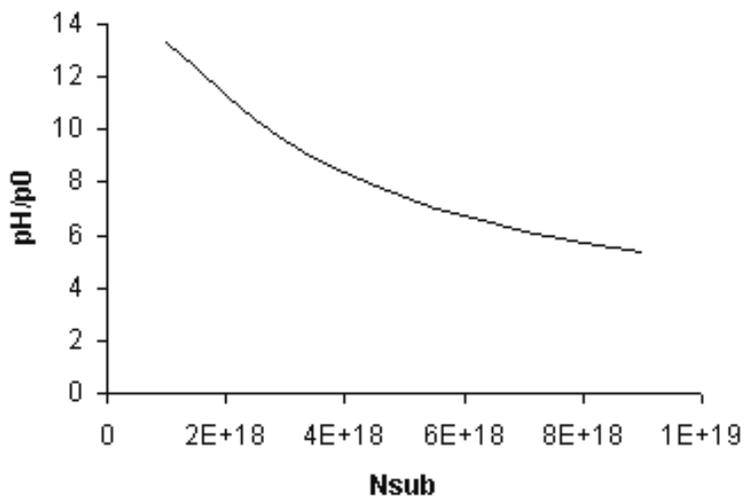


Figure 7-10. Ratio of hole density at the top of QW to hole density at the top of the cap layer as a function of  $N_{sub}$ . Device is 3,5,5,5/0.3/5e18, $N_{sub}$ .  $V_{gs} = -1.5$  V,  $V_{ds} = -0.05$  V,  $L_g = 500$  nm.

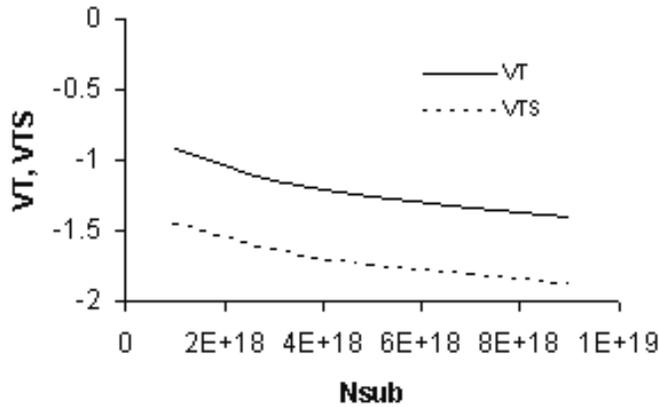


Figure 7-11. Substrate doping dependence of  $V_T$  and  $V_{TS}$  for the device 3,5,5,5/0.3/5e18,  $N_{sub}$ .  $L_g = 500$  nm

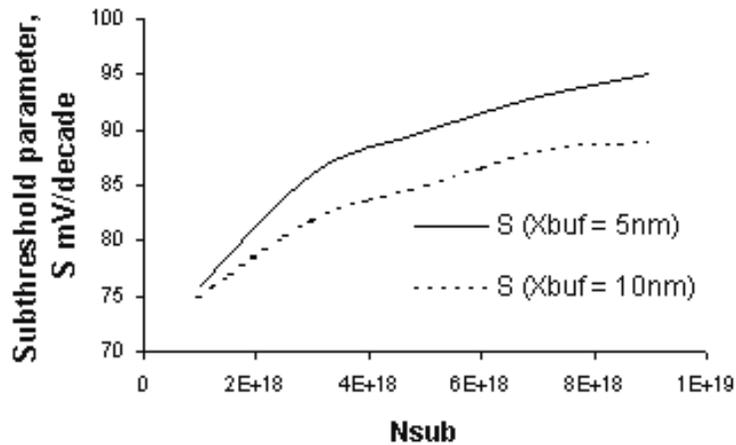


Figure 7-12. Sub threshold parameter (S) dependence on substrate doping concentration for the device 3,5,5,5/0.3/5e18,  $N_{sub}$ .  $L_g = 500$  nm

Sub threshold parameter  $S$  is defined as,

$$S = 2.3m \frac{kT}{q} \quad (7.1)$$

$m$  is the body parameter defined in terms of the coupling efficiency  $\eta$  as follows,

$$\eta = \frac{1}{1 + \frac{C_D}{C_O}} \quad (7.2)$$

$$m = 1/\eta \quad (7.3)$$

$C_O$  is the capacitance from gate to channel which is the series combination of oxide and cap layer capacitances.  $C_D$  is the series combination of capacitances of QW layer, buffer layer, p-layer and the depletion region in the substrate. The depletion thickness is given by,

$$W_{dm} = \sqrt{\frac{4\epsilon_{Si}\Phi_B}{qN_{sub}}} \quad (7.4)$$

$\Phi_B$  is the doping parameter given by,

$$\Phi_B = \frac{kT}{q} \ln\left(\frac{N_{sub}}{n_i}\right) \quad (7.5)$$

The variation of sub threshold parameter (S) with variation of substrate doping is given in Figure 7-12 for two buffer layer thicknesses. There is a large variation from 85 to 101 for buffer layer thickness of 5nm and from 81 to 92 for buffer layer thickness of 10nm. The depletion capacitance increases with increasing substrate doping since the depletion thickness in the substrate reduces. This increases  $C_D$  and the body coefficient  $m$  which increases the value of S as given in equation ( 7.1 ). When the depletion capacitance is smaller (wider depletion thickness in the substrate), as in the case of low doping, an increase in buffer layer thickness do not make much difference in  $C_D$  since the

series combination is dominated by the depletion capacitance. But when the depletion capacitance is larger (small depletion width) a change in buffer layer thickness makes a larger difference since the series combination is now dominated by the buffer layer capacitance. This is seen in figure as an increasing difference between the two curves as  $N_{\text{sub}}$  is increased. At higher doping levels the series combination of  $C_D$  is dominated by the QW, buffer and p-layer capacitances. Hence the value of  $S$  saturates. Since a low value of  $S$  is desirable to reduce sub threshold currents lower  $N_{\text{sub}}$  seems to be the choice. Having a thicker buffer layer also seems desirable. However as will be seen later in chapter 7 the sub threshold currents in HMOSFETs are much lower than those in MOSFETs and hence the sub threshold parameter may not be of concern.

Overall the doping density of the substrate may not be a good parameter of choice. The strongest effects are upon threshold voltages and the sub threshold parameter. In both cases a low doping density is desirable as seen in Figure 7-11 and Figure 7-12. As mentioned earlier sub threshold parameter may not be of importance. In order to avoid subsurface punch through higher doping densities may be needed. So it is better to control the threshold voltages by changing the Ge fraction  $x$  or by changing the p-layer thickness.

### 7.1.3 p-doped Layer Thickness, $X_p$

Changing p-layer doping or thickness is expected to produce almost similar results since the total number of dopant atoms is given by the product  $N_p X_p$  where  $N_p$  is the doping density and  $X_p$  is the thickness. There will be a slight difference due to the shift in the centroid of the ionic charge when the thickness is varied. As mentioned in

section 5.3.2 a lower p-layer doping provides a higher barrier at the source - p-layer junction. This reduces the off currents through the p-layer. The hole population in the p-layer could also contribute to  $I_{OFF}$ . In the case of  $I_{ON}$  the QW hole density is much higher compared to the p-layer hole density. A doping density of  $5 \times 10^{18}/\text{cm}^3$  gives a doping parameter  $\phi_B$  of 0.5 eV giving a barrier at the source over twice the thermal voltage. So this may be a maximum p-layer doping allowed. Hence the doping density is kept fixed at  $5 \times 10^{18}/\text{cm}^3$ . Keeping the acceptor doping at  $5 \times 10^{18}/\text{cm}^3$  the thickness of the p-layer is changed from 3 to 10 nm for a device 3,5,5, $X_p/0.3/5 \times 10^{18}, 5 \times 10^{18}$  (see footnote 1).

Figure 7-13 shows the variation of hole density at the top of the SiGe QW as the p-layer thickness is changed. It is to be noted that the  $N_{sub}$  doping tilts the bands upwards whereas the p-layer doping tilts the bands downwards. Changing the thickness amounts to increasing the total dopant atoms. This tilts the band edge downwards. Downward tilt of band edge and the fixed  $V_{gs}$  moves the QW upwards closer to the Fermi level. This leads to an increase in the hole density. The density changes from  $2.2 \times 10^{19}/\text{cm}^3$  to  $2.9 \times 10^{19}/\text{cm}^3$  or by 32%. Also the increase in density seems to saturate at about 9 nm. This is because the field in the buffer layer is reaching zero with no potential drop in the buffer layer. So an increase in p-layer thickness will not move the QW upwards, instead will move the QW downwards. So a thick p-layer is not desirable. Also a thick p-layer will have more potential drop and will contribute to  $I_{OFF}$ . From these points the p-layer design is more complex. As a note at this stage the Boltzman approximation of the carrier density is not valid and Fermi statistics has to be used. Figure 7-15 shows the variation of

the ratio of the hole density at the top of the QW to the density at the top of the cap layer. As  $X_p$  increases there is a variation from 4.2 to 13.6. This is due to the slight tilting down of the valence band edge in the cap layer and a rise of the band edge in the QW as seen in Figure 7-14. At the top of the cap layer this slight lowering of valence band edge for  $X_p$  of 8 nm can be seen.

Figure 7-16 gives the variation of  $V_T$  and  $V_{TS}$  with increasing p-layer thickness.  $V_T$  drops from  $-1.4$  to  $-1$  V and  $V_{TS}$  drops from  $-1.9$  to  $-1.6$  V. But the difference between them stays fixed. This difference is determined by the Ge x and the cap layer thickness which are kept fixed. The reduction in  $V_T$  can be explained as follows. The p-dopant field bends the band edge downwards. Since the gate potential is fixed this field bending causes the QW to move upwards. This increases the hole density in the QW.

Variation of sub threshold parameter (S) is given in Figure 7-17. It varies from 103 to 90 and the variation is nearly similar for p-layer and buffer layer thickness variation. Both  $X_p$  and  $X_{buf}$  change  $C_D$ , the depletion capacitance, the same way giving similar variation of S. From the above p-layer doping or thickness can be independently used to control threshold voltages whereas buffer layer thickness can be used for varying S, if needed.

It is to be noted that the threshold voltage increases with increasing  $N_{sub}$  whereas they drop by almost the same amount for increasing  $X_p$  for the range considered, 2 to 8 nm. This however suggests that the p-layer thickness can be used to control the value of  $V_T$ , a higher thickness giving a lower  $V_T$ . This is useful because  $N_{sub}$  can be chosen

independently to resist subsurface punch through while  $X_p$  or  $N_p$  can be used to adjust threshold voltages.

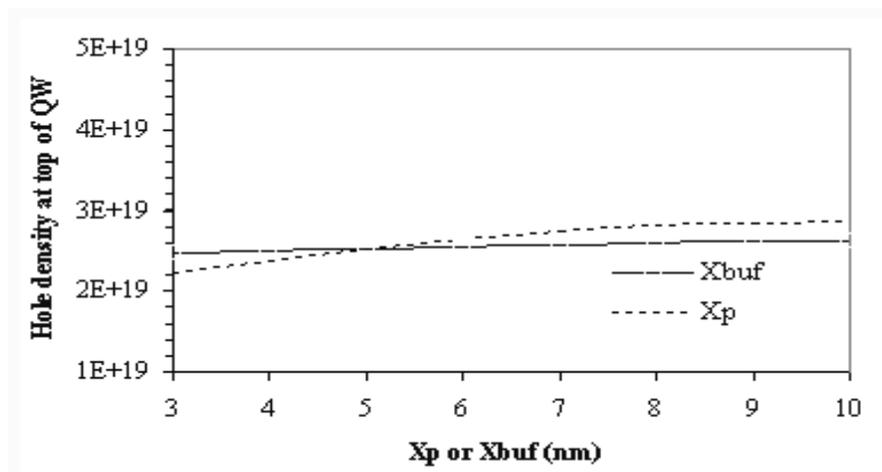


Figure 7-13. Dependence of hole density at the top of the QW upon the thickness of the p-layer and the buffer layer. The device is 3,5, $X_{buf}$ , $X_p/0.3/5e18,5e18$  and 500nm channel length. While varying one of them the other is kept fixed at 5 nm

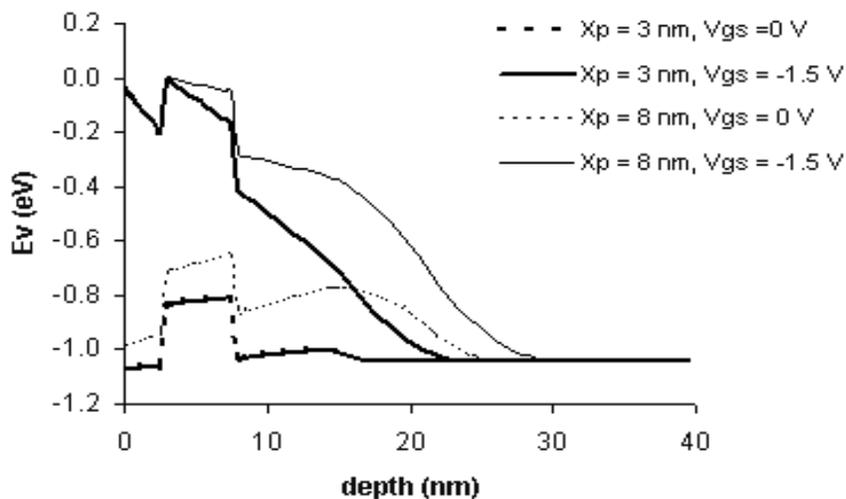


Figure 7-14. Valence band edge along depth starting from the top of the cap layer ( $x = 0$ ) for two p-layer thicknesses.  $X_{buf} = 5$  nm. Device is 3,5, $X_{buf}$ , $X_p/0.3/5e18,5e18$  and 500nm channel length

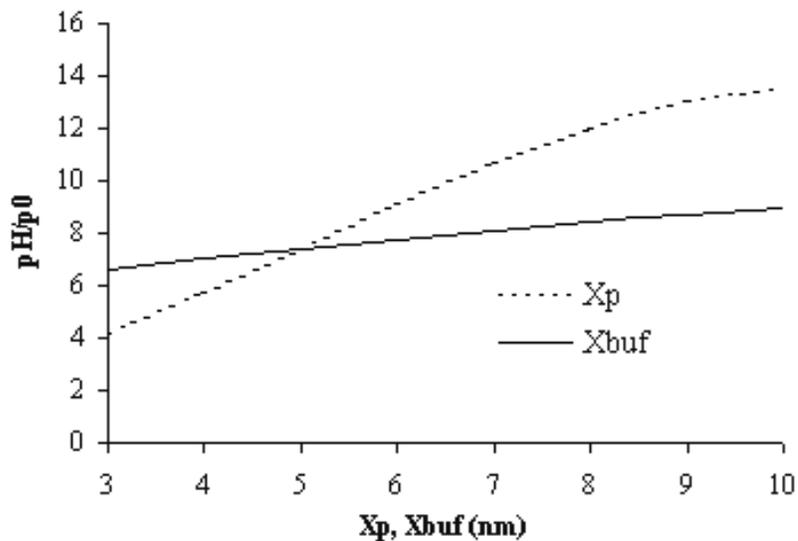


Figure 7-15. Dependence of the ratio of the hole density at the top of the QW to the density at the top of the cap layer with variation of p-layer and buffer layer thickness. While varying one of them the other is kept fixed at 5 nm. Device is 3,5,Xbuf,Xp/0.3/5e18,5e18 and 500nm channel length

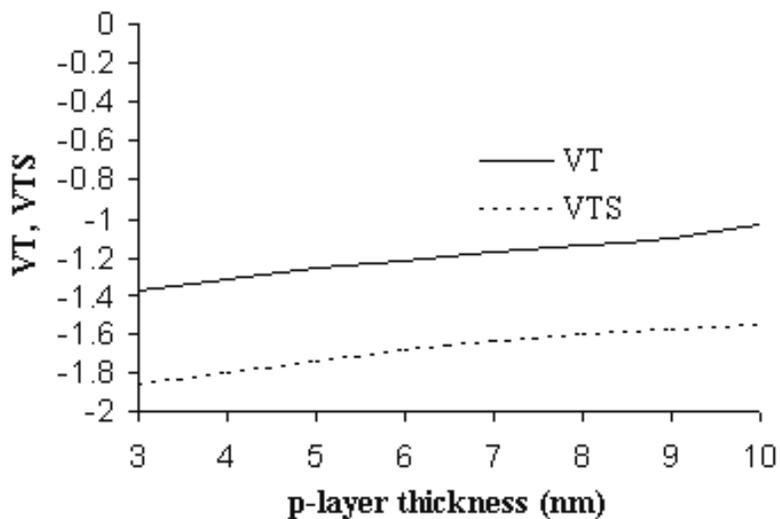


Figure 7-16. Dependence of  $V_T$  and  $V_{TS}$  upon the thickness of the p-layer  $X_p$ . Device is 3,5,5, $X_p$ /0.3/5e18,5e18 and 500nm channel length.

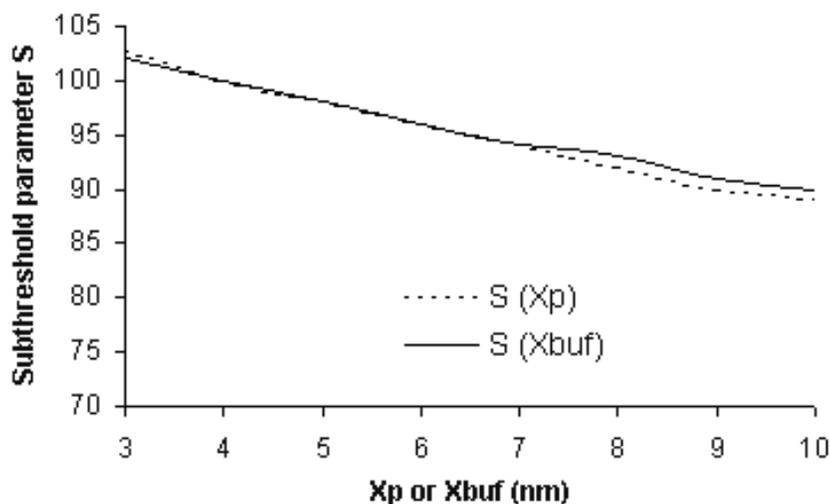


Figure 7-17. Variation of the sub threshold parameter S with variation of p-layer thickness and buffer layer thickness. One is varied at a time keeping the other fixed at 5 nm. Device is 3,5,Xbuf,Xp/0.3/5e18,5e18 and 500nm channel length

#### 7.1.4 Buffer Layer Thickness, Xbuf

Same as the p-layer thickness variation in the previous subsection the buffer layer thickness is varied from 3 to 10 nm while keeping the rest of the parameters fixed. Referring to Figure 7-15 the hole density shows only very little increase, below 7% . The buffer layer simply causes a linear potential drop in itself due to the field at the top of the p-layer. In Figure 7-14 the buffer layer is located between 8 and 13 nm where this linear drop can be observed. For the  $N_{\text{sub}}$  and  $X_p$  chosen here this drop which is very low, due to the thin buffer layer, causes the valence band edge to move up towards the quasi Fermi level increasing the hole density in the QW slightly. The ratio of hole densities at top of QW and top of the cap layer also shows a slight increase, as seen in Figure 7-15 due to the same reason. The variation of threshold voltage also seems minimal as seen in

Figure 7-18.  $V_T$  varies from  $-1.3$  to  $-1.2$  V and  $V_{TS}$  varies from  $-1.8$  to  $-1.7$  V. The difference between them remains unaltered since  $V_T$  depends on  $X_{cap}$  and  $x$  which are fixed. So the buffer layer seems to have negligible effect on device properties if it is thin and hence the potential drop in it is very low. This is true only if the field in the buffer layer is low. The field in the buffer layer depends on the difference of charge in the substrate depletion layer and the p-layer. This field can be made zero by proper choice of p-layer thickness. This is discussed in chapter 8. For low  $N_{sub}$  and high p-layer doping or thickness the potential drop in the buffer can be higher compared to the structure used here. Buffer layer thickness does not alter the hole density, hole density ratio or the threshold voltages significantly. It however can be used to adjust the sub-threshold slope parameter if needed as mentioned in the previous sub-section.

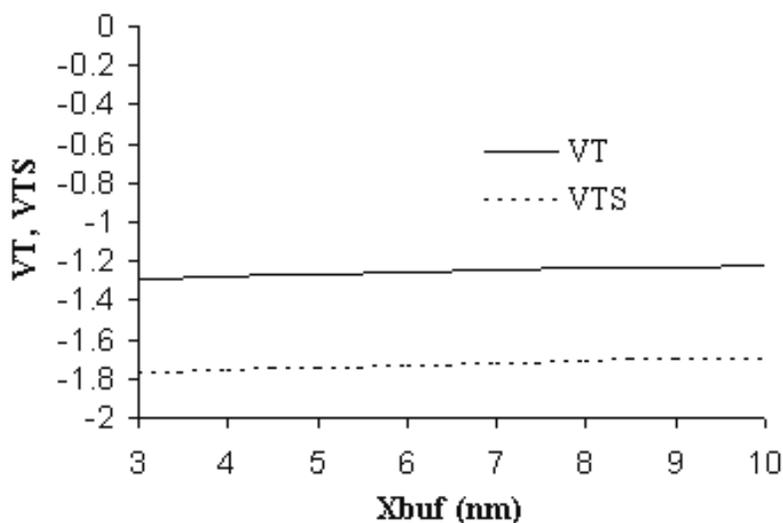


Figure 7-18. Threshold voltage variation with buffer layer thickness.  $X_p = 5$  nm. The device is 3, 5,  $X_{buf}$ , 5/0.3/5e18, 5e18 and 500 nm channel length.

## 7.2 Long Channel Device Design Windows

From the above section the following observations are made, Xbuf has only very minimal effect on hole density, density ratio and threshold voltages  $N_{\text{sub}}$  variation has only slight effect on QW hole density.  $N_{\text{sub}}$  variation from  $1 \times 10^{18}/\text{cm}^3$  to  $1 \times 10^{19}/\text{cm}^3$  reduces QW hole density from  $2.9 \times 10^{19}$  to  $2.3 \times 10^{19}/\text{cm}^3$  or about -20%. However substrate doping has moderate effect on hole density ratio which varies from 13 to 5. The lower limit of  $N_{\text{sub}}$  is  $2 \times 10^{18}/\text{cm}^3$  for the structure chosen. Below  $2 \times 10^{18}/\text{cm}^3$   $V_{\text{TS}}$  falls below -1.5 V.  $N_{\text{sub}}$  variation from  $2 \times 10^{18}/\text{cm}^3$  to  $1 \times 10^{19}/\text{cm}^3$  increases  $V_{\text{T}}$  from -1 to -1.4 V  $X_{\text{p}}$  variation from 3 to 10 nm varies the hole density in the QW from about  $2 \times 10^{19}/\text{cm}^3$  to  $3 \times 10^{19}/\text{cm}^3$  and tends to saturate at about 8 nm.  $X_{\text{p}}$  variation has a moderate effect on the hole density ratio. Variation of  $X_{\text{p}}$  from 3 to 10 nm varies the hole density ratio from 4 to 14, about the same as the case of  $N_{\text{sub}}$ . However the variation is in opposite direction. Increasing  $X_{\text{p}}$  increases the ratio while increasing  $N_{\text{sub}}$  decreases the ratio.  $X_{\text{p}}$  variation has a moderate effect on threshold voltages. Variation from 3 to 10 nm reduces  $V_{\text{T}}$  from -1.4 to -1 V, again almost the same as the  $N_{\text{sub}}$  case but opposite in nature. Increasing  $X_{\text{p}}$  reduces  $V_{\text{T}}$  whereas increasing  $N_{\text{sub}}$  increases  $V_{\text{T}}$ . Ge fraction  $x$  has a moderate effect on QW hole density. Changing  $x$  from 0.3 to 0.6 increases the hole density from  $2.5 \times 10^{19}/\text{cm}^3$  to  $4.5 \times 10^{19}/\text{cm}^3$ . Ge fraction  $x$  has a huge effect on the hole density ratio. Variation from 0.3 to 0.6 increases the hole density ratio from 7 to over 6100. This could be due to the lowering of band edge in the cap layer which will reduce the surface hole density. Increasing  $x$  reduces  $V_{\text{T}}$  and increases  $V_{\text{TS}}$ .  $V_{\text{T}}$  varies from -1.3 V to -1 V for a variation

in  $x$  from 0.3 to 0.6, the useful range discussed in section 6.1.1. The rate of increase in  $V_{TS}$  is larger than the same for  $V_T$ .  $V_{TS}$  is larger than  $-1.5$  V in this range.

Based on the above  $X_{buf}$  is not considered as an effective design parameter. It is needed to reduce Coulomb scatter of carriers due to the negative ions in the p-layer. So it can be chosen as 5 nm for reasons given in section 4.3.3. If the buffer layer is thicker and the field in it is large it will have more effect on  $V_T$ ,  $V_{TS}$  and hole density. So the Ge fraction,  $N_{sub}$  and  $X_p$  are varied in the simulations keeping  $X_{cap}$ ,  $X_{SiGe}$  and  $X_{buf}$  fixed at 3 nm, 5 nm and 5 nm respectively.

Since the device is to operate at a voltage of  $-1.5$  V the surface threshold  $V_{TS}$  should be higher than 1.5 in magnitude. A QW channel threshold  $V_T$  around  $-1$  V is desired for high  $I_{ON}$ . This gives an overdrive of 0.5 V. So these values are set as the boundary of the design space. Figure 7-19, Figure 7-20 and Figure 7-21 give design spaces of  $N_{sub}$  vs  $x$  for three p-layer thicknesses of 2, 5 and 8 nm. The design space is indicated with an arrow pointing into the space. The two curves are for  $V_T = -1$  V and  $V_{TS} = -1.5$  V respectively or they represent the boundaries. For example region to the top of the  $V_T$  curve has  $V_T$  magnitudes higher than 1 V. Similarly the region to the bottom of  $V_{TS}$  curve has magnitude of  $V_{TS}$  less than  $-1.5$  V. So the design space lies between the two curves where  $V_T > -1$  V and  $V_{TS} < -1.5$  V. As an example in Figure 7-19 the range of  $N_{sub}$  for an  $x$  of 0.4 is about  $3 \times 10^{17}/\text{cm}^3$  to  $1 \times 10^{18}/\text{cm}^3$ . In all three figures the design space expands with increasing  $x$  and is wider for smaller  $X_p$  values. Referring to equation (4-4) an increase in  $x$  leads to a higher  $\Delta E_v$ , which makes  $V_T$  more positive. In order to restore  $V_T$  to  $-1.0$  V the magnitude of the doping parameter  $\phi_B$  has to be increased. This can be done

by an increase in  $N_{\text{sub}}$ , This can be inferred from Figure 7-19 to Figure 7-21. Referring to equation An increase in  $X_p$  adds a positive term, which has to be countered by an increase of the negative term  $\phi_4$ . This can be done by an increase of  $N_{\text{sub}}$ . This is also seen in Figure 7-19 to Figure 7-21. From Figure 7-19 it can be observed that an  $x$  value above

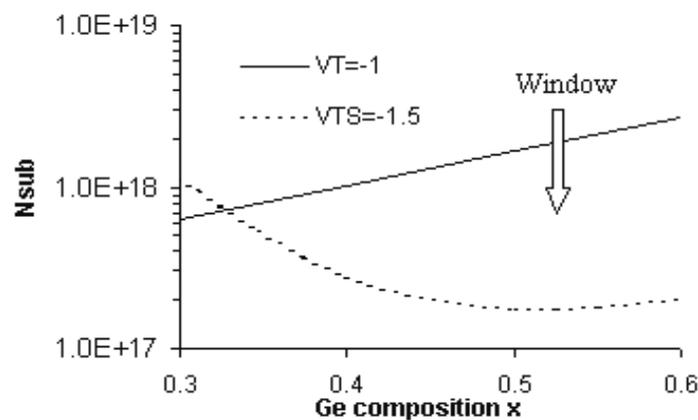


Figure 7-19.  $N_{\text{sub}}$  vs  $x$  design space for  $X_p = 2$  nm. Device is 3,5,5,2/x/5e18,5e18. Combinations of  $x$  and  $N_{\text{sub}}$  to get  $V_T = -1$  V is given by the solid line. Above this line magnitude of  $V_T > 1$  V. Combinations of  $x$  and  $N_{\text{sub}}$  to get  $V_{\text{TS}} = -1.5$  V is given by the dashed line. Below this line magnitude of  $V_{\text{TS}} < -1.5$  V. So the window lies in between the two lines.

0.33 is essential to have a design space at all. A wider design space allows wider choice of  $N_{\text{sub}}$  to avoid subsurface punch through. Also this allows wider choice of  $V_T$  to get a higher overdrive ( $V_{\text{gs}} - V_T$ ). It can be inferred from Figure 7-19 to Figure 7-21 that a high doped p delta layer or a thick p-layer is not desirable since it constricts the design space.

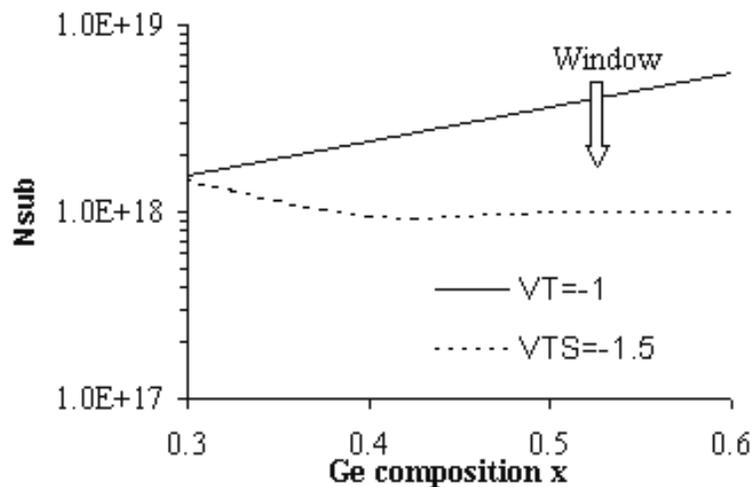


Figure 7-20.  $N_{sub}$  vs  $x$  design space for  $X_p = 5$  nm. Device is 3,5,5,5/ $x$ /5e18,5e18. Combinations of  $x$  and  $N_{sub}$  to get  $V_T = -1$  V is given by the solid line. Above this line magnitude of  $V_T > 1$  V. Combinations of  $x$  and  $N_{sub}$  to get  $V_{TS} = -1.5$  V is given by the dashed line. Below this line magnitude of  $V_{TS} < -1.5$  V. So the window lies in between the two lines.

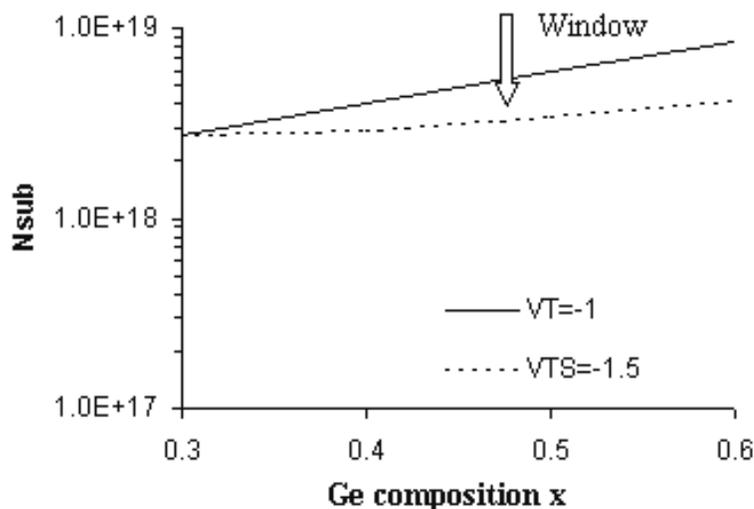


Figure 7-21.  $N_{sub}$  vs  $x$  design space for  $X_p$  of 8 nm. Device is 3,5,5,8/ $x$ /5e18,5e18. Combinations of  $x$  and  $N_{sub}$  to get  $V_T = -1$  V is given by the solid line. Above this line magnitude of  $V_T > 1$  V. Combinations of  $x$  and  $N_{sub}$  to get  $V_{TS} = -1.5$  V is given by the dashed line. Below this line magnitude of  $V_{TS} < -1.5$  V. So the window lies in between the two lines.

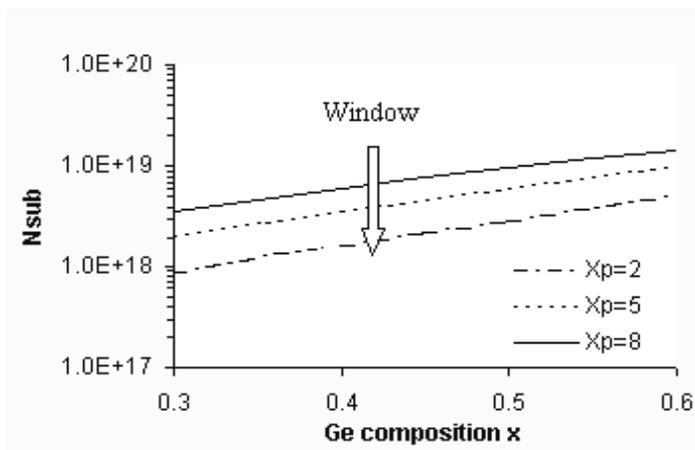


Figure 7-22. Design space for  $I_{ON} = 100 \mu\text{A}$  for the device 3,5,5, $X_p/x/5e18,5e18$  for p-layer thickness  $X_p$  of 2, 5 and 8 nm.  $X_{buf} = 5$  nm. The area below the respective lines have  $I_{ON}$  more than  $100 \mu\text{A}$ .

In Figure 7-22 the boundary for  $V_T = -1$  V and  $V_{TS} = -1.5$  V is plotted for  $X_p$  values of 2, 5 and 8 nm. The design space is below the respective lines and is shown by the arrow. In this region any combination of  $x$  and  $N_{apt}$  will satisfy the above requirement on threshold voltages and will give a device which can operate at  $-1.5$  V. But  $I_{ON}$  and  $I_{OFF}$  will depend on the particular combination of  $x$  and  $N_{apt}$ .

Figure 7-22 gives design space for  $I_{ON} > 100 \mu\text{A}$  for three values of  $X_p$  and for an  $X_{buf}$  of 5 nm. Figure 7-23 gives design space for  $I_{ON} > 100 \mu\text{A}$  for three values of  $X_p$  and for an  $X_{buf}$  of 10 nm.  $I_{ON}$  is calculated based on a charge sheet approximation, the gate voltage and the capacitance from gate to the centroid of the hole distribution in the QW. As can be seen in the above two figures the curves are almost the same in both figures. This shows the weak effect  $X_{buf}$  has on hole density and hence  $I_{ON}$ .

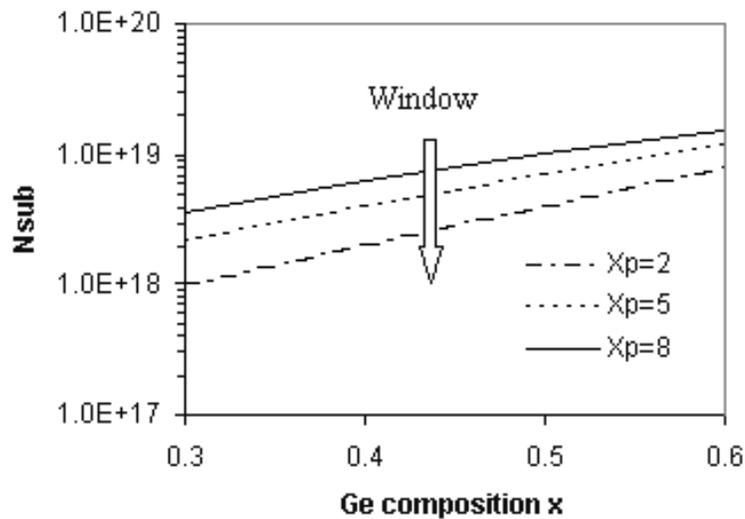


Figure 7-23. Design space for  $I_{ON} = 100 \mu\text{A}$  for the device 3,5,10, $X_p/x/5e18,5e18$  for p-layer thicknesses of 2, 5 and 8 nm.  $X_{buf} = 10$  nm. The area below the respective lines have  $I_{ON}$  more than  $100 \mu\text{A}$ .

Figure 7-22 shows the design space combination of  $x$  and  $N_{sub}$  for a device with a buffer layer thickness of 5 nm and p-layer thickness of 2, 5 and 8 nm. The device is 3,5,5, $X_p/x/5e18,N_{sub}$ . The area below the respective lines pointed by the arrows gives an  $I_{ON}$  above  $100 \mu\text{A}$ . The same design parameter combination of  $x$  and  $N_{sub}$  is given for a buffer layer thickness of 10 nm in Figure 7-23. In Figure 7-24 design space combination of  $X_p$  and  $N_{sub}$  is given for Ge fraction  $x$  of 0.2 and 0.5 with an  $X_{buf}$  of 5 nm. Combinations of  $X_p$  and  $N_{sub}$  below the respective curves have an  $I_{ON}$  value of  $100 \mu\text{A}$ . These are shown by window 1 and window 2 respectively. From this figure the selection range for  $N_{sub}$  is more for higher value of  $x$ . The same can be inferred from Figure 7-19 to Figure 7-21.

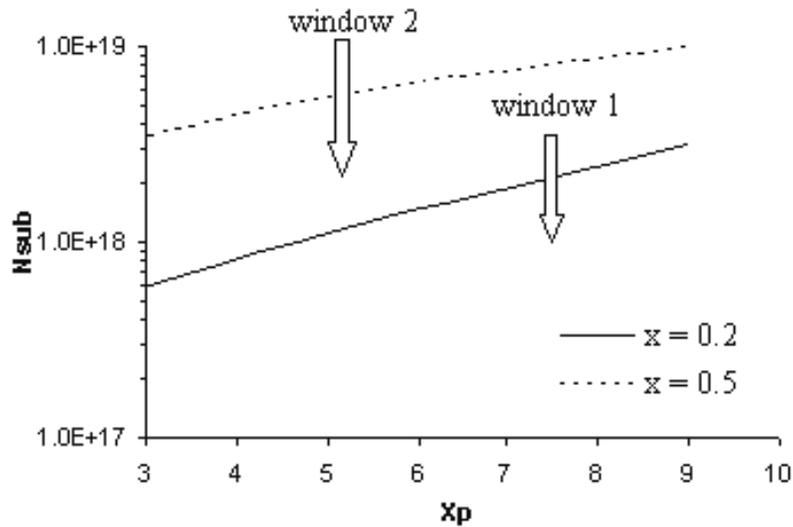


Figure 7-24. Design space for  $I_{ON} = 100 \mu\text{A}$  for the device  $5,5,X_p/x/5e18,N_{sub}$  for QW Ge fraction  $x = 0.2$  and  $0.5$ .  $X_{buf} = 5 \text{ nm}$ . The area below the respective lines shown by the arrows satisfies the requirement that  $I_{ON}$  is more than  $100 \mu\text{A}$

### 7.3 An Optimized Device

From Figure 7-7 it is found that the difference between  $V_T$  and  $V_{TS}$  increases for increasing Ge fraction. Since the device is to be operated at  $-1.5 \text{ V}$   $V_{TS}$  should be more negative than  $-1.5 \text{ V}$ . So a device was formulated from Figure 7-19 from near the center of the design space. The device is described as  $3,5,5,2,/0.5/,5e18,3e17$ . Figure 7-25 gives the terminal characteristic  $I_{DS}$  vs  $V_{ds}$  from analytical and numerical simulations. The match between analytical and numerical solutions are close with a maximum error less than 5%. The difference in  $I_{Dsat}$  is due to the omission of channel length modulation in the analytical program. The device has a  $V_T$  of  $-0.7$ , a  $V_{TS}$  of  $-1.7$ , an  $I_{ON}$  of  $228 \mu\text{A}$ ,  $I_{OFF}$  of  $7.6e-15 \text{ A}$  and an S parameter value of 82. These values are far better than those achievable in a MOSFET.

It can be seen in the above paragraph that the off current computed by the numerical simulation is very low for this device, in the sub fA range. These low currents can be explained by noting two facts. First the cap layer, SiGe QW and buffer layer are all intrinsic. Secondly the source-p-layer junction potential is high, compared to a delta-doped p-layer, due to the moderate doping of the p-layer. It is to be noted here that the thermionic currents have to be included since there is a band discontinuity at the source-QW junction. The thermionic currents are not included in above simulations. Thermionic currents are of consideration In  $I_{OFF}$  simulations. Since  $I_{OFF}$  is not calculated in the analytic program this is not taken into account. The error due to this omission in the numerical simulation is found to be less than 5% in  $I_{OFF}$  for Ge fractions above 0.3. Useful range of  $x$  values is found to be 0.3 to 0.6 as discussed in section 6.1.1.

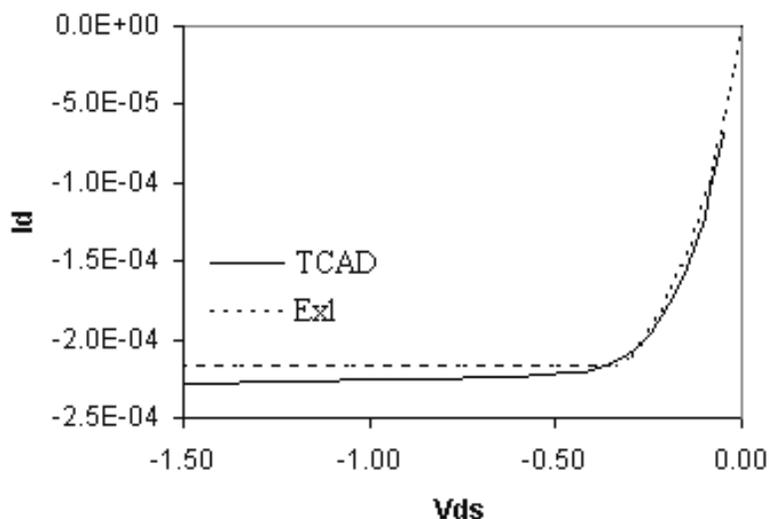


Figure 7-25. Analytical and numerical simulation comparison of drain characteristic for an optimized device. The device is described as 3,5,5,2/0.5/5e18,3e17. It has an  $I_{ON}$  of 228  $\mu\text{A}$ .

## 7.4 Summary of Long Channel Device Simulations

In this chapter a long channel device with  $L_g$  of 500 nm is simulated using one dimensional analytical program and results are compared with two dimensional numerical simulations using ISE-TCAD. The variable design parameters of importance are found to be the Ge fraction of the SiGe QW, the doping of the substrate and the thickness of the p-layer, in that order of importance. Buffer layer thickness is found to be not an influential parameter when the field in the buffer is not very high. It shields the QW from negative ions in the p-layer. In order to keep the evanescent wave functions of the QW coupling to the p-layer a minimum thickness of 5 nm is needed. Hence  $X_{buf}$  of 5 nm is used as a fixed parameter. The buffer layer however could be used for adjustment of sub threshold parameter and hence the  $I_{off}$ . But as seen from the device characteristics  $I_{OFF}$  for these devices are very low, around fA range, and hence  $S$  is not of importance. A minimum width of 5 nm is recommended for the QW to avoid quantization. A small thickness is desirable from critical thickness during growth. To hold the strain the thickness has to be below critical thickness. This critical thickness reduces very fast with Ge fraction. Hence the thickness of the QW  $X_{sig}$  is kept fixed at 5 nm. Low  $N_{sub}$  is found desirable to get  $V_T$  around  $-1$  V. But this may lead to subsurface punch through. Thickness of p-layer has the opposite effect of  $N_{sub}$  upon  $V_T$ ,  $V_{TS}$  and ratio of hole density. So this allows independent choice of  $N_{sub}$  and adjustment of threshold voltages using the thickness of the p-layer. However thicker p-layer over 9 nm will not give higher  $I_{ON}$ , for the structure used, as mentioned in section 6.1.3. In both cases of  $N_{sub}$  and  $X_p$  the

difference between  $V_T$  and  $V_{TS}$  remains the same as their values are varied. Variation of Ge content in the SiGe QW has good effect upon the hole density and a huge effect on the hole density ratio mentioned in subsection 6.1.1. Contrary to the case of  $N_{sub}$  and  $X_p$  the difference between  $V_T$  and  $V_{TS}$  increases very much with increasing Ge content. This gives a wider design space.

Based on the above observations appropriate design windows are constructed for  $V_T = -1$  and  $V_{TS} = -1.5$  for devices to operate at  $-1.5$  V. Further design windows are constructed for  $I_{ON}$  of  $100 \mu A$ . Samples from design windows are taken and checked with numerical simulations. The results from both analytical and numerical simulations compare well. To form an optimum device a set of parameters were chosen from the center of the design space and numerical simulations were carried out. These results are compared to the analytical simulations. It is found that the terminal characteristics from both simulations match well. The simulated device has an  $I_{ON}$  of  $228 \mu A$ ,  $I_{OFF}$  of  $7.6e-3$  pA,  $V_T$  of  $-0.7$  and  $V_{TS}$  of  $-1.7$ . This device is clearly superior to similar long channel p-HMOSFET devices found in publications cited in chapter 3. This device is also much better compared to the MOSFET. The  $I_{ON}$  is equal to or larger than the best published results of MOSFET and  $I_{OFF}$  is more than two orders of magnitude less than MOSFET. However the  $I_{OFF}$  behavior need to be watched as the channel length is shrunk to short channel range and also with the inclusion of the thermionic terms at the source-channel junction.

## **CHAPTER 8**

### **SHORT CHANNEL DEVICE SIMULATIONS**

In chapter 7 long channel HMOSFET is simulated using the analytical analysis and design program developed in chapter 6. Design windows which enables choice of structural parameters for selected performance values are constructed. The channel length was kept fixed at the long channel value of 500 nm. In order to increase operating speed of the device, to increase the output drive currents and to increase the packing density scaling down of the channel length is resorted to. The channel length cannot be arbitrarily reduced as discussed in chapter 2. The shortening of the channel length leads to short channel effects (SCE) such as drop in the threshold voltage and DIBL, to name two of the most important SCE.

In this chapter the short channel HMOSFET characteristics are compared to the short channel conventional MOSFET characteristics in order to evaluate the feasibility of scaling down the HMOSFET. The comparisons of short channel behavior of both MOSFET and HMOSFET are carried out using numerical simulator ISE-TCAD[109]. For a quick and simple analytical comparison the analytical analysis and design program developed in chapter 6 and a simple MOSFET design program developed in this chapter are used. The analytical analysis and design program is modified to include short channel HMOSFET. These two analytical results are compared with the numerical simulator results. Plots of both numerical and analytical results are given in section 8.3.

In order to accomplish the above short channel behavior comparison a MOSFET and a HMOSFET with the same  $N_{\text{sub}}$  and the same long channel (500 nm) saturation current,  $I_{\text{Dsat}}$ , are required to start with. For this purpose a simple MOSFET design program is developed using long channel MOSFET formulas. An appropriate  $N_{\text{sub}}$  and  $I_{\text{Dsat}}$  are first selected for the long channel MOSFET. Once the MOSFET is designed using the above program the modified analysis and design program of chapter 6 is used to design a long channel HMOSFET with the same  $N_{\text{sub}}$  and  $I_{\text{Dsat}}$  as the MOSFET. Now there is a long channel MOSFET and a long channel HMOSFET with the same  $N_{\text{sub}}$  and  $I_{\text{Dsat}}$ . Now the channel length is shortened in steps for both MOSFET and HMOSFET and comparisons of  $\text{Log}(\text{mag}(I_{\text{ds}}))$  vs  $V_{\text{gs}}$  are carried out for each channel length. Comparisons of  $V_{\text{T}}$  and drain characteristics are also made. These results are discussed in section 8.3.

In chapter 2 scaling down the channel length,  $L_{\text{g}}$ , is shown to reduce the source – channel barrier. This is true since the operating voltages may not get scaled down at the same rate as  $L_{\text{g}}$ . When this barrier reduces to the thermal voltage ( $kT/q$ ) carrier injection over this barrier becomes significant and the off current increases, leading to DIBL. To suppress this to some extent the substrate doping is increased thereby increasing the barrier at the source-channel junction. Hence to make a fair comparison of the SCE in the two devices, the Si MOSFET and the  $\text{Si}_{1-x}\text{Ge}_x$  HMOSFET, the substrate doping density  $N_{\text{sub}}$  has to be the same. This is because the drain depletion encroachment into the channel region and the associated DIBL depends on the value of  $N_{\text{sub}}$ . Apart from this one of the monitoring parameters has to be the same in either device to start with. Threshold voltage  $V_{\text{T}}$ , sub-threshold parameter  $S_{\text{t}}$ ,  $I_{\text{off}}$  or  $I_{\text{on}}$  can be used as a monitoring parameter.

$I_{off}$  or  $I_{on}$  is normally used for this purpose. Since the  $I_{off}$  in HMOSFET is found to be a few orders of magnitude lower than the  $I_{off}$  in conventional MOSFET  $I_{off}$  is not a good choice. So  $I_{on}$  is chosen as a monitoring parameter.  $I_{on}$ , as discussed earlier is chosen to be the same for both the long channel MOSFET and the long channel HMOSFET.  $V_T$  and  $S_t$  are also monitored to assess the short channel effects in both devices. In a well designed p-HMOSFET magnitude of  $V_T$  will be higher than the magnitude of  $V_T$  for the MOSFET, for the same  $I_{on}$ . To start with the comparison requires both long channel devices, MOSFET and HMOSFET, to have the same  $N_{sub}$  and the same  $I_{on}$  at 500 nm channel length. Since it is time consuming to achieve this through numerical simulations a program is written in Excel to design a MOSFET with  $N_{sub}$  as the design parameter. This is used to design the starting MOSFET with a channel length of 500 nm. In chapter 6 an elaborate analytical analysis and design program is developed in Excel for the design of HMOSFET. This is used for the design of the long channel (500 nm) HMOSFET. Comparison of results from the numerical simulation and from the above two analytical programs is carried out for both the long channel starting devices, the MOSFET and the HMOSFET. The MOSFET design program will be discussed first in the section below.

Having designed the 500 nm MOSFET and HMOSFET with the same  $N_{sub}$  and  $I_{on}$  their gate characteristics,  $\log(\text{amp}(I_{ds}))$  vs  $V_{gs}$ , is plotted using the numerical simulator. Their  $V_T$  and  $S_t$  are also calculated. Then the channel length is reduced from 500 nm to 100 nm in steps. At each step  $\log(\text{amp}(I_{ds}))$  vs  $V_{gs}$  is plotted as before.  $V_T$  and  $S_t$  are also noted. The gate length  $L_g$  at the onset of SCE is compared for both devices. The sub-

threshold slope of the above  $\log(\text{amp}(I_{ds}))$  vs  $V_{gs}$  plots in each device is compared at each step in  $L_g$ . The onset of SCE is indicated by one of the following: (1) a reduction in sub-threshold slope. (2) a reduction in magnitude of  $V_T$  (3) an increase in  $I_{off}$  even before a change in sub-threshold slope occurs. These results are given in sub-section 8.3.2.

In short the procedure adopted here is as follows,

1. Choose an appropriate value for  $N_{sub}$  such that the MOSFET  $L_g$  can be scaled down to about 100nm. This is done by using an empirical formula that uses (1) junction depth,  $x_j$  (2) gate oxide thickness,  $t_{OX}$  (3) substrate doping density,  $N_{sub}$  and (4) drain voltage. These are used to compute the possible minimum value of  $L_g$  before the onset of SCE [185]. This formula was developed for MOSFET. The HMOSFET structure is very different and this formula may not be applicable. So this formula is used only for the MOSFET to find an  $N_{sub}$  that allows scaling down to 100 nm.
2. Corresponding to the  $N_{sub}$  chosen a value of  $I_{on}$  is computed for the MOSFET using the numerical simulator ISE-TCAD.
3. Using the selected  $N_{sub}$  the QW channel Ge fraction,  $x$ , of the HMOSFET can be adjusted easily in the analytical analysis and design program for HMOSFET, given in chapter 6, to match the  $I_{on}$  to that of the MOSFET. This result,  $I_{on}$ , is also cross checked with the numerical simulator ISE-TCAD.
4. The  $\log(\text{mag}(I_{ds}))$  vs.  $V_{gs}$  is plotted using the numerical simulator ISE-TCAD.  $V_T$  and  $S_t$  are also measured using the numerical simulator and the analytical programs.

5. Channel length  $L_g$  of both the MOSFET and the HMOSFET are reduced in steps and the  $\log(\text{mag}(I_{ds}))$  vs.  $V_{gs}$  is plotted using the numerical simulator ISE-TCAD at each step. Also  $V_T$  and  $S_t$  are also noted at each step.
6. The beginning of DIBL causes a drop in  $V_T$ . This is compared in the case of either device. The beginning of SCE also causes an increase in  $I_{off}$ . This is also noted.
7. In the plots of the  $\log(\text{mag}(I_{ds}))$  vs.  $V_{gs}$  sub-threshold slope will remain the same until the onset of SCE. A reduction in slope is an indication of DIBL. At this point the  $I_{off}$  increases.  $I_{off}$  can also increase before a visible drop in sub-threshold slope occurs. So the value of  $L_g$  at this point, where  $I_{off}$  increases, can be taken as the limit of scaling with the chosen values of  $N_{sub}$  and  $I_{on}$ . Either device may reach this point at a different value of  $L_g$ . Further, how fast this degradation occurs while scaling down is an indication of SCE susceptibility of either device.

This procedure is adopted in short channel simulation comparisons of section 8.3.

### **8.1 Design Program for MOSFET with $N_{sub}$ as Parameter**

The aim of this section is to find out ranges of  $I_{Dsat}$  vs  $N_{sub}$  for a MOSFET and select a value of  $N_{sub}$  that gives a reasonable  $I_{Dsat}$  and minimum short channel gate length.

Figure 8-1 to Figure 8-4 are used for this purpose. So only MOSFET simulations are given in these figures. The short channel comparisons between MOSFET and HMOSFET are given in section 8.3. In order to have a design program in Excel for the MOSFET the charge sheet approximation model [3], [89] is convenient and easy. This is discussed in appendix A. The full expression for  $I_{ds}$  is given below. It is re-written by taking  $C_{OX}$  inside the brackets. This is done to make it clear that the term inside the bracket is the

difference between the total silicon charge and the depletion charge. This difference gives the inversion charge.

$$I_{ds} = \mu_{eff} \frac{W}{L} \left[ -C_{OX} \left( V_{gs} - V_{fb} - 2\psi_B - \frac{V_{ds}}{2} \right) V_{ds} + \frac{2\sqrt{2\epsilon_{si} q N_a}}{3} \left\{ (2\psi_B + V_{ds})^{3/2} - (2\psi_B)^{3/2} \right\} \right] \quad (8.1)$$

For high  $V_{ds}$ , beyond  $V_{Dsat}$ , the above expression can be simplified to find  $I_{Dsat}$  [4].

$$I_{Dsat} = \mu_{eff} C_{OX} \frac{W(V_{gs} - V_T)^2}{2mL} \quad (8.2)$$

Here the channel width  $W$  is  $1\mu\text{m}$ ,  $L$  is the channel length and  $C_{OX}$  is the gate capacitance.  $V_{gs} - V_T$  is the overdrive parameter.  $m$  is the body effect coefficient given by,

$$m = 1 + \frac{3t_{ox}}{W_{dmax}} \quad (8.3)$$

$t_{ox}$  is the gate oxide thickness and  $W_{dmax}$  is the substrate depletion width at  $V_{gs} = V_T$ . The effective mobility is found by a linear interpolation of published results [4], [106].

$$\mu_{eff} = \frac{166(0.1)^m}{\xi^m} = \frac{47}{\xi^m} \quad (8.4)$$

where  $\xi$  is the transverse electric field at the oxide-silicon interface. This is found from the depletion and inversion charge density.  $m$  in equation (8.4) is a linear slope parameter with a value of 0.539.

$$\xi = \frac{1}{\epsilon_{si}} \left( Q_D + \frac{Q_I}{3} \right) \quad (8.5)$$

$Q_D$  and  $Q_I$  are depletion and inversion layer charges. The MOSFET device has a p+ polysilicon gate. The thickness of the gate oxide is kept the same as in the HMOSFET case, which is 3 nm.

Using equations ( 8.1) to ( 8.5) and the Excel spreadsheet an analytical MOSFET design program is written to find  $I_{Dsat}$  and  $V_T$  for the MOSFET. For the short channel MOSFET the numerical simulator is used to generate the  $I_{Dsat}$  plots given in section 8.3.1. On top of these plots the results from the simple analytical MOSFET design program are also shown. Similar results for the HMOSFET, derived using the analysis and design program of chapter 6, are also given. As mentioned in the beginning of the chapter a long channel MOSFET and HMOSFET of the same  $N_{sub}$  and  $I_{Dsat}$  have to be designed first. The analysis and design program of chapter 6 is already tested with the numerical simulator ISE-TCAD [109]. To check the accuracy of the simple MOSFET design program developed in this chapter numerical simulation results are compared to the results of this simple analytical MOSFET design program. Since  $N_{sub}$  and  $I_{Dsat}$  relation is of interest these are compared in Figure 8-1. It was found that the simple design program and numerical simulator plots are parallel with a dc shift of 23.25  $\mu A$ . This may be due the inaccuracy in the effective mobility calculated using the equation ( 8.4) The calculated mobility may be lower than the value computed by the numerical simulator. Adding the above offset the analytical plot overlaps the numerical plot over a  $N_{sub}$  range of  $1 \times 10^{17}$  to  $1 \times 10^{18}/cm^3$ .

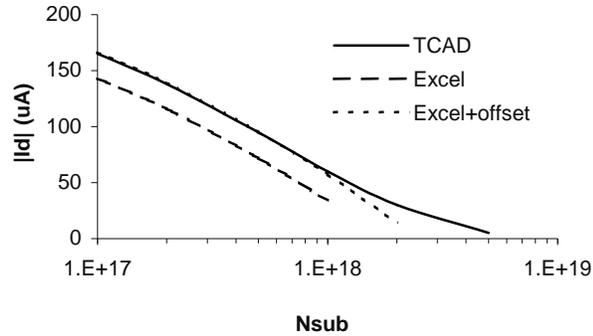


Figure 8-1.  $I_{Dsat}$  comparison of p-MOSFET using analytical MOSFET design program and ISE-TCAD numerical simulations.  $L_g = 500$  nm,  $t_{OX} = 3$  nm,  $x_J = 30$  nm, p+ polysilicon gate.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V. The Excel computed values are lower by a constant value

It is good to know an approximate value of gate length  $L_g$  at which the MOSFET begins to have short channel effects, specifically DIBL. Knowing this the numerical simulations can be extended to this value and beyond. The following empirical formula (Equation ( 8.6)) gives an estimate of the minimum channel length at which the MOSFET enters SCE [185]. This formula is for MOSFET structures. HMOSFET structure is very different and this formula may not apply. But as seen from the comparison of results (sub-section 8.3.1) it is found that both MOSFET and HMOSFET shows SCE at almost the same  $L_g$ . HMOSFET shows SCE at a little bit shorter  $L_g$  compared to MOSFET.

$$L_{\min} = A[x_j t_{OX} (W_S + W_D)^2]^{1/3} \quad (8.6)$$

$$W_S = \sqrt{\frac{2\epsilon_{si} V_{bi}}{qN_d}} \quad (8.7)$$

$$W_D = \sqrt{\frac{2\epsilon_{si}(V_{bi} + V_{ds})}{qN_d}} \quad (8.8)$$

$$V_{bi} = \frac{E_g}{2} + \frac{kT}{q} \ln\left(\frac{N_d}{n_i}\right) \quad (8.9)$$

The constant A is a proportionality factor and has a value of 0.41,  $x_j$  is the junction depth of source and drain regions,  $t_{OX}$  is the gate oxide thickness in angstroms,  $V_{bi}$  is the source junction built in potential,  $W_S$  is the source depletion width and  $W_D$  is the drain depletion width at full drain voltage. Using equation ( 8.6 ) the minimum value of  $L_g$  vs  $N_{sub}$  is plotted in Figure 8-2. This plot is to get an idea of an approximate value of minimum  $L_g$  for the range of  $N_{sub}$  values chosen in Figure 8-1. Figure 8-3 and Figure 8-4 give the threshold voltage and off currents respectively, obtained from numerical simulations.

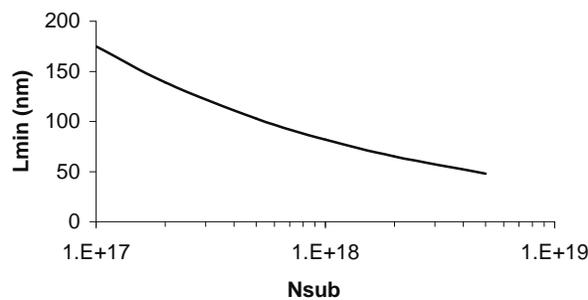


Figure 8-2. Estimated minimum value of  $L_g$ , at the onset of short channel effects, as a function of substrate doping density for the MOSFET. junction depth  $x_j = 30$  nm,  $t_{OX} = 3$  nm,  $V_{ds} = -1.5$  V This is using the equation ( 8.6).

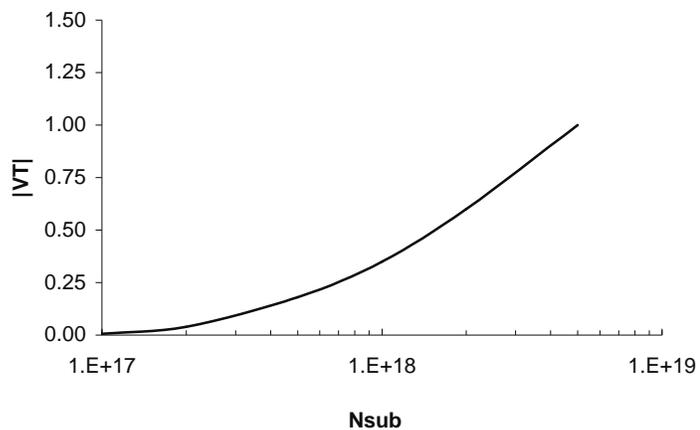
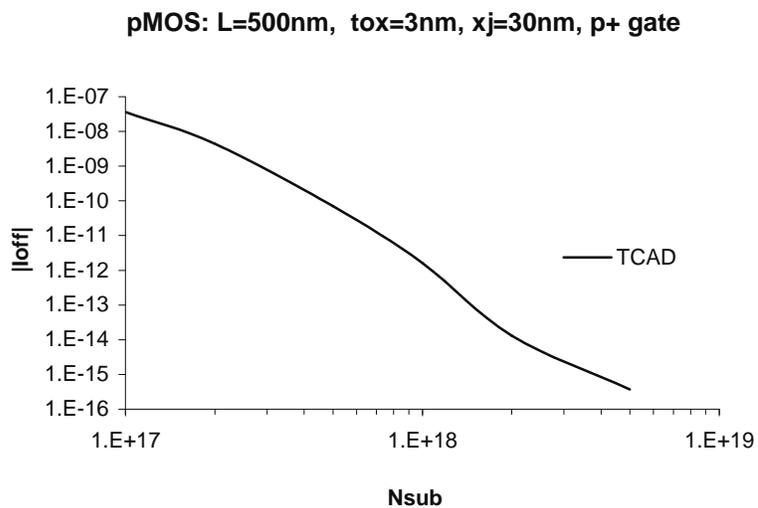


Figure 8-3. Threshold voltage  $V_T$  for a MOSFET as a function of substrate doping  $N_{sub}$  obtained using the numerical simulator ISE-TCAD.  $L_g = 500$  nm,  $t_{ox} = 3$  nm,  $x_j = 30$  nm, p+ gate.



For the  $N_{\text{sub}}$  of  $5 \times 10^{17}/\text{cm}^3$  the off current,  $I_{\text{off}}$ , is approximately  $1 \times 10^{-10}$  A for the MOSFET which is a reasonable value. For the same  $N_{\text{sub}}$  the minimum  $L_g$  is around 100 nm as seen in Figure 8-2. This gives an  $I_{\text{Dsat}}$  of approximately  $100 \mu\text{A}/\mu$ . So an  $N_{\text{sub}}$  of  $5 \times 10^{17}/\text{cm}^3$  is chosen for the MOSFET. A HMOSFET with the same values of  $N_{\text{sub}}$  and  $I_{\text{Dsat}}$  is designed as explained below. This is carried out using the analysis and design program of chapter 6 and is confirmed using the numerical simulator ISE-TCAD.

## 8.2 Modified Analytical Design Program for Short Channel HMOSFET

The analytical design and analysis program developed in chapter 6 using the one dimensional Poisson equation solution and Excel spread sheet is valid for long channel devices. For long channel devices potential contours are parallel to the current direction and are one dimensional as discussed in Appendix A. Due to the two dimensional nature of potential contours in short channel devices development of an analytical solution is more difficult. To make use of the analytical program of chapter 6 for short channel simulations, as the channel length is reduced from 500 nm to 100 nm, some form of short channel correction is needed. It has been proposed in literature that a voltage-doping transformation and correction of the substrate doping  $N_{\text{sub}}$  can achieve this very easily [184]. Since in short channel devices the source and drain depletion widths consist of a good percentage of the channel region the gate controls charges only in a reduced volume of the region under it. So the effective substrate doping can be written as a reduced value of the actual doping. This reduced  $N_{\text{sub}}$  is proportional to the drain voltages and the built in potential as given below [146],

$$N^* = N_d - \frac{2\epsilon_{si} V_{ds}^*}{qL_g^2} \quad (8.10)$$

$$V_{ds}^* = V_{ds} - 2(V_{bi} + \phi_{TH}) - 2\sqrt{(V_{bi} + \phi_{TH})(V_{bi} + \phi_{TH} + V_{ds})} \quad (8.11)$$

where  $N^*$  is the modified  $N_{sub}$  and  $V_{ds}^*$  is a modified drain voltage. It is to be noted that  $\phi_{TH}$ , the quantum well potential at the threshold voltage, is similar to the surface potential at inversion in MOSFET.  $V_{bi}$  is the built in potential at the drain-substrate junction. Using these formulas the long channel  $N_{sub}$  is modified. Now the same analytical HMOSFET analysis and design program is made use of for short channel HMOSFET. The modified analytical program can be used for both long and short channel simulations since the  $N_{sub}$  is automatically reduced to an effective  $N^*$  as the channel length reduces.

In the case of MOSFET  $I_{Dsat}$  and  $V_T$  for 500 nm down to 100 nm are computed using the analytical MOSFET design program discussed in section 8.1. It uses the full expression in equation (8.1). Now there are two analytical programs (1) MOSFET design program and (2) the HMOSFET design program modified with the inclusion of short channel HMOSFET devices. Simulations are also carried out using the numerical simulator ISE-TCAD. The  $I_{Dsat}$  and  $V_T$  results from the numerical simulator are compared to the above mentioned analytical programs. For plots of  $\log(\text{mag}(I_{ds}))$  vs.  $V_g$  only the numerical simulator is used.

### 8.3 Short Channel simulations

Following the procedure given in section 8.1 a MOSFET and a HMOSFET with the same substrate doping density,  $N_{sub}$ , and the same saturation current,  $I_{Dsat}$ , are designed. From section 8.1 it was found that a conventional MOSFET with a substrate doping of

$5 \times 10^{17}/\text{cm}^3$ , p+ gate and oxide thickness of 3 nm gives an  $I_{D\text{sat}}$  of approximately 100  $\mu\text{A}$ . To compare effects of scaling both MOSFET and HMOSFET should be designed to have the same  $I_{D\text{satt}}$  at 500 nm or long channel starting channel length. To make the comparison fair both the MOSFET and HMOSFET devices should also have the same  $N_{\text{sub}}$ , as mentioned at the beginning of the chapter. Using the HMOSFET analytical design program of chapter 6 a HMOSFET is designed with the same  $N_{\text{sub}}$  of  $5 \times 10^{17}/\text{cm}^3$ . This HMOSFET also should have 100 $\mu\text{A}$   $I_{D\text{sat}}$  at a  $L_g$  of 500 nm. The easiest way to get this is to adjust the Ge fraction  $x$  of the  $\text{Si}_{1-x}\text{Ge}_x$  QW channel. It was found from the analytical HMOSFET design program that a fraction  $x$  of 0.24 gives an  $I_{\text{on}}$  of 96 $\mu\text{A}$ . The HMOSFET structure is 3,5,5,2/0.24/5e18,5e17. Now there are two devices, a MOSFET and a HMOSFET, of equal  $I_{D\text{sat}}$  and same  $N_{\text{sub}}$  to compare.

Drain induced barrier lowering (DIBL) occurs with high drain voltage in the off state of the device ( $V_{\text{gs}} = 0 \text{ V}$ ). This is a SCE and is manifested by the encroachment of the drain-substrate junction depletion region into the source-substrate depletion region. This leads to a reduction of the potential barrier at the source-channel junction leading to injection of carriers from source over this reduced barrier into the channel. This causes higher leakage currents in the sub-threshold region. A parameter of interest in the sub-threshold region of a MOSFET device is the slope parameter  $S_t$  defined as [4],

$$S_t = \left[ \frac{d(\log(I_{ds}))}{dV_{gs}} \right]^{-1} = 1 + \frac{C_{dm}}{C_{ox}} = 2.3 \frac{kT}{q} \left( 1 + \frac{3t_{ox}}{W_D} \right) \quad (8.12)$$

Where  $C_{dm}$  is the depletion capacitance,  $C_{ox}$  is the gate oxide capacitance,  $k$  is the Boltzman constant,  $t_{ox}$  is the gate oxide thickness,  $q$  is the hole charge and  $W_D$  is the

depletion depth at the onset of inversion. As can be seen this is a constant for a particular MOSFET. This equation should be valid for HMOSFET also with a modification of the value of  $C_{dm}$  due to the undoped cap layer, undoped  $\text{Si}_{1-x}\text{Ge}_x$  layer, undoped buffer layer and the p-layer giving,

$$\frac{1}{C_{dm}} = \frac{X_{CAP}}{\epsilon_{Si}} + \frac{X_{SiGe}}{\epsilon_{SiGe}} + \frac{X_{BUF}}{\epsilon_{Si}} + \frac{W_d N_d - X_p N_p}{\epsilon_{Si} (N_D + N_P)} \quad (8.13)$$

Here it is assumed that the p-layer is fully depleted.  $W_d$  is the depletion width in the n-substrate.  $N_d$  is same as  $N_{sub}$ .  $X_p$  and  $N_p$  are thickness and doping density of the p-layer.

### 8.3.1 Short channel comparison

Figure 8-5 has numerically simulated plots of  $I_{Dsat}$  for MOSFET and HMOSFET as the channel length is reduced from 500 nm to 100 nm, keeping other parameters the same as at 500 nm. The open and closed circles indicate the output values for the MOSFET and the HMOSFET simulated using the analytical programs of section 8.1 and the analysis and design program of chapter 8.2 respectively. Up to about 200 nm in case of HMOSFET and 300 nm in case of MOSFET the increase in  $I_{Dsat}$  is nearly linear. Below these values of  $L_g$ ,  $I_{Dsat}$  rises faster indicating short channel effects due to the threshold voltage drop. This indicates that the MOSFET suffers from SCE slightly earlier in scaling than the HMOSFET.

Figure 8-6 has numerically simulated threshold voltages  $V_T$  for both MOSFET and HMOSFET for reducing values of  $L_g$  from 500 nm to 100 nm. Again other parameters were kept fixed at their 500 nm values. The dark circles indicate analytical analysis and

design program output values for HMOSFET. In the case of MOSFET the analytical values, as given in Figure 8-3, are not comparable to numerical results and hence not plotted. The reason may be the difference in the definition of  $V_T$ . In the MOSFET design program of section 8.1  $V_T$  is defined as the  $V_{gs}$  at which the surface hole density becomes equal to the substrate electron density. But in the numerical simulation  $V_T$  is defined as the value of  $V_{gs}$  corresponding to an  $I_{ds}$  of  $0.1 \mu A/\mu$ . From the numerical simulation results  $V_T$  begins to drop at about 250 nm for MOSFET and 200 nm for HMOSFET respectively. So these values can be taken as the beginning of SCE. In the case of HMOSFET the analytical values are about 0.1 V lower and drops off slowly compared to numerical results.  $V_T$  for HMOSFET in the case of analytical program is defined as the gate voltage at which the QW potential change by twice the doping parameter (Equation (4.5)) whereas  $V_T$  from numerical simulations are calculated for an  $I_{ds}$  of  $0.1 \mu A/\mu$ . This could be the reason for this difference.

Figure 8-7 plots numerical results for  $I_{ds}$  vs.  $V_{ds}$  for MOSFET and HMOSFET for channel lengths of 500 nm and 150 nm. The other parameters are again kept fixed at their 500 nm values. For long channel devices (500 nm)  $I_{Dsat}$  for both MOSFET and HMOSFET are the same. But for the short channel device at 150 nm HMOSFET current reaches a constant plateau whereas MOSFET current keeps rising. This shows that MOSFET has higher short channel susceptibility compared to HMOSFET.

It is also noticed that HMOSFET has higher  $I_{ds}$  for low  $V_{ds}$  values for both long and short channel cases and reaches constant values at a lower  $V_{ds}$  compared to MOSFET. This can be attributed to the higher hole mobility in the undoped strained  $Si_{1-x}Ge_x$

channel layer compared to the mobility in the doped Si channel in the case of MOSFET. This can bring the drain end to saturation for lower  $V_{ds}$  in case of HMOSFET. This can cause pinch off at a lower drain field. Pinch off and saturation are described in section 2.2.2. Pinch off can cause  $I_{ds}$  to saturate but the fields in the channel region can however remain low. So the channel region need not be in velocity saturation. This can be confirmed only through numerical simulation. Figure 8-25 shows a plot of velocity in the channel vs distance from the source. The velocity is less than  $3 \times 10^6$  cm/s whereas  $v_{sat}$  for holes is  $8.37 \times 10^6$  cm/s. This shows that the HMOSFET channel is not velocity saturated. Hence the high low field mobility is applicable. This gives an advantage over MOSFET whose low field mobility is inferior to that of HMOSFET.

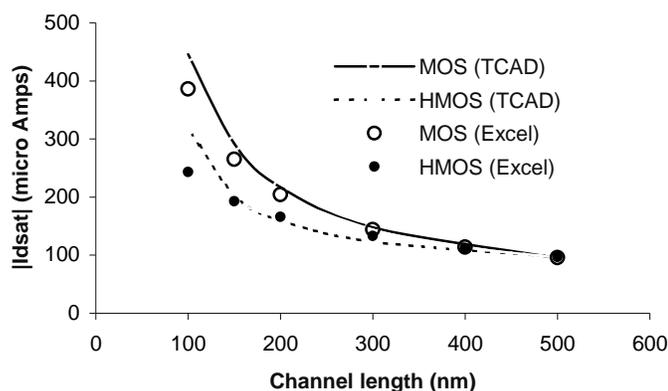


Figure 8-5. Saturation values  $I_{Dsat}$  for 500, 400, 300, 200, 150 and 100 nm channel lengths. Up to 200 nm  $I_{Dsat}$  for HMOSFET shows almost linear variation but at 150 nm  $I_{Dsat}$  is much higher indicative of short channel effects. In the case of MOSFET the linear variation is only upto about 300 nm. MOSFET has  $t_{ox} = 3$  nm,  $N_{sub} = 5e17/cm^3$ ,  $x_J = 30$  nm, p+ gate. HMOSFET is 3,5,5,2/0.24/5e18,5e17 with  $t_{ox} = 3$  nm,  $x_J = 30$  nm, p+gate.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V

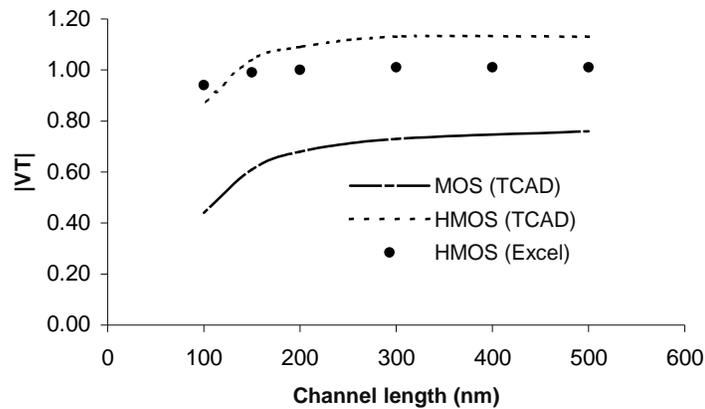


Figure 8-6. Threshold voltage  $V_T$  vs channel length. Lines are from numeric simulator ISE-TCAD and the dark circles are from the analytical analysis and design program of Section 8.2. MOSFET has  $t_{ox} = 3$  nm,  $x_j = 30$  nm,  $N_{sub} = 5e17/cm^3$ , p+ gate.. HMOSFET is 3,5,5,2/0.24/5e18,5e17 with  $t_{ox} = 3$  nm,  $x_j = 30$  nm, p+gate. Around 200 nm the threshold voltage begins to drop for both devices.

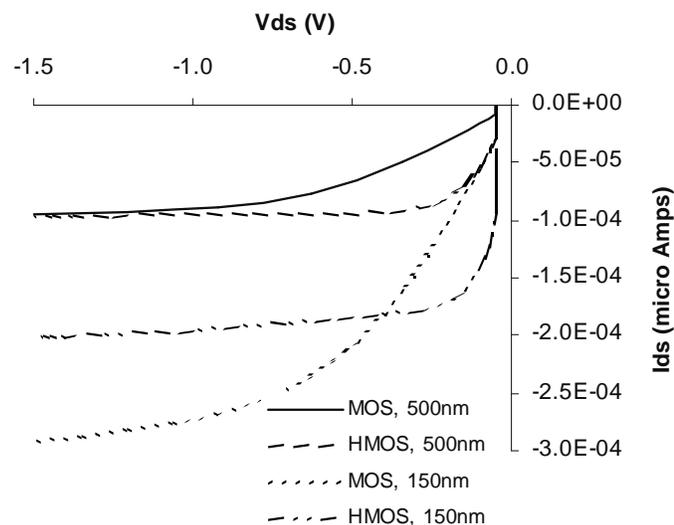


Figure 8-7.  $I_{ds}$  for MOSFET and HMOSFET for 500 nm and 150 nm channel lengths. For 150 nm the MOSFET do not show any saturation indicating SCE. MOSFET has  $t_{ox} = 3$  nm,  $x_j = 30$  nm,  $N_{sub} = 5e17/cm^3$ . HMOSFET is 3,5,5,2/0.24/5e18,5e17 with  $t_{ox} = 3$  nm,  $x_j = 30$  nm, p+gate.  $V_{gs} = -1.5$  V. These plots are from the ISE-TCAD.

In Figure 8-7 the long channel  $I_{ds}$  for both devices saturate to the same value of 100  $\mu$ A.

But for short channel device at 150 nm HMOSFET current saturates whereas MOSFET

current does not show saturation. Instead it steadily increases, indicating once again, that MOSFET has higher short channel susceptibility compared to HMOSFET.

It is also noticed that HMOSFET has higher  $I_{ds}$  for low  $V_{ds}$  values for both long and short channel cases. This can be attributed to the higher hole mobility in undoped strained  $Si_{1-x}Ge_x$  channel layer compared to the mobility in doped Si channel in the case of MOSFET. HMOSFET enters saturation at around 0.4 V whereas MOSFET enters saturation at around 0.9 V. The lateral field corresponding to 0.4 V in HMOSFET is only 8000 V/cm much lower than the critical field for velocity saturation. In the case of MOSFET this field is  $1.8 \times 10^4$  V/cm, again much below the critical field. All these arguments lead to the fact that neither MOSFET nor HMOSFET are in velocity saturation and the low field mobility applies.

### 8.3.2 Sub-threshold Comparison

Numerically simulated values of  $I_{ds}$  vs. channel length  $L_g$ , for the MOSFET is plotted in Figure 8-8 for 500, 300, 200, 150 and 100 nm. Since  $S_t$  is the inverse slope of these curves in the sub-threshold region the sub-threshold region should have the same slope as given by equation ( 8.12 ). A reduced sub-threshold slope will move the off current to higher values (intercept with the vertical axis). So does a translation of threshold voltage  $V_T$  to lower values ( $V_T$  roll off). Both of the above indicate SCE. Figure 8-8 does not indicate a reduced slope down to a  $L_g$  of 100 nm but at 100 nm the above figure shows a translation or reduction of  $V_T$  to lower values. Down to a  $L_g$  of 150 nm there is very little change in slope or reduction in  $V_T$ . Similar numerically simulated plots for the HMOSFET are given in Figure 8-9. This set of plots also does not suffer from DIBL

down to 150 nm. This indicates that both MOSFET and HMOSFET have similar limitations on  $L_g$  as far as short channel effects are concerned or both devices scale down similarly. These plots are generated by using the numerical simulator ISE-TCAD.

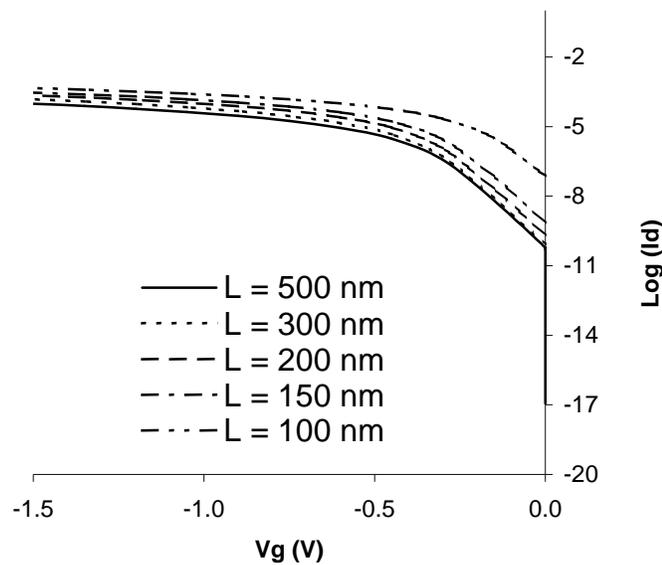


Figure 8-8.  $\text{Log}(I_{ds})$  vs  $V_{gs}$ . MOSFET has  $t_{ox} = 3$  nm,  $x_j = 30$  nm,  $N_{sub} = 5e17/\text{cm}^3$ , p+ gate. Channel lengths  $L_g$  of 500, 400, 300, 200, 150 and 100 nm. At 100 nm  $I_{off}$  increased by more than two orders of magnitude indicating SCE.  $V_{ds} = -1.5$  V. These plots are from the numerical simulator ISE-TCAD.

Figure 8-10 plots  $\log(\text{mag}(I_{ds}))$  vs  $V_{gs}$  for the MOSFET and HMOSFET simulated using the numerical simulator. The gate length is 500 nm. Notice that both devices have the same  $I_{on}$  whereas the  $I_{off}$  values are very different. HMOSFET has an  $I_{off}$  of almost seven orders of magnitude lower than that of MOSFET. This is a very important observation. In chapter 4 it was mentioned that the thermal conductivity of  $\text{Si}_{1-x}\text{Ge}_x$  is nearly fifteen times lower than that of Si. With such lower  $I_{off}$  the HMOSFET generates

far less standby power dissipation compared to Si MOSFET. With scalability almost the same, as discussed before, this indicates that HMOSFET can have a much denser packing density on the dice since the thermal designs can be easier due to reduced off state currents..

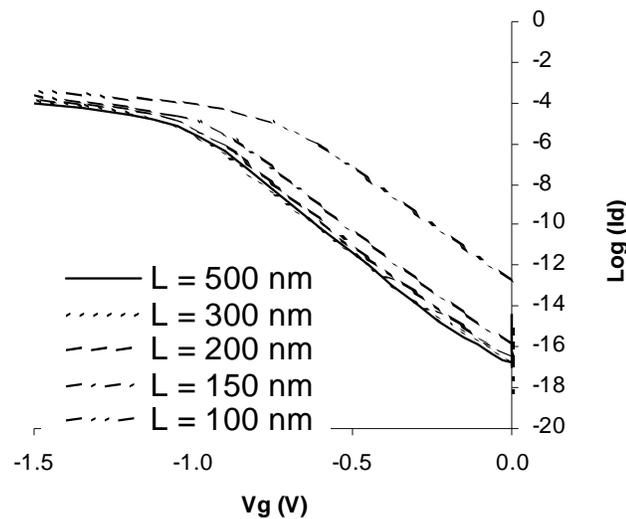


Figure 8-9.  $\text{Log}(\text{mag}(I_{ds}))$  vs  $V_{gs}$  for a HMOSFET 3,5,5,2/0.24/5e18,5e17 with  $t_{OX} = 3$  nm,  $x_J = 30$  nm, p+ gate.  $L_g$  of 500, 400, 300, 200, 150 and 100 nm. At 100 nm  $I_{off}$  is much higher indicating the onset of SCE.  $V_{ds} = -1.5$  V. These plots are from the numerical simulator ISE-TCAD.

From the above results it can be argued that the low off current in HMOSFET is due to the difference in  $V_T$ . From the numerically simulated Figure 8-10 it can be observed that the MOSFET has a  $V_T$  around  $-0.4$  V and the HMOSFET has a  $V_T$  around  $-1.0$  V. In the numerically simulated Figure 8-10 the sub-threshold slope is nearly the same for both devices. So if the threshold voltages are made the same both curves will overlap in the sub-threshold region or the  $I_{off}$  value for HMOSFET and MOSFET will become the same. But it has to be understood that the ON current  $I_{on}$  of HMOSFET will increase very

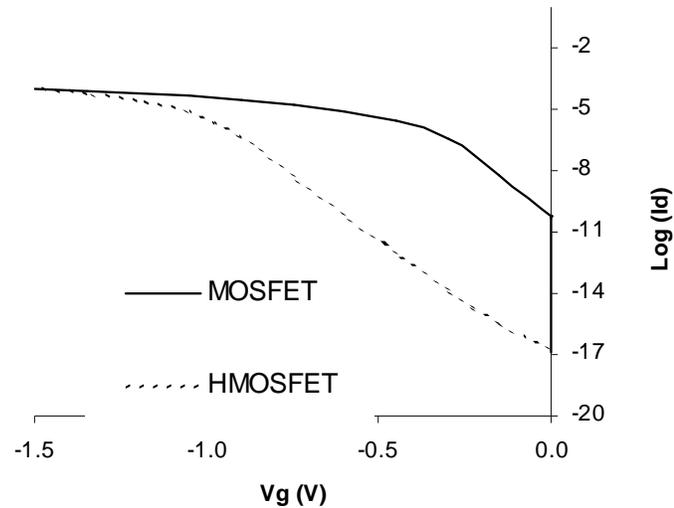


Figure 8-10. Comparison of  $\text{Log}(\text{mag}(I_d))$  vs  $V_{gs}$  for MOSFET and HMOSFET. Both devices have the same  $L_g = 500$  nm,  $N_{\text{sub}} = 5 \times 10^{17}/\text{cm}^3$ ,  $t_{\text{OX}} = 3$  nm,  $x_j = 30$  nm, p+ gate. Both have the same  $I_{\text{Dsat}}$  at  $V_{gs} = -1.5$  V and  $V_{ds} = -1.5$  V. HMOSFET is  $3,5,5,2/0.24/5e18,5e17$ . Notice that  $I_{\text{off}}$  in the case of HMOSFET is almost seven orders of magnitude less compared to that of MOSFET. These plots are from ISE-TCAD.

much with such low  $V_T$ , in comparison to the MOSFET [ 177 ]. This is sketched in Figure 8-11. Since sub-threshold slopes are nearly the same for both MOSFET and HMOSFET the curves overlap in the sub-threshold region as shown in Figure 8-11. But in the ON region the current in the HMOSFET is much higher. This implies a steeper slope for HMOSFET in the ON region. This is possible only if the channel for HMOSFET is not velocity saturated so that the high low field mobility of the  $\text{Si}_{(1-x)}\text{Ge}_x$  QW channel can provide a steeper slope.

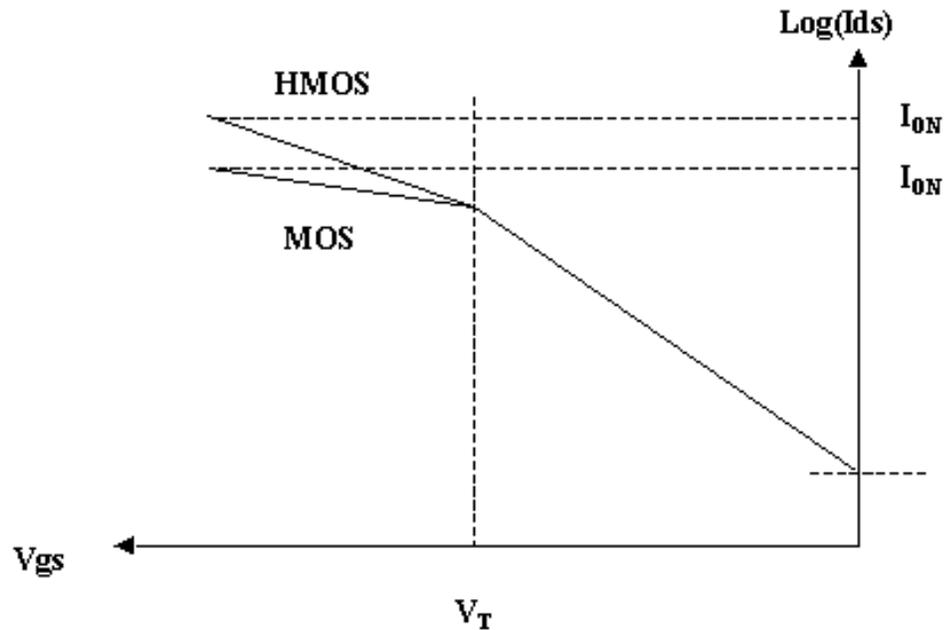


Figure 8-11.  $\text{Log}(I_{ds})$  vs  $V_{gs}$  sketch when both MOSFET and HMOSFET have the same  $V_T$ .  $I_{on}$  is higher for HMOSFET at  $V_{gs} = -1.5$  V and  $V_{ds} = -1.5$  V. This is an approximate sketch only to convey the idea.

When both devices are set to have the same  $I_{on}$ , as sketched in Figure 8-12, the HMOSFET has a very low  $I_{off}$ . The HMOSFET can have a higher  $V_T$  and still give the same ON current as in the MOSFET due to the high mobility. Again it has to be assumed that the channel is not velocity saturated and the high low field mobility of the  $\text{Si}_{(1-x)}\text{Ge}_x$  QW channel can provide a steeper slope in the ON region for the HMOSFET.

Due to the high  $V_T$  HMOSFET now has a very low OFF current. It has to be noted here that Figure 8-11 and Figure 8-12 are mere sketches to convey the comparison with the same  $V_T$  or the same  $I_{on}$  respectively. It should be noted here that the HMOSFET was found to be not in velocity saturation but in pinch off state in sub-section 8.3.2.1 and in Figure 8-7.

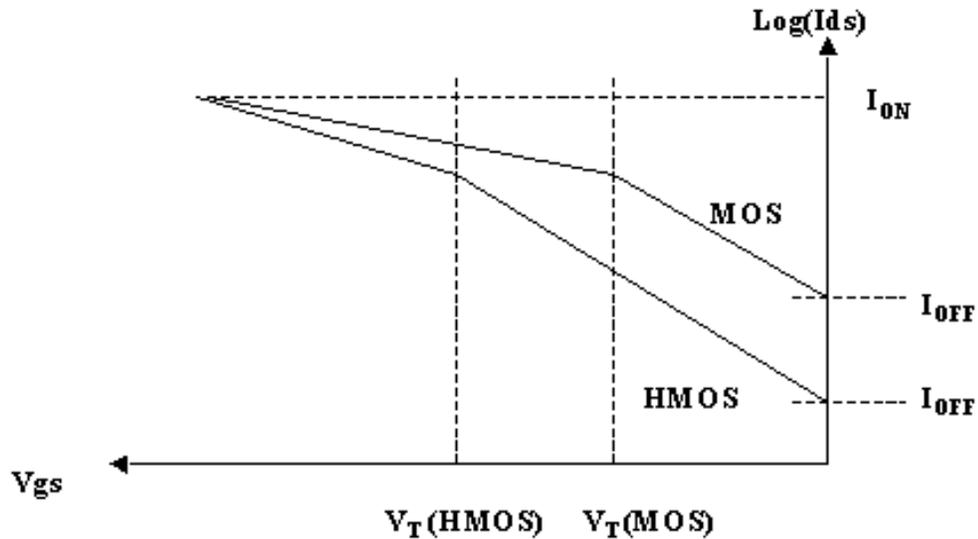


Figure 8-12.  $\text{Log}(I_{ds})$  vs  $V_{gs}$  sketch when the HMOSFET has a higher  $V_T$ .  $V_T$  for either devices is chosen such that  $I_{on}$  is the same for both MOSFET and HMOSFET at  $V_{gs} = -1.5$  V.  $V_{ds} = -1.5$  V. This is a mere sketch to convey the idea. Numerically simulated Figure 8-10 gives the correct values.

### 8.3.2.1 ON Current Calculation

(1) For vsat case: (HMOSFET)

$$v_{sat} = 8.37 \times 10^6 \text{ cm/s}$$

$$\text{Gate length } L_g = 500 \text{ nm}$$

$$V_g = -1.5 \text{ V}$$

$$V_T = -1.0 \text{ V (approximate)}$$

$$\text{Overdrive} = 1.5 - 1.0 = 0.5 \text{ V}$$

$$\text{Total charge } Q = (V_g - V_T)C_{ox}$$

$$\text{Transit time } t = L_g/v_{sat}$$

$$I = \frac{Q}{t}$$

$$C_{ox} = \frac{\epsilon_{ox} L_G * 1\mu}{t_{ox}} = \frac{8.85 \times 10^{-14} \times 3.9 \times 100 \times 10^{-7}}{3 \times 10^{-7}} = 11.5 \times 10^{-16} F \quad (8.14)$$

$$I_{DS} = \frac{0.5 \times 11.5 \times 10^{-16} \times 8.4 \times 10^6}{500 \times 10^{-7}} = 96 \mu A \quad (8.15)$$

The designed value of  $I_{ds}$  for 500 nm channel length is 100  $\mu A$ . So the above results are reasonable. The above results suggest that the HMOSFET channel is velocity saturated. A note of caution here. Even though the assumption of velocity saturation seem to give correct result the reason does not seem right. In sub-section 8.3.2.2 numerical simulations show that channel saturation is absent.

For the same overdrive MOSFET also should give the same current since  $v_{sat}$  is the same for both Si and  $Si_{(1-x)}Ge_x$ . The designed ON current in both devices is 100  $\mu A$ .

(2) For non-saturated case (HMOSFET)

$$\mu_{SiGe} = 470 + (1900 - 470)x = 470 + (1900 - 470)0.24 = 813 \text{ cm}^2/V \text{ sec} \quad (3)$$

Assuming uniform field in the channel field is given by,

$$\xi = \frac{1.5}{500 \times 10^{-7}} = 3 \times 10^4 \text{ V/cm} \quad (4)$$

$$\text{Velocity} = 813 \times 3 \times 10^4 = 2.4 \times 10^7 \text{ cm/s} > v_{sat}$$

This is much above the saturation velocity which cannot happen That means the field in the channel is not as high as estimated, which could be accounted for by a non-uniform potential drop along the channel. For example the potential drop and field near the drain are much greater. The current can be limited due to pinch off mechanism explained more

elaborately in sub-section 2.2.2. So the HMOSFET may not be in velocity saturation.

This aspect is further evidenced under sub-section 8.3.2.2.

When pinch off occurs the channel operates between zero at the source and  $V_{Dsat}$  at the other end closer to the drain (pinch off point). Rest of  $V_{ds}$  drops in the pinch off region near the drain.  $V_{Dsat}$  is much lower than  $V_{ds}$  and normally a fraction of a volt. In Figure 8-7 the HMOSFET saturates around 0.3 V. From these discussions it can be inferred that the HMOSFET channel potential is near zero at source end and approximately 0.3 V near the drain end. So the field in the channel region is going to be very low. Hence HMOSFET cannot be in velocity saturation.

So greater part of the channel is in the low field regime. From Figure 8-12 this implies lower carrier density in the high mobility HMOSFET to give the same  $I_{on}$  as in the low mobility MOSFET. The carrier density is proportional to the overdrive. From Figure 8-10  $V_T$  for HMOSFET and MOSFET are approximately  $-1.0$  V and  $-0.4$  V respectively. Hole mobilities for  $Si_{(1-x)}Ge_x$  ( $x = 0.25$ ) and Si (doped to  $5 \times 10^{17}$ ) are 813 and 200  $cm^2/Vs$  respectively. Since the calculated field need to be reduced to a third the potential drop for most of the channel length can be assumed to be a third of  $V_{ds}$  or  $-0.5$  V. This will give a field of,

$$\xi = \frac{0.5}{500 \times 10^{-7}} = 10^4 \text{ V/cm} \quad (8.16)$$

This value of the field gives an  $I_{Dsat}$  of approximately 100  $\mu A$ . The numerically simulated Figure 8-10 also gives an  $I_{Dsat}$  of around 100  $\mu A$ .

(3) For vsat case: (MOSFET)

$$v_{sat} = 8.37 \times 10^6 \text{ cm/s}$$

$$\text{Gate length } L_g = 500 \text{ nm}$$

$$V_g = -1.5 \text{ V}$$

$$V_T = -0.4 \text{ V (approximate)}$$

$$\text{Overdrive} = 1.5 - 0.4 = 1.1 \text{ V}$$

Since the overdrive is more than double in this case the total charge also will be more than double compared to HMOSFET in case (1). In order to maintain the current at  $100 \mu\text{A}$  the velocity has to be less than half of  $v_{sat}$ . There must be some other mechanism limiting the field in the channel or reducing the potential drop in the channel. One such mechanism is pinch off as discussed in the previous calculation..An elaborate discussion of pinch off is given sub-section 2.2.2. Now the lack of charges near the drain end increases the resistivity and a voltage equal to approximately  $V_{ds} - V_{Dsat}$  drops in the pinch off region near the drain. This limits the potential available for the channel region thus reducing the field in the channel. So in spite of the fact that the charge density away from the drain is more than double (due to higher over drive) the current will not be double. As noted above the velocity has to be less than half of  $v_{sat}$ . So the MOSFET cannot be in saturation.

(4) For non-saturated case (MOSFET)

$$\mu_{Si} = 200 \text{ cm}^2/\text{V sec (doping at } 5 \times 10^{17}/\text{cm}^3)$$

$$\xi = \frac{1.5}{500 \times 10^{-7}} = 3 \times 10^4$$

$$\text{Velocity} = 200 \times 3 \times 10^4 = 6 \times 10^6 \text{ cm/sec} = 0.71 \times v_{\text{sat}}$$

Since the overdrive and total charge are double compared to HMOSFET the velocity should be less than half of  $v_{\text{sat}}$  (comparing with case 1) to give 100  $\mu\text{A}$ . So the chances are most of the channel is not velocity saturated. Most of the potential drops near the drain junction where the fields are many times the rest of the channel. To make the velocity less than half of  $v_{\text{sat}}$  the channel can have a potential of  $(0.5/0.71)V_{\text{ds}} = 0.7V_{\text{ds}}$  only.

The potential plot from source to drain along the channel for MOSFET is given in Figure 8-13. The MOSFET has  $t_{\text{OX}} = 3 \text{ nm}$ ,  $x_{\text{J}} = 30 \text{ nm}$ ,  $N_{\text{sub}} = 5 \times 10^{17}/\text{cm}^3$  and p+gate. The structure of the MOSFET in Figure 8-13 to Figure 8-18 are the same. The source-channel barrier is almost missing in the 100 nm channel length case. For the 500 nm case there is significant curvature indicating the MOSFET may be in saturation mode.

The potential along the QW channel for the HMOSFET is given in Figure 8-14. An enlarged view near the source end is given in Figure 8-15. The potential is nearly linear up to the vicinity of the drain. Note also that the potential drop in most of the channel is only about 0.2 V. So most of the channel is in the low field regime. This will explain Figure 8-10. Since most of the channel is in low field regime the ON region of the HMOSFET can be steeper due to the high low field mobility. For the short channel HMOSFET, in Figure 8-12, the source-channel barrier is almost the same as the long

channel case indicating that the SCE has not yet developed. Lateral electric field for the MOSFET is given in Figure 8-16 and for the HMOSFET is given in Figure 8-17.

Figure 8-18 plots the electric field for 500 nm MOSFET and HMOSFET together.

Compared to the MOSFET the electric field is nearly six times in HMOSFET near the drain end as seen in Figure 8-18. For the HMOSFET the fields in most parts of the channel are very low as indicated by the small drop (approximately 0.2 V) in Figure 8-14 for the 500 nm HMOSFET. So the rest of the potential (1.5 – 0.2) V has to drop in a small region near the drain end. This leads to very high fields near the drain end of the HMOSFET. The two dimensional potential contours for 100 nm channel MOSFET and HMOSFET are given in Figure 8-19 and Figure 8-20 respectively. Both devices seem to have nearly the same two-dimensional spread in the contours. Only below the channel region there is an inconspicuous difference.

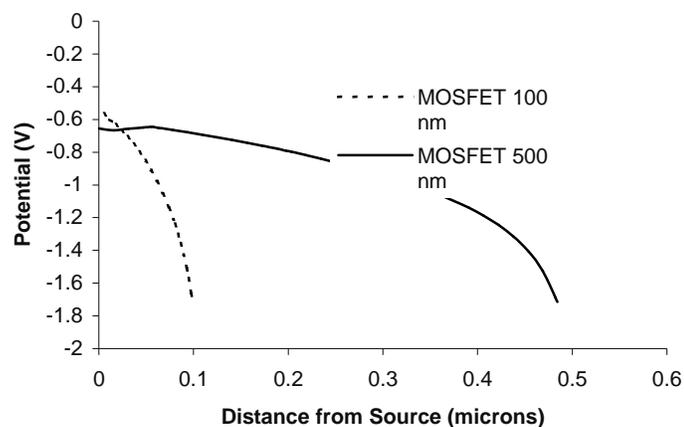


Figure 8-13. Potential plots from source to drain for 100 and 500 nm MOSFET.  $t_{OX} = 3$  nm,  $x_J = 30$  nm,  $N_{sub} = 5 \times 10^{17}/cm^3$ , p+ gate.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V. This is from TCAD.

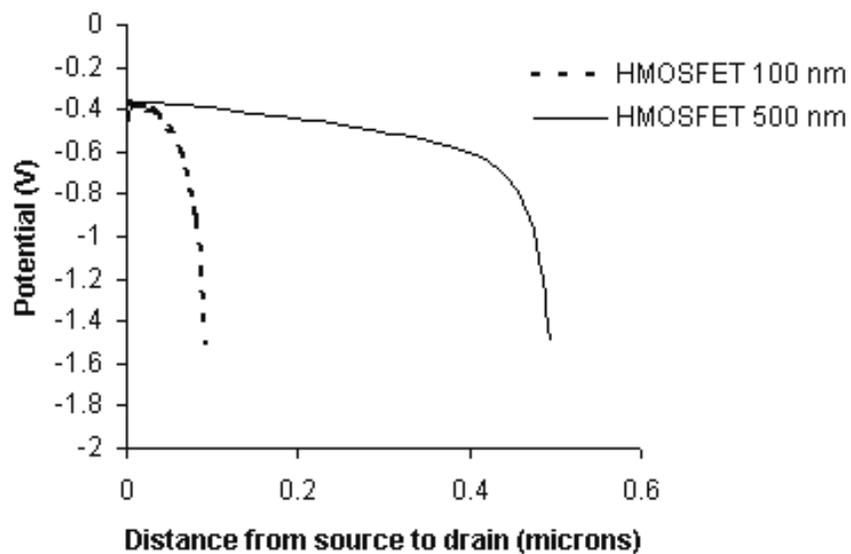


Figure 8-14. Potential along the channel for 100 nm and 500 nm HMOSFET.

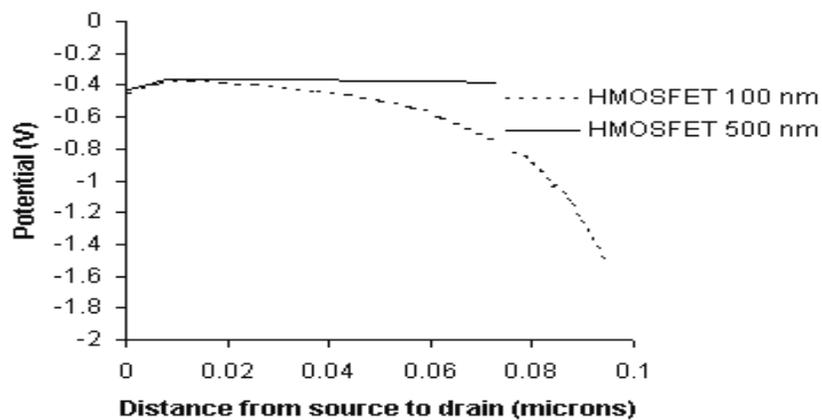


Figure 8-15. An enlarged view of Figure 8-14..  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V.

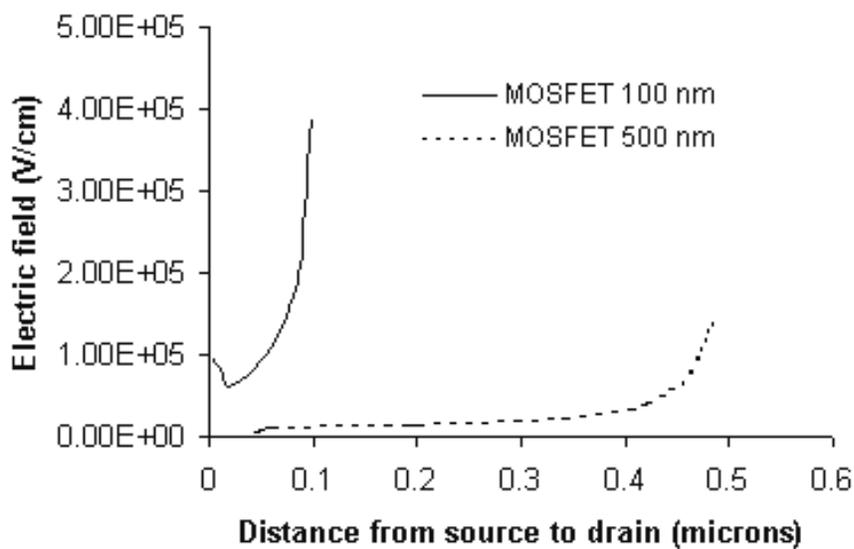


Figure 8-16. Lateral electric field along the channel for long and short channel MOSFET.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V.

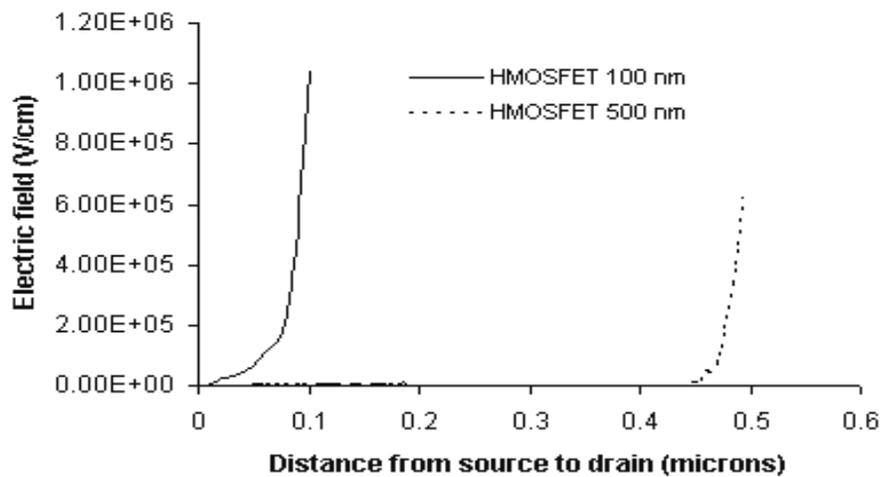


Figure 8-17. Lateral electric field along the QW channel for long and short channel HMOSFET.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V.

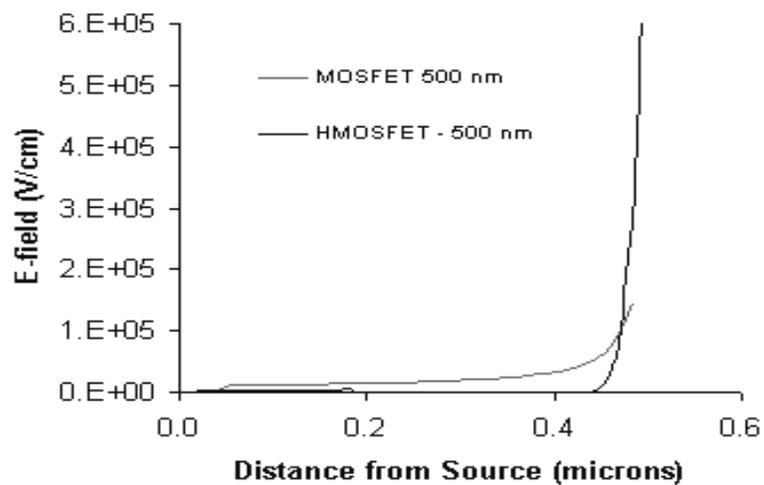


Figure 8-18. Lateral electric field along the channel for MOSFET and HMOSFET.  $L_g = 500$  nm. HMOSFET is 3,5,5,2/0.24/5e18,5e17.  $x_j = 30$  nm,  $t_{ox} = 3$  nm, p+ gate. MOSFET has  $t_{ox} = 3$  nm,  $x_j = 30$  nm,  $N_{sub} = 5 \times 10^{17}/\text{cm}^3$ , p+ gate.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V.

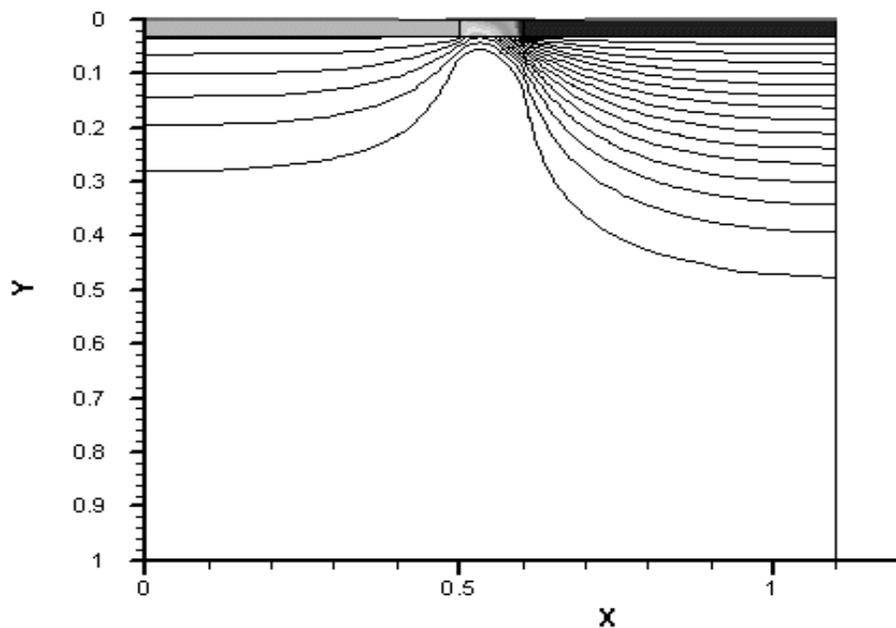


Figure 8-19. Potential contours for the 100 nm MOSFET.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V.

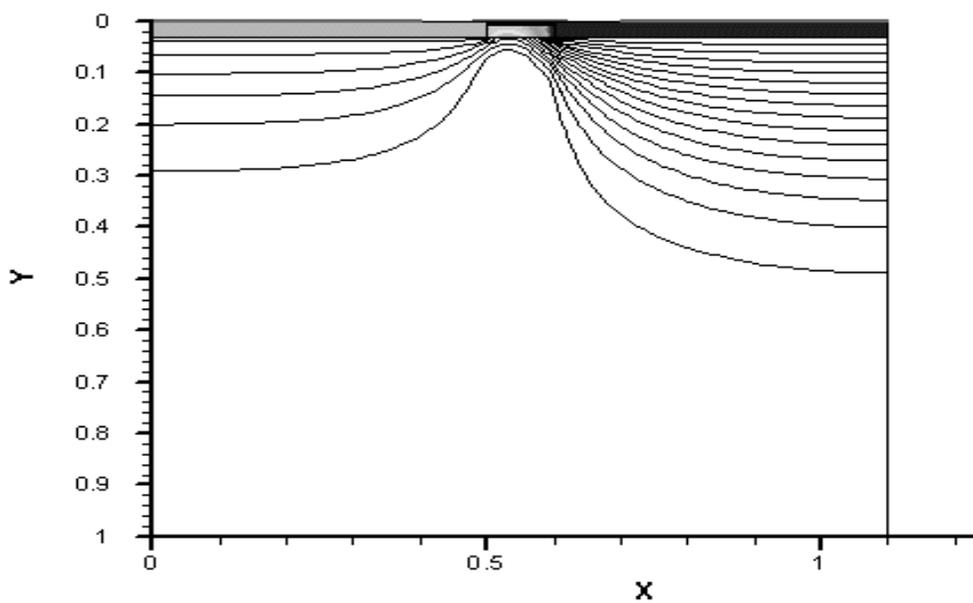


Figure 8-20. Potential contours for 100 nm channel HMOSFET.  $V_{gs} = -1.5$  V,  $V_{ds} = -1.5$  V.

### 8.3.2.2 Riddle of the large threshold voltage of HMOSFET

In the previous section an attempt is made to calculate the drain saturation currents  $I_{Dsat}$  by assuming the channel is velocity saturated and non-velocity saturated. The assumptions seem to give inconsistencies. It looked like HMOSFET is operating in velocity saturated mode and MOSFET is not. Assumptions were introduced arguing that the MOSFET channel is only partially velocity saturated. All these guess work are introduced to get the designed  $I_{Dsat}$  value of 100  $\mu$ A. In any of the case there was no experimental or simulated data to back up the assumptions. An idea was also put forth such that the troublesome portion of the  $V_{ds}$  must be dropping near the high field drain junction.

In this sub-section the data from the numerical simulator is used to probe into the assumptions and validities. Quite curiously it is found that the arguments have qualitative value in most cases but there were no data sets to support the arguments. Before getting into calculation using simulator data (1) it will be worthwhile to take a look at the parameters used and the ways in which ISE-TCAD computes certain important values, eg. mobilities under different conditions, (2).the numerically simulated data can be tested with known theoretical relationships to ensure the validity of the data, (3) how to control and design both MOSFET and HMOSFET with pre-determined  $I_{Dsat}$  relationships. Table 8.1 gives some of the important parameters used in the numerical simulator ISE-TCAD.

Table 8-1. Some of the important material constants and other values used by the numeric simulator ISE-TCAD for the simulation of  $\text{Si}_{1-x}\text{Ge}_x$  HMOSFET and conventional silicon MOSFET.  $\text{Si}_{1-x}\text{Ge}_x$  has an x value of 0.24 and the n-type silicon substrate doping is  $5 \times 10^{17}/\text{cm}^3$ .

	$\text{Si}_{1-x}\text{Ge}_x, x = 0.24$	Si, $N_{\text{sub}} = 5e17/\text{cm}^3$	
x	0.24	0	
mu_low	813	200	
beta	1.213	1.213	Cauchy-Thomas
vsat	8.37e6	8.37e6	default
$\epsilon(x)$	12.4	11.7	

Ge fraction x of 0.24 was chosen for the HMOSFET in simulations. This fraction of Ge gives a HMOSFET with a  $N_{\text{sub}}$  of  $5e17/\text{cm}^3$  and an  $I_{\text{Dsat}}$  of  $100 \mu\text{A}$ , same values as the MOSFET. The low field mobility,  $\mu_{\text{low}}$ , calculation is given in chapter 4. It is reproduced here for convenience.

$$\mu_{\text{low}} = 470 + (1900 - 470)x \quad (8.17)$$

Here 470 is the hole mobility in undoped silicon, 1900 is the hole mobility in undoped Ge and x is the Ge fraction. A plot of the low field mobility is given in Figure 8-21.. Using this low field mobility and  $v_{\text{sat}}$  high field mobility is calculated using Cauchy-Thomas formula.

$$\mu(F) = \frac{\mu_{\text{low}}}{\left[ 1 + \left( \frac{\mu_{\text{low}} F}{v_{\text{sat}}} \right)^\beta \right]^{1/\beta}} \quad (8.18)$$

The high field mobility for  $\text{Si}_{1-x}\text{Ge}_x$  and silicon are plotted in Figure 8-22. . Knowing the high field mobility and the field the velocity is calculated. The velocity calculated this

way is used in TCAD. The whole process of finding velocity this way is called Canoli model. This is plotted in Figure 8-23 for both  $\text{Si}_{1-x}\text{Ge}_x$  and silicon.

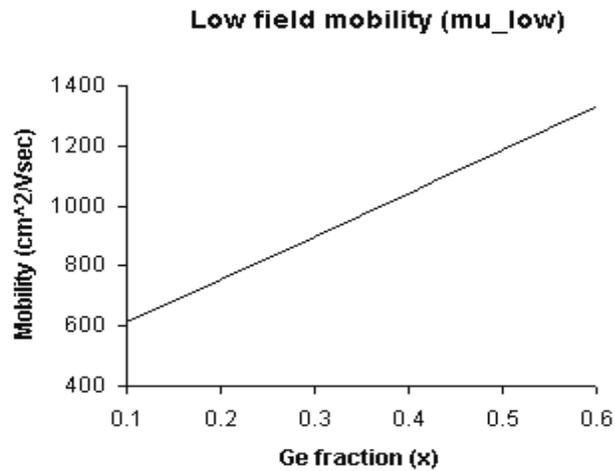


Figure 8-21. Low field hole mobility for strained  $\text{Si}_{1-x}\text{Ge}_x$  vs Ge fraction

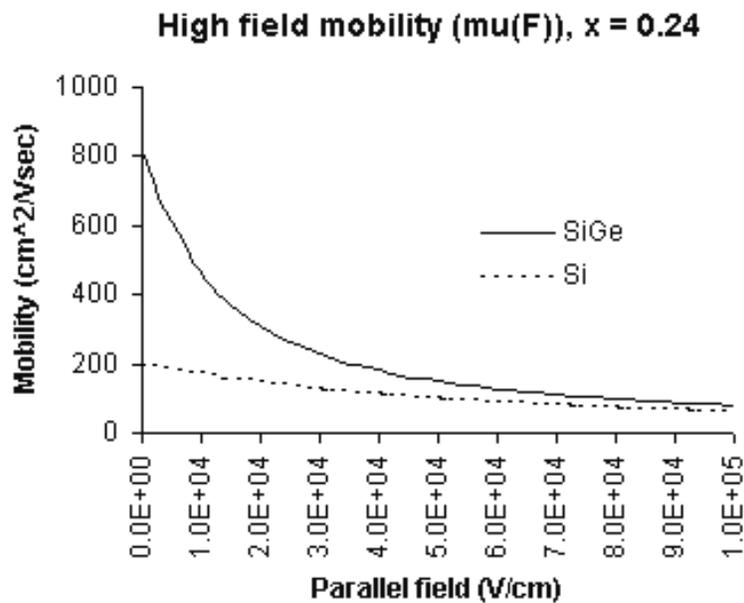


Figure 8-22. High field mobility for strained  $\text{Si}_{1-x}\text{Ge}_x$  with  $x = 0.24$  and silicon substrate with  $N_{\text{sub}}$  of  $5 \times 10^{17}/\text{cm}^3$

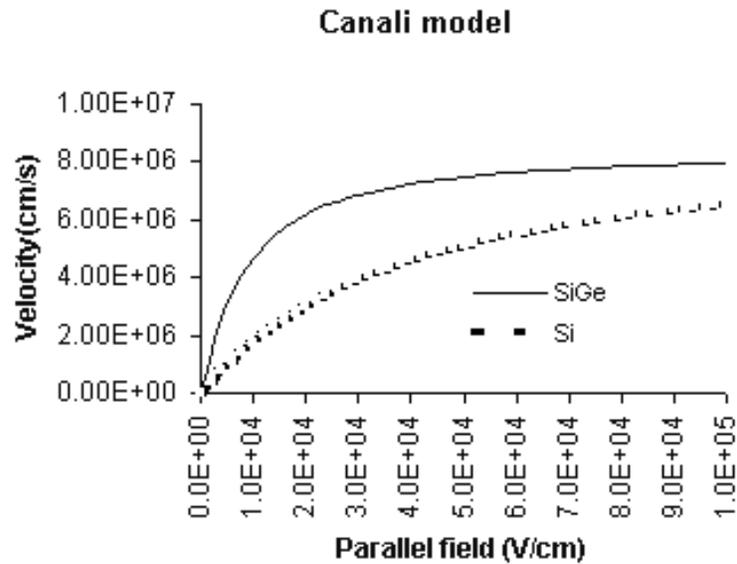


Figure 8-23. Velocity vs field for  $\text{Si}_{1-x}\text{Ge}_x$  with  $x = 0.24$  and silicon substrate with  $N_{\text{sub}}$  of  $5e17/\text{cm}^3$

The aim is to calculate the drain saturation currents for both MOSFET and HMOSFET and make sure they are equal. The two devices were designed to have  $100 \mu\text{A}$  before inputting to the simulator. It must be possible to use the numerical output data and back calculate  $I_{\text{Dsat}}$ . If charge per unit area and the average velocity of holes in the channel region are known  $I_{\text{Dsat}}$  calculation is straight forward. Assume  $Q$  is the total charges under unit area and  $v_{\text{avg}}$  is the average velocity of holes in the channel region.

$$I_{\text{Dsat}} = Qv_{\text{avg}}$$

$v_{\text{avg}}$  has to be found by averaging the velocity from source to drain through the channel region. In a MOSFET the charge density is given in terms of gate capacitance and the gate overdrive.

$$Q = C_{OX} (V_{gs} - V_T)$$

In the numerically simulated plots of Figure 8-10 the point of intersection of the two straight line segments is located approximately at the threshold voltage. So  $V_T$  and  $V_{TH}$  for MOSFET and HMOSFET respectively can be found from this figure. Numerical simulator also gives velocity distributions. The velocity along the channel region can be plotted and  $v_{avg}$  can be approximately found from this velocity vs source-to-drain distance plot.

Now if the currents in both devices is to be  $100 \mu A$ ,

$$Q_{avg} (mos) v_{avg} (mos) = Q_{avg} (hmos) v_{avg} (hmos) = I_{Dsat} = 100 \mu A$$

This gives the relation,

$$\frac{Q_{avg} (mos)}{Q_{avg} (hmos)} = \frac{v_{avg} (hmos)}{v_{avg} (mos)} \quad (8.19)$$

From Figure 8-10,

$$V_T \text{ (MOSFET)} \approx -0.4 \text{ V}$$

$$V_{TH} \text{ (HMOSFET)} \approx -1 \text{ V}$$

Magnitudes of charges are given by,

$$Q_{avg} \text{ (MOSFET)} = C_{OX} (V_{gs} - V_T) = 1.5 - 4 = 1.1 \text{ (normalized with } C_{OX})$$

$$Q_{avg} \text{ (HMOSFET)} = 1.5 - 1 = 0.5$$

The velocity profile for the MOSFET and the HMOSFET are plotted in Figure 8-24 and Figure 8-25 respectively. From the plot approximate average velocities for MOSFET and HMOSFET are,

$$v_{avg} \text{ (MOSFET)} \approx 0.8 \times 10^6$$

$$v_{avg}(\text{HMOSFET}) \approx 2.5 \times 10^6$$

The ratios of charge and average velocity should match equation ( 8.19)

$$\frac{Q_{avg}(mos)}{Q_{avg}(hmos)} = 3.2 \quad (8.20)$$

$$\frac{v_{avg}(hmos)}{v_{avg}(mos)} = 3.1 \quad (8.21)$$

These two ratios are quite close showing that the approximate theory and numerical simulation are in agreement.

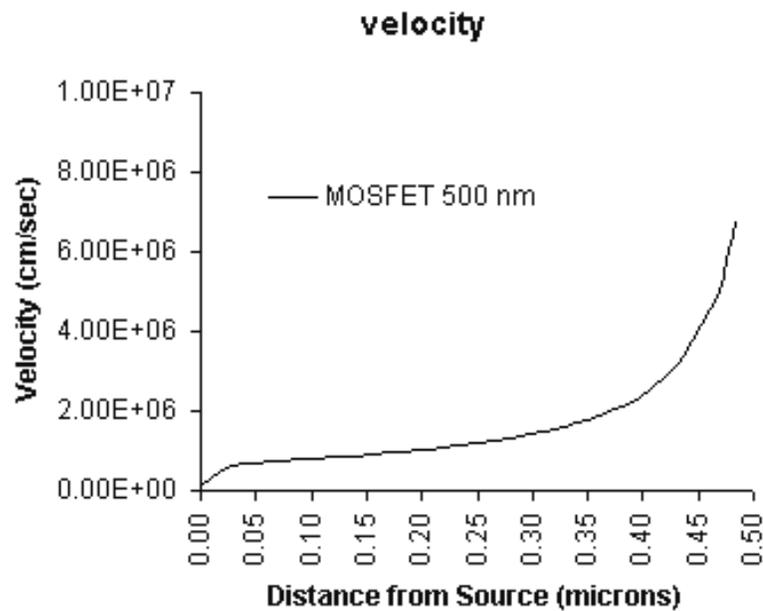


Figure 8-24. Plot of hole velocity vs source-to drain distance. This is for the 500 nm channel length MOSFET.

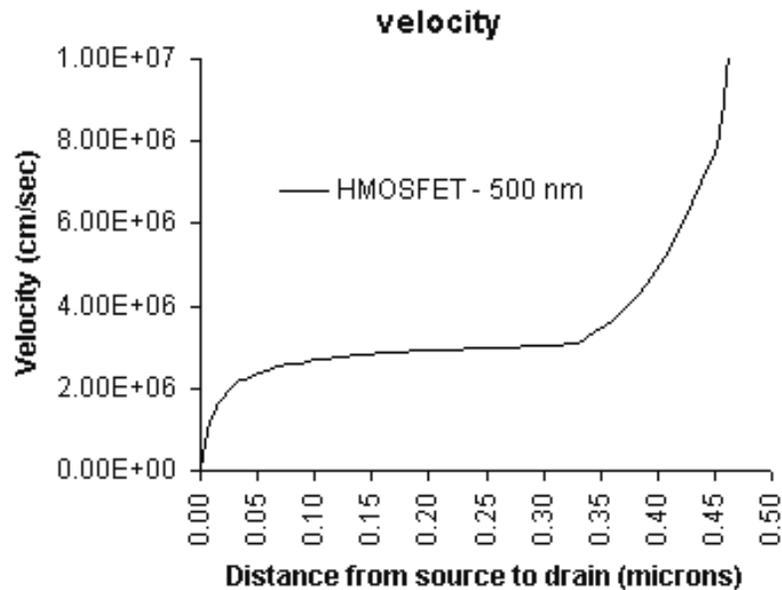


Figure 8-25. Plot of velocity vs source-to-drain distance. This is for the 500 nm channel length HMOSFET.

A careful examination of the numerically simulated results in Figure 8-24 and Figure 8-25, for 500 nm channel length devices, indicates that indeed neither MOSFET nor HMOSFET are velocity saturated. As a consequence, the higher mobility in the HMOSFET structure allows a lower carrier density than does the MOSFET structure. Hence the threshold voltage of the HMOSFET can be increased beyond the threshold voltage of the MOSFET. This greatly reduces the off-current in the HMOSFET compared to the MOSFET. So for the same on-current,  $I_{on}$ , the off-current,  $I_{off}$ , in HMOSFET is few orders of magnitude lower than the  $I_{off}$  in MOSFET. This can be seen in Figure 8-10.

#### 8.4 Scalability of $\text{Si}_{1-x}\text{Ge}_x$ HMOSFET

From the previous sections it is clear that the HMOSFET can be scaled down to the same limits of MOSFET or even slightly more. In addition  $I_{off}$  in the HMOSFET is a few

orders of magnitude less compared to MOSFET. This will reduce the overall power dissipation and compensate for the poor thermal conductivity of SiGe compared to Si (see Table 4.1). With orders of magnitude lower  $I_{off}$  HMOSFET can have much higher packing density compared to Si MOSFET.

$Si_{1-x}Ge_x$  can be introduced into the Si foundry with very few modifications, as is done in SiGe BiCMOS process. Fabricating all  $Si_{1-x}Ge_x$  CMOS structures is advantageous from the following aspects. First very high speed CMOS devices will be useful for low noise RF front ends [ 100 ], it is radiation tolerant, its overall power dissipation is lower than the same in silicon CMOS and it can be scaled down with methods similar to that of silicon CMOS.

## 8.5 Summary of Short Channel Simulations

In this chapter the short channel behavior of MOSFET and HMOSFET with the same  $N_{sub}$  and the same long channel  $I_{Dsat}$  are studied. Two devices, a MOSFET and a HMOSFET, with the same long channel saturation currents  $I_{Dsat}$  and the same substrate doping density  $N_{sub}$  are designed using analytical programs.  $N_{sub}$  is selected in such a way that the MOSFET can be scaled down to 100 nm. This is determined using an empirical formula. Both devices have 500 nm long channel length to begin with. The MOSFET has  $t_{ox} = 3$  nm,  $x_j = 30$  nm,  $N_{sub} = 5e17/cm^3$  and p+ gate. The HMOSFET structure used is 3,5,5,2/0.24/5e18,5e17 with  $t_{ox} = 3$  nm and p+gate. The long channel  $I_{Dsat}$  for both devices is around  $100 \mu A/\mu$ . Then the channel lengths of both devices are reduced in steps and a comparison is made of their sub-threshold gate characteristics, threshold voltages and saturation currents. This is carried out at each step of the scaling down. By

plotting  $\log(\text{mag}(I_{ds}))$  vs  $V_{gs}$  side by side for all channel lengths it was found that below 150 nm both MOSFET and HMOSFET show short channel effects indicated by a rise in off state currents. These plots are generated using the numerical program. Next threshold voltage drop off was compared. MOSFET starts showing threshold voltage drop around 200 nm whereas HMOSFET starts to show threshold voltage drop at 150 nm.

The drain characteristics of 500 nm and 150 nm MOSFET and HMOSFET are plotted together in Figure 8-7. For the long channel, 500 nm, both devices saturates to the same  $I_{Dsat}$ . But for short channel, 150 nm, the HMOSFET saturates whereas the  $I_{ds}$  of the MOSFET do not show saturation but keeps increasing with  $V_{ds}$ . This is an indication of SCE in MOSFET. From Figure 8-5 to Figure 8-7 it is very clear that the MOSFET shows SCE at a slightly early stage of scaling down than the HMOSFET. So it can be stated that the HMOSFET can be scaled down a bit further than the MOSFET. This will reduce the die area and improve operational speed. Another factor of importance is that for the same ON currents the HMOSFET has few orders of magnitude lower OFF currents compared to MOSFET. This is very important since it can compensate for the poorer thermal conductivity, compared to silicon, and the device layout density can be made higher without adverse thermal limitations. It also help save power in battery operated systems.

The results  $I_{Dsat}$  and  $V_T$  from the analytical HMOSFET design program and the numerical simulation program closely matches down to a channel length of 150 nm. Below this channel length short channel effects are prevalent and the analytical program is not designed to accommodate severe SCE.

## CHAPTER 9

### OPTIMIZATION OF THE P-HMOSFET STRUCTURE

In the previous chapters p-HMOSFET structures were explained. Design of the same was also discussed. Specific long channel structures were designed. Simulations were carried out on these long channel devices using the analytical analysis and design program developed in this project. These results were compared with numerical simulation results using ISE-TCAD [109]. Further the channel lengths were reduced. Analytical and numerical simulations were carried on these short channel devices. A comparison is made between the short channel behavior of CMOS and p-HMOSFET. Both showed nearly same behavior even though p-HMOSFET showed the ability to be scaled down a little bit more.

#### 9.1 Limitations of the p-HMOSFET Structure

In this chapter the structure is studied further and some modifications are introduced to the structure discussed in the previous chapters. Figure 5.1 gives the device structure and Figure 5.4 gives the band diagram in chapter 5. The structures discussed in the previous chapters have two drawbacks.

1. As can be seen in Figure 5.4 the buffer layer cannot be made much thicker due to the potential drop in that layer. Since the device is designed to operate at 1.5 V this is a limitation.
2. Because of the above reason the p-layer is bridging the source and the drain. This is because bringing the p-layer down will make the buffer layer thicker.

Limitation of the device is shown in Figure 9-1. The figure shows the band diagram from the source to the drain through the p-layer. The doping in the p-layer cannot be increased to higher values since the barrier at the source junction will reduce. This will increase the off state leakage currents.

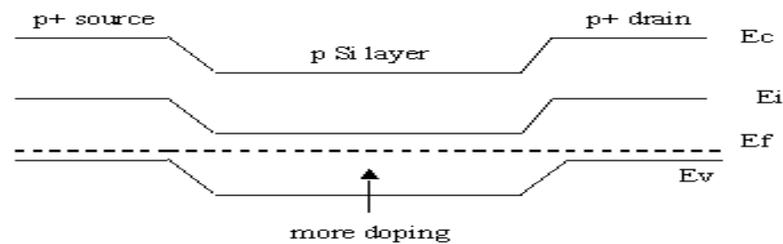


Figure 9-1. Band diagram from source to drain through the p-layer. This is copied here from Figure 5.2 for convenience.

## 9.2 A Modified Structure

These limitations can be overcome if the potential drop in the buffer layer is reduced to zero or the field is made zero. This can be accomplished if the negative ions in the p-layer are made equal to the positive ions in the depletion region of the substrate. This gives two advantages and a disadvantage.

1. The p-layer can be pushed down to below the source. This will avoid the p-layer bridging the source and drain thus reducing off state leakage currents.
2. If the p-layer is pushed down below the source the doping of the p-layer can be increased as desired to make the field in the buffer layer zero. So a delta-doping layer can be used.

3. The regions above the p-layer are undoped. When the p-layer is moved down there will be more undoped regions between source and drain. This can widen the depletion widths and lead to leakage currents.

Such a modified device structure is shown in Figure 9-2. The valence band diagram for the same is shown in Figure 9-3.

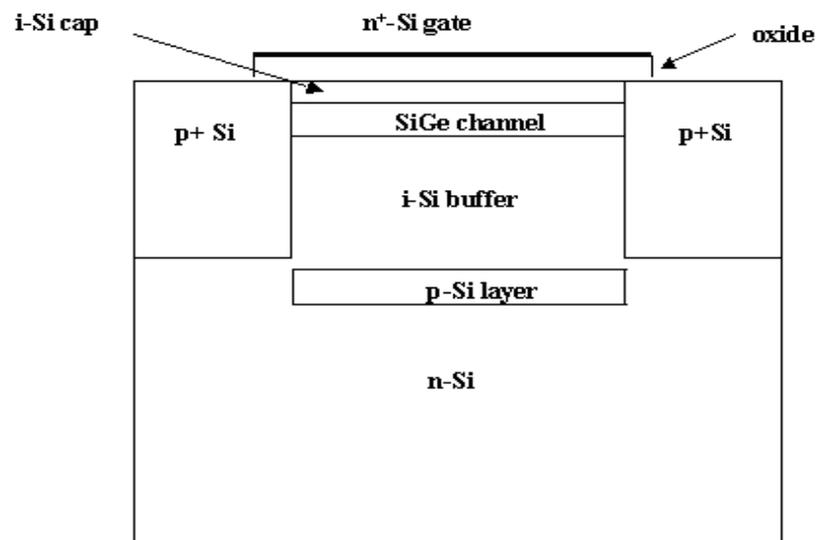


Figure 9-2. Structure of a modified p-HMOSFET. Here the p-layer doping is adjusted such that the field in the buffer layer goes to zero.

### 9.3 A Feasibility Analysis of the Modified Structure

In order to do a quick analysis of the modified structure and to study the feasibility of this structure a coarse analysis is done using an Excel program that calculates the potentials at the junctions [ 159]. The program uses the potential at the substrate junction as an input variable. Once that is known the potentials in the p-layer and buffer can be calculated. Then the potential in the quantum well is obtained by adding the well height. The cap layer potential needs a correction due to the holes in the quantum well. This

correction is applied while calculating the potential in the cap layer. The input to the program and the calculated values for one particular structural choice are given in Appendix B. This simple program is used only for demonstration since it has the following shortcomings,

1. The potential at the substrate depletion edge is assumed and it is not controlled by a gate voltage as in the real case
2. Since the substrate potential drop is an assumed value the gate voltages vary with different structures.
3. Due to above reasons hole densities calculated for different structures are for different gate potentials. So comparison is difficult.

Finding the substrate potential drop with an applied gate voltage needs iteration. The analytical analysis and design program developed in this project does that. Similar structures with flat band buffer layer are analyzed using this analytical design program and the results are presented following this demonstrative analysis

#### **9.4 Analysis Using the Analytical HMOSFET design Program**

In the analytical program it is fast and easy to vary the p-layer location as well as doping. As mentioned before the analytical program does the computation by iteration for a fixed gate voltage. A 3 nm p-layer doped to a density of  $5 \times 10^{18}/\text{cm}^3$  is used. This value gives a nearly flat band in the buffer layer for the structural parameters chosen here. The cap layer and quantum well layer are kept fixed at 3 nm and 5 nm respectively. The substrate doping is  $5 \times 10^{17}/\text{cm}^3$ . Ge fraction in QW is 0.5. So the device is  $3,5, X_{\text{buf}}, 3/0.5/5 \times 10^{18}, 5 \times 10^{17}$ . First a thin buffer layer of 5 nm was used. This places the p-

layer in between the source and drain. Figure 9-4 gives the valence band edge. As can be seen in Figure 9-4 the 5 nm buffer layer just below the quantum well has a nearly flat band edge. Now the buffer layer is made 35 nm which puts the p-layer below the source-substrate junction. Rest of the parameters are kept the same. Figure 9-5 shows the valence band edge which is nearly flat. Figure 9-6 shows the analytically computed drain characteristics.

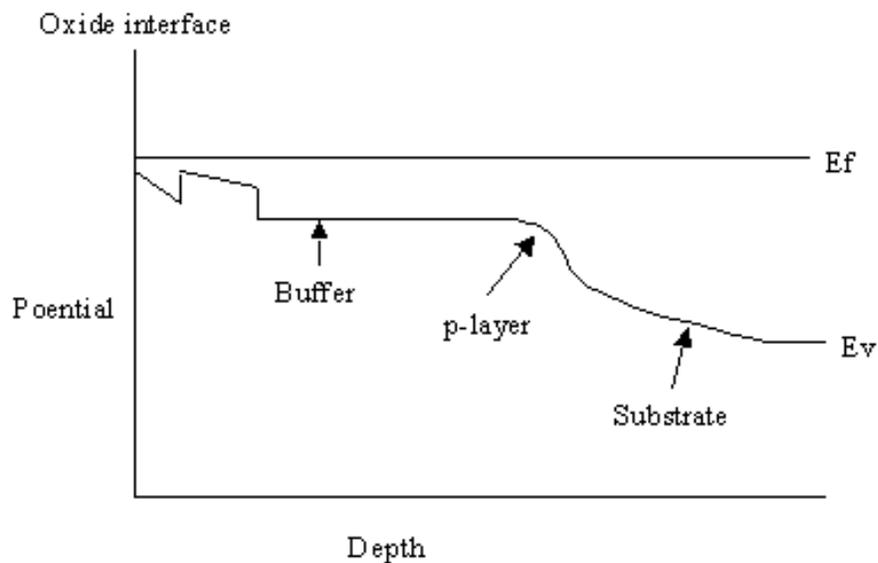


Figure 9-3. The valence band diagram for the structure in Figure 9-2. p-layer doping is adjusted to make the field in the buffer zero.

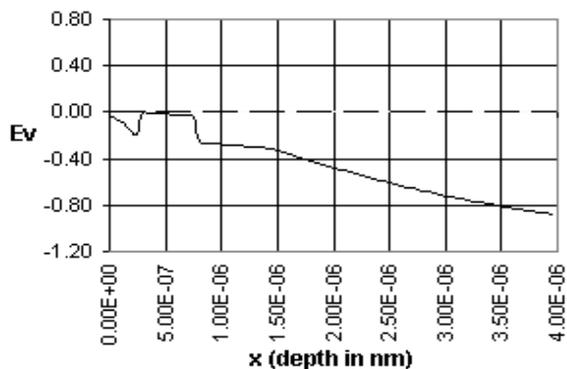


Figure 9-4. Valence band edge of a 5 nm buffer layer device. The band edge in the buffer layer below the quantum well is nearly flat. The device is 3,5,5,3/0.5/5e18,5e17.

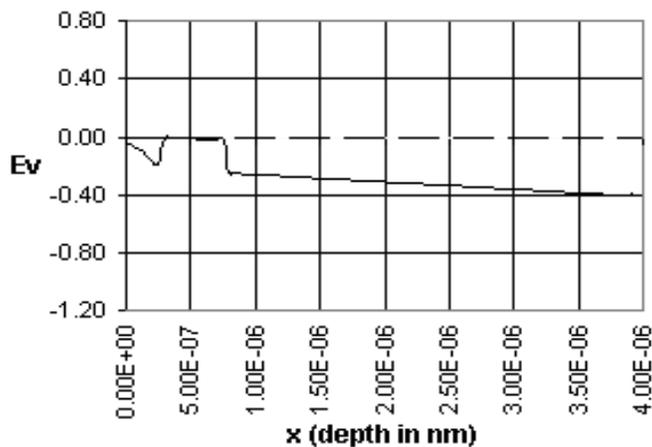


Figure 9-5. Plot of valence band edge. The buffer layer is 35 nm which puts it below the source which is 30 nm deep. p-layer doping and rest of the values are same as in Figure 9-4. The device is 3,5,35,3/0.5/5e18,5e17.

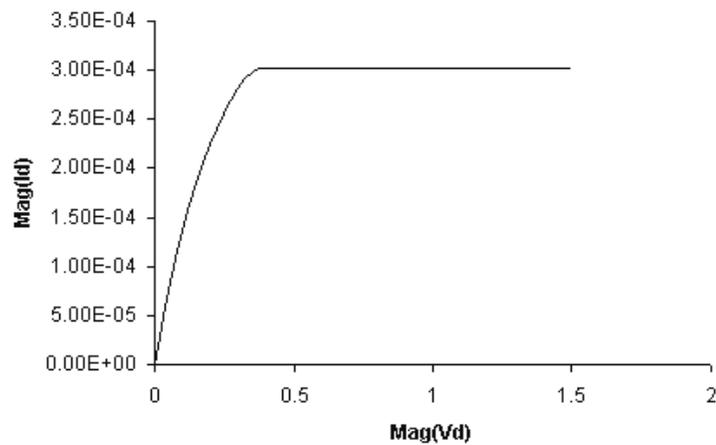


Figure 9-6. Analytically computed drain characteristic for the device in Figure 9-5. The buffer is 35 nm thick. The device is 3,5,35,3/0.5/5e18,5e17. Analytical design program used.

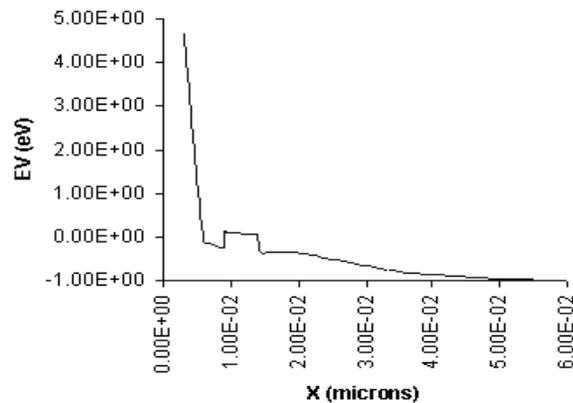


Figure 9-7. Valance band of a 5 nm thick buffer device. Notice the flat region over 5 nm below the quantum well. The result is from numerical simulation. The left side sharp rise is in the oxide.  $V_{ds} = -1.5$  V,  $V_{gs} = -1.5$  V

## 9.5 Analysis Using Numeric Simulator

The analytic program gave the ON currents. But one of the advantages, as discussed earlier, of moving the p-layer down is the reduction of leakage currents. In order to

evaluate the same the same structures are simulated using the numerical simulator ISE-TCAD [109]. Figure 9-7 shows the valence band edge of a 5 nm buffer device. The buffer region to the back of the quantum well is flat. Figure 9-8 shows the drain characteristic of the same device. The gate characteristic is given in Figure 9-9. It does not show any SCE.

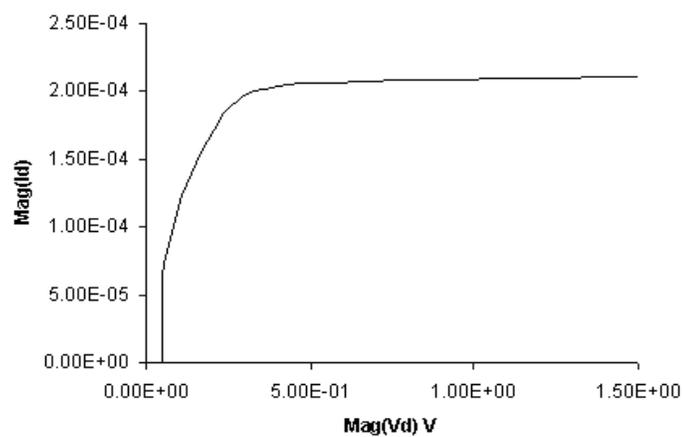


Figure 9-8. Numerically simulated drain characteristic of the 5 nm buffer device.

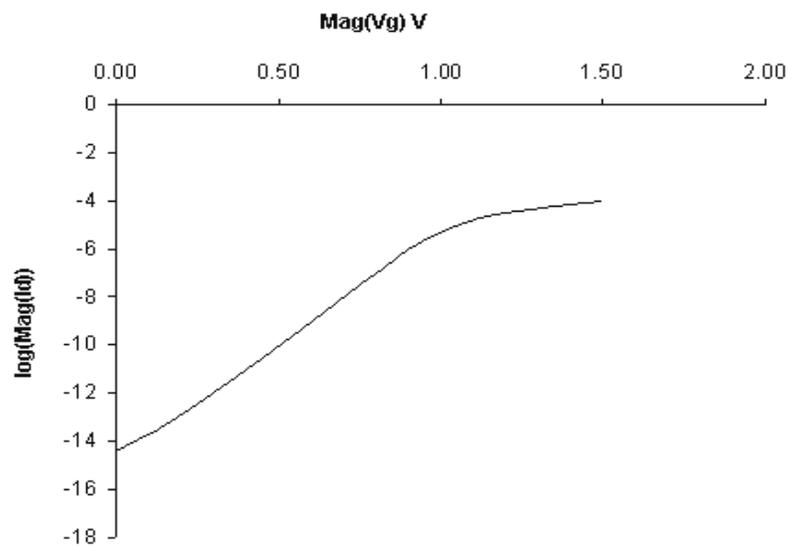


Figure 9-9. Id vs  $V_{gs}$  of the 5nm buffer device from numerical computation

Now the p-layer is moved down to below the source. The buffer layer was made 35 nm thick. To compensate for the n-type substrate doping the p-layer doping is increased by the substrate dopant concentration. Figure 9-10 shows the valence band edge. The buffer layer band edge is not flat all over. The buffer layer extends to 43 nm. 25 to 43 nm the band edge is flat. The rising edge above 25 nm could be due to the fact that the n-type dopant is present in that region. Figure 9-11 shows the drain characteristic. The saturation current  $I_{Dsat}$  in this case is higher than the 5 nm buffer layer case given in Figure 9-8. Figure 9-12 has the gate characteristic. As speculated at the beginning there is leakage happening at sub-threshold region. Even though the off currents are not much different from the 5 nm buffer case the advantage of the higher slope is lost at low gate voltage.

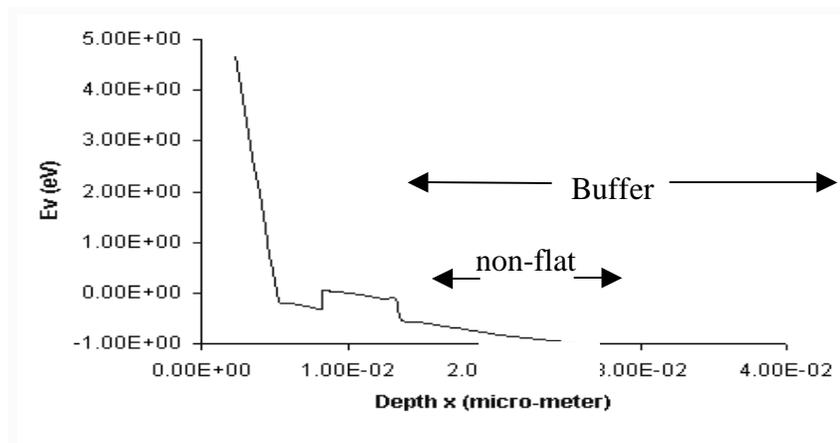


Figure 9-10. Valence band edge of a 35 nm thick buffer layer device. It is not flat all over the 35 nm width. It has a slight slope up to 25 nm. This plot is from numerical simulation. The left side sharp rise is in the oxide.  $V_{ds} = -1.5$  V,  $V_{gs} = -1.5$  V. Buffer layer extends to 43 nm. But buffer band edge is not flat below 25 nm.

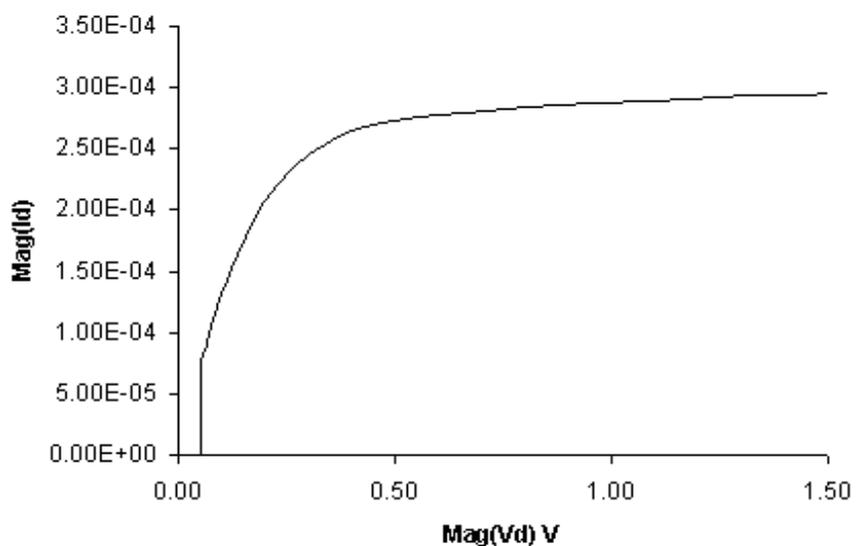


Figure 9-11. Numerically simulated drain characteristic of the 35 nm buffer device.

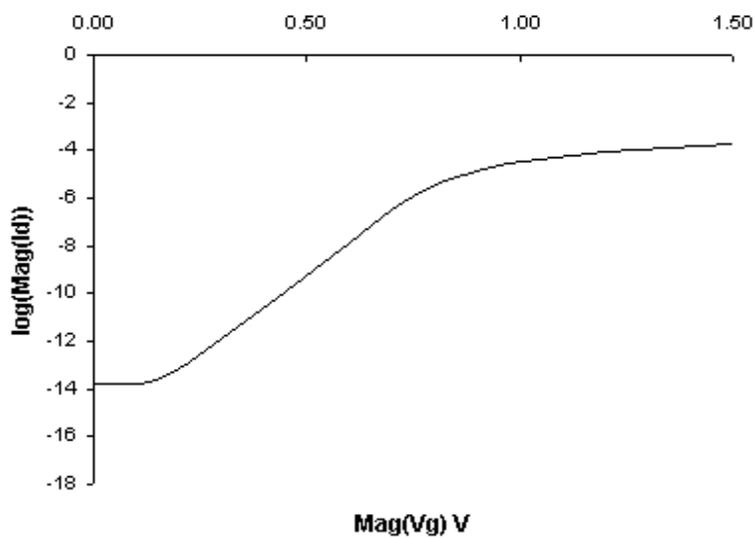


Figure 9-12.  $I_d$  vs  $V_{gs}$  of the 35 nm thick buffer device. Notice the leakage at low gate voltage. This result is from numerical simulation.

## 9.6 Conclusion

In this chapter a modification was introduced to the p-HMOSFET analyzed in chapter 7. The idea is to make the band flat in the buffer layer so that the p-layer can be moved down below the source level. This will allow higher doping in the p-layer without reducing the barrier at the source junction that occurs if the p-layer is located between the source and drain. It was found that the above modification gives more  $I_{on}$ . This could possibly be due to the source and drain depletion regions coming close while the drain voltage is present. The fact that at low sub-threshold region there seems to be leakage happening justifies this speculation. Figure 9-13 shows the potential contours for the 35 nm buffer structure. The sharp kinks are located at the p-layer. It can be imagined that the kink will move to right side (drain) if the p-layer is raised upwards.

Raised source and drain designs could prevent the p-layer from bridging across them. As a concluding remark it looks like there is nothing to gain by moving the p-layer deep into the substrate. Even though it will still allow low voltage operation due to the zero field in the buffer layer the undoped regions above seem to be contributing to the leakage.

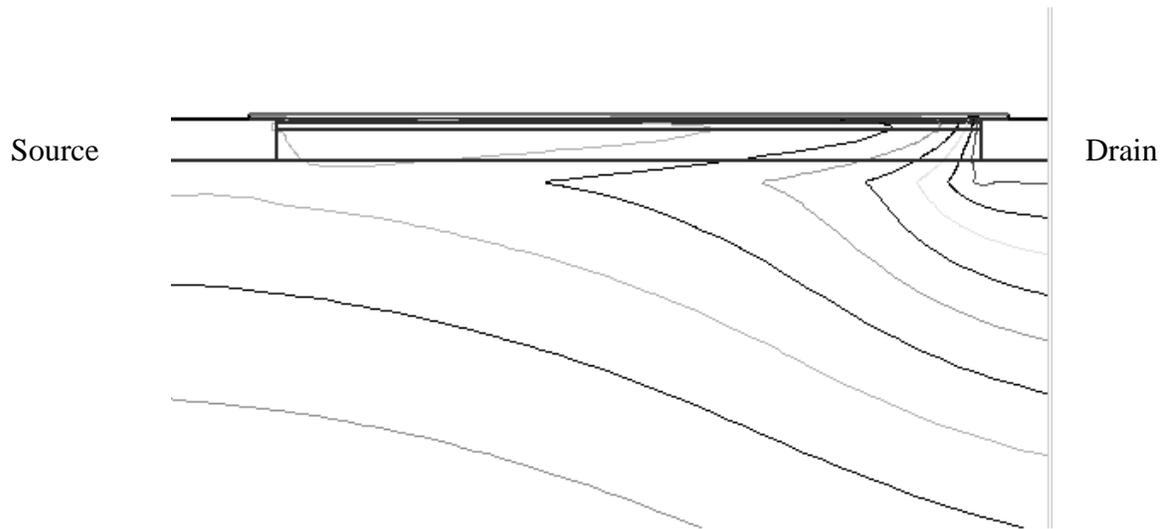


Figure 9-13. Two dimensional potential contours for the case of 35 nm buffer. p-layer is located where the notches are seen. It looks like there is a tendency for the notches to spread out to the left (source) if the p-layer is moved down.

## CHAPTER 10

### CONCLUSION

Conclusions are given in this chapter. Contribution of this project, an overview of the design tool or the analytical HMOSFET analysis and design program, limitations of the analytical HMOSFET design program and future improvements to this work are discussed. In this first trial program only dc characteristics are studied. Structural modifications to make the HMOSFET operate at high frequencies are also discussed.

Silicon germanium alloy semiconductor seems to hold promise to the silicon semiconductor industry which is at the verge of reaching the limits of its consistent down scaling of devices. Since  $\text{Si}_{1-x}\text{Ge}_x$  is compatible with silicon processing introduction of  $\text{Si}_{1-x}\text{Ge}_x$  to the silicon integrated circuits industry is not difficult. In fact  $\text{Si}_{1-x}\text{Ge}_x$  based bipolar transistors are already in production in the BiCMOS process lines of many industries. In the dominating market of CMOS integrated circuits the p-MOSFET takes up over two times the foot print of an n-MOSFET due to the low mobility of holes. Hence it is reasonable to start with a p-channel MOSFET device and make it at least as fast as the n-channel device. Hence a p-channel device is chosen as the focus of this dissertation.

Band gap engineering used to be an exclusive tool in the domain of III-V compound devices which has produced MODFET which can operate to speeds corresponding to less than 10 psec delay. This is achieved by confining carriers in an undoped QW layer and transporting them through this undoped channel layer. Addition of Ge to Si reduces the

band gap as well as increases the mobility. Using this to advantage band gap engineering can be implemented in the all silicon industry by introducing SiGe alloy. Two decades back  $\text{Si}_{1-x}\text{Ge}_x$  based device designs were based on the MODFET and did not have much success. These designs were not suitable for silicon industry. To name a few reasons (1) the MODFET uses Schottky metal-semiconductor gates whereas Si technology uses  $\text{SiO}_2$  insulated gates, (2) The doped layer is on top of the GaAs channel layer in III-V based MODFET whereas in Si technology it has to be below the channel layer to improve transconductance, (3) With complex processing steps and many layers the designs were not suitable for economical fabrication in the silicon industry. But now a days many MOSFET based hetero junction device designs, making use of the superior  $\text{SiO}_2$  gate insulator, are in publications. The p-channel hetero MOSFET (p-HMOSFET) presented in this work is also based on MOSFETs. A recessed undoped pseudomorphically grown strained QW channel layer is used to confine holes within and transport them through it. The mobility is enhanced due to (1) lack of dopant ions in the QW, (2) very low defect pseudomorphically grown QW walls, (3) recessed QW keeps the holes away from the inferior Si/ $\text{SiO}_2$  interface. Simulations have shown that these p-HMOSFET devices have about three orders of magnitude less OFF currents compared to MOSFET while keeping the ON currents the same in both devices.

## **10.1 Contributions of this Dissertation**

The main contribution of this work is in the development of an easy to use analytical analysis and design program suitable for quick design of p-HMOSFET using a minimum layer sequence and parameter set sufficient to achieve a good deal of improvement over

the conventional p-MOSFET. This program can be used for optimization of long channel p-HMOSFET. The one dimensional Poisson equation is solved first. The results are programmed into an Excel spreadsheet as a tool to come out with an easy to use analytical program to analyze the internal characteristics of the device structure used. More details of the testing of the analytical program is given at the end of section 10.2. Using an easy to use input worksheet of the Excel spread sheet the variable parameters can be varied interactively and very quickly devices meeting required specifications like threshold voltage,  $I_{Dsat}$  or  $I_{off}$  can be arrived at. In addition internal parameters like potentials, hole distributions or hole population centroid can be monitored to find out how they change in response to structural changes and to understand the physics involved. This gives a good deal of insight compared to numerical simulators. Optimization can be achieved by monitoring the responses to the input parameters. Internal parameters like potentials and hole density distributions can be monitored and near ideal devices can be quickly arrived at. Afterwards only a few numerical simulations may be needed to fine tune the design. On the contrary numerical device simulators like ISE-TCAD requires devices to be drawn, modified and takes time to give out results which in most cases need further processing. The development of a rapid design tool is the objective set out in chapter 1 and this goal has been achieved. A generalized analytical program of this nature that can assist in generalized design of a p-HMOSFET is not found in the literature.

Most of the published works fall into two categories. In one set devices are fabricated, tested and results are given. In many cases improvements in the on currents  $I_{ON}$  has been claimed. But how the test device structure was chosen is unknown. In some cases

simulations are carried out prior to fabrication. The second set is purely simulations. In either group of publications how or why a particular structure is chosen for fabrication or simulation is unknown. Most of these works in both categories focus around few parameter studies. Some examples are on current  $I_{on}$  dependence on cap layer thickness, threshold voltage  $V_T$  and surface threshold voltage  $V_{TS}$  dependence on cap layer thickness and so on. Both the fabricated device and analytical simulations are partial studies of the device. That is because they report some terminal characteristics like  $I_{on}$ ,  $V_T$ ,  $V_{TS}$ , short channel behavior and so on. There is hardly any mention of internal potentials or hole density distributions. Curiously the low off currents  $I_{off}$  found in simulations with the analytical analysis and design program developed here are mentioned only in a couple of publications. The analytical program presented here displays internal parameters like potentials and hole densities and their distributions. This gives the designer a better idea of the functioning of the device and guides in the directions to vary the structural parameters to arrive at an optimum device. The feedback from the Excel spread sheet analytical program is almost instantaneous. Out of the published works only a couple are presented as analytical methods. The vast majority of simulations are carried out in time consuming numerical simulators. From this discussion it is clear that the analysis tool developed herein is a step forward in two directions – information and design time.

The following are achieved in the work presented in this dissertation,

1. Generalized p-channel hetero junction MOSFET, named p-HMOSFET, is analyzed using one dimensional Poisson equation. This general structure chosen

has the basic necessary layers and parameters to achieve a good deal of improvements over the conventional p-MOSFET

2. The result of the analysis is transformed into an analytical program using Excel spreadsheet as a tool. The analytical program takes user inputs for device structure and other parameters and very quickly produces internal potential distributions and the terminal characteristics of the device without the need for the time consuming and tedious numerical algorithms.
3. This analytical program has been tested using numerical programs and the accuracy is found to be good. Devices can be analyzed and designed with this program up to the onset of short channel effects. Short channel modifications are also implemented in this program and the short channel support automatically takes effect in this program. So the program can be used for both long and short channel device analysis and design, the limit being the onset of severe short channel effects.
4. The analysis and design program discussed here can be a very efficient interface between a successful design and the numerical simulators. The designer can quickly come up with a near optimum design using this analytical program. Afterwards only a few numerical simulation runs would be enough to confirm or fine tune the device parameters.
5. Using this program the useful range of device parameters as well as design windows for long channel p-MOSFET devices have been produced.

6. A comparative study of the short channel MOSFET and p-HMOSFET also has been conducted. It was found that both devices can be scaled down to similar limits before the onset of short channel effects.

A few important aspects associated with the p-HMOSFET are brought out and studied.

1. One aspect associated with HMOSFET is that it has two threshold voltages, (1) for the recessed QW channel, (2) an undesirable surface channel one. All devices analyzed are for  $-1.5$  V operation. Hence while maintaining the lowest threshold magnitude for the SiGe QW channel to obtain maximum currents it is important to keep the magnitude of surface threshold greater than  $1.5$  V.
2. Another important aspect of Si/Si<sub>1-x</sub>Ge<sub>x</sub> p-HMOSFET is that the off currents in these devices are few orders of magnitude less than those in conventional MOSFET for the same on current  $I_{on}$ . This is attractive for battery operated systems and for thermal management.
3. It has been found that it is not desirable to vary all design parameters. It is better to keep some of them as fixed parameters, at their optimum values. The design variables have limitations or bounds and are interdependent in many cases

Long channel simulations were carried out using the analytical program as well as numerical simulator with various combinations of parameters. These results are organized into design windows. These are given in chapter 7. This is useful in design because for certain fixed parameters range of other parameters are available and

appropriate values can be selected and simulated in numerical simulators for checking the functional accuracy. This saves time since the numerical program need to be run only a few times since the parameter values are known to some extent. For example there is a design window for  $V_T$  and  $V_{TS}$  for substrate doping  $N_{sub}$  vs Ge composition  $x$ .  $V_{TS}$  has to be kept above the operating voltage magnitude to avoid surface channel. At the same time a low  $V_T$  magnitude is desirable to get more  $I_{on}$ . From this window appropriate combinations of  $N_{sub}$  and  $x$  can be chosen and simulated.

It was found that one of the design modification with zero field in the buffer layer may have advantages. This will change the design considerably. The number of negative ions in the p-layer has a strong influence in performance of the device. The p+ source and the p-layer will form a junction with reduced junction potential or barrier, particularly if the p-layer doping is increased. This will produce more leakage currents. With zero field in the buffer layer the p-layer can be moved down below the source and drain junctions thus avoiding the above junction altogether. Simulations were carried out to assess the implications of this modified design. It was found that there is a slight increase in the  $I_{ON}$  with the p-layer placed below the source region. But it was also found that there is SCE and hence leakage in the sub-threshold region. This could be due to the undoped cap, QW and buffer layers at the top. With the p-layer moved far down there could be wide depletion regions in these layers which may be causing SCE.

## 10.2 Outline of the Tool

The device structure of the p-HMOSFET chosen here is simple and general in nature including all desirable structural choices described in chapter 1 and chapter 4. From the

literature survey listed in section 4.5 a pattern for the necessary input parameters in a HMOSFET structure seem to evolve. They are listed below,

1. Thin undoped cap layer
2. Undoped  $\text{Si}_{1-x}\text{Ge}_x$  QW channel layer
3. Ge composition  $x$  in the QW
4. Undoped buffer layer
5. p-layer thickness
6. p-layer doping
7. n-type substrate. doping

All of the above are included in the structure considered here. The thicknesses of these layers and doping profiles are all variables. Out of all these few parameters are kept at their optimum values. For example it is desirable to have the thinnest gate oxide for low voltage application with good gate control of the drain current. But at the current technological level this thickness is limited to 3 nm for a good quality  $\text{SiO}_2$  gate insulator. In addition the tunneling currents will become high at oxides thinner than 3 nm. So there is no point in varying this thickness and is kept fixed at 3 nm. However it is possible to vary all parameters in the analytical program presented in this work, if needed.

Taking the full structure as above the one dimensional Poisson equation is solved starting from the substrate and proceeding towards gate. The details are presented in Appendix C. From these solutions the threshold voltages, hole distribution and potential distribution are evaluated. A complex analytical program involving all of the above parameters is written in Excel spread sheet. The program is complex since it involves

iterations to arrive at desired output parameters like hole density distribution and potential distribution. Few of the chosen structures are analyzed using this program and analytical results are compared to the numerical simulation results. The details are given in chapter 6. The analytical program output matches with numerical simulation results very closely upto the onset of short channel effects in both MOSFET and p-HMOSFET simulated. Many structures mentioned in the literature survey of chapter 4 can in fact be tested easily using this program provided those devices do not have any additional layers or graded regions.

After checking the above analytical program outputs with numerical simulations [109] the parameters chosen as variables are stepped through one at a time. The gate oxide thickness, cap layer thickness, QW thickness and setback buffer layer thickness are kept fixed at their optimum values. The p layer doping is also kept at a high value of  $5 \times 10^{18}/\text{cm}^3$  and the thickness is varied instead to increase the effective number of holes in this layer. From these simulations appropriate range of the variables of practical use are identified. For example the useful range of Ge fraction in the QW layer is from 0.3 to 0.5 contrary to the intuitive assumption that higher values will give more carrier confinement and better performance. Similarly the substrate doping range is limited to about  $10^{17}/\text{cm}^3$  to  $10^{18}/\text{cm}^3$  These are detailed in chapter 6. Apart from this design windows are also evaluated with the choice of two parameters at a time keeping the others at fixed values. To name a few the required characteristic for a specific HMOSFET of interest could be on current  $I_{\text{on}}$  value or off current  $I_{\text{off}}$  value. These design windows are useful as a guide to design a desired type of p-HMOSFET using the analytical program. This process is

very easy. The specific structure thus arrived at can be simulated numerically, if desired, for verification purposes.

### **10.3 Summary**

To sum up the research an analytical analysis and design program has been developed and tested taking into account all necessary structural parameters of a generalized p-HMOSFET structure. Using this program as a tool useful ranges of the structural, compositional and doping density are arrived at. Further design windows are constructed to aid as a guide in choosing specific parameters for a desired performance of the p-HMOSFET. The analytical program takes user input parameter values and provides with internal potential distributions and hole density distributions as well as terminal characteristics in an extremely short time compared to the time and efforts involved in defining the device and running time consuming numerical simulations. This analytical analysis and design program is a very efficient interface between the designer and cumbersome time consuming numerical simulators.

### **10.4 Limitations**

Analytical program developed is rigorously tested and the accuracy has been confirmed. Yet it has certain limitations.

1. Being complex with iterative routines convergence may be an issue for some combination of parameters, particularly substrate doping concentration. Then the computations have to be started with seed values in few of the functions and a better choice of parameters should be made.

2. Short channel effects are not included. Some of the effects, particularly confined to short channel devices, like velocity overshoot are not included. Modern day nano meter MOSFET uses many different SCE prevention techniques such as channel engineering, halo profiling, raised source and drain and many other techniques to suppress SCE. Those techniques are beyond the scope of this simple simulations. The assessment of scalability presented is very primitive and did not use any of these modern day techniques. This short channel study is only a preliminary assessment of the strained  $\text{Si}_{(1-x)}\text{Ge}_x$  p-channel HMOSFET.
3. Quantization of energy levels in the SiGe QW is not included. Quantization will cause a reduction in  $I_{\text{on}}$  because the carrier occupancies are split into discrete energy sub-bands in the direction perpendicular to the device plane reducing carrier density because of reduced density of states.
4. Mobility dependence on well width is not considered. The side wall reflections of carriers will reduce carrier mobility. This is particularly important in narrow QW. The maximum QW width may be less than about 7 nm for a maximum Ge fraction of 0.5. This is due to the limitations in growing pseudomorphic interfaces.
5. Density of states dependence on Ge composition  $x$  is not included. Published relation of DOS to  $x$  are available to a limited extent.
6. The velocity saturation and channel length modulation has been taken into account in computing the on state current  $I_{\text{on}}$ .

## 10.5 Recommendations for Future Work

1. The simple analysis and results produced here are a first step in the design of

p-HMOSFET. To improve the analysis few more modifications and additions could be introduced. Carriers, in this case holes, are confined in a 5 nm  $\text{Si}_{1-x}\text{Ge}_x$  QW layer. At high  $V_{gs}$  the upper edge of this QW is lifted higher above the bottom end forming a triangular quantum well where most of the holes get confined. Published analysis show that the lowermost energy band wave functions extend only up to about 5 nm. But in the above condition at least the second energy band also need to be included. So it is more accurate to add quantization in the analysis. This can be done by iteratively solving Poisson equation along with Schrodinger equation.

2. To make the analyses near perfect there are factors like interface roughness scattering, inter-band scattering particularly with quantization, phonon scattering at the hetero interfaces and acoustic bulk phonons in the barrier as well as the channel are to be taken into account. These will have effects on mobility.

In concluding it could be stated that strained Si/ $\text{Si}_{1-x}\text{Ge}_x$  hetero junction HMOSFET has come to a technical level where it could be designed methodically to have high  $I_{on}$  and very low  $I_{off}$ . The transition of these devices into Si integrated circuits or by itself with its own complementary devices seem to be a near future reality.

## APPENDIX - A

### REVIEW OF CONVENTIONAL MOSFET

A brief review of MOSFET operation, its physics and associated formulas are presented in this chapter. The operation of both MOSFET and HMOSFET are similar in principle. The difference is in the fact that the carrier transport is at the surface, close to the gate oxide interface, in the case of MOSFET whereas the carrier transport is at a submerged undoped quantum well  $\text{Si}_{1-x}\text{Ge}_x$  layer in the case of HMOSFET. For the HMOSFET formulas used to find the hole density distribution, potential distribution, on current and off current are similar to that of a MOSFET. These will be used and/or compared while deriving corresponding derivations for the p-HMOSFET in chapter 4. In both cases one-dimensional Poisson equation is solved along the midline of a long channel device to arrive at the parameters of interest. However in the case of HMOSFET the procedure is more involved due to the many layers of different properties involved.

In CMOS systems the p-MOSFET has hole mobility less than half of electrons in n-MOSFET. Hence p-MOSFET has to be made more than double the size of n-MOSFET. This takes up more chip area and increases parasitics that affect the circuit performance. Realizing this problem design of a p-HMOSFET having higher speed of operation is defined as the aim of the project here. Hence a p-MOSFET will be discussed in this chapter. This makes the comparisons and derivations in the case of p-HMOSFET much easier.

A review of the essential physics is given in section 2.1. Energy band diagrams and the different modes of operation of a MOSFET are given in section 2.2. In section 2.3 one-dimensional Poisson equation solutions are discussed and surface potential and depletion width are derived. Then the gradual channel approximation and chart sheet model are used to find threshold voltage, on current and off currents.

### **A.1 Basic semiconductor Physics**

Silicon has a completely filled valence band and a completely empty conduction band at absolute zero temperature. At room temperature some of the electrons from valence band are excited to conduction band leaving behind holes in the valence band. Since there are empty higher states available for electrons in the conduction band and empty lower states available for holes in the valence band both can be accelerated by applying an electric field. The probability of occupation of an energy level by an electron or hole is given by Fermi-Dirac distribution,

$$f(E) = \frac{1}{1 + e^{(E-E_f)/kT}} \quad (10.1)$$

where  $k$  is the Boltzmann constant and  $E_f$  is the Fermi energy level. Fermi level is defined such that at absolute zero all levels below  $E_f$  are filled with electrons and all levels above are empty. So the Fermi level has an abrupt transition at absolute zero. Above absolute zero and particularly at room temperature some of the electrons are excited to levels above  $E_f$  leaving the same number of levels empty (holes) below  $E_f$ . So the distribution function  $f(E)$  makes a smooth transition from unity to zero above  $E_f$ , being 0.5 at  $E_f$ . The

width of the transition region is a function of the thermal energy. Further if the energy level  $E$  is located a few thermal voltage ( $V_{th}$ ) away from  $E_f$  the distribution function can be approximated to Boltzman distribution.

$$f(E) \approx e^{-(E-E_f)/kT} \quad (10.2)$$

$$f(E) = 1 - e^{-(E_f-E)/kT} \quad (10.3)$$

for  $E > E_f$  and  $E < E_f$  respectively. These distributions are sketched in Figure 10-1. It can be seen that for  $E$  2 – 3 times  $V_{th}$  away from  $E_f$  Boltzman approximation is quite close to Fermi distribution.

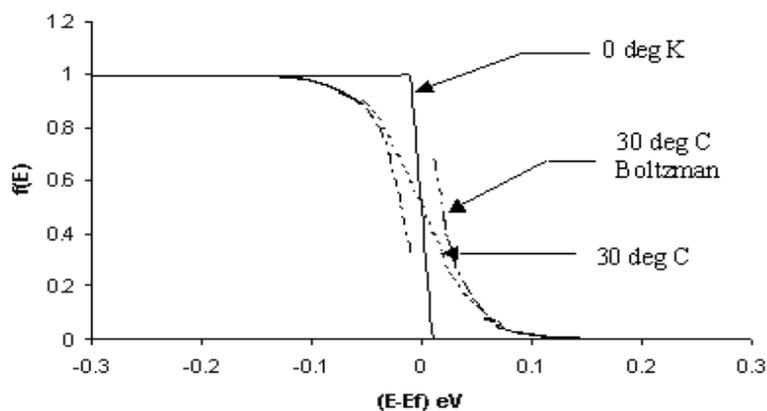


Figure 10-1. Fermi distribution at  $0^\circ$  K (solid) and at room temperature  $30^\circ$  C (dash). The chain line is the Boltzman distribution.

Equation( 10.2) gives the probability of finding an electron in the conduction band and equation ( 10.3 ) gives the probability of not finding an electron in the conduction band or probability of finding a hole in the valence band. Based on these equations the electron

concentration in the conduction band and the hole concentration in the valence band respectively are given by,

$$n = N_C e^{-(E_C - E_f)/kT} \quad (10.4)$$

$$p = N_V e^{-(E_f - E_V)/kT} \quad (10.5)$$

where  $N_C$  and  $N_V$  are the effective density of states of conduction and valence bands respectively. The intrinsic or undoped concentration of electrons and holes  $n_i$  and  $p_i$  can be found if Fermi level  $E_f$  is substituted with its intrinsic value,

$$E_i = E_f = \frac{E_C + E_V}{2} - \frac{kT}{2} \ln\left(\frac{N_C}{N_V}\right) \quad (10.6)$$

This gives  $n_i = n = p$  as,

$$n_i = \sqrt{N_C N_V} e^{-E_g/2kT} \quad (10.7)$$

From equation (10.4), equation (10.5) and Figure 10-2 electron and hole concentration can be written as,

$$n = n_i e^{(E_f - E_i)/kT} = n_i e^{q(\phi_i - \phi_f)/kT} \quad (10.8)$$

$$p = n_i e^{(E_i - E_f)/kT} = n_i e^{q(\phi_f - \phi_i)/kT} \quad (10.9)$$

This allows electron and hole concentrations to be written in terms of local potential.

Si is doped with donor  $N_d$  (e.g.: Phosphorous) to make n-type or acceptor  $N_a$  (eg. Boron) to make p-type At room temperature it is reasonable to assume that all donors and acceptors are ionized for the common dopants used. This gives  $n = N_d$  for n-type and  $p = N_a$  for p-type. These are majority carriers. The minority carriers are  $p$  in n-type and  $n$  in p-type. These are given by,

$$p = \frac{n_i^2}{N_d} \quad (10.10)$$

$$n = \frac{n_i^2}{N_a} \quad (10.11)$$

where  $n_i$  is the intrinsic carrier concentration, which is approximately  $1.4 \times 10^{10}/\text{cm}^3$  at room temperature. Figure 10-2 shows the band diagram for n and p type doped silicon.

The doping parameter in figure is given by,

$$\Phi_b = \frac{kT}{q} \ln\left(\frac{N_A}{n_i}\right) \quad (10.12)$$

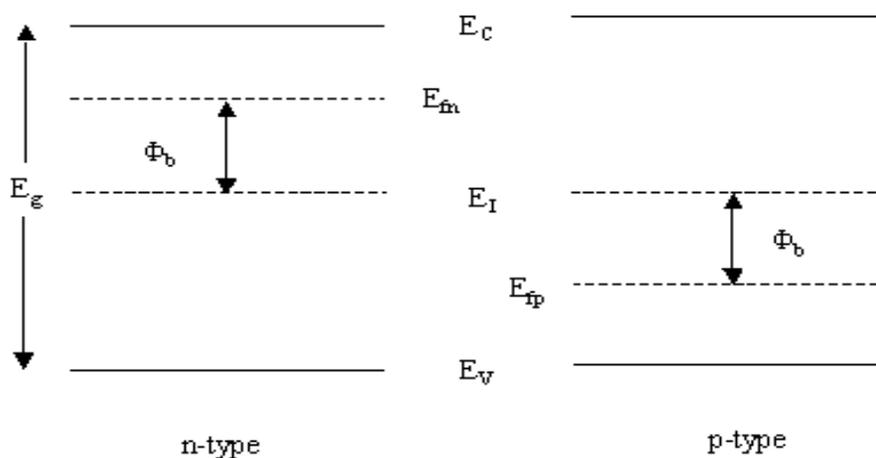


Figure 10-2. Band diagram of n and p doped silicon.  $E_g = 1.12$  eV is the band gap,  $E_C$  and  $E_V$  are conduction and valence band respectively,  $E_{fn}$  and  $E_{fp}$  are the electron quasi Fermi potential and hole quasi Fermi potential respectively and  $\Phi_b$  is the doping parameter

## A.2 Structure and operation of MOSFET

Contrary to the text book way of going through the theory and then discussing a general device we will introduce the specific device structure used in this project first and

then explain how it works and analyze it. The idea behind this is that it will suffice if the particular MOSFET structure used here, as a standard comparison device, is understood. Other variations and choices of MOSFET structures are irrelevant here. A sketch of the p-MOSFET is given in Figure 10-3. It has a n-type substrate as usual and p+ source and drain diffusions. A n+ polysilicon gate is chosen to get a threshold voltage  $V_T$  around  $-0.5$  V. It was discussed in chapter 1 that the p-HMOSFET design here is for low voltage applications with a maximum supply voltage of  $-1.5$  V. The choice of p+ polysilicon will

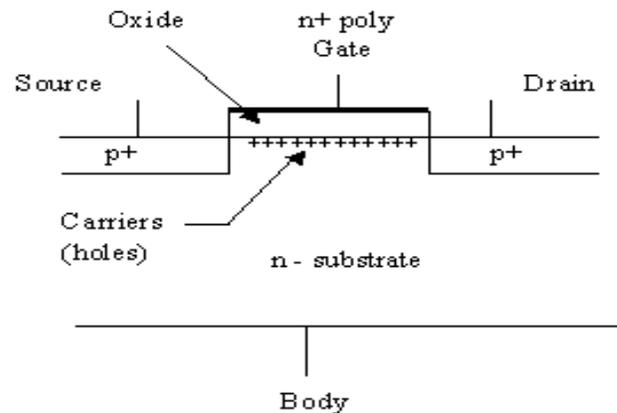


Figure 10-3. The p-MOSFET structure used in this project as a comparison reference device. It has a n+ polysilicon gate.

give very low  $V_T$  and more sub threshold leakage. As will be seen in later chapters the very low sub threshold current is one of the attractive point of the p-HMOSFET. The channel length  $L_g$  and the substrate doping  $N_{sub}$  are the only parameters varied during the simulations to set the channel length and  $I_{on}$  equal to the p-HMOSFET being simulated

### A.2.1. Formation of the p-MOSFET

Figure 10-4. shows the band gap details of the n+ polysilicon gate, SiO<sub>2</sub> gate oxide and silicon substrate before forming metallurgical contacts. The Fermi level of polysilicon is at conduction band E<sub>C</sub> and is higher than that in silicon. On metallurgically combining the three the silicon bands bend down by q (φ<sub>Si</sub> - φ<sub>G</sub>) eV. So to make the bands flat again a potential φ<sub>G</sub> - φ<sub>Si</sub> need to be applied. This is termed flat band voltage,

$$V_{fb} = \phi_G - \phi_{Si} = \chi_{Si} - \Phi_{Si} = -\left(\frac{E_g}{2} - \phi_b\right) \quad (10.13)$$

where φ<sub>b</sub> is the doping parameter described earlier and E<sub>g</sub> is the silicon band gap. When a negative voltage equal to V<sub>fb</sub> is applied to the gate the bands of polysilicon gate and silicon substrate align. This condition is shown in Figure 10-6. It is to be noted here that in the band diagrams the energy is positive going upwards but the potential is positive going downwards. Also note that in this analysis the oxide interface charge and the polysilicon depletion are neglected.

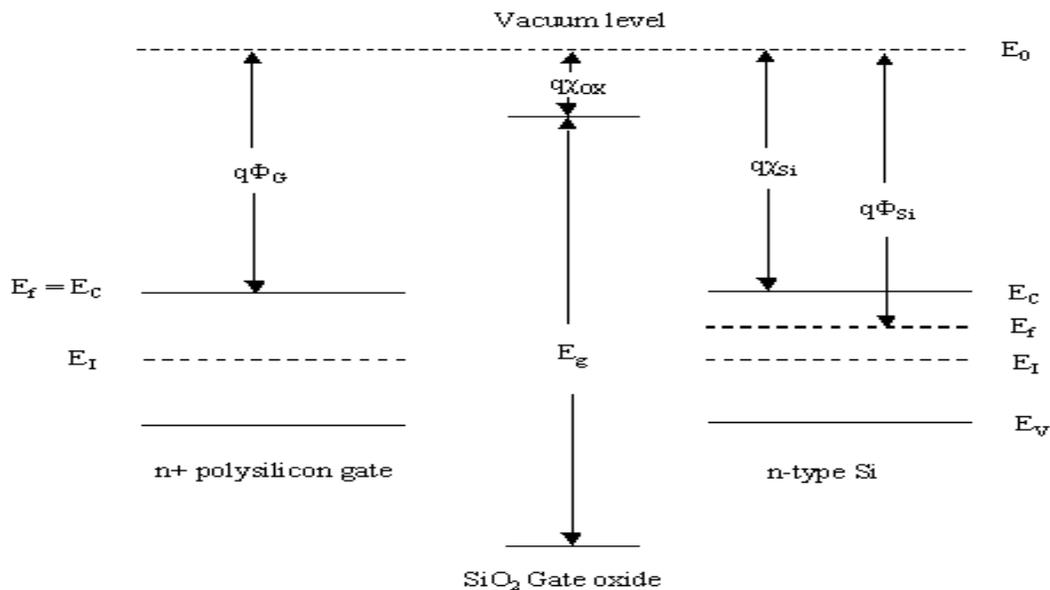


Figure 10-4. Band structure of n+ polysilicon gate, SiO<sub>2</sub> gate oxide and silicon substrate prior to forming device by metallurgically joining them.  $\phi_G$  and  $\phi_{Si}$  are work functions of n+ polysilicon and silicon substrate respectively.  $\chi_{OX}$  and  $\chi_{Si}$  are the electron affinities of oxide and silicon respectively.  $\phi_G = \phi_{Si} = 4.05$  eV,  $\chi_{OX} = 0.95$  eV and  $\chi_{Si} = 4.05$  eV

### A.2.2. Threshold Voltage

After reaching the flat band condition if the gate voltage is made more negative the bands will bend upwards and the surface potential  $\phi_s$  will become negative. This will cause accumulation of holes at the surface. When the surface potential  $\phi_s$  becomes  $-\phi_b$  the surface concentration of holes will become equal to the electron concentration in the n-type substrate. This condition is termed inversion. The gate voltage required to reach this condition is termed the threshold voltage  $V_T$ . Figure 10-7 shows this condition. It could be seen that the surface is now at a potential of  $-q\phi_b$ . The oxide has a field within and this causes a potential drop of  $-\phi_{OX}$ . From the surface up to the flat band the mobile carriers

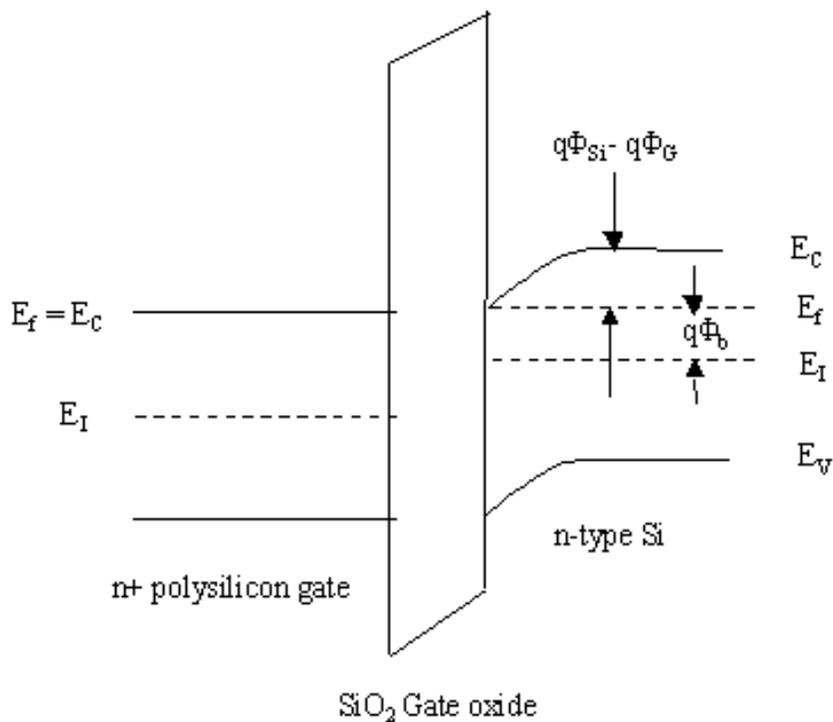


Figure 10-5. Thermal equilibrium band diagram of polysilicon, oxide and silicon after forming the p-MOSFET.  $V_{gs} = 0V$ . Note that the silicon conduction band is bent down by  $qV_{fb}$

are depleted. This depth is now occupied by the positive immobile ions and is termed depletion depth  $W_D$ . The application of negative gate voltage causes depletion to a depth of  $W_D$  and thin layer of accumulation of holes near the surface. The gate now has a negative charge equal to the sum of both these charges. Since the oxide has no charges in it the field is constant and is given by  $E_{OX} = \phi_{OX}/t_{ox}$  where  $t_{OX}$  is the oxide thickness.

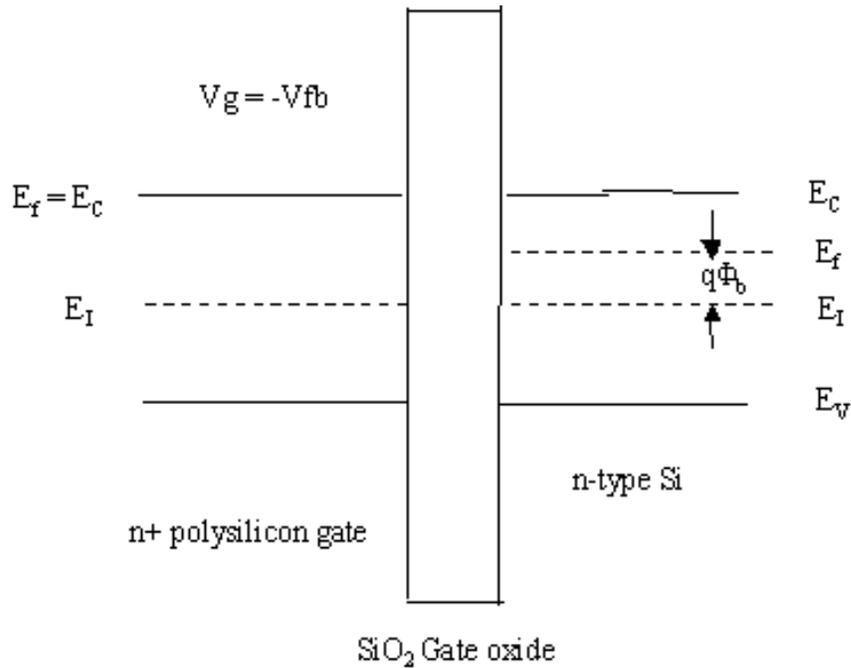


Figure 10-6. Flat band condition of the p-MOSFET with  $V_{gs} = -V_{fb}$

$$Q_G = -q(N_I + N_D) = -q(N_I + N_d W_D) = \epsilon_{ox} E_{ox} = \epsilon_{si} E_s = C_{ox} (V_G - \phi_s) \quad (10.14)$$

where  $N_d$  is the donor concentration. At the onset of inversion when  $V_{gs} = V_T$  the accumulated charge  $N_I$  is negligible compared to the depletion charge. Also the surface potential  $\phi_s = -2\phi_b$ . Hence including the flat band voltage the threshold voltage can be found to be,

$$V_T = 2\phi_b + V_{fb} - \frac{qN_d W_D}{C_{ox}} \quad (10.15)$$

### A.2.3. Depletion Width

The depletion width referred to in the earlier section need to be computed to find the value of  $V_T$ . A generalized derivation taking into account the mobile carriers at the edge of the depletion width is given below. The edge of the depletion width is not well defined

at room temperature since holes with energy greater than the thermal energy  $kT/q$  can penetrate into the depletion region. In the derivation the substrate is assumed to have both donors  $N_d$  and acceptors  $N_a$ . Two parameters  $\beta$  and Debye length  $L_B$  are defined first and then the Poisson equation is solved in the silicon substrate for the generalized case. Refer to Figure 10-7 for the co-ordinate detail.

$$\beta = \frac{q}{kT} \quad (10.16)$$

$$L_B = \sqrt{\frac{\epsilon}{q\beta N_d}} \quad (10.17)$$

$$\frac{d^2\phi}{dx^2} = -\frac{q}{\epsilon_{Si}} [p(x) - n(x) + N_d(x) - N_a(x)] \quad (10.18)$$

Assuming full ionization and using equations ( 10.10) and ( 10.11),

$$\frac{d^2\phi}{dx^2} = -\frac{q}{\epsilon_{Si}} \left[ N_a (e^{-q\phi/kT} - 1) - \frac{n_i^2}{N_a} (e^{q\phi/kT} - 1) \right] \quad (10.19)$$

Multiply by  $\beta$ , take  $\beta$  in the exponents and take  $N_a$  outside,

$$\frac{d^2(\beta\phi)}{dx^2} = \frac{1}{L_B^2} \left[ \left( \frac{n_i}{N_a} \right)^2 (e^{\beta\phi} - 1) - (e^{-\beta\phi} - 1) \right] \quad (10.20)$$

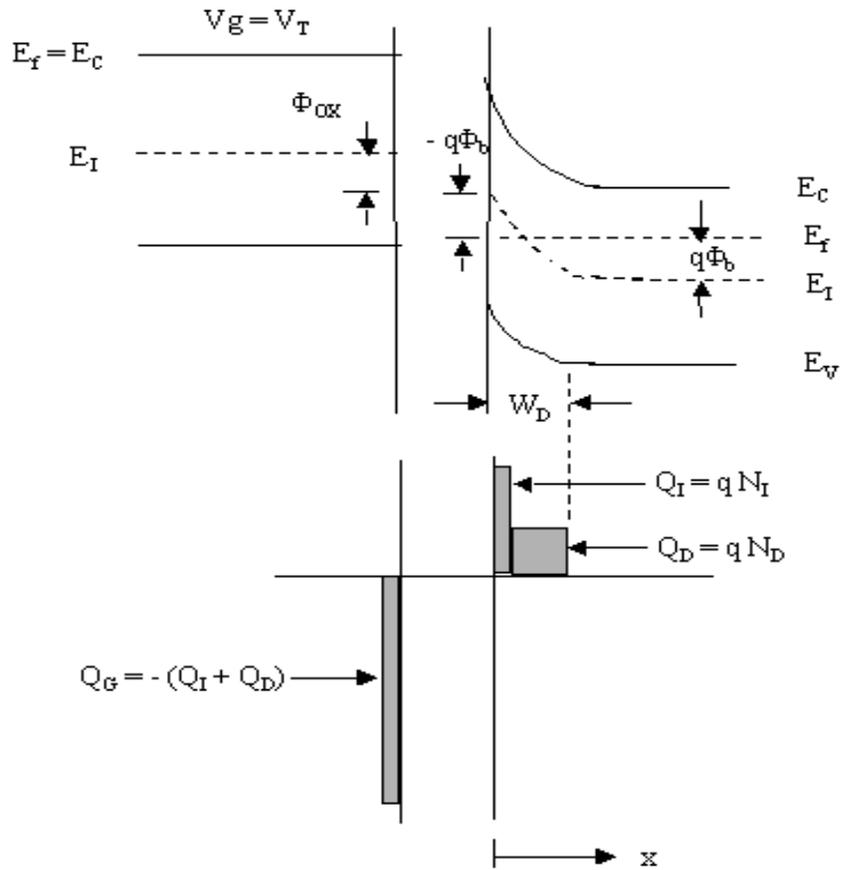


Figure 10-7. Band diagram at  $V_{gs} = V_T$  for the p-MOSFET (top). Charge accumulation at  $V_{gs} = V_T$  (bottom)

Multiply by  $d\beta\phi/dx$  and integrate from 0 to  $d\phi/dx$ ,

$$\frac{\beta_2}{2} \left( \frac{d\phi}{dx} \right)^2 = \frac{1}{L_B^2} \left[ \left( \frac{n_i}{N_a} \right)^2 (e^{\beta\phi} - \beta\phi - 1) + (e^{-\beta\phi} + \beta\phi - 1) \right] \quad (10.21)$$

Noting  $E(x) = d\phi/dx$ ,

$$E(x) = \frac{\sqrt{2}}{\beta L_B} \left[ (e^{-\beta\phi} + \beta\phi - 1) + \left( \frac{n_i}{N_a} \right)^2 (e^{\beta\phi} - \beta\phi - 1) \right]^{1/2} \quad (10.22)$$

For surface potential  $\phi_s$  between  $\phi_b$  and  $2\phi_b$   $n_i/N_a \ll 1$  and  $e^{-\beta\phi}$  is negligible,

$$E(x) = -\frac{d\phi}{dx} = \frac{\sqrt{2}}{\beta L_B} (\beta\phi - 1)^{1/2} \quad (10.23)$$

Integrating,

$$\int_0^\phi \frac{\beta}{(\beta\phi - 1)^{1/2}} = \frac{-\sqrt{2}}{L_B} \int_{W_D}^x dx \quad (10.24)$$

$$2\sqrt{\beta\phi - 1} = \frac{\sqrt{2}}{L_B} (W_D - x) \quad (10.25)$$

At  $x = 0$   $\phi = \phi_S$  giving,

$$W_D = \sqrt{2} L_B \sqrt{\beta\phi_S - 1} \quad (10.26)$$

Substituting back into equation (10.25),

$$\phi(x) = \frac{1}{\beta} \left[ (\beta\phi_S - 1) \left(1 - \frac{x}{W_D}\right) + 1 \right] \quad (10.27)$$

In the approximation  $\sqrt{\beta\phi_S - 1} \approx \sqrt{\beta\phi_S}$

$$W_D = \sqrt{\frac{2\epsilon_{Si}\phi_S}{qN_a}} \quad (10.28)$$

$$\phi(x) = \phi_S \left(1 - \frac{x}{W_D}\right)^{1/2} \quad (10.29)$$

#### A.2.4. Drain Saturation Current, $I_{Dsat}$

The charge sheet model assumes that the inversion charges are located like a sheet of charges at the silicon surface next to the  $\text{SiO}_2$  interface [188]. Further the depletion approximation assumes that after the onset of inversion the surface potential do not change and stays pinned at  $\phi_S = 2\phi_b + V(y)$  where  $\phi_b$  is the doping parameter defined earlier. For a p-MOSFET the depletion charge  $Q_D$ , total charge in silicon  $Q_S$  and the inversion charge  $Q_I$  are given by,

$$Q_D = qN_d W_D = \sqrt{2\epsilon q N_d (2\phi_b + V)} \quad (10.30)$$

$$Q_S = C_{OX} (V_g - V_{fb} - \phi_s) = C_{OX} (V_g - V_{fb} - 2\phi_b - V) \quad (10.31)$$

$$Q_i = Q_S - Q_D = C_{OX} (V_g - V_{fb} - 2\phi_b - V) - \sqrt{2\epsilon_{Si} q N_d (2\phi_b + V)} \quad (10.32)$$

$I_{ds}$  is given by the following equation,

$$I_{ds} = \mu_{eff} \frac{W}{L} \int_0^{V_{ds}} [-Q_i(V)] dV \quad (10.33)$$

Substituting equation (10.32) into (10.33) a good approximation to the drain current is given by,

$$I_{ds} = \mu_{eff} C_{OX} \frac{W}{L} [(V_g - V_T)V_{ds} - \frac{m}{2} V_{ds}^2] \quad (10.34)$$

with  $m = 1 + \frac{3t_{OX}}{W_D}$ . The equation for  $I_{ds}$  is a parabolic equation. Finding its derivative and

equating to zero  $V_{ds} = V_{dsat} = (V_g - V_T) / m$ . Using this the saturation value of  $I_{ds}$  is given by,

$$I_{dsat} = \mu_{eff} C_{OX} \frac{W(V_g - V_T)^2}{2mL} \quad (10.35)$$

Some of the details are omitted here. Only the final form of the saturation drain current form is of importance since the same is to be compared with similar form in the case of p-HMOSFET in chapter 5.

### A.2.5. Sub-threshold Slope Parameter (S)

Sub threshold slope parameter S is a metric used to represent the off state currents or how fast the device turns off with lowering  $V_{gs}$ . At lower  $V_{gs}$  the diffusion component of the current is much higher than its drift component. Hence  $I_{ds}$  is independent of  $V_{ds}$  for

$V_{ds} > 2 - 3 kT/q$ . This parameter is defined as the inverse of the slope of the  $V_{gs} - I_{ds}$  curve below  $V_T$ . A lower value of  $S$  indicates a good turn off characteristic. Typically  $S$  is 70 – 100.

$$S = \left[ \frac{d(\log(I_{ds}))}{dV_g} \right]^{-1} = 2.3 \frac{mkT}{q} \quad (10.36)$$

### A.3. Short Channel Effects (SCE)

Channel length reduction of MOSFET is desirable to increase speed and packing density. As the channel length is reduced below a few hundred nanometers the characteristics of the device changes in many ways. These effects together are termed under short channel effects (SCE). One of the prominent effects is the reduction of the threshold voltage  $V_T$ . A drop in  $V_T$  indicates poor sub threshold slope or more off state leakage. It is particularly true when the channel length  $L_g$  dropped below 100 nm. These devices have such high leakage currents compared to longer channel length (couple of hundred nm) devices that they do not have an off state.

Reduction in  $V_T$  happens due to the fact that the potential contours in these short channel length devices are no more parallel to the interface but are two-dimensional. This happens due to the encroachment of the source and drain depletion layers, which becomes significant compared to the channel length. A significant amount of charge in the silicon substrate under the gate is now controlled by the source and drain. This means the gate has control only over a portion of the charge. In the case of a p-MOSFET the threshold voltage given by equation (10.15) can be re-written as,

$$V_T = V_{fb} - 2\phi_B - \frac{Q_D}{C_{ox}} \quad (10.37)$$

This equation shows that any reduction of depletion charge reduces the threshold voltage. Another important SCE is drain induced barrier lowering (DIBL). This leads to degraded sub threshold parameter or more off state leakage. This is again due to the penetration of source and drain depletion regions into the channel region. In the long channel case the source – channel junction, which is a p-n junction, has a barrier over which the holes need to be injected from the source. But this barrier is reduced by the above phenomenon that it becomes easier for holes to go over the barrier from source to drain.

Velocity saturation,  $v_{sat}$ , is another effect that can happen in short channel MOSFET due to high fields. Both carriers electrons and holes have a bulk saturation velocity beyond which they cannot be accelerated. In silicon  $v_{sat}$  for holes is about  $8.37 \times 10^6$  cm/s [4]. Due to this  $I_{ds}$  saturates at a  $V_{ds} = V_{Dsat}$  lower than the value given earlier.  $V_{Dsat}$  is given by,

$$V_{Dsat} = (V_{gs} - V_T)/m.$$

Body-effect coefficient  $m = 1 + \frac{3t_{ox}}{W_{dm}}$ ,  $W_{dm}$  is the maximum depletion width. Typically  $m$  lies between 1.1 and 1.4 [4]

## APPENDIX – B

### EXCEL SPREAD SHEET – BAND BENDING WITH P-DELTA

#### B.1 Input:

Bandbending at bulk (V)	Psi_N	0.64
Bandgap in bulk (V)	Psi_G	1.12
D of States in bulk (/cm <sup>3</sup> )	Nc_N	2.80E+19
VB density of states in bulk	Nv_N	2.65E+19
D of states in well (/cm <sup>3</sup> )	Nv_W	1.48E+19
N-doping in bulk (/cm <sup>3</sup> )	N_D	4.00E+18
Dielectric constant in bulk	k_Si	11.7
Dielectric constant in well	k_W	13.05
Dielectric constant oxide	k_OX	3.90E+00
Permittivity empty space (F/cm)	e_0	8.85E-14
Electron charge ( C)	q	1.60E-19
Boltzman k (eV/K)	k_B	8.62E-05
Temperature (K)	T_K	290
p+ delta layer doping/cm <sup>2</sup>	D_p	<b>5.64E+12</b>
Well height (V)	Psi_H	0.4
Well width (nm)	w_W	3
Buffer width (nm)	w_B	3
Cap width (nm)	w_C	3
Oxide width (nm)	t_OX	3

#### B.2 Calculated:

P+delta layer location (nm)	x_p	9
Thermal votlage	V_th	0.02499046
Bulk Debye length (cm)	L_D	2.00987E-07
Fermi level relative to VB (V)	Psi_F	1.07137081
Depletion width N-layer (cm)	W_N	1.41006E-06
Depletion width N-layer (nm)	W_N_nm	14.10059308
Carriers in N-layer (/cm <sup>2</sup> )	Q_N	<b>5.6402E+12</b>
Field at n- bulk (V/cm)	F_N	8.7232E+05
Field at buffer side of delta layer	F_D	4.0047E-08
Potential at well side of buffer (V)	Psi_B	0.64

Potential at buffer side of well (V)	Psi_W	1.04
Poisson coeff/cm <sup>2</sup>	a	1.97569E-05
Cap side of well (thermal units)	u_C	42.93121238
Buffer side of well (thermal units)	u_B	41.61588062
Field at buffer side of well (thermal units)	F_B	1.4367E-06
Well width (nm) (Calculated)	w_W_0	3.000000001
Error in well width (nm)	Error	-5.41214E-10
Potential at cap side of well	Psi_C0	1.072870746
Average field in well (V/cm)		1.0957E+05
Field at cap side of well (V/cm)	F_C	2.8228E+05
Field at buffer side of well (V/cm)		3.5904E-08
Hole density /cm <sup>2</sup>	N_h	<b>2.0358E+12</b>

### B.3 Computational panel:

	Well width	x_W (nm)	
	<b>Target</b>	<b>5</b>	← Input
Bulk doping	<b>n_B3</b>	<b>2.00E+18</b>	← Input
Number of ions in delta layer/cm <sup>2</sup>	Q_d_set3	1.30E+12	
Potential drop in bulk n-layer (kT/q)	u_nB_on3	26.02815027	
Potential drop in bulk n-layer (kT/q)	<b>u_nB_off3</b>	<b>3</b>	← Input
Well height (kT/q)	<b>u_W_set3</b>	<b>10</b>	← Input
potential at front of well (on)	u_on3	41.7766699	
potential at front of well (off)	u_off3	14.4354091	

#### B.3.1 Computed values:

	well width (nm)	Gate V	Q_nB (/cm <sup>2</sup> )	p/cm <sup>2</sup>	S mV/dec
ON	5.00032	1.51E+00	2.897E+12	1.06E+12	1.33E+02
OFF	5	2.04E-01	8.188E+11	5.96E+00	1.78E+11

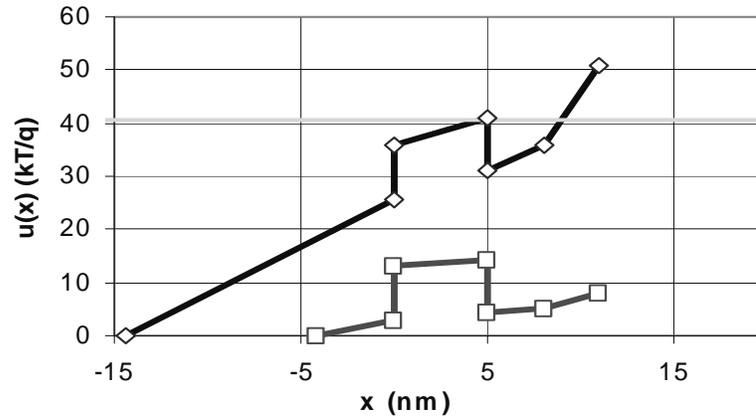


Figure B- 1. ON and OFF state potentials from substrate to gate

## B.4 Computational Formulas:

### B.4.1 Formulas for basic computations on inputs

P+delta layer location (nm)	$x_p$	$=w_C+w_W+w_B$
Thermal voltage	$V_{th}$	$=k_B T_K$
Bulk Debye length (cm)	$L_D$	$=1/\text{SQRT}(q N_D / (k_{Si} e_0 V_{th}))$
Fermi level relative to VB (V)	$\Psi_{F0}$	$=\Psi_G + V_{th} \ln(N_D / N_c)$
Depletion width N-layer (cm)	$W_N$	$=L_D \text{SQRT}(2(\Psi_N / V_{th} - 1))$
Depletion width N-layer (nm)	$W_{N\_nm}$	$=W_N * 10000000$
Carriers in N-layer (#/cm <sup>2</sup> )	$Q_N$	$=N_D * W_N$
Field at n- bulk (V/cm)	$F_N$	$=q * Q_N / (k_{Si} e_0)$
Field at buffer side of delta layer	$F_D$	$=F_N - q * D_p / (k_{Si} e_0)$
Potential at well side of buffer (V)	$\Psi_B$	$=\Psi_N + F_D * w_B * 0.0000001$
Potential at buffer side of well (V)	$\Psi_W$	$=\Psi_B + \Psi_H$
Poisson coeff/cm <sup>2</sup>	$a$	$=(k_{Si} / k_W) * (N_V * W + N_c * N_D / N_D^2) * (1 / L_D^2) * \text{EXP}(-\Psi_G / V_{th})$
Cap side of well (thermal units)	$u_C$	42.9312123783872
Buffer side of well (thermal units)	$u_B$	$=\Psi_W / V_{th}$
Field at buffer side of well (thermal unit)	$F_B$	$=(k_{Si} / k_W) * F_D / V_{th}$
Well width (nm) (Calculated)	$w_{W\_0}$	$=w_{Well}(u_C, u_B, F_B, a) * 10000000$
Error in well width (nm)	Error	$=w_W - w_{W\_0}$
Potential at cap side of well	$\Psi_{C0}$	$=u_C * V_{th}$
Average field in well (V/cm)		$=(\Psi_{C0} - \Psi_W) / (w_W * 0.0000001)$
Field at cap side of well (V/cm)	$F_C$	$=\text{SQRT}(F_B^2 + 2 * a * \text{EXP}(u_B) * (\text{EXP}(u_C - u_B) - 1)) * V_{th}$
Field at buffer side of well (V/cm)		$=F_B * V_{th}$
Hole density /cm <sup>2</sup>	$N_h$	$=k_{Si} e_0 * (k_W / k_{Si} * F_C - F_D) / (q)$

```
Function w_Well(u_C As Double, u_B As Double, F_B As Double, a As Double) As
```

```
Double
```

```
Dim z_C As Double
```

```
Dim z_B As Double
```

```
Dim B As Double
```

```
Dim x_B As Double
```

```
Dim x_C As Double
```

```
Dim ratio1 As Double
```

```
Dim ratio2 As Double
```

```
Dim diff As Double
```

```
Dim delta As Double
```

```
Dim mult As Double
```

```
delta = u_C - u_B
```

```
If (delta < 0.01) Then
```

```
    mult = delta + delta ^ 2 / 2 + delta ^ 3 / 6 + delta ^ 4 / 24
```

```
Else
```

```
    mult = (Exp(delta) - 1)
```

```
End If
```

```
z_C = Sqr(F_B * F_B + 2 * a * Exp(u_B) * mult)
```

```
z_B = F_B
```

```
B = 2 * a * Exp(u_B) - F_B * F_B
```

If (B > 0) Then

$$B = \text{Sqr}(B)$$

$$w_{\text{Well}} = (2 / (B)) * (\text{Atn}(z_{\text{C}} / B) - \text{Atn}(z_{\text{B}} / B))$$

ElseIf (B < 0) Then

$$B = \text{Sqr}(-B)$$

$$x_{\text{B}} = (z_{\text{B}} / B)$$

$$x_{\text{C}} = (z_{\text{C}} / B)$$

$$\text{ratio1} = (1 + x_{\text{B}}) / (1 + x_{\text{C}})$$

$$\text{ratio2} = (1 - x_{\text{C}}) / (1 - x_{\text{B}})$$

$$\text{diff} = 0.5 * \text{Log}(\text{ratio1}) + 0.5 * \text{Log}(\text{ratio2})$$

' arctan =  $\text{Log}((1 + X) / (1 - X)) / 2$

$$w_{\text{Well}} = 2 * (1 / (B)) * \text{diff}$$

Else

$$w_{\text{Well}} = 2 * \text{Sqr}(2 / a) * \text{Exp}(-u_{\text{B}} / 2) * (1 - \text{Exp}(-(u_{\text{C}} - u_{\text{B}}) / 2))$$

End If

End Function

#### **B.4.2 Formulas for device analysis:**

$$\text{(ON) well width} = x(u_{\text{on3}}, u_{\text{B}}, F_{\text{B}}, L_{\text{B}})$$

$$\text{(OFF) well width} = x L_{\text{B}} 107$$

$$\text{(ON) gate V} = \text{cap drop} + \text{ox drop} + u_{\text{on3}} - u_{\text{W}}$$

$$\text{(OFF) gate V} = \text{cap drop} + \text{ox drop} + u_{\text{off3}} - u_{\text{W}}$$

$Q_{\text{nB}}$  given under 'calculated' above

(ON)  $p = p_B L_B^2 \text{ABS}(P(u_{on3}, u_B, F_B, L_B))$

(OFF)  $p = p_B L_B^2 \text{ABS}(P(u_{off3}, u_B, F_B, l_B))$

$S = 1000 V_{gs} / \log(p_{on}/p_{off})$

Function  $x(u, u_B, F_B, L_B)$

'Poisson's equation in well

Dim a As Double

Dim xi As Double

Dim eps As Double

$a = F_B * L_B$

$a = a * a$

$a = a - 1$

If (a = 0) Then

$xi = 2 * (\text{Exp}((u - u_B) / 2) - 1)$

ElseIf (a > 0) Then

$a = \text{Sqr}(a)$

$eps = \text{Exp}(u - u_B) / (a * a)$

If (eps < 0.0000001) Then

$xi = (1 / a) * (\text{Log}(a * eps / 2) - \text{Log}(\text{Sqr}(\text{Exp}(u - u_B) + a * a) + a) + \text{Log}(\text{Sqr}(a * a + 1) + a) - \text{Log}(\text{Sqr}(a * a + 1) - a))$

Else

$xi = (1 / a) * (\text{Log}(\text{Sqr}(\text{Exp}(u - u_B) + a * a) - a) - \text{Log}(\text{Sqr}(\text{Exp}(u - u_B) + a * a) + a) + \text{Log}(\text{Sqr}(a * a + 1) + a) - \text{Log}(\text{Sqr}(a * a + 1) - a))$

```

    End If

Else

    a = -a

    xi = (2 / a) * (Atn((Sqr(Exp(u - u_B) - a * a) / a)) - Atn(Sqr(1 - a * a) / a))

End If

If (u > u_B) Then

    x = xi

Else

    x = -xi

End If

End Function

Function P(u, u_B, F_B, L_B)

'Gauss's law for fields

Dim eps As Double

eps = (Exp(u - u_B) - 1) / (F_B * L_B) ^ 2

If (Abs(eps) < 0.000001) Then

    P = Abs(F_B) / 2 * (eps - eps * eps / 2 + eps * eps * eps / 4)

Else

    P = -F(u, u_B, F_B, L_B) + F_B

End If

End Function

```

## APPENDIX - C

### ONE DIMENSIONAL POISSON EQUATION SOLUTION

Poisson equation is solved starting from the substrate layer which contains the depletion layer completely.

#### Symbols:

$\Phi_x$	:	Potential at any point along the depth
$\Phi_H$	:	Potential at the top of SiGe layer
$\Phi_L$	:	Potential at the bottom of SiGe layer
$\Phi_{TH}$	:	Potential at the top of SiGe layer at threshold
$H(\Phi)$		Hole contribution to the electric field
$H(\Phi_H)$		Total hole contribution to the electric field, at the top of SiGe layer
$\Delta E_v$		Valence band discontinuity between SiGe channel and the Si buffer
$X_d$	:	Depletion thickness in the APT layer
$X_{dmax}$	:	Max depletion thickness in the APT layer
$X_{ox}$	:	Oxide thickness
$N_d$		n-doping density in the APT layer
$N_a$		p-doping density of the p-delta layer ( $N_p$ )
$X_p$		p-delta layer thickness
$\epsilon_{Si}$		Dielectric constant of silicon
$\epsilon_{SiGe}$		Dielectric constant of $Si_xGe_{1-x}$
$\epsilon_{ox}$		Dielectric constant of oxide
$\Phi_{ms}$		Metal semiconductor work function difference
$R$		Ratio of permittivity, $\epsilon_{Si}/\epsilon_{SiGe}$
$V_T$		Threshold voltage

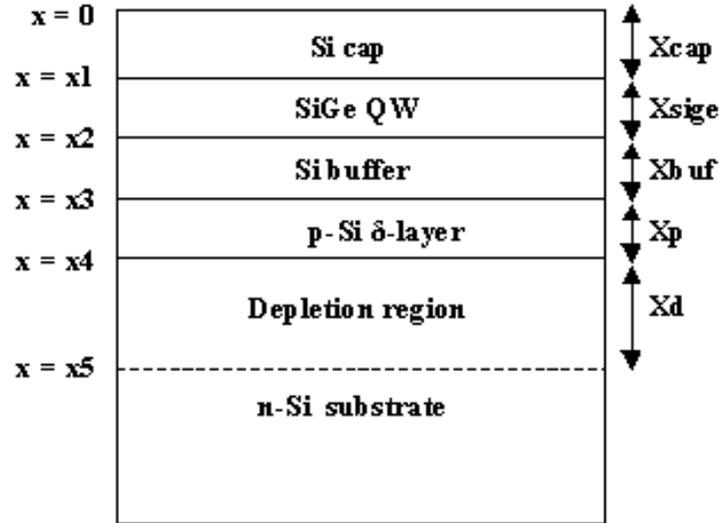


Figure C. 1. p-HMOSFET sketch giving details of layers and dimensions

(C1) Within  $X_d$ , the depletion region in the substrate:

$$d^2\Phi/dx^2 = -qNd/\epsilon_{si} \quad \text{BC: } \Phi(X_5) = 0, \quad d\Phi/dx(X_5) = 0$$

$$d\Phi/dx = -(qNd/\epsilon_{si})x + C_1$$

$$= -(qNd/\epsilon_{si})x + (qNd/\epsilon_{si})X_5$$

$$\Phi = -(qNd/2\epsilon_{si})x^2 + (qNdX_5/\epsilon_{si})x + C_2$$

$$C_2 = -(qNdX_5^2/\epsilon_{si}) + (qNdX_5^2/2\epsilon_{si})$$

$$= -(qNdX_5^2/2\epsilon_{si})$$

$$\begin{aligned} \Phi &= -(qNd/2\epsilon_{si})x^2 + (qNdX_5/\epsilon_{si})x - (qNd/2\epsilon_{si})(X_5)^2 \\ &= -(qNd/2\epsilon_{si})(X_5 - x)^2 \end{aligned}$$

(C. 1)

(C2) Within  $X_p$ , the p-doped layer.:

$$d^2\Phi/dx^2 = qNa/\epsilon_{si} \quad \text{BC: } \Phi \text{ and } d\Phi/dx \text{ continuous at } X_4$$

$$d\Phi/dx = (qNa/\epsilon_{si})x + C_3$$

$$(qNa/\epsilon_{si})X_4 + C_3 = -(qNd/\epsilon_{si})X_4 + (qNd/\epsilon_{si})X_5$$

$$C3 = (q Nd/\epsilon_{si}) Xd - (q Na/\epsilon_{si}) X4$$

$$d\Phi/dx = (qNa/\epsilon_{si})(x-X4) + (qNd/\epsilon_{si})Xd$$

$$\Phi = (q Na/2\epsilon_{si}) x^2 + (q Nd Xd/\epsilon_{si}) x - (q Na X4/\epsilon_{si}) x + C4$$

At  $x = X4$ ,

$$\begin{aligned} (q Na/2\epsilon_{si}) (X4)^2 + (q Nd Xd/\epsilon_{si}) X4 - (q Na)/\epsilon_{si} (X4)^2 + C4 \\ = -(q Nd/2\epsilon_{si})(Xd)^2 \end{aligned}$$

$$\Phi = (q Na/2\epsilon_{si})x^2 + (q NdXd/\epsilon_{si})x - (q NaX4/\epsilon_{si})x - (q Nd/2\epsilon_{si})(Xd)^2$$

$$(q Ndxd/\epsilon_{si})X4 + (q Na/2\epsilon_{si})(X4)^2$$

$$= (q Na/2\epsilon_{si})(X4 - x)^2 + (q Nd Xd/\epsilon_{si}) x - (q Nd/2\epsilon_{si}) (Xd)^2 - (q Nd Xd/\epsilon_{si}) X4$$

$$= (qNa/2\epsilon_{si})(X4 - x)^2 + (qNdXd/\epsilon_{si})(x - X4) - (qNd/2\epsilon_{si})Xd^2 \quad (C. 2)$$

(C3) Within Xbuf:

$$d^2\Phi/dx^2 = 0 \quad \text{BC: } \Phi \text{ and } d\Phi/dx \text{ continuous at } x3$$

$$d\Phi/dx = C5$$

$$C5 = (q Na/\epsilon_{si}) X3 + (q Nd/\epsilon_{si})Xd - (q Na/\epsilon_{si})X4$$

$$= (q Nd/\epsilon_{si})Xd - (q Na/\epsilon_{si})Xp$$

$$\Phi = (q Nd Xd/\epsilon_{si}) x - (q Na Xp/\epsilon_{si}) x + C6$$

$$(q NdXd/\epsilon_{si})X3 - (q NaXp/\epsilon_{si})X3 + C6 = (q Na/2\epsilon_{si})(Xp)^2 + (q NdXd/\epsilon_{si})X3$$

$$(q Nd/2\epsilon_{si})(Xd)^2 - (q NdXd/\epsilon_{si})X4$$

$$Cbuf = C6 = \frac{qNa(Xp)^2}{2\epsilon_{si}} + \frac{qNa(Xp)(X3)}{\epsilon_{si}} - \frac{qNd(Xd)^2}{2\epsilon_{si}} - \frac{qNd(X4)Xd}{\epsilon_{si}}$$

Add and subtract  $(qNaX3^2)/2\epsilon_{si}$  and  $(qNdX4^2)/2\epsilon_{si}$

$$\begin{aligned}
C_{buf} &= C6 = \frac{qNa}{2\epsilon_{si}}(Xp + X3)^2 - \frac{qNa}{2\epsilon_{si}}X3^2 - \frac{qNd}{2\epsilon_{si}}(Xd + X4)^2 + \frac{qNd}{2\epsilon_{si}}X4^2 \\
&= \frac{qNa}{2\epsilon_{si}}X4^2 - \frac{qNa}{2\epsilon_{si}}X3^2 - \frac{qNd}{2\epsilon_{si}}X5^2 + \frac{qNd}{2\epsilon_{si}}X4^2 \\
&= \frac{qNa}{2\epsilon_{si}}(X4^2 - X3^2) - \frac{qNd}{2\epsilon_{si}}(X5^2 - X4^2) \\
&= \frac{qNaXp}{2\epsilon_{si}}(X4 + X3) - \frac{qNdXd}{2\epsilon_{si}}(X5 + X4)
\end{aligned}$$

This is like two delta layers at the center of  $X_p$  and  $X_d$  respectively.

(C. 3)

$$\Phi = \left( \frac{qNdXd}{\epsilon_{si}} - \frac{qNaXp}{\epsilon_{si}} \right) x + C_{buf}$$

(C4) Within  $X_{sig}$ :

$$\epsilon_{sig} \frac{d^2\Phi}{dx^2} = -qp(x)$$

$$p(x) = Nd \exp[(\Phi_{TH} - \Phi_x)(q/kT)]$$

$p(x)$  above is obtained by using the quasi Fermi potential definition in the bulk and taking into account the valence band discontinuity in the SiGe layer. It is explained later.

Multiply by  $\frac{d\Phi}{dx} dx$  and integrating,

$$\epsilon_{sig} \xi_x^2 / 2 = qNd(kT/q) \exp[(\Phi_{TH} - \Phi_x)(q/kT)] + C7$$

At  $x = x_2$  matching fields and noting the discontinuity in dielectric constant,

$$\xi_x^2 = \frac{2kTNd}{\epsilon_{sig}} \exp[(\Phi_{TH} - \Phi_{x_2})(q/kT)] + C7 = \frac{(qNdXd - qNaXp)^2}{\epsilon_{sig}^2}$$

$$C7 = \frac{(qNdXd - qNaXp)^2}{\epsilon_{sig}^2} - \frac{2kTNd}{\epsilon_{sig}} \exp[(\Phi_{TH} - \Phi_{x_2})(q/kT)]$$

$$\begin{aligned}\xi_x^2 &= \frac{2kTNd}{\mathcal{E}_{sige}} \exp[(\Phi_{TH} - \Phi_x)(q/kT)] + \frac{(qNdXd - qNaXp)^2}{\mathcal{E}_{sige}^2} \\ &- \frac{2kTNd}{\mathcal{E}_{sige}} \exp[(\Phi_{TH} - \Phi_{x2})(q/kT)] \\ &= \frac{(qNdXd - qNaXp)^2}{\mathcal{E}_{sige}^2} [1 + H(\Phi_x)]\end{aligned}$$

$$\xi_x = \frac{(qNdXd - qNaXp)}{\mathcal{E}_{sige}} [\sqrt{1 + H(\Phi_x)}]$$

$$H(\Phi_x) = \frac{2kTNd\mathcal{E}_{sige}}{(qNdXd - qNaXp)^2} \{ \exp[(\Phi_{TH} - \Phi_x)(q/kT)] - \exp[(\Phi_{TH} - \Phi_{x2})(q/kT)] \}$$

Further integration seems difficult. But the  $H(\Phi_x)$  term, which is small below  $V_T$  (or at  $\Phi_H < \Phi_{TH}$ ), can be neglected. The case while retaining  $H(\Phi_x)$  term when  $p(x)$  in SiGe is higher is given later.

$$\frac{d\Phi}{dx} = \frac{RqNdXd}{\mathcal{E}_{si}} - \frac{RqNaXp}{\mathcal{E}_{si}}$$

$$\Phi = \frac{RqNdXd}{\mathcal{E}_{si}} x - \frac{RqNaXp}{\mathcal{E}_{si}} x + C8$$

Applying BC at X2,

$$\frac{RqNdXd}{\mathcal{E}_{si}} X2 - \frac{RqNaXp}{\mathcal{E}_{si}} X2 + C8 = \frac{qNdXd}{\mathcal{E}_{si}} X2 - \frac{qNaXp}{\mathcal{E}_{si}} X2 + Cbuf$$

$$C8 = Csigex = \frac{qNdXd}{\mathcal{E}_{si}} x2(1-R) - \frac{qNaXp}{\mathcal{E}_{si}} x2(1-R) + Cbuf$$

$$\Phi = \frac{Rq(NdXd - NaXp)}{\mathcal{E}_{si}} x + \frac{X2(1-R)q}{\mathcal{E}_{si}} (NdXd - NaXp) + Cbuf$$

To solve for Xd all terms are inserted to give,

$$\begin{aligned}\Phi &= \frac{Rq(NdXd - NaXp)}{\mathcal{E}_{si}} x + \frac{qNdXd}{\mathcal{E}_{si}} X2(1-R) - \frac{qNaXp}{\mathcal{E}_{si}} X2(1-R) + \frac{qNa}{2\mathcal{E}_{si}} X4^2 \\ &+ \frac{qNd}{2\mathcal{E}_{si}} X4^2 - \frac{qNa}{2\mathcal{E}_{si}} X3^2 - \frac{qNd}{2\mathcal{E}_{si}} X5^2\end{aligned}$$

At top of SiGe, at x1, substituting  $X2 = X1 + Xsige$ ,  $X4 = X3 + Xp$ ,  $X5 = X4 + Xd$  and

simplifying,

$$\begin{aligned} \Phi_H = & -\frac{qNd}{2\epsilon_{si}} [Xd^2 + 2Xd(Xbuf + Xp + RXsige)] \\ & + \frac{qNa}{2\epsilon_{si}} [Xp^2 + 2Xp(Xbuf + RXsige)] \end{aligned} \quad (C.4)$$

The second term is a constant and named constant A

Solving for Xd,

$$Xd = \sqrt{\frac{2\epsilon_{si}}{qNd} (A - \Phi_H) + (Xbuf + Xp + RXsige)^2} - (Xbuf + Xp + RXsige) \quad (C.5)$$

Following the criteria for  $V_T$  similar to that in MOS, but in the SiGe layer, and including the valence band discontinuity,

$$\text{at threshold } \Phi_H = \Phi_{TH} = -\frac{2kT}{q} \ln(Nd / ni) + \Delta E_v / q$$

Substituting  $\Phi_{TH} = \Phi_H$  gives the maximum depletion depth ( $X_{dmax}$ ) in the APT layer.

Since the electric field for  $x < x_3$  is given by  $d\Phi/dx$  at  $X_3$  in Xbuf potential drop in oxide and cap layer is given by,

$$V_{ox} + V_{cap} = -q \frac{(NdX_{dmax} - NaXp)}{\epsilon_{si}} \left( \frac{X_{cap}}{\epsilon_{si}} + \frac{X_{ox}}{\epsilon_{ox}} \right), \text{ giving threshold voltage}$$

$$V_T = \Phi_{ms} + \Phi_{TH} + V_{ox} + V_{cap} \quad (C.6)$$

(C5) Within Xcap:

$$\frac{d^2\Phi}{dx^2} = 0, \text{ BC: } \Phi \text{ and } d\Phi/dx \text{ continuous at } x_1$$

Integrating and taking into account the discontinuity in dielectric constant,

$$\frac{d\Phi}{dx} = C9 = \frac{qNdXd}{\epsilon_{si}} - \frac{qNaXp}{\epsilon_{si}}$$

$$\Phi = \frac{qNdXd}{\epsilon_{si}} x - \frac{qNaXp}{\epsilon_{si}} x + C10$$

At x1,

$$\frac{qNdXd}{\epsilon_{si}} x1 - \frac{qNaXp}{\epsilon_{si}} x1 + C10 = \frac{Rq}{\epsilon_{si}} (NdXd - NaXp)x1$$

$$+ \left[ \frac{qNdXd}{\epsilon_{si}} x2(1-R) - \frac{qNaXp}{\epsilon_{si}} x2(1-R) \right]$$

$$+ \left[ \frac{qx4^2}{2\epsilon_{si}} (Na + Nd) - \frac{q}{2\epsilon_{si}} (Nax3^2 + Ndx5^2) \right]$$

The square bracket term is Csig, giving

$$\Phi = \frac{qNdXd}{\epsilon_{si}} x - \frac{qNaXp}{\epsilon_{si}} x + \frac{q}{\epsilon_{si}} x1(1-R)(NaXp - NdXd) + Csig \quad (C. 7)$$

If the contribution of holes in SiGe to the potential drop in the same layer can be neglected and the second exponent term being negligible,

$$H(\Phi_H) = \frac{2kTNd\epsilon_{sige}}{(qNdXd - qNaXp)^2} \exp[(\Phi_{TH} - \Phi_H)(q/kT)] \quad (C. 8)$$

where  $\Phi_H$  is the potential at the top of the SiGe layer.

Now equation for  $V_{gs}$  is modified with,

$$V_{OX} + V_{CAP} = \frac{-q(NdXd_{max} - NaXp)}{\epsilon_{si}} \sqrt{1 + H(\Phi_H)} \left( \frac{Xcap}{\epsilon_{si}} + \frac{Xox}{\epsilon_{OX}} \right)$$

$$V_g = \Phi_{MS} + \Phi_H + V_{OX} + V_{CAP}$$

This equation has a circular relation between  $\Phi_H$  and  $H(\Phi_H)$ , other values being known.

(C6) Within cap layer:

Integrating Poisson equation and applying boundary condition at X1,

$$\begin{aligned}
& \frac{qNdXd - qNaXp}{\epsilon_{si}} \sqrt{1 + H(\Phi_H)} X1 + C11 = \Phi_H \\
\Phi &= \left[ \frac{qNdXd - qNaXp}{\epsilon_{si}} \sqrt{1 + H(\Phi_H)} \right] x + \Phi_H - \left[ \frac{qNdXd - qNaXp}{\epsilon_{si}} \sqrt{1 + H(\Phi_H)} \right] X1 \\
&= \Phi_H - \frac{qNdXd_{\max} - qNaXp}{\epsilon_{si}} \sqrt{1 + H(\Phi_H)} (X1 - x)
\end{aligned}$$

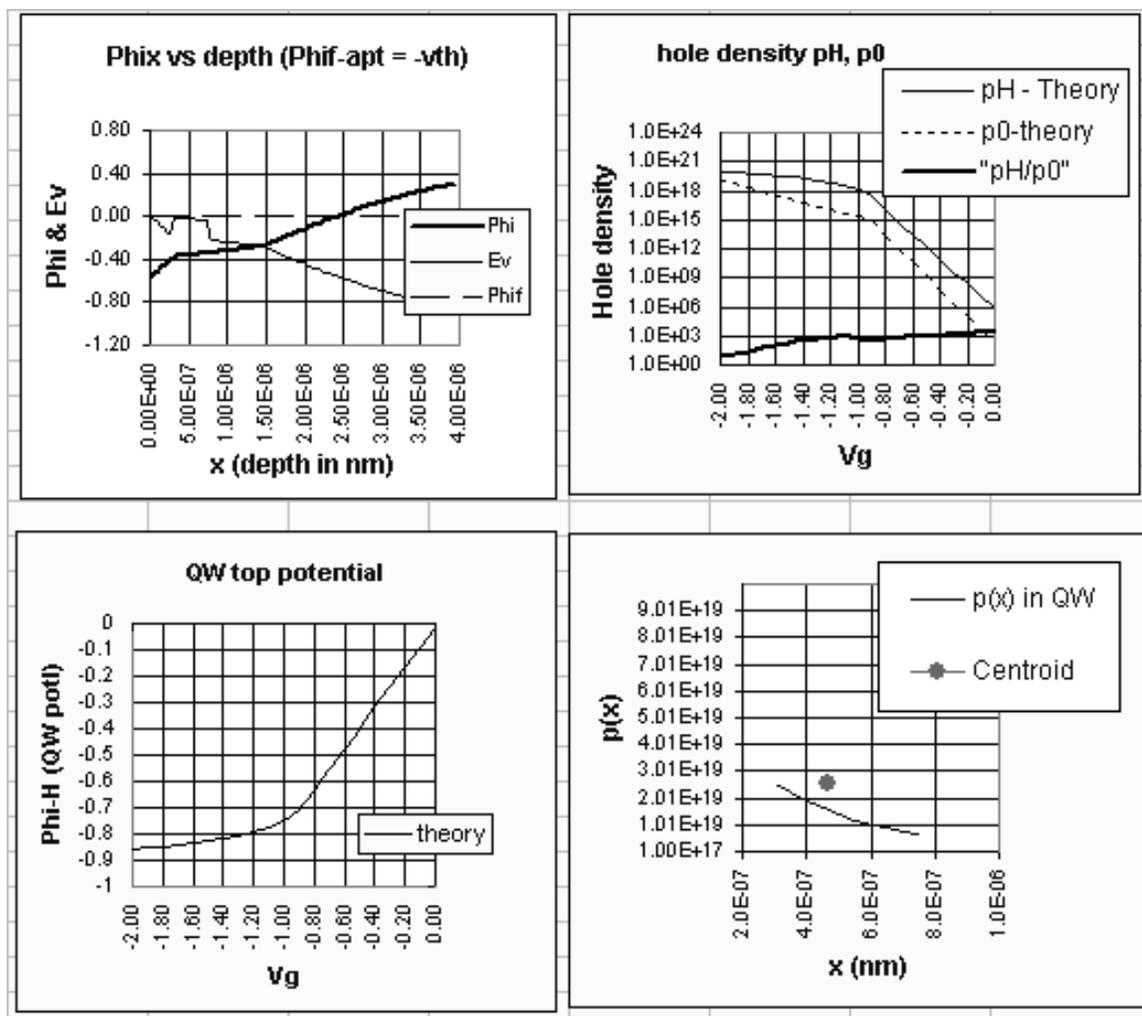
## APPENDIX – D

### EXCEL ANALYTICAL P-HMOSFET DESIGN PROGRAM

#### D.1 Parameter Design Panel

INPUTS			COMPUTED VALUES	
Xox	3.00E-07	*	<b>Excel-Analytical</b>	
Xcap	3.00E-07	*	VT	-0.69
Xchl	5.00E-07	*	VTS	-1.04
Xbuf	5.00E-07		pH (x1)	9.49E+17 QW hole density
Xp	3.00E-07		p0	2.60E+12 Surface hole density
chlx	0.5	QW Ge fraction	pH/p0	3.65E+05
Np	5.00E+18	p-delta doping density	psheet	3.80E+11 QW hole sheet density
Nsub	5.00E+17	Sub doping density	S	84.1
Lg	5.00E-05	Channel length (cm)	Vdsat	-0.35
Vg	-1.500		Idsat	-2.35E-04
Vd	-1.500			
Note:	* fixed parameters			

### D.1.1 Internal parameters



## D.2. Computational panel

STRUCTURE INPUTS			CALCULATED CONSTANTS		PHYSICS CONSTANTS	
Xox_	3E-07	cm	beta	38.61004	Eg-chl	0.924
Xcap_	3E-07	cm	Lb	5.83E-07	Eg-si	1.124
Xchl_	5E-07	cm	Cbuf	0.034752	Chi-si	4.05
Xbuf_	5E-07	cm	Csige	0.150591	eps-ox	3.45E-13
Xp_	3E-07	cm	Ccap	0.24916	eps-si	1.04E-12
x_chl_	0.5		C9	1.23E-06	eps-chl	1.123E-12
Np_	5E+18		C10	-0.78573	ni	1.45E+10
Nsub_	4.93E+17		tsi	1.53E-06	q	1.60E-19
Lgate	0.00005	cm	tsiox	2.43E-06	R	9.22E-01
CALCULATED:			tcapox	1.2E-06	vth	0.0259
PhiBsub_	0.394605	eV	x1_	3E-07	eps0	8.854E-14
Phi-sub	4.217395	eV	x2_	8E-07	epr-ox	3.90E+00
delEv_	0.375	eV	x3_	1.3E-06	epr-Si	11.7
Phims	-0.11287	V	x4_	1.6E-06	epr_Ge	15.8
			x5_	5.32E-06	epr-chl	12.684

Formulas for calculated constants are given in a table under the computational formulas section. This table is not given here to avoid distraction and thus make it easier to understand the overall computational details.

## D.2.1 Definitions of Symbols

---

### Definitions of symbols:

x	Depth in cm from the top of cap layer
Phi <sub>x</sub>	Potential at x
Phi <sub>f</sub>	Fermi level at potential 0 eV
R	Ratio of eps(Si)/eps(SiGe)
Phi_TH_	Potential at top of QW at V <sub>gs</sub> = V <sub>T</sub>
Xdmax_	Substrate depletion at V <sub>gs</sub> = V <sub>T</sub> , maximum
V <sub>T</sub> _	Threshold voltage for top of QW
V <sub>TS</sub> _	Threshold voltage for top of cap layer
Phi_TS	Potential at top of cap at V <sub>gs</sub> = V <sub>TS</sub>
Phi <sub>H</sub>	Potential at top of QW
HPhiHVT	H_PhiH at V <sub>gs</sub> = V <sub>T</sub>
PhiH(V <sub>TS</sub> )	QW potential at V <sub>gs</sub> = V <sub>TS</sub>
Phi4(V <sub>T</sub> )	Substrate potential at V <sub>gs</sub> = V <sub>T</sub>
E <sub>fd</sub> 4(V <sub>T</sub> )	Substrate field at V <sub>gs</sub> = V <sub>T</sub>
V <sub>g</sub> _	V <sub>gs</sub>
PhiH1	Qw potential for V <sub>gs</sub> < V <sub>T</sub>
PhiH2	Qw potential for V <sub>gs</sub> > V <sub>T</sub>
Xd_	Substrate depletion for V <sub>gs</sub> < V <sub>T</sub>
H_PhiH	Correction term in cap due to holes in QW
p <sub>H</sub> _	Hole density at QW top
p <sub>0</sub> _	Hole density at top of cap
Phi <sub>0</sub>	Potential at top of cap (x = 0)
Phi4(V <sub>g</sub> _)	Substrate potential at V <sub>gs</sub>
E <sub>fd</sub> 4(V <sub>g</sub> _)	Substrate field at V <sub>gs</sub>
C15	Constant field in cap
CH	A factor used in the iteration of Phi_TH_
Ev	Valence band energy in eV

## D.2.2 Computations

```
>>>VT calculation: (Xdmax from PhiTH, then Phi4 and Efld4 from Xdmax)
Phi_TH_ -7.24E-01      Phi4(VT) -0.60658   Phi4 from Xdmax
Xdmax_  3.91E-06      Efld4(VT) 2.97E+05   Efld4 from Xdmax
VT_     -0.927        CH(VT)   0.958
HPhiHVT 0.3           C15(VT)  6.56E+04
Phi_TS  -0.89829
VTS_    -1.0705
PhiH(VTS) 9E+18
```

```
>>>Potl and hole density from Vgs: Efld4(Vg_) from approximate formula,
      then Phi4(Vg_), Xd and PhiH1. Then PhiH2 from PhiH1.
```

```
Vg_     -1.5          Phi4(Vg_) -0.60658
PhiH1   -0.82621     Efld4(Vg_) 297316.2
PhiH2   -0.82599     Phi0      -9.68E-01
Xd_     4.16E-06     C15       65637.63
H_PhiH  49.34009     CH        0.95763
pH_     2.52E+19     p0_       7.26E+18
```

```

<<< ----- Potential vs depth for one Vgs input ----- >>>
x          Phix      Ev      psheet= 7.10E+12
4.00E-06  nil        nil        p(x)      Integ(p(x))
3.95E-06  2.99E-01  -8.61E-01    0         0
3.90E-06  2.92E-01  -8.54E-01    0  0.00E+00
3.85E-06  2.85E-01  -8.47E-01    0  0.00E+00
3.80E-06  2.77E-01  -8.39E-01    0  0.00E+00
3.75E-06  2.70E-01  -8.32E-01    0  0.00E+00
3.7E-06   0.261905  -0.8239      0         0

```

```

<<< --- PhiH & hden vs Vgs (Vgs as input) --->>>
Vg      E4      PhiH      pH
-2.00E+00 297316.2 -0.85767 8.55E+19
-1.90E+00 297316.2 -0.85265 7.05E+19
-1.80E+00 297316.2 -0.84713 5.69E+19
-1.70E+00 297316.2 -0.84099 4.49E+19
-1.60E+00 297316.2 -0.83409 3.44E+19
-1.50E+00 297316.2 -0.82621 2.54E+19
-1.4      297316.2 -0.81706 1.78E+19

```

```

<<<----- Surface density vs Vgs ----->>>
Vg      Phi0      p0      pH/p0
-2.00E+00 -9.82E-01 1.26E+19 6.782974
-1.90E+00 -9.59E-01 5.06E+18 13.93821
-1.80E+00 -9.35E-01 2.04E+18 27.93106
-1.70E+00 -9.12E-01 8.28E+17 54.28198
-1.60E+00 -8.89E-01 3.39E+17 101.5398
-1.50E+00 -8.66E-01 1.40E+17 180.8933
-1.4      -0.84332 5.9E+16 302.1809

```

### D.3 Computational Formulas

>>>VT calculation:

Phi_TH_	=-2*PhiBsub_+delEv_-\$V\$15
Xdmax_	=-J7+SQRT(J7*J7-(Phi_TH_+vth-J6)*2*J2*J3*J3)
VT_	=VTnew(Phims,Phi_TH_,\$B\$9,\$F\$10,\$J\$11)
HPhiHVT	0.3
Phi_TS	=-2*PhiBsub_
VTS_	=VTSnew(\$B\$10,\$B\$12,Phi_TH_,Phims,Xox_,\$F\$10,\$F\$9)
PhiH(VTS)	90000000000000000000

Phi4(VT)	F8-Xdmax_*Xdmax_/(2*J2*J3*J3)-vth
Efld4(VT)	Xdmax_/(J2*J3*J3)
CH(VT)	Ch(Nsub_,eps_chi,\$J\$2,\$F\$8,Np_,Xp_)
C15(VT)	\$F\$8-q*Np_*Xp_/eps_si

>>>Potl and hole density from Vgs:

Vg_	-1.5
PhiH1	E4toPhiH(Vg_,VT_,Phi_TH_,B16,Phims,\$J\$3,D16,Nsub_,Np_,Xp_,eps_chi,\$J\$7,\$J\$11,\$J\$6)
PhiH2	Phi_TH_-((Phi_TH_-PhiH1)-2*vth*LOG(SQRT(1+1/H_PhiH)))
Xd_	IF(H_PhiH<0.1,-J7+SQRT(J7*J7-(PhiH1+vth-J6)*2*J2*J3*J3),-B13+SQRT(J7*B20-(PhiH2+vth-J6)*2*J2*J3*J3))
H_PhiH	hphih(PhiH1,Phi_TH_,\$D\$19)
pH_	pH(Nsub_,Phi_TH_,\$B\$18)

Phi4(Vg_)	(-1/\$J\$2)*((D16*\$J\$2*\$J\$3)^2/2+1)
Efld4(Vg_)	IF(Vg_>VT_,VglowtoE4(Vg_,Phims,\$J\$3,\$J\$6,Np_,Xp_,Xcap_,Xox_,\$J\$10),\$F\$8)
Phi0	Phi0(G16,Phi_TH_,B17,J3,J6,J7,Nsub_,Np_,Xp_,Xcap_,eps_chi)
C15	\$D\$16-q*Np_*Xp_/eps_si
CH	Ch(Nsub_,eps_chi,\$J\$2,\$D\$16,Np_,Xp_)
p0_	Nsub_*EXP((\$B\$10-\$D\$17)/vth)

<<< ----- Potential vs depth for one Vgs input ----- >>>	
Phix	Phix(A28,\$G\$15,\$G\$16,Np_Xp_Xbuf_Xchl_Xcap_Xd_eps_chl,\$J\$7,\$G\$18,\$J\$4,\$J\$5,\$J\$6,H_PhiH)+\$M\$12
Ev	Ev(A28,x1_x2_B28,Eg_si,delEv_)
p(x)	IF((A28<=(Xcap_+Xchl_)),pH(Nsub_,Phi_TH_,(B28-PhiBapt_),0)*IF(A28<Xcap_,0,1)
<<< --- PhiH & hden vs Vgs (Vgs as input) --->>>	
E4	IF(F27>VT_VglowtoE4(F27,Phims,\$J\$3,\$J\$6,Np_Xp_Xcap_Xox_\$J\$10),\$F\$8)
PhiH	E4toPhiH(F27,VT_Phi_TH_H27,Phims,\$J\$3,G27,Nsub_Np_Xp_eps_chl,\$J\$7,\$J\$11,\$J\$6)
pH	Nsub_*EXP((Phi_TH_-H27)/Mh)
<<<----- Surface density vs Vgs ----->>>	
Phi0	Phi0(G27,Phi_TH_H27,\$J\$3,\$J\$6,\$J\$7,Napt_Np_Xp_Xcap_eps_chl)
p0	Nsub_*EXP((B\$10-J27)*\$J\$2)
pH/p0	I27/K27
<<< --- PhiH & hden vs Vgs (Vgs as input) --->>>	
E4	IF(F27>VT_VglowtoE4(F27,Phims,\$J\$3,\$J\$6,Np_Xp_Xcap_Xox_\$J\$10),\$F\$8)
PhiH	E4toPhiH(F27,VT_Phi_TH_H27,Phims,\$J\$3,G27,Napt_Np_Xp_eps_chl,\$J\$7,\$J\$11,\$J\$6)
pH	Napt_*EXP((Phi_TH_-H27)/Mh)
<<<----- Surface density vs Vgs ----->>>	
Phi0	Phi0(G27,Phi_TH_H27,\$J\$3,\$J\$6,\$J\$7,Napt_Np_Xp_Xcap_eps_chl)
p0	Napt_*EXP((B\$10-J27)*\$J\$2)
pH/p0	I27/K27

As can be seen from the above lots of variables are calculated. These calculations refer to a large number of cells in Excel spreadsheet. They also call a lot of functions. Even though, all relevant equations are listed above without an understanding of the cells and the functions it is difficult to understand how the computational scheme works. Hence the cell references and functions are listed below. First cell references are listed since it is a long list and often appears in the formulas appearing in the above computations. The cells could have been listed almost in the order they appear in the above computations. But it

may be easier to track them if they are given in alphabetical order. Hence they are given in the alphabetical order.

Table D-1. Excel spreadsheet cell references in the calculations given in the section of computational formulas

Cell	Details	Comments
A28	Depth from top of cap, x	
B9	HPhiHVT	
B10	Phi_TS	
B12	PhiH(VTS)	
B16	PhiH1	
B20	pH_	
B28	Phix	Potential at any depth x from top of cap
D16	Efld4(Vg_)	$\xi_4$ for any Vgs
D19	CH	Iteration const, for HphiH, to account for holes in QW
F8	Efld4(VT)	$\xi_4$ for $V_{gs} = VT$
F9	CH(VT)	CH for $V_{gs} = VT$
F27	Vgs	
G15	Efld4 from approximation	
G16	Phi4 from Efld4	Potential and field in the substrate
G18	C15	$\xi_4 - q N_p X_p / \epsilon_{Si}$ . Field at the bottom of p-layer – the field due to holes in p-layer
G27	E4 (Vg)	
H27	PhiH	
I27	pH	
J2	beta	See table D-2 for J2 – J11
J3	Lb	
J4	Cbuf	
J5	Csig	
J6	Ccap	
J7	C9	
J8	C10	
J9	tsi	
J10	tsiox	
J11	tcapox	
J27	Phi0	
K27	p0	
M16	Phisub or doped Si potential	
V15	vth(ln(Nvsi/Nvsige))	Nv: Valece band DOS

### D.3.1 Cell Reference:

The Excel references used in the formulas used for computations are listed in Table D.1. Short descriptions of them are also included. This table comes as an easy reference in the understanding of the computations.

### D.3.2 Calculated constants:

Details of Excel cells referred in the calculated constants can be seen in a table at the bottom of computational formulas section. It is given towards the end to reduce

Table D-2. Some of the constants used in the section on computational formulas

Constant	Details	Comments
beta	=1/vth	vth is kT/q
Lb	=SQRT(eps_si/(q*\$J\$2*Nsub_))	Cell J2 has beta
Cbuf	=q*Np_*Xp_*Xp_/(2*eps_si)	Integration const.
Csig	=\$J\$4+q*Np_*Xp_*Xbuf_/eps_si	Integration const. J4 = Cbuf
Ccap	=\$J\$5+q*Np_*Xp_*Xchl_/eps_chl	Integration const. J5 = Csig
C9	=Xp_+Xbuf_+R_*Xchl_	Equivalent Si thickness of p-layer + buffer + channel
C10	=Phi_TH_- \$J\$6	
tsi	=Xp_+Xbuf_+R_*Xchl_+Xcap_	Equivalent Si thickness of p-layer + buffer + channel
tsox	=\$J\$9+(eps_si/eps_ox)*Xox_	J9 = tsi. So equivalent Si thickness including p-layer to the gate
tcapox	=Xcap_+(eps_si/eps_ox)*Xox_	
x1_	=Xcap_	Cap thickness
x2_	=x1_+Xchl_	Cap + Si <sub>1-x</sub> Ge <sub>x</sub> channel
x3_	=x2_+Xbuf_	Cap + channel + buffer
x4_	=x3_+Xp_	The above + p-layer
x5_	=x4_+Xd_	The above + sub depletion
PhiBsub_	=vth*LN(Nsub_/ni)	Doping parameter
Phisub	=Chi_si+Eg_si/2-PhiBsub_	Chi for n-doped silicon
delEv	=0.75*x_chl_	Valence band discontinuity
Phims	=Chi_si-\$M\$16	Work function difference

distraction and give more emphasis to the main computational scheme. In the above table most of the entries has some meaning even though they can be considered as coming out of the integration. For example,

$$C_{buf} = \frac{qN_p X_p^2}{2\epsilon_{st}} = \frac{qN_p X_p}{2} X_p$$

So this can be looked upon as the potential drop in the p-layer which is a boundary condition at the bottom of the buffer layer. Similarly  $C_{sig}$  is this potential plus the potential drop in the undoped buffer layer due to the full charge in the p-layer etc.

### D.3.3 Functions:

Function VTnew(Phims, PhiTH, hph, C15, tcapox)

Const q = 1.6E-19

Const epssi = 1.04E-12!

Const epsox = 3.45E-13!

Ht = (1 + hph) ^ 0.5

VTnew = Phims + PhiTH - C15 \* tcapox \* Ht

End Function

Function VTSnew(PhiTS, PhiHVTS, PhiTH, Phims, Xox, C15, Ch)

Const q = 1.6E-19!

Const epssi = 1.04E-12!

Const epsox = 3.45E-13!

Const ni = 1.4E+10!

Const vth = 0.0259!

Bta = 1 / vth

hph = Ch \* Exp((PhiTH - PhiHVTS) / vth)

VTSnew = PhiTS + Phims - (epssi / epsox) \* C15 \* ((1 + hph) ^ 0.5) \* Xox

End Function

Function Ch(Napt, epssige, Beta, Efld4, Np, Xp)

Const q = 1.6E-19!

Const epssi = 1.04E-12!

Ch = (2 \* q \* Napt \* epssige) / (Beta \* (epssi \* Efld4 - q \* Np \* Xp) ^ 2)

End Function

Function E4toPhiH(Vg, VT, PhiTH, PhiH, Phims, Lb, Efld4, Napt, Np, Xp, epssige, C9, tcapox, Ccap)

Const q = 1.6E-19!  
 Const epssi = 1.04E-12!  
 Const epsox = 3.45E-13!  
 Const ni = 1.4E+10!  
 Const vth = 0.0259!  
 Bta = 1 / vth

C15 = Efld4 - q \* Np \* Xp / epssi  
 C1 = C15 \* tcapox  
 Ch1 = Ch(Napt, epssige, Bta, Efld4, Np, Xp)  
 Phi4 = E4toPhi4(Efld4, Bta, Lb)

If (Vg > VT) Then 'Vglow, no HPhiH term  
 E4toPhiH = Phi4 - C9 \* Efld4 + Ccap  
 Else 'PhiH to PhiH iterate, HPhiH >> 1  
 E4toPhiH = PhiTH - 2 \* vth \* Log(Abs(PhiH - Vg + Phims) / (C1 \* Ch1 ^ 0.5))  
 'E4toPhiH = -0.555 'SEED VALUE if iteration fails  
 End If  
 End Function

Function hphih(PhiH, PhiTH, Ch)  
 'Jan10, includes carriers in depln, valid for all Vg

Const q = 1.6E-19!  
 Const epssi = 1.04E-12!  
 Const epsox = 3.45E-13!  
 Const ni = 1.4E+10!  
 Const vth = 0.0259!  
 Bta = 1 / vth

hphih = Ch \* Exp((PhiTH - PhiH) / vth)  
 'hphih = 0.555 ' SEED VALUE if frozen  
 End Function

Function PH(Napt, PhiTH, PhiH)  
 Const vth = 0.0259  
 PH = Napt \* Exp((PhiTH - PhiH) / vth)  
 End Function

Function VglowtoE4(Vg, Phims, Lb, Ccap, Np, Xp, Xcap, Xox, tsiox)

'Jan10, includes carriers in depln, valid for Vg below VT only

Const q = 1.6E-19!

Const epssi = 1.04E-12!

Const epsox = 3.45E-13!

Const ni = 1.4E+10!

Const vth = 0.0259!

Bta = 1 / vth

C13 = tsiox \* (2 / (Bta \* Lb \* Lb))

Lhs = Vg - Phims - Ccap - q \* Np \* Xp \* (Xcap / epssi + Xox / epsox) + 1 / Bta

C14 = Lhs \* (2 / (Bta \* Lb \* Lb))

VglowtoE4 = 0.5 \* (-C13 + (C13 \* C13 - 4 \* C14) ^ 0.5)

End Function

Function Phi0(Efld4, PhiTH, PhiH, Lb, Ccap, C9, Napt, Np, Xp, Xcap, epssige)

'surface potl for any Vg input

'calculates E4 first and input, uses this further

Const q = 1.6E-19!

Const epssi = 1.04E-12!

Const epsox = 3.45E-13!

Const ni = 1.4E+10!

Const vth = 0.0259!

Bta = 1 / vth

Phi4 = E4toPhi4(Efld4, Bta, Lb)

C15 = Efld4 - q \* Np \* Xp / epssi

Hcon = Ch(Napt, epssige, Bta, Efld4, Np, Xp)

hph = Hcon \* Exp((PhiTH - PhiH) \* Bta)

Ht = (1 + hph) ^ 0.5

Phi0 = Phi4 + Ccap - Efld4 \* C9 - C15 \* Xcap \* Ht

End Function

Function Ch(Napt, epssige, Beta, Efld4, Np, Xp)

Const q = 1.6E-19!

Const epssi = 1.04E-12!

Ch = (2 \* q \* Napt \* epssige) / (Beta \* (epssi \* Efld4 - q \* Np \* Xp) ^ 2)

End Function

Function Ev(x, x1, x2, Phi, Egsi, delEv)

Phif = 0 'bulk Fermi level zero

Evsf = Phif - Phi - Egsi / 2

Ev = Evsf

If ((x >= x1) And (x <= x2)) Then Ev = Ev + delEv

End Function

```
Function PH(Napt, PhiTH, PhiH)
Const vth = 0.0259
PH = Napt * Exp((PhiTH - PhiH) / vth)
End Function
```

```
Function Phix(x, Phi4, Efld4, Np, Xp, Xbuf, Xsige, Xcap, Xd, epssige, C9, C15, Cbuf,
Csige, Ccap, hphih)
```

```
'Jan 10, includes mobile carriers in depln
'Gives Phi(x) vs x for a given Vg (E4 and Phi4 need to be found
'from Vg and input). Valid for all Vg
Const q = 1.6E-19!
Const epssi = 1.04E-12!
Const epsox = 3.45E-13!
Const ni = 1.4E+10!
Const vth = 0.0259!
Bta = 1 / vth
```

```
x1 = Xcap
x2 = x1 + Xsige
x3 = x2 + Xbuf
x4 = x3 + Xp
x5 = x4 + Xd
Ht = (1 + hphih) ^ 0.5 'used only in cap layer
```

```
If (x > x5) Then 'APT layer beyond depln, APT potl is more +ive than bulk by jn potl
Phi = (-1 / Bta) * ((-Bta * Phi4 - 1) * (((x5 - x) / Xd) ^ 2) + 1)
Phi = -vth 'carriers not at zero energy but +vth, so potl -vth
ElseIf ((x > x4) And (x <= x5)) Then 'APT depln region
Phi = (-1 / Bta) * ((-Bta * Phi4 - 1) * (((x5 - x) / Xd) ^ 2) + 1)
ElseIf ((x > x3) And (x <= x4)) Then 'p-delta region
Phi = Phi4 - Efld4 * (x4 - x) + (q * Np / (2 * epssi)) * ((x4 - x) ^ 2)
ElseIf ((x > x2) And (x <= x3)) Then 'buffer region
Phi = Phi4 - Efld4 * (x4 - x) + (q * Np * Xp / epssi) * (x3 - x) + Cbuf
ElseIf ((x > x1) And (x <= x2)) Then 'SiGe chnl region
Phi = Phi4 - Efld4 * (Xp + Xbuf) - Efld4 * (epssi / epssige) * (x2 - x) + (q * Np * Xp /
epssige) * (x2 - x) + Csige
ElseIf ((x >= 0) And (x <= x1)) Then 'cap region
Phi = Phi4 + Ccap - Efld4 * C9 - C15 * Ht * (x1 - x)
```

```
End If
Phix = Phi
End Function
```

```

Function PhiHhighVg(Vg, PhiH, PhiTH, Phims, C15, tcapox, Ch)
'Jan10, includes carriers in depln, valid for Vg above VT
'need iteration with PhiH itself
'for VTS calculation feed Vg = VTS

```

```

Const q = 1.6E-19!
Const epssi = 1.04E-12!
Const epsox = 3.45E-13!
Const ni = 1.4E+10!
Const vth = 0.0259!
Bta = 1 / vth

```

```

C1 = C15 * tcapox
PhiHhighVg = PhiTH - 2 * vth * Log((PhiH - Vg + Phims) / (C1 * Ch ^ 0.5))
'PhiHhighVg = -0.5555 'SEED VALUE if iteration fails
End Function

```

#### **D.4 Notes on simulations:**

The computation of internal potentials is the most important function of this analysis and design program. Other parameters, like hole density for example, are calculated using the potential distribution. The program is thoroughly tested w. r. to the potential distribution and is found to be quite accurate. The program is written as an interface between the designer and the time consuming numeric simulators. Unlike the numeric simulators this program can quickly display outputs, like potentials and hole density distribution for example. Hence the designer can vary parameters and arrive at a close enough design that can be cross-checked or fine tuned with very little numerical simulations.

Altogether there are eight device parameters those determine the device performance. Please refer to chapter 6 and appendix – C for more details of the details and formulas. It has been found that there are some boundary conditions of the parameter choice that can

cause problems with the iterations involved in computing potentials above the threshold voltage. This may cause a lot of errors and computation of many variables may freeze. By exploring with simulations it has found that there could be any of three functions that could cause this problem. If this occurs the parameter which was varied prior to the failure should be reset to the prior value. Then the following functions should be explored and the seed values should be used to bring the program back to normal. The functions to be looked through are: (1) HPhiH (2) PhiHhighVg (3) E4toPhiH. The seed values are given towards the end of these functions. While turning off the normal output value the seed value can be made the active output.

The beginning of this appendix gives the displayed output parameters as well as few of the plots of potentials, hole density distribution, surface hole density distribution and so on. These are vital parameters of interest in the design and can assist in arriving at a good design in a very efficient manner.

## APPENDIX – E

### A HIGH FREQUENCY STRUCTURE OF P-HMOSFET

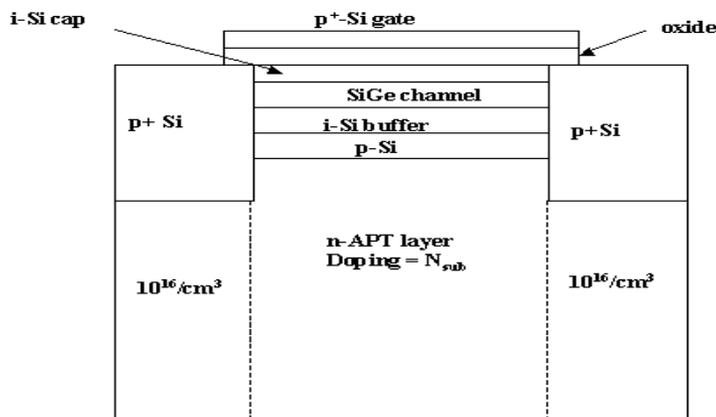


Figure E.1. Only the central region is doped with the  $N_{sub}$  used in the program. The rest of the substrate is doped with a low doping density of  $10^{16}/cm^3$ . This will reduce the source and drain capacitance to substrate. There is not a significant difference in the on current between this and doping the full substrate with  $N_{sub}$ .

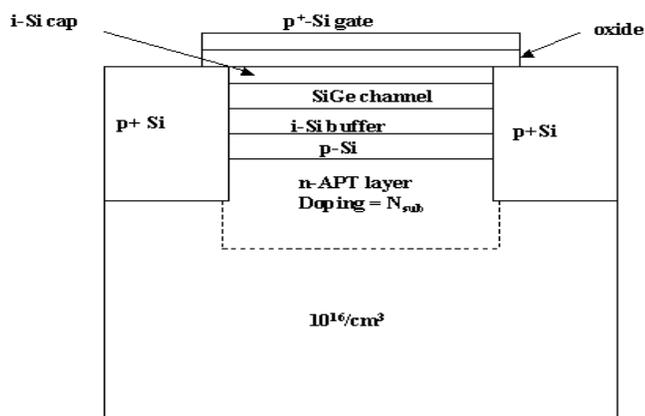


Figure E.2. The central highly doped layer need not extend all the way down. It can be made about as deep as the drain depletion width corresponding to  $N_{sub}$  and high  $V_d$ . This may be easier to fabricate

In Figure 8-10  $I_{off}$  for HMOSFET is about seven orders of magnitude less compared to that of MOSFET. While keeping the long channel  $I_{on}$  the same  $V_T$  for the HMOSFET is much higher than the  $V_T$  for the MOSFET, -1.0 V and -0.4 V respectively. This may be stated as the reason for the low  $I_{off}$  in HMOSFET, The higher  $V_T$  for HMOSFET is possible because of the higher low field hole mobility in  $Si_{1-x}Ge_x$ . Since it was found that most of the channel length was not in velocity saturation the velocity of holes in HMOSFET is higher than the velocity in MOSFET. Due to this fact the HMOSFET need much less number of holes compared to the MOSFET. This allows reduced overdrive for HMOSFET thus allowing higher  $V_T$ . It is to be mentioned here that the above will not be possible if the channel is velocity saturated.

## REFERENCES

- [1] Julius Edgar Lilienfeld, U.S. patent USP 1,745,175, 1928, patent filed in 1925
- [2] D. Kahng and M. M. Atalla, "Silicon Dioxide Field Surface Devices", Presented at the IEEE Device Research Conf., Pittsburgh, 1960.
- [3] F. Wanlass and C. T. Sah, "Nanowatt logic using field-effect metal oxide semiconductor triodes", Tech. Digest of the 1963 Int. Solid-State Circuit Conf., IEEE, pp. 32-33, 1963
- [4] Yuan Taur and Tak H.Ning, Fundamentals of Modern VLSI Devices, Cambridge University Press, New York, NY, USA, 1999
- [ 5 ] Mark Horowitz, Elad Alon, Dinesh Patil, Samuel Naffziger, Rajesh Kumar, Kerry Bernstein, "Scaling, Power and the Future of CMOS", IEDM Tech. Digest, pp. 7 – 15, 2005
- [ 6 ] Ed Sperling, "IBM to roll out 45 nm immersion process by end of year", Electronic News, 2/23/2007
- [ 7 ] Scott Hamilton, "Taking Moore's Law Into the Next Century", Computer, pp. 43 – 48, January 1999
- [8] Samir S. Rofail and Kiat Seng Yeo, Low Voltage Low Power Digital BiCMOS Circuits, Prentice Hall, New Jersey, USA, 2000
- [9] B. Davari, R. H. Dennard and G. G. Shahidi, "CMOS Scaling for High Performance and Low Power – The Next Ten Years", Proceedings of the IEEE, vol. 83, pp. 595-606, Apr. 1995
- [10] Scott Thompson, Paul Packan and Mark Bohr, "MOS Scaling: Transistor Challenges for the 21st Century", Intel Technology Journal, Q398, 1998
- [11] Hitoshi Wakabayashi, Toyoji Yamamoto, Toru Tatsumi, Ken ichi Tokunaga, Takao Tamura, Tohru Mogami and Takemitsu Kunio, "A High Performance 0.1 $\mu$ m CMOS with Elevated Salicide using Novel Si-SEG process", IEEE Electron Device Meeting, pp. 99-102, 1997
- [12] K. F. Lee, R. H. Yan, D. Y. Jeon, G. M. Chin, Y.O. Kim, D. M. Tennant, B. Razavi, H. D. Lin, Y. G. Wey, E. H. Westerwick, M. D. Morris, R. W. Johnson, T. M. Liu,

- M.Tarsia, M. Cerullo, R. G. Swartz and A. Ourmazd, "Room Temperature 0.1 $\mu$ m CMOS Technology with 11.8 ps Gate Delay", IEEE Electron Device Meeting, pp. 131-134, 1993
- [13] Mizuki Ono, Masanobu Saito, Takashi Yoshitomi, Claudio Fiegna, Tatsuya Ohguro and Hiroshi Iwai, "A 40 nm Gate Length n-MOSFET", IEEE Trans. Electron Devices, vol. 42, no. 10, October 1995
- [14] Hisao Kawaura, Toshitsugu Sakamoto, Toshio Baba, Yukinori Ochiai, Jun ichi Fujita, Shinji Matsui and Jun ichi Sone, "Transistor operation of 30 nm Gate Length EJ-MOSFETs", IEEE Electron Device Letters, vol. 19, no. 3, March 1998
- [15] Paul Vande Voode, "MOSFET Scaling into the Future", Hewlett Packard Journal, pp. 1-6, August 1997
- [16] S. T. Ma, A. Keshavazi, V. De and J. R. Brews, "A Statistical Model for Extracting Geometric Sources of Transistor Performance Variation", IEEE Trans. Electron Devices, vol. 51, no. 1, pp 36 – 41, 2004
- [17] Bijan Davari, "CMOS Technology Scaling, 0.1 $\mu$ m and Beyond", IEEE Electron Device Meeting, pp. 555-558, 1996
- [18] C. P. Chao, M. Mehrotra and I. C. Chen, "Optimum Junction Depth Design of the S/D Extension Regions (MDD) for Sub 0.18 $\mu$ m CMOS Technologies", SPIE Conf. Microelectronic Device Tech. II, Santa Clara, pp. 281-291, Sept. 1998
- [19] Clement H. Wann, Kenji Noda, Tetsu Tanaka, Makoto Yoshida and Chemming Hu, "A Comparative Study of Advanced MOSFET Concepts", IEEE Trans. Electron Devices, vol. 43, no.10, pp. 1742-1753, October 1996
- [20] Yuan Taur and Edward J. Mowak, "CMOS Devices below 0.1 $\mu$ m : How High Will Performance Go?", IEEE Electron Device Meeting, pp. 215-218, 1997
- [21] Hyunsang Hwang, Dong Hoon Lee and Jeong Mo Hwang, "Degradation of MOSFETs Drive Current due to Halo Ion Implantation", IEEE Electron Device Meeting, pp.567-570, 1996
- [22] E. J. Nowak, "Maintaining the Benefits of CMOS Scaling when Scaling Bogs Down", IBM Journal of Research and Development, vol. 46, no. 2-3, 2002
- [23] Douglas J. Paul, "Silicon-germanium Strained Layer Materials in Microelectronics", Advanced Materials, vol. 11, no. 3, pp. 191-204, 1999

- [ 24 ] Scott E. Thompson, Mark Armstrong, Chis Auth et. al., “A 90 nm Logic Technology Featuring Strained Silicon”, IEEE Trans. Electron Devices, pp. 1 – 8, 2004
- [ 25 ] Max Schulz, “The End of the Road for Silicon”, Nature, vol. 399, pp. 729 – 730, June 1999
- [ 26 ] 2001 Technology Roadmap for Semiconductors, IEEE Computer Society, vol. 35, no.1, pp. 42-53, 2002
- [ 27 ] International Technology Roadmap for Semiconductor (ITRS) 2002 Update
- [ 28 ] International Technology Roadmap for Semiconductor (ITRS) 2007 Update
- [ 29 ] E. Y. Wu, E. J. Nowak, A. Vayshenker, W. L. Lai and D. L. Hamon, “CMOS Scaling Beyond the 100 nm Mode with Silicon Dioxide based Gate Dielectrics”, IBM J. Res & Dev. Vol. 46, no.23, March – May 2002
- [ 30 ] Hitoshi Wakabayashi, Makoto Ueki, Mitsuru Narihiro et. al., “Sub 50 nm Physical Gate Length CMOS Technology and Beyond Usins Steep Halo”, IEEE Trans. Electron Devices, vol. 49, no. 1, pp. 89 – 93, 2002
- [ 31 ] W. Haensch, E. J. Nowak, R. H. Dennard et. al., “Silicon CMOS Devices Beyond Scaling”, IBM J. Res & Dev, vol. 50, no. 45, pp. 339 - 361, July – September 2006
- [ 32 ] Xuejue Huang, Wen-Chin Lee, Charles Kuo et. al., “Sub 50 nm P-Channel FinFet”, IEEE Trans. Electron Devices, vol 48, no. 5, pp. 880 – 886, 2001
- [ 33 ] P. G. Pikus and K. K. Likharev, “Nanoscale Field Effect Transistors: An Ultimate Size Analysis”, Appl Phys. Lett, vol. 71, no. 25, pp. 3661 – 3663, 1997
- [ 34 ] Y. Naveh, K. Likharev, “Modeling of 10 nm scale Ballistic MOSFETs, IEEE Electron. Device. Lett., vol. 21, pp. 242 – 244, 2000
- [ 35 ] Leland Chang, Yang-Kyu Choi, Daewon Ha et. al., “Extremely Scaled Silicon Nano CMOS Devices”, Proc. IEEE, vol. 91, no. 11, pp. 1860 – 1873, 2003
- [ 36 ] Yang-Kyu Choi, Nick Lindert, Peiqi Xuan et. al., “Sub 20 nm CMOS FinFet Technologies”, IEDM 01, pp. 421 – 424, 2001
- [ 37 ] “Double Gated Transistor Aims to Reduce Chip Processing Costs”, Technology Advances, MRS Bulletin, pp. 916 – 917, December 2004

- [ 38 ] S. Parke, K. DeGregorio, D. Hackler and D. Wilson, "Ultra Low Power Dynamic Threshold Digital Circuits in the FlexFET Independently Double Gated SOI CMOS Technology", IEEE Electron Device Lett., pp. 33-34, 2006
- [ 39 ] Eric Pop, Sanjiv Sinha and Kenneth E. Goodson, "Heat Generation and Transport in Nanometer Scale Transistors", Proc. IEEE, vol. 94, no.8, pp. 1587 – 1601, 2006
- [ 40 ] The National Technology Roadmap for Semiconductors, Semiconductor Industry Association, San Jose, CA, 1994
- [ 41 ] James B. Kuo and Jea-Hong Lou, "Low Voltage CMOS VLSI Circuits", John Wiley & Sons, Inc., New York, 1999
- [ 42 ] C. Papavassiliou, C. Toumazou and K. Foblets, "SiGe Hetero FETs Potential for Micropower Applications", IEICE Trans. Electron., vol. E00-A, no. 1, pp. 1 – 8, 1995
- [ 43 ] J. B. Roldan, F. Gamiz, J. A. Lopez-Villanueva and J. E. Carceller, "Understanding the Improved Performance of Strained Si/Si<sub>1-x</sub>Ge<sub>x</sub> Channel MOSFETs", Semicond. Sci. Technol, vol. 12, pp. 1603 – 1608, 1997
- [ 44 ] Y. Taur, D. A. Buchanan, W. Chen et. al., Proc. IEEE, vol. 85, pp. 486, 1997
- [ 45 ] Linda Gepert, Technology 1999 Analysis & Forecast, IEEE Spectrum, pp. 52 – 56, January 1999
- [ 46 ] David C. Ahlgren and James Dunn, "SiGe Comes of Age", IBM Micronews, vol. 6, no. 1, pp. 1 – 5, 2000
- [ 47 ] James M. Moniz, "Is SiGe the Future of GaAs for RF Applications ?", GaAs Integrated Circuits Symposium Technical Digest, pp. 229 – 231, 1997
- [ 48 ] M. V. Fischetti, "Scaling MOSFETs to the Limit: A Physicist's Perspective", Journal of Computational Electronics, vol. 2, pp. 73 – 79, 2003
- [ 49 ] Scott E. Thompson, Mark Armstron, Chis Auth et. al., "A Logic Nanotechnology Featuring Strained Silicon", IEEE Electron Device Lett., vol. 25, no. 4, pp. 191 – 193, 2004
- [ 50 ] Krol, E. J., "Silicon Germanium", Potentials IEEE, vol. 18, no. 2, pp. 17 – 19, 1999
- [ 51 ] S. Tyagi, C. Auth, P. Bai et. al., "An Advanced Low Power, High Performance Strained Channel 65nm Technology", Electron Device Meeting, IEDM Digest, pp. 245 – 247, December, 2005

- [ 52 ] 2003 Technology Roadmap for Semiconductors, IEEE Computer, vol. 37, no. 1, pp. 47 – 56, 2004
- [ 53 ] The International Roadmap for Semiconductors (ITRS), 2005
- [ 54 ] Shekhar Borkar, “Electronics Beyond Nano-scale CMOS”, Proceedings of 43<sup>rd</sup> Design Automation Conference, San Francisco, California, July, 2006
- [ 55 ] R. Jacob Baker, CMOS Circuit Design, Layout and Simulation, Second edition, Wiley Interscience, 2005
- [ 56 ] Daniel Foty, “Mosfet Modeling with Spice”, Prentice Hall, New Jersey, 1997
- [ 57 ] Laszlo B. Kish, “End of Moore’s law: thermal (noise death of integration in micro and nano electronics”, Physics Letters A, pp. 144 – 148, Elsevier Science, 2002
- [ 58 ] R. Kamankar, H. Bu, C. Bowen et. al., “An Enhanced 90 nm High Performance Technology with Strong Performance Improvements from Strain and Mobility Increase through Simple Process Changes”, Symposium on VLSI Technology Digest of Technical Papers, pp. 162 – 163, 2004
- [ 59 ] A. Ono, K. Fukasaku, T. Matsuda et. al., “A 70 nm Gate Length CMOS Technology with 1.0 V Operation”, 2000 Symposium on VLSI Technology. Digest of Technical Papers, pp.14 – 15
- [ 60 ] JS Zhao, A. Chatterjee, S. Tang et. al., “Transistor Optimization for Leakage Power Management in a 65 nm CMOS Technology for Wireless and Mobile Applications”, Symposium on VLSI Technology Digest of Technical Papers, pp. 14 – 15, 2007
- [ 61 ] A. Oishi, T. Komoda, Y. Morimasa et. al., “High Performance CMOSFET Technology for 45 nm generation”, Symposium on VLSI Technology Digest of Technical Papers, pp. 166 – 167, 2004
- [ 62 ] N. Yasutake, K. Ohuchi, M. Fujiwara et. al., “A hp 22 nm Node Low Operating Power (LOP) Technology with Sub – 10 nm Gate Length Planar Bulk CMOS Devices”, Symposium on VLSI Technology Digest of Technical papers, pp. 84 – 85, 2004
- [ 63 ] G. Timp, J. Bude, K. K. Bourdelle et. al., “The Ballistic Nano-transistor”, IEDM, pp. 55 – 58, 1999
- [ 64 ] Akira Nishiyama, Kazuya Matsuzava and Shin-ichi Takagi, “SiGe Source/Drain Structure for the Suppression of the Short-Channel Effect of Sub-0.1 $\mu$ m p-Channel MOSFETs”, IEEE Transactions on Electron Devices, vol. 48, no. 6, 2001

- [ 65 ] I. S. Gradshteyn and I.M. Ryzhik, "Table of Integrals, Series and Products", Academic Press, San Diego, 2000
- [ 66 ] M. Chudzik, B. Doris, R. Mo et. al., "High Performance High- $\kappa$ /Metal Gates for 45 nm CMOS and Beyond with Gate First Processing", VLSI Technology 2007 IEEE Symposium, pp. 194 –195
- [ 67 ] C. Isheden, P. E. Hellstrom, H. H. Radamson et. al., "MOSFETs with Recessed SiGe Source/Drain Junctions Formed by Etching and Growth", Electrochem. Solid-State Lett., vol. 7, issue. 4, pp. G53 – G55, 2004
- [ 68 ] Asen Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1  $\mu\text{m}$  MOSFETs: A 3-D "Atomistic" Simulation Study", IEEE Transactions on Electron Devices, vol. 45, no. 12, 1998
- [ 69 ] Dennis Sinitsky, Fariborz Assaderanghi, Chemmin Hu and Jeffrey Boker, "High Field Hole Velocity and Velocity Overshoot in Silicon Inversion Layers", IEEE Electron Device Letters, vol. 18, no. 2, pp. 54 –56, 1997
- [ 70 ] J.W. MacPherson, "Reliability Challenges for 45 nm and Beyond", Proc. 43<sup>rd</sup> Design Automation Conf., San Francisco, 2006
- [ 71 ] F. Nouri, P. Verheyen, L. Washington et. al., "A Systematic Study of Trade offs in Engineering a Locally Strained pMOSFET", IEDM 04, pp. 1055 – 1058, 2004
- [72] U. Konig and H. Dambkes, "Si<sub>1-x</sub>Ge<sub>x</sub> HBTs and HFTs", Solid State Electronics, vol. 38, no. 9, pp. 1595 – 1602, 1995
- [73] A. K. Goel and H. Shaw, "Gallium Arsenide-on-Silicon Technology – an Overview", Proc. 32nd Midwest Symposium on Circuits and Systems, August 1989
- [ 74 ] K. W. Goossen, J. A. Walker, L. A. D'Asaro et. a., "GaAs MQW Modulators Integrated with Silicon CMOS", IEEE Photonics Tech. Lett., vol. 7, no. 4, pp. 360-362, 1995
- [75] Mark LaPedus and Peter Clarke, "Motorola Backs off GaAs-on-Silicon Efforts", Silicon Strategies, January 2003
- [76] A. Levitas et. al., "Electrical Properties of Germanium Silicon Alloys", Phys. Rev., vol. 99, 1810-1814, 1955
- [77] M. Glickman, Phys. Rev., vol. 100, pp. 1146, 1955

- [ 78] Loi D. Nguyen, Paul J. Tasker, David C. Radulescu and Lester F. Eastman, “Characterization of Ultra High Speed Pseudomorphic AlGaAs/InGaAs (on GaAs) MODFETs”, IEEE Trans. Electron Devices, vol. 36, no. 10, pp 2243 – 2248, 1989
- [79] Erich Kasper, H. J. Herzog and H. Kibbel, Appl. Phys., vol 8, pp. 199, 1975
- [80] Erich Kasper, “Properties of Strained and Relaxed Silicon Germanium”, EMIS Datareview Series, no. 12, INSPEC Publications, London, 1995
- [ 81] S. Kaya, Y. P. Zhao, J.R. Watling et al, “Indication of Velocity Overshoot in Strained Si<sub>0.8</sub>Ge<sub>0.2</sub>p-channel MOSFET”, Semiconductor Science and Technol., Issue 6, 2000
- [ 82 ] M. J. Kearney and A. I. Horrell, “The effect of alloy scattering on the mobility of holes in a Si<sub>1-x</sub>Ge<sub>x</sub> quantum well”, Semicond. Sci. Technol, vol. 13, pp. 174 – 180, 1998
- [ 83 ] Tanjinder Manku, Joel M. McGregor, Arokia Nathan, David J. Roulston, J. P. Noel and D. C. Houghton, “ Drift Hole Mobility in Strained and Unstrained Doped Si<sub>1-x</sub>Ge<sub>x</sub> Alloys”, IEEE Trans.Electron Devices, vol. 40, no. 11, 1993
- [ 84 ] P Gaworzewski, K. Tittelbach, U. Penner and N. V. Abrosimov, “Electrical properties of lightly doped p-type silicon-germanium single crystals”, Journal of Applied Physics, vol. 83, no. 10, pp. 5258 – 5263, 1998
- [ 85] M. V. Fischetti and S. E. Laux, “ Band Structure, Deformation Potentials and Carrier Mobility in Strained Si, Ge and SiGe Alloys”, J. Appl. Phys., vol. 80, no.4, pp. 2234 – 2252, 1996
- [ 86] Martin M. Rieger and P. Bogt, “Electronic Band Parameters in Strained Si<sub>1-x</sub>Ge<sub>x</sub> Alloys on Si<sub>1-y</sub>Ge<sub>y</sub> Substrates”, Physical Rev. B, vol. 48, no. 19, pp. 276 – 287, 1993
- [ 87 ] Hiroshi Nakatsuji, Yoshinari Kamakura and Kenji Taniguchi, “Full Band Monte Carlo Study for Two Dimentional Hole Transport in Strained Si p-MOSFETs”, Journal of Computational Electronics, vol. 2, pp. 109 – 112, 2000
- [88] Kaushik Bhaumik, Yosi Shacham Diamond, J. P. Noel, Joze Bevk and L. C. Feldman, “Theory and obervation of enhaced, high field hole transport in Si<sub>1-x</sub>G<sub>x</sub> quantum well p-MOSFETs, IEEE Trans. Electron Devices, vol. 43, no. 11, pp.1965-1971, Nov 1996
- [ 89] S. Krishnan, M. Fischetti and D. Vasileska, “Self Consistent Full Band Two Dimensional Monte Carlo Two Dimensional Poisson Device Solver for Modeling SiGe p-Channel Devices”, J. Vac. Sci. Technol., vol. 24, nr. 4, pp. 1197 – 2003, 2006

- [ 90] Borna Obradovic, Philippe Matagne, Lucian Shifren, Everett Wang, Mark Stettler, Jun He and Martin D. Giles, “A Physically based Analytical Model for Strain Induced Hole Mobility Enhancement”, *Journal of Computational Electronics*, vol. 3, pp 161 – 164, 2004
- [ 91 ] Ying Fu, Kaj J. Grahn and Magnus Willander, “Valence Band Structure of  $\text{Ge}_s\text{Si}_{1-x}$  for Hole Transport Calculation”, *IEEE Trans. Electron Devices*, vol. 41, no. 1, pp. 26 – 31, 1994
- [ 92] Lianfeng Yang, Jeremy R Watling, Richard C W Wilkins, Mirela Boriçi, John R Barker, Asen Asenov and Scott Roy, “Si/SiGe heterostructure parameters for device simulations”, Device Modelling Group, Department of Electronics and Electrical Engineering, University of Glasgow, Glasgow G12 8LT, UK
- [ 93] J. S. Dunn, D. C. Ahlgren, D. D. Coolbaugh et. al., “Foundation of rf CMOS and SiGe BiCMOS Technologies”, *IBM J. Res & Dev.* Vol. 47, no. 2/3, 2003
- [ 94] J. Bock, H. Schafer, H. Knapp et. al., “3.3 ps (225 GHz) SiGe Bipolar Technology”, *IEDM 04*, pp255 – 258, 2004
- [ 95] B. Jagannathan, M. Khater, F. Pagette et. al., “Self Aligned SiGe NPN Transistors with 285 GHz  $f_{\text{MAX}}$  and 207 GHz  $f_{\text{T}}$  in a Manufacturable Technology”, *IEEE Electron Device Lett.*, vol 23, no. 5, pp. 258 – 260, 2002
- [ 96] “Georgia Tech/IBM Team Demonstrates first 500 GHz Silicon-Germanium Transistors”, *Georgia Institute of Technology Research News*, June 2006
- [ 97] Nanoengineered Silicon Germanium Improves Chips, *Georgia Institute of Technology News*, Atlanta, Georgia, December 10, 2005
- [98] S. S. Iyer, G. L. Patton, S. S. Delage, S. Tiwari and J. M. C. Stork, “Silicon-Germanium Base Heterojunction Bipolar Transistors by Molecular Beam Epitaxy”, *IEDM Tech. Dig.*, pp. 874-876, 1987
- [ 99] Suraj J. Mathew, Guofu Niu, Waded B. Dubbelday et. al., “Hole Confinement and its Impact on Low Frequency Noise in SiGe pFETs on Sapphire”, *IEDM-97*, pp. 815 – 818, 1997
- [ 100] Maxim Semiconductor Corp., “Silicon Germanium (SiGe ) Technology Enhances Radio Front End Performance”, *Application Note 697*, March 2000
- [101] David L. Hame, David C. Ahlgren, Douglas D. Coolbaugh, James S. Dunn, Gregory G. Freeman, John D. Gillis, Robert A. Groves, Gregory N. Hendersen, Robb A. Johnson, Alvin J. Joseph, Seshardi Subbanna, Alan M. Victor, Kimball M. Watson,

- Charles S. Webster and Peter J. Zampardi, “Current Status and Future Trends of SiGe BiCMOS Technology”, IEEE Trans. Electron Devices, vol. 48, no. 11, Nov. 2001
- [102] Katsuya Oda, Eiji Ohue, Masamichi Tanabe, Hiromi Shimamoto, Takahiro Onai and Katsuyoshi Washio, “130 GHz fT SiGe HBT Technology”, IEEE IEDM Tech. Dig., pp. 791-794, 1997
- [103] Katsuyoshi Washio, Eiji Ohue, Hiromi Shimanoto, Katsuya Oda, Reiko Hayami, Yukihiro Kiyota, Masamichi Tanabe, Masao Kondo, Takashi Hashimoto and Takashi Harada, “A 0.2  $\mu\text{m}$  180 GHz  $f_{\text{max}}$  6.7 ps ECL SOI/HRS Self Aligned SEG SiGe HBT/CMOS Technology for Microwave and High Speed Digital Applications” IEEE Trans. Electron Devices, vol. 49, no. 2, pp. 271-278, Feb. 2002
- [104] Tak H. Ning, “History and Future Perspective of the Modern Silicon Bipolar Transistor”, IEEE Trans. Electron Devices, vol. 48, no. 11, pp. 2485-2491, Nov. 2001
- [105] Robert F. Pierret, Advanced Semiconductor Fundamentals, Modular Series on Solid State Devices, vol. VI, Addison-Wesley Publishing Co., 1989
- [106] S. Takagi, M. Iwase and A. Toriumi, “On the Universality of Inversion Layer Mobility in N and P Channel MOSFETs”, IEEE Device Meeting, pp. 398 – 401, 1988
- [107] Ken Yamaguchi, “A Mobility Model for Carriers in the MOS Inversion Layer”, IEEE Trans. Electron Devices, vol. 30, no. 6, June 1983
- [108] R. People, “Indirect Bandgap of Coherently Strained  $\text{Ge}_x\text{Si}_{1-x}$  Bulk Alloys On  $\langle 001 \rangle$  Silicon Substrates”, Physical Review B, vol. 22, no. 2, pp. 1405 – 1408, 1985
- [109] ISE TCAD user’s manual, release 6.0, ISE – Integrated Systems Engineering AG, Technoparkstrasse 1, CH-8005, Switzerland
- [110] Lianfeng Yang, Jeremy R. Watling, Richard C. W. Wilkins, Mirela Borici, John R. Barker, Asen Asenov and Scott Ray, “Si/SiGe Heterostructure Parameters for Device Simulations”, Semicond. Sci. Technol., vol. 19, no. 10, pp. 1173-1182, 2004
- [111] John D. Cressler and Gnofu Niu, Silicon Germanium Bipolar Trnasistors, Artech House, 2003
- [112] Haugeard, Becca Mary, “Radiation and Strain Effects in Silicon-Germanium Bipolar Complementary Metal Oxide Semiconductor Technology
- [113] Hadis Morkoc and Paul M. Solomon, “The HEMT: A Superfast Transistor”, IEEE Spectrum, vol. 21, pp. 28 – 35, 1984

- [ 114] T. P. Pearsall, J. C. Bean, R. People and A. T. Fiory, “GeSi modulation doped p-channel field effect transistors”, Proc. 1<sup>st</sup> Intl. Symp. Silicon Molecular Beam Epitaxy, pp. 400 – 405, 1985
- [ 115] X. Xiao, G. W. Liu and J. C. Stum, “Quantum Confinement Effects in Strained Silicon Germanium Alloy Quantum Wells”, Appl. Phys. Lett., vol. 60, no. 17, pp. 2135-2137, 1992
- [ 116 ] Dingle R., Stormer H. L. Gossard A. C. and Wiegmann W., “Electron Mobilities in Modulation Doped Semiconductor Heterojunction Superlattices”, Appl. Phys. Lett., vol. 33, pp. 665 - 669, 1978
- [ 117] Adam H. Pawlikiewicz and David Hess, “Choosing RF CMOS or SiGe BiCMOS in Mixed Signal Design”, Mixed Signal Design, pp. 36 – 44, March 2006
- [ 118] H. Morkoc, H. Unlu and G. Ji, “Principles and Technology of MODFETs”, vol. 1, John Wiley & Sons 1991
- [ 119 ] H. Morkoc, H. Unlu and G. Ji, “Principles and Technology of MODFETs”, vol. 2, John Wiley & Sons 1991
- [120] H. J. Herzog, “Crystal Structure, Lattice Parameters and Liquidus-Solidus Curve of the SiGe System”, Properties of Strained and Relaxed Silicon Germanium, Erich Kasper ed., EMIS Datareviews Series, INSPEC, 1995
- [121] F. Schaffler, Silicon-germanium (Si<sub>1-x</sub>Ge<sub>x</sub>)”, Properties of Advanced Semiconductor Materials, Edited by Michael E. Levinstein, Sergey L. Rumyantsev and Michael S. Shur, John Wiley & Sons, 2001
- [ 122] Friedrich Schaffler, “High Mobility Si and Ge structures”, Semicon. Sci. Technology, vol. 12, pp. 1515-1549, 1997
- [123] B. S. Meyerson, Proc. IEEE, vol. 80, pp. 1592, 1992
- [124] M. R. Caymax and W.Y. Leong, Advanced Silicon and Semiconducting Silicon-alloy Based Materials and Devices, ed. J. F. A. Nijs, IOP Publishing, Bristol, pp. 141, 1993
- [125] S. M. Mokler, N. Ohtani, M. H. Xie, J. Zhang and B. A. Joyce., Appl. Phys. Lett., no. 61, pp. 2548, 1992
- [126] T. I. Kamins and D. J. Meyer, Appl. Phys. Lett., no 59, pp. 178, 1991

- [127] D. Kinosky, R. Qian, J. Irby, T. Hsu, B. Anthony, S. Manerjee, A. Tasch, C. Magee and C. L. Grove, *Appl. Phys. Lett.* No.59, pp. 817, 1991
- [128] G. Theodorou and P. C. Kelires, "The Si/Ge Interface, energy and Interdiffusion", *Properties of Strained and Relaxed Silicon Germanium*, Erich Kasper ed., EMIS Datareviews Series, INSPEC, 1995
- [129] J. W. Mathews and A. E. Blakeslee, *J. Crystal Growth*, vol. 27, pp. 118, 1974
- [130] J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara and I. K. Robinson, *J. Vac. Sci. Technol. A*, vol. 2, pp. 436, 1984
- [131] M. L. Green et. al., *J. Appl. Phys.*, vol. 69, pp.745, 1991
- [132] D. C. Houghton, *J. Appl. Phys.* No. 70, pp. 2136, 1991
- [133] F. K. LeGoues, B. S. Meyerson and J. Morar, *Phys. Rev. Lett.* No 66, pp. 2903, 1991
- [ 134 ] E. Murakami, K. Nakagawa, A. Nishida and M. Miyao, " Fabrication of a Strain Controlled Si Ge/Ge MODFET with Ultrahigh Hole Mobility", *IEEE Trans. Electron Devices*, vol. 41, no.5, pp. 857 – 861, 1994
- [135] G. Kissinger, T. Morgenstein, G. Morgenstern and H. Richter, "Stepwise equilibrated  $\text{Ge}_x\text{Si}_{1-x}$  buffer with low threading dislocation", *Appl. Phys. Lett.*, Vol. 66, No. 16, pp. 2083, 1995
- [136] T. Fromherz and G. Bauer, "Energy Gaps and Band Structure of  $\text{Si}_{1-x}\text{Ge}_x$  and their Temperature Dependence", *Properties of Strained and Relaxed Silicon Germanium*, Erich Kasper ed., EMIS Datareviews Series, INSPEC, 1995
- [137] Katsuya Oda, Eiji Ohue, Masamichi Tanabe, Hiromi Shimamoto, Takahiro Onai and Katsuyoshi Washio, "130 GHz fT SiGe HBT Technology", *IEEE Electron Devices Meeting*, pp. 791-794, 1997
- [138] Kiyota. Y, udo. T, Hashmoto. T, Kodama. A, Shimamoto. H, Hayami. R, Ohue. E and Wa. K, "HCl free selective epitaxial Si-Ge growth by LPCVD for high frequency HBTs", *IEEE Trans. Electron Devices*, vol. 49, no. 5, pp. 739-745, May 2002
- [ 139] Basanth Jagannathan, Mounir Meghelli, Kevin Chan, Jai-Sung Rieh, Kathryn Schonenberg, David Ahlgrin, Seshadri Subbamma and Greg Freeman, "3.0 ps Silicon Germanium HBT ECE Ring Oscillator and Transistor Design for Minimum Gate Delay", *IEEE Electron Device Lett.*, vol. 24, no. 5, pp. 324-326, 2003

- [140] B. Jagannathan et al., Self-aligned SiGe p-channel transistors with 285 GHz  $f_{max}$  and 207 GHz  $f_T$  in a manufacturable technology, *IEEE Electron Device Letters*, 23, No. 5, pp. 258-260, 2002.
- [141] George Halkias and Aliko Vegiri, "Device parameter optimization of strained Si channel SiGe/Si n-MOSFETs using one dimensional charge control model", *IEEE Trans. Electron Devices*, vol. 45, no. 12, p.2430-2435, 1998
- [142] Sogetes T. Pantelides and Stefan Zoller, eds., *Silicon Germanium Carbon Alloys – Growth, Properties and Applications*, Taylor and Francis, vol. 15, 2002, chap. 4
- [143] K. Ismail, S. Rishton, J.O. Chu, K. Chan and B.S. Meyerson, "High performance Si/SiGe n-type modulation doped transistors", *IEEE Electron Device Lett.*, vol. 14, pp. 348-350, 1993
- [144] Kern Rim, Judy L. Hoyt and James F. Gibbons, "Fabrication and analysis of deep submicron strained-Si n-MOSFETs", *IEEE Trans. Electron Devices*, vol. 47, no.7, July 2000
- [145] M. Rashed, S. Jallepalli, R. Zaman, W. Shih, T. J. T. Kwan and C. M. Maziar, "Simulation of electron transport in strained silicon on relaxed Si<sub>1-x</sub>Ge<sub>x</sub> substrates", *Proc. 11th Biennial University/Government/Industry Microelectronics Symposium*, pp. 168-171, May 1995
- [146] Tomasz Skotnicki, Gerard Merckel and Thierry Pedron, "The Voltage Doping Transformation: A New Approach to the Modeling of MOSFET Short Channel Effects", *IEEE Electron Device Lett.*, vol. 9, no. 3, pp. 109 – 112, 1988
- [147] U. Konig and A. Gruhle, "High frequency SiGe heterostructure devices", *Electron Device Lett.*, I-5, pp. 14-23, 1997
- [148] T. E. Whall, E. H. C. Parker, "Growth and characterization of pseudomorphic Si/SiGe/Si heterostructures for p-channel field effect transistors", *Thin Solid Films*, vol. 306, pp. 338-345, 1997
- [149] S. Kal and B. Chakrabarti, "Strained silicon-SiGe devices using Germanium implantation", *IETE Journal of Research*, vol. 43, no. 2&3, pp. 185-192, 1997
- [150] Robert J. P. Lander, Yuri V. Ponomarev, Jurgen G. M. van Berkum and Wiebe B. de Boer, "High hole mobilities in fully strained Si<sub>1-x</sub>Ge<sub>x</sub> layers (0.3 < x < 0.4) and their significance for SiGe pMOSFET performance, *IEEE Trans. Electron Devices*, vol. 48, no. 8, pp.1826-1832, Aug 2001

- [151] Deepak Nayak et. Al., “Modulation doped p-channel MOSFET”, Proc. 48th Device Research Conf., pp., 1990
- [152] D. K. Nayak, “SiGe p-MOSFETs: SiGe strained 2%”, Properties of Strained and Relaxed Silicon Germanium, Erich Kasper ed., EMIS Datareviews Series, INSPEC, 1995
- [153] Qiqing Ouyang, Xiangdong Chen, Sivakumar P. Mudanai, Xin Wang, David L. Kencke, Al F. Tasch, Leonard F. Register and Sanjay Kumar Banerjee, “A novel Si/SiGe heterojunction pMOSFET with reduced short channel effects and enhanced drive current”, IEEE Trans. Electron Devices, vol. 47, no. 10, pp.1943-1949, Oct 2000
- [ 154] Hang Jiang Na Robert Elliman, “Electrical Properties of GeSi Surface and Buried Channel p-MOSFETs Fabricated by Ge Implantation”, IEEE Transactions on Electron Devices, vol. 43, no. 1, pp. 97 – 103, 1996
- [155] V. P. Kesan et. Al., “High performance 0.25 micron p-MOSFETs with SiGe channels for 300K and 77K operation”, Technical Digest IEEE Electron Device Meeting pp. 25-28,1991
- [ 156] P. M. Garone, V. Venkataraman and J. C Sturm, “Mobility Enhancement and Quantum Mechanical Modeling in GeSi Channel MOSFET from 90 to 300 K”, IEDM – 91, pp. 29 – 32, 1991
- [ 157] F. Aniel, Enciso-Aguilar, N. Zerunian et. al., “SiGe Hetero FETs on Silicon at Cryogenic Temperature”, J. Phys. IV France, vol. 12, pp. 3-3 – 3-10, 2002
- [ 158] Zhonghai Shi, David Onsongo, Raghaw Rai, Srikanth B. Samavedam and Sanjay K. Banerjee, “ Hole Mobility Enhancement and Si Cap Optimization in Nanoscale Strained  $\text{Si}_{1-x}\text{Ge}_x$  pMOSFETs”, Solid State Electronics, vol. 48, pp. 2299 – 2306, 2004
- [ 159] J. R. Brews, Private communications
- [ 160] Bhongki Mheen, Young-Joo Song, Jin-Young Kang and Songcheal Hang, “Strained SiGe Complementary MOSFETs Adopting Different Thicknesses of Silicon Cap Layers for Low Power and High Performance Applications”, ETRI Journal, vol. 27, nr. 4, 2005
- [ 161 ] D. J. Paul, “Silicon Germanium Heterostructures in Electronics: the Present and the Future”, Thin Solid Films, vol. 321, no. 1-2, pp. 172 – 180, 1998
- [ 162] Steve Hall and Bill Eccleston, “Silicon-Germanium for ULSI”, Journal of Telecommunications and Information Technology, pp. 3 – 8, 2000

- [ 163] Akira Nishiyama, Osamu Arisumi, Mamuru Terauchi et. al., "Formation of SiGe Source/Drain Using Ge Implantation for Floating Body Effect Resistant SOI MOSFETs", *Jpn. J. Appl. Phys.*, vol. 35, Part I, no.2B, 1996
- [ 164] M. Horstmann, A. Wei, T. Kammler et. al., "Integration and Optimization of Embedded SiGe Compressive and Tensile Strained Linear Films and Strain Memorization in Advanced SOI CMOS Technologies", *IEEE Electron Dev. Lett.*, 2005
- [ 165] Toshinori Numata, Toshifumi Irisawa, Tsutomu Tezuka et. al., "Performance Enhancement of Partially and Fully Depleted Strained SOI MOSFETs", *IEEE Trans. Electron Devices*, vol 53, nr. 5, 2006
- [166] D. K. Nayak and S. K. Chun, "Low field hole mobility of strained Si on (100) Si<sub>1-x</sub>Ge<sub>x</sub> substrate", *Appl. Phys. Lett*, vol. 64, p2514, 1994
- [167] Deepak Nayak, K. Goto, A. Yutani, J. Murota and Yasuhiro Shiraki, "High-Mobility strained-Si PMOSFETs", *IEEE Trans. Electron Devices*, vol. 43, no. 10, pp. 1709-1715, Oct 1996
- [168] Sophie Verdonckt, Vandebroek et al., High mobility modulation doped SiGe channel p-MOSFETs, *IEEE Electron Device Lett.*, vol 12, pp. 444-446, 1991
- [169] S. Subbanna et. al., Si/SiGe p-channel MOSFETs, *VLSI Tech. Symp.* pp. 103-104, 1991
- [170] E. Murakami et. al., "Strain controlled SiGe modulation doped FET with ultrahigh hole mobility, *IEEE Device Lett.*, vol. 12, pp. 71-73, 1991
- [171] K. L. Wang, S. G. Thomas and M. O. Tanner, "SiGe band engineering for MOS, CMOS and quantum effect devices, *J. Material Sc: materials for electronics*, vol. 6, pp. 311-323, Oct. 1995
- [172] G. O'Neill and D. A. Antoniadis, "Deep submicron CMOS based on Silicon Germanium technology", *IEEE Trans. Electron Devices*, vol 43, no. 6, pp. 911-918, June 1996
- [173] Sadek, K. Ismail, M. A. Armstrong, D. A. Antoniadis and F. Stern, "Design of Si/SiGe heterojunction complementary metal-oxide-semiconductor transistors, *IEEE Trans. Electron Devices*, vol. 43, no. 8, Aug 1996
- [174] Sophie Verdonckt-Vandebroek, Emmanuel F. Crabb, Bernard S. Meyerson, David Harame, Philip J. Restle, Johannes M. C. Stork and Jeffrey B. Johnson, "SiGe-Channel Heterojunction p-MOSFETs, *IEEE Trans. Electron Devices*, vol. 41, no. 1, pp. 90-101, Jan 1994

- [ 175 ] Tejas Krishnamohan, Christoph Jungemann and Krishna C. Saraswat, “A Novel, Very High Performance, Sub 20nm Depletion Mode Double Gate (DMDG) Si/Si<sub>1-x</sub>Ge<sub>x</sub>/Si Channel pMOSFET”, IEDM 03, pp. 687 – 690, 2003
- [ 176 ] Brian K. Ip and John R. Brews, “Quantum Effects upon Drain Current in a Biased MOSFET”, IEEE Trans. Electron Devices, vol. 45, no. 10, pp. 2213 – 2221, 1998
- [ 177 ] C. Jungemann, N. Subba, J. S. Goo et. al., “Investigation of Strained Si/SiGe Devices by MC Simulation”, Solid State Electronics, vol. 48, no. 8. pp. 1417 – 1422, 2004
- [ 178 ] P. W. Li and W. M. Liao, “Analysis of Si/SiGe Channel pMOSFETs for Deep-submicron Scaling”, Solid State Electronics, vol. 46, no. 1, pp. 39 – 44, 2002
- [ 179 ] Marijan Persun, Branimir Pejcinovic and Sida Zhou, “ Design of Silicon on Insulator (SOI) SiGe p-MOSFET for CMOS Applications”, Proceedings of the Twelfth Biennial University/Government/ Industry Microelectronics Symposym, pp. 103 – 106, 1997
- [180] M. Aoki, T. Ishi, T. Yoshimura, Y. Kiyota, S. Iijma, T. Yamanaka, T. Kure, K. Ohyu, T. Nishida, S. Okazaki, K. Seki and K. Shimohigashi, “0.1 μm CMOS devices using low impurity channel transistors (LICT), IEDM Tech. Dig., pp.939-941, 1990
- [181] Sean T. Ma and John R. Brews, “Comparison of deep submicrometer conventional and retrograde n-MOSFETs”, IEEE Trans. Electron Devices, vol. 47, no.8, August 2000
- [182] S. P. Voinigescu, P. B. Rabkin, C. A. T. Salama and P. A. Blakey, “ 2D Numerical investigation of the impact of compositional grading on the performance of submicrometer Si-SiGe MOSFETs”, IEEE Trans. Electron Devices, vol. 42, no. 6, pp.1039-1046, Jun 1995
- [183] S. Voinigescu, Electron Technology, Institute of Electron Technology, vol. 26, no. 4, 1993, chap. 6
- [184] K. Iniewski, S. Voinigescu, J. Atcha and C. A. T. Salama, “Analytical modeling of threshold voltages in p-channel Si/SiGe/Si MOS structures”, Solid State Electronics, vol. 36, no.5, pp. 775-783, 1993
- [185] J. R. Brews, W. Fichtner, E. H. Nicollian and S. M. Sze, “Generalized guide for MOSFET miniaturization”, IEEE Electron Device Lett. vol. EDL-1, no.1, January 1980

- [186] R. Hull, "Metastable strained layer configurations in the SiGe/Si system", Properties of Strained and Relaxed Silicon Germanium, Erich Kasper ed., EMIS Datareviews Series, INSPEC, 1995
- [187] Stanley Wolf, Silicon Processing for the VLSI Era, chapter. 3, Lattice Press, Sunset Beach, California, 1995
- [188] J. R. Brews, "A charge sheet model of the MOSFET", Solid State Electronics, vol. 21, pp. 345-355, 1978
- [ 189 ] S. Tagaki, M. Iwase and A. ToriUmi, "On Universality of Inversion Layer Mobility in n and p Channel MOSFETs", IEDM Tech. Digest, pp. 398 – 401, 1988
- [ 190 ] Karl E. Fritz, Barbara A. Randall, Gregg J. Fokken, Wayne L. Walters, Michael J. Lorsung, Ann D.Nielsen, Jason F. Prairie, Devon J. Post, David R. Greenberg and Barry K. Gilbert, "High Speed Low Power Digital and Analog Circuits Implemented in IBM SiGe BiCMOS Technology", IEEE Gallium Arsenide Integrated Circuit Symposium, 21st Annual, pp. 41 – 44, 1999
- [ 191] A. M. Stoneham and S. C. Jain, "GeSi Strained Layers and their Applications", Institute of Physics Publishing, Institute of Physics, London, 1995
- [ 192] U. L. Rhode, "A Novel RFIC for UHF Oscillator", IEEE Radio Frequency Integrated Circuits Symposium, Digest of Papers, pp. 53 – 56, June 2000
- [ 193] S. Subbanna, G. Freeman, D. Ahlgren, B. Jagannathan, D. Greenberg, J. Johnson et. al., "Review of Silicon-Germanium BiCMOS Technology after four years of Production and Future Directions", IEEE Gallium Arsenide Integrated Circuit Symposium, 22nd Annual, pp. 7 – 10, Nov. 2000
- [ 194] T. E. Whall and E. H. C. Parker, "SiGe Heterostructures for FET Applications", J. Phys. D: Appl. Phys. Vol. 31, pp. 1397 – 1416, 1998
- [ 195] V. I. Kuznetsov, K. Werner, S. Radelaar, J. W. Metselaar, "Characterisation of an n-type Si/SiGe Modulation Doped Fiedl Effect Transistor", Thin Solid Films, vol. 294, pp. 263 – 266, 1997
- [ 196] Fazal Ali and Aditya Gupta, "HEMTs and HBTs: Devices, Ffbrication and Circuits", Artech House, 685 Canton Street, Norwood, MA02062, USA, 199 State Devices, vol. VI, Addison-Wesley Publishing Co., 1989