STRAINED SILICON AND SILICON-GERMANIUM

QUANTUM DEVICES BY CHEMICAL VAPOR DEPOSITION

JIUN-YUN LI

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
ELECTRICAL ENGINEERING

ADVISOR: JAMES C. STURM

SEPTEMBER 2013
Abstract

Strained SiGe band-to-band tunneling (BTBT) devices and strained Si two-dimensional electron gases (2DEGs) are promising for low-power and quantum computing applications. The objective of this dissertation is to pursue the fundamental understanding of BTBT in strained SiGe films and electron transport properties in strained Si.

We report the first quantitative study of BTBT in strained p⁺-SiGe/n⁺-Si heterojunctions and p⁺-SiGe/n⁺-SiGe homojunctions at forward and reverse biases. Negative differential resistance (NDR) at forward bias is clearly observed for each device, with the highest observed peak current density of $10^4$ A/cm². In reverse bias, a BTBT current density of $10^6$ A/cm² is measured and a model comparison with good agreement is also presented. Furthermore, we demonstrate that the precise modeling of reverse-biased BTBT devices requires the observation of NDR in forward bias.

The surface segregation of phosphorus in relaxed SiGe films is studied with an extremely sharp phosphorus turn-off slope of 6 nm/decade reported. This enables effective Schottky gating on a depletion-mode device of a Si two-dimensional electron gas (2DEG). We also investigate the effect of surface hydrogen on phosphorus segregation. A phenomenological model for this segregation is proposed to explain the experimental results with good agreement.

A 2DEG with a record high mobility of 522,000 cm²/V-s in an isotopically enriched $^{28}$Si quantum well is presented. The estimated electron dephasing time of $\sim 2$ μs is presented. We investigate the effects of different layers in a Si 2DEG structure on
electron mobility and conclude that the remote impurity charges are the dominant source for electron scattering. The reduced segregation of phosphorus enables an inverted modulation-doped Si 2DEG with extremely high mobility of 470,000 cm²/V-s. For the first time second subband occupancy was achieved in a Si quantum well.
Acknowledgements

First, I would like to thank my thesis advisor, Professor James C. Sturm for his guidance and support throughout my doctoral study in Princeton University. His encouragement and constant supports were essential to the completion of this thesis. Furthermore, his approach to advising students has been an invaluable asset for us. I am also very thankful to Professor Sigurd Wagner and Professor Jason Petta for spending time to read this thesis and providing many valuable comments to improve the level of this thesis. I would also like to thank Professor Stephen Chou and Professor Claire Gmachl for their time to serve on my FPO committee.

Special thanks to Professor Daniel Tsui for his encouragement, which was enlightening and inspiring. I greatly appreciate Professor Leonid Rokhinson at Purdue University for his help on magneto-resistance measurements of Si 2DEG samples and valuable inputs. I would also like to thank Professor Jason Petta for stimulating discussions on Si 2DEG work, and Dr. Steve Koester, Dr. Amlan Majumdar, and Dr. Isaac Lauer at IBM Watson Research Center for the stimulating discussions on tunneling work. I also like to thank Drs. Malcolm Carroll, James Ohlhausen, and Tzu-Ming Lu at Sandia National Laboratories for their support on SIMS analysis of isotopically enriched $^{28}$Si 2DEG. Special thanks to Dr. Tzu-Ming Lu for his inputs on the work of Si 2DEGs.

I would like to thank the PRISM staff: Dr. Pat Watson, Dr. Mikhail Gaevski, and Joe Palmer for their excellent support. Special thanks to Pat for taking his time to read this thesis while he was on a trip! It’s been fun and pleasure to work in PRISM cleanroom. I also appreciate the great assistance from Bob Hengstebeck at EAG to offer
very efficient SIMS analysis. This work cannot be done without his support.

I would like to thank the members in Professor Sturm’s group: Kun Yao, Hongzheng Jin, Bahman Hekmatshoar, Sushobhan Avashi, Yifei Huang, Ting Liu, Chiao-Ti Huang, Amy Wu, Ken Nakamatsu, Bhadri Lalgudi Visweswaran, Warren Rieutort-Louis, Josue Sanz-Robinson, Janam Jhaveri, and Yasmin Afsar. Special thanks to Chiao-Ti Huang because in most of my phd life in Princeton, he was the only helper on RTCVD maintenance.

Finally, I would like to thank my family, especially to my wife. Without her tireless and continuous support, I would not have any chance to finish this thesis. She spends all her time to raise our two adorable daughters Irene and Leah together and always comforts me. She is not only a mother of two, a wife, but also the soul mate to me.

To my parents, thanks for being so supportive. Thank you brother to take care of them while I have been ~ 8,000 mile away from home. To my cute daughters, dad kept promise and finally graduated. Let’s go home! Grandparents, we’re coming home!
Contents

Abstract ......................................................................................................................... iii
Acknowledgements ....................................................................................................... v
List of Tables ................................................................................................................ x
List of Figures ............................................................................................................... xii

1 Introduction .............................................................................................................. 1
  1.1 Motivation for Strained Si and SiGe Quantum Devices ................................. 1
  1.2 Thesis Organization ............................................................................................ 2

2 Band-to-Band Tunneling in Strained p⁺-SiGe/n⁺-Si Heterojunctions and
   p⁺-SiGe/n⁺-SiGe Homojunctions ........................................................................... 4
  2.1 Motivation .......................................................................................................... 4
  2.2 Fundamentals of Band-to-Band Tunneling ......................................................... 6
    2.2.1 Negative Differential Resistance ................................................................. 6
    2.2.2 Defect Assisted Tunneling .......................................................................... 10
  2.3 Band-to-Band Tunneling in p⁺-SiGe/n⁺-Si Heterojunctions ............................ 13
    2.3.1 Device Fabrication ...................................................................................... 13
    2.3.2 Effects of Annealing Temperature on Negative Differential
        Resistance ........................................................................................................... 16
    2.3.3 Effects of Electric Field and Bandgap Energy on Band-to-Band
        Tunneling ............................................................................................................ 18
2.4 Band-to-Band Tunneling in $p^+\text{-SiGe}/n^+\text{-SiGe}$ Homojunctions. ........ 21
  2.4.1 Device Fabrication. ........................................... 21
  2.4.2 Forward-Biased Band-to-Band Tunneling. ................... 22
  2.4.3 Reverse-Biased Band-to-Band Tunneling ...................... 27
    2.4.3.1 Effects of Defect-Assisted Tunneling .................. 27
    2.4.3.2 Effects of Series Resistance. ......................... 29
    2.4.3.3 Comparison of Experimental Results and Models. .... 30
  2.5 Summary. ................................................................... 33

3 Surface Segregation of Phosphorus in Relaxed Si$_{0.7}$Ge$_{0.3}$ Layers Epitaxially Grown by Chemical Vapor Deposition. .................. 35
  3.1 Introduction .......................................................... 35
  3.2 Two-State Model ...................................................... 36
  3.3 Failure of Two-State Model ....................................... 40
    3.3.1 Experimental Results. ........................................ 40
    3.3.2 Effects of Hydrogen on Phosphorus Segregation ....... 44
  3.4 Summary ............................................................... 46

4 High Mobility in Modulation-Doped Si Two-Dimensional Electron Gases 48
  4.1 Introduction .......................................................... 48
  4.2 Characteristics of a Si 2DEG ....................................... 49
    4.2.1 Band Offset between Strained Si and Relaxed SiGe Layers ... 49
    4.2.2 Layer Structure and Epitaxial Growth ...................... 50
    4.2.3 Electrostatics. .................................................. 55
    4.2.4 Mobility Model. ................................................ 57
4.3 Efforts Towards High Mobility in Si 2DEGs ............................... 59
  4.3.1 Effects of Phosphorus Background Impurity on 2DEG Mobility . 59
  4.3.2 Reduction of Phosphorus Background Impurity Levels. .......... 61
  4.3.3 Transport Properties of Si 2DEGs with Low Phosphorus Background Impurity Levels. ........................................ 64
4.4 Effects of Layer Structure on 2DEG Mobility ............................ 66
  4.4.1 SiGe Buffer Layer. .................................................. 66
  4.4.2 Si Quantum Well Layer. ........................................... 68
  4.4.3 SiGe Setback Layer ................................................. 70
4.5 Extremely Low Electron Density by Effective Schottky Gating in Single 2DEG Devices. ......................................................... 71
4.6 Summary. ........................................................................... 78

5 Isotopically Enriched $^{28}\text{Si}$ 2DEGs and Inverted Modulation-Doped 2DEGs 80
  5.1 Introduction. ...................................................................... 80
  5.2 Isotopically Enriched $^{28}\text{Si}$ 2DEGs ................................. 81
    5.2.1 Reduction of Spin-Carrying Isotope $^{29}\text{Si}$ by $^{28}\text{Si}$ Enrichment ... 83
    5.2.2 Magneto-Transport Properties of a Modulation-Doped Enriched $^{28}\text{Si}$ 2DEG ................................................................. 85
    5.2.3 Gating of Enhancement-Mode Enriched $^{28}\text{Si}$ 2DEG ........... 87
  5.3 Inverted Modulation-Doped Si 2DEGs ................................. 90
    5.3.1 Effects of Phosphorus Turn-Off Slope on 2DEG Mobility ....... 91
    5.3.2 Effects of Remote Impurity at the Si/Al$_2$O$_3$ Interface on Mobility 94
    5.3.3 Second Subband Occupancy. ......................................... 97
5.3.4 Comparison of 2DEG Mobility in Different Structures .......................... 104
5.4 Summary .......................................................................................... 105
6 Conclusions and Future Work ................................................................. 107
  6.1 Conclusions .................................................................................. 107
  6.2 Future Work .................................................................................. 109
Bibliography ........................................................................................... 111
Appendix Publications and Conference Presentations Resulting from this Thesis ........................................................................................................ 121
Lists of Tables

2.1 Steps of ion implantation of phosphorus for n⁺-Si layer .................. 13
2.2 Growth parameters of p⁺-SiGe layers .......................................... 15
2.3 Growth parameters of p⁺-SiGe and n⁻-SiGe layers ......................... 22
4.1 Layer structures of sample #5858, #5838, and #5862 for study of the effects of SiGe buffer layer thickness on the 2DEG mobility ....................... 67
4.2 Layer structures of test samples for study of the effects of Si QW layer thickness on the 2DEG mobility .............................................. 68
4.3 Layer structures of test samples for study of the effects of SiGe setback layer thickness on the 2DEG mobility ........................................... 70
5.1 Layer structures of isotopically enriched ²⁸Si 2DEGs ....................... 82
5.2 Layer structures of sample #5457, #5850, and #5630 for study of the effects of P turn-off slope on 2DEG characteristics in an inverted modulation-doped structure ......................................................... 92
5.3 Layer structures of sample #5877 and #5630 for study of the effects of impurity charges at the Si/Al₂O₃ interface on 2DEG mobility ........... 94
## Lists of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Schematics of (a) n-type MOSFET and (b) n-type TFET in the ON state</td>
<td>4</td>
</tr>
<tr>
<td>2.2</td>
<td>Band diagrams of n-type TFET at (a) OFF state and (b) ON state.</td>
<td>5</td>
</tr>
<tr>
<td>2.3</td>
<td>A degenerately doped p-n junction in Si at a small forward bias. $eV_n$ and $eV_p$ represent the energy difference between the Fermi level in the n-type region and the conduction band edge, and the Fermi level in the p-type region and the valence band edge, respectively.</td>
<td>7</td>
</tr>
<tr>
<td>2.4</td>
<td>Band diagrams (top) and I-V curves (bottom) of a tunneling diode at different biases.</td>
<td>8</td>
</tr>
<tr>
<td>2.5</td>
<td>Band diagram in a degenerately doped p-n junction, which shows the possible paths of electron tunneling from the n-type conduction band to the p-type valence band via defects in the bandgap [26]. The shadowing areas represent the energy bands filled with electrons.</td>
<td>11</td>
</tr>
<tr>
<td>2.6</td>
<td>Three main current components in a tunneling diode in practice: BTBT current (blue), DAT current (red), and thermal current (green) with their sum also shown (black). There is no NDR because DAT dominates over BTBT due to high defect density in the p-n junction.</td>
<td>12</td>
</tr>
<tr>
<td>2.7</td>
<td>Fabrication steps of p$^+$-SiGe/n$^+$-Si heterojunction tunneling diodes.</td>
<td>14</td>
</tr>
<tr>
<td>2.8</td>
<td>SIMS analysis of a p$^+$-Si$<em>{0.73}$Ge$</em>{0.27}$/n$^+$-Si heterojunction tunneling diode (sample #5150). The slopes of B trailing edge and P leading edge are 13 and 14 nm/decade, respectively.</td>
<td>15</td>
</tr>
<tr>
<td>2.9</td>
<td>J-V curves of p$^+$-Si$<em>{0.73}$Ge$</em>{0.27}$/n$^+$-Si heterojunction tunneling diodes at different</td>
<td>xii</td>
</tr>
</tbody>
</table>
annealing temperatures. ........................................ 17
2.10 $J_{\text{peak}}$ and $J_{\text{DAT}}$ (at 0.3 V) of $p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si}$ heterojunction tunneling
   diodes vs. annealing temperature. ............................ 18
2.11 J-V curves of $p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si}$ heterojunctions tunneling diodes of three
different boron concentrations.................................. 19
2.12 J-V curves of $p^+\text{-Si}_{1-x}\text{Ge}_x/n^+\text{-Si}$ tunneling diodes of different Ge fractions .................. 20
2.13 $J_{\text{peak}}$ vs. $\Delta V_{\text{peak}}$ for different Ge fractions. The linear dash line represents the
   spreading resistance of $n^+\text{-Si}$ substrate....................... 20
2.14 J-V curves of $p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si}_{0.73}\text{Ge}_{0.27}$ homojunction tunneling diodes
   of three different boron levels ................................ 23
2.15 J-V curves of $p^+\text{-Si}_{1-x}\text{Ge}_x/n^+\text{-Si}_{1-x}\text{Ge}_x$ homojunction diodes of four different
   Ge fractions of 14, 21, 27, and 35 % ......................... 24
2.16 Correction procedure of peak current density for each Ge fraction to a single
   set of fixed doping levels ........................................ 25
2.17 $J_{\text{peak}}$ vs. Ge fraction. Blue squares are the experimental data, red squares
   represent corrected data to a single set of fixed doping levels ($N_A = 1.2 \times 10^{20}$ cm$^{-3}$
   and $N_D = 0.7 \times 10^{20}$ cm$^{-3}$), and solid line is the model prediction based on Eqs.
   (2.4), (2.7), and (2.8) ........................................... 26
2.18 Schematic of band energy diagram of $p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si}$ heterojunction
to show the processes of band-to-band tunneling and defect-assisted tunneling
   in reverse bias. ................................................... 27
2.19 J-V curves of $p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si}$ heterojunction tunneling diodes with three
   annealing temperatures to show the importance of the presence of NDR
in forward bias on BTBT current in reverse-bias. ...................... 28

2.20 J-V curves of Si$_{0.73}$Ge$_{0.27}$ homojunction tunneling diodes with different mesa widths .................................................. 29

(a) An enlarged view of J-V curves in Fig. 2.20 at small reverse biases, and (b) reverse-biased current density vs. $W^{-1}$ at $V = -0.1$, -0.3, -0.5, -0.7, and -0.9 V, respectively ........................................ 30

2.22 Reverse-biased BTBT current density per volt ($J_{BTBT, RB}/V$) vs. electric field in SiGe BTBT. Symbols are the experimental data and multiple lines represent model prediction based on Eqs. (2.10) and (2.12). An inset provides an enlarged view at high electric field of 2 to 3 $\times$ $10^8$ V/m. ......................... 31

2.23 $J_{BTBT, RB}$ vs. Ge fraction at $E_{field}$ of 2 $\times$ $10^8$ V/m. Points are the experimental data with error bars, representing the effects of the variations of electric field (the horizontal error bars in Fig. 2.22) ......................... 33

3.1 Schematic of atomic layer structures near the surface during the epitaxial growth 38

3.2 Schematic of phosphorus energy near the surface during the epitaxial growth in a two-state model. Layer 2 represents the surface layer and layer 1 (sub-surface layer) represents the next layer below the surface ......................... 39

3.3 Phosphorus profiles in strained and relaxed Si$_{0.7}$Ge$_{0.3}$ layers grown at 575 °C under the same growth conditions [56]. The growth rates of strained and relaxed Si$_{0.7}$Ge$_{0.3}$ layers are 8 nm/min and 6 nm/min, respectively .................. 41

3.4 Phosphorus profiles in relaxed Si$_{0.7}$Ge$_{0.3}$ layers grown at 500, 550, and 600 °C. Phosphorus supply was turned off at the depth of 45 nm. P turn-off slopes were 127, 40, and 9 nm/decade for 600, 550, and 500 °C, respectively 42
3.5 Phosphorus turn-off slope vs. growth rate for different temperatures.
Experimental results (points) and the theoretical prediction (lines) are presented for comparison [56] 43

3.6 Comparison of experimental data and modified TSM of phosphorus turn-off slope vs. growth rate by including the effect of surface hydrogen on the segregation energy ($E_{\text{surf}}$) [56] 45

3.7 Segregation energy ($E_{\text{surf}}$) vs. growth temperature. At 575 °C, $E_{\text{surf}}$ for hydrogen pressure of 2, 6, 23 torr are also plotted to confirm the effect of surface hydrogen on P surface segregation [56] 46

4.1 (a) Band diagram of a relaxed-SiGe/strained-Si/relaxed-SiGe heterostructure. The energy levels in the conduction band of strained Si are split into two states: $\Delta_2$ and $\Delta_4$; (b) under tensile strain, electrons reside in the two states of $\Delta_2$ along (001) in Si with an in-plane effective mass of $m^* = 0.19 m_0$[75] 50

4.2 (a) Layer structure of a Si 2DEG and (b) the associated band diagram with a Ge fraction of 0.27 in a Si/SiGe heterostructure without gate bias 51

4.3 SIMS analysis of a typical Si 2DEG (sample #5737). Ge, P, B, C, and O were measured from the surface to below the growth interface at the depth of 255 nm 53

4.4 Top view of a typical Hall bar device. Hall bar was first mesa-etched by RIE followed by alloyed AuSb (1 % Sb) for electrical contacts 54

4.5 Conduction band diagram of a Si 2DEG at 4 K with two boundary conditions for the Fermi level ($E_F$) to be pinned at (i) the midgap of Si surface and (ii) the donor level 56

4.6 SIMS results of a Si 2DEG (sample #5144). The phosphorus background levels in
Si and SiGe layers are $6 \times 10^{17}$ cm$^{-3}$ and $2 \times 10^{17}$ cm$^{-3}$, respectively.

4.7 New gas supplying system with a separation of process gases (H$_2$, SiH$_4$, GeH$_4$, etc.) from PH$_3$ to reduce the memory effect of phosphorus.

4.8 B, P, Ge profiles of a test structure of multiple Si and SiGe layers grown. Both B and P levels are down to the SIMS detection limits of $5 \times 10^{15}$ and $3 \times 10^{14}$ cm$^{-3}$, respectively except a phosphorus doped layer grown at the depth of 360 nm.

4.9 P and B profiles in Si layers grown at different temperatures (sample #5823).

4.10 Hall mobility vs. electron density at low temperatures (4 K or 0.3 K). Highest mobility of 522,000 cm$^2$/V-s was measured at 0.3 K [84]. Note each data point was taken from different samples with different layer structures.

4.11 Longitudinal ($R_{xx}$) and transverse ($R_{xy}$) magneto-resistances vs. magnetic field at 0.3 K [84].

4.12 Electron Mobility at 4 K vs. SiGe buffer layer thickness.

4.13 Mobility at 4 K vs. Si QW layer thickness grown at 575 °C and 625 °C. The arrow indicates the predicted critical thickness of strained Si on relaxed Si$_{0.73}$Ge$_{0.27}$ [87].

4.14 (a) Electron density and (b) mobility vs. SiGe setback layer thickness at 4 K.

4.15 Top view of the Hall bar devices. The contacts were made by AuSb alloying (1 % Sb) and top Pd Schottky gate covers the Hall bar.

4.16 SIMS results of Ge and P profiles of (a) sample #5414 and (b) sample #5613. The phosphorus level at the surface are $2 \times 10^{18}$ and $2 \times 10^{16}$ cm$^{-3}$ for sample #5414 and #5613, respectively.
4.17 Gate leakage current density at 4 K vs. gate voltage for sample #5414 and #5613, respectively. Pd Schottky gate covered the entire Hall bar shown in Fig. 4.15.

4.18 Layer structures of (a) sample #5737 and (b) sample #5747. The phosphorus doping levels are $1 \times 10^{18}$ and $3 \times 10^{18}$ cm$^{-3}$ for (a) and (b), respectively.

4.19 SIMIS results of Ge and P profiles for (a) #5737 and (b) #5747. The bump of P in the Si QW layer is thought the artifact of SIMS measurement.

4.20 (a) Gate leakage current ($I_g$) and (b) electron density vs. gate voltage ($V_g$) for sample #5737 at 4 K. At $V_g = 0.5$ V, a normal forward-bias Schottky diode operation was observed. The drive current for Hall measurement was 100 nA.

4.21 Hall mobility vs. density at 4 K for sample #5737 (setback: 50 nm) and #5747 (setback: 20 nm).

5.1 Three isotopes $^{28}$Si, $^{29}$Si, and $^{30}$Si vs. depth in 2DEG structure with Ge as an indicator by SIMS measurements. The growth started at the depth of 185 nm and Si QW is at the depth 75 nm.

5.2 Dephasing time vs. $^{29}$Si fraction. Solid line is the model prediction, the solid square is the experimental result in Si QD [13], and the vertical dash line represents the fraction of $^{29}$Si in our enriched $^{28}$Si 2DEGs. Dephasing time of GaAs QDs [12] is also presented for comparison.

5.3 Magneto-resistances of depletion-mode enriched $^{28}$Si 2DEG device measured at 0.3 K. Electron density ($4.02 \times 10^{11}$ cm$^{-2}$) and mobility (522,000 cm$^2$/V-s) were extracted from the periods of Shubnikov-de Haas oscillations in longitudinal resistance ($R_{xx}$) vs. (1/B) and its value at zero field.
5.4 (a) The layer structure of enhancement-mode devices of Si 2DEGs, and
(b) the top view of the Hall bar geometry ........................................ 87

5.5 Electron density vs. gate voltage by Hall measurement at 4 K for
enhancement-mode devices of enriched $^{28}$Si 2DEG with a SiGe setback
layer of 60 and 150 nm. The slopes represent the effective capacitance
between a Cr/Au gate and 2DEG .................................................. 88

5.6 Hall mobility vs. density for enhancement-mode devices of enriched $^{28}$Si
2DEG with a SiGe setback layer of 60 and 150 nm at 4 K ................. 89

5.7 Typical layer structure of an inverted modulation-doped Si 2DEGs. The
SiGe supply layer of phosphorus modulation-doping is below the 2DEG
layer (Si QW) ................................................................. 90

5.8 P and Ge profiles of sample #5457, #5850, and #5630. P turn-off slopes
of these three samples are 40, 14, and 8 nm/dec, respectively. The setback
layer thicknesses between the phosphorus peak level and the lower Si/SiGe
heterojunction are 20, 20, and 33 nm for those three samples ............ 92

5.9 Hall density and mobility vs. P turn-off slope for sample #5457 (40 nm/dec),
#5850 (14 nm/dec), and #5630 (8 nm/dec). A corrected electron density was
calculated by Poisson’s equation to compensate the effect of the thicker SiGe
setback layer in sample # 5630 .................................................. 93

5.10 SIMS profiles of (a) #5877 and (b) #5630 of different top SiGe cap thicknesses 95

5.11 Hall mobility vs. density at 4 K for sample #5630 (50 nm from the 2DEG
to the surface) and #5877 (26 nm from the 2DEG to the surface) for a
comparison of the effect of upper SiGe cap layer thickness ............ 96
5.12 (a) Hall density and (b) Hall mobility at 0.3 K vs. \( V_g \) of sample #5630.

The onsets of the intersubband scattering and second subband occupancy occurs at \( V_g = 3.2 \) V and 3.5 V, respectively. In region (i), only the first subband is populated. For region (ii), some electrons in the first subband scatter into and are trapped in the localized states, so the mobility drops.

In region (iii), electrons reside in both the first and second subbands, so that the measured Hall mobility increases with electron density because of stronger screening.

5.13 Hall mobility vs. density at 0.3 K of sample #5630.

5.14 (a) Hall density and (b) Hall mobility at 0.3 K vs. \( V_g \) of sample #5630. In region (i) and (ii), \( n_{\text{Hall}} \) and \( \mu_{\text{Hall}} \) represent \( n_1 \) and \( \mu_1 \), respectively. For region (iii), assuming \( n_2 \) increases from zero with \( V_g \) (at \( V_g = 3.5 \) V) by a simple parallel capacitor model, and \( n_1 \) is constant, \( \mu_1 \), and \( \mu_2 \) were solved by Eqs. (5.1) and (5.2).

5.15 Magneto-resistances (\( R_{xx} \) and \( R_{xy} \)) vs. magnetic field at 0.3 K of sample #5630 at (a) \( V_g = 3 \) V and (b) \( V_g = 4 \) V.

5.16 Mobility vs. density at 4 K for different structures of Si 2DEGs in this thesis.
Chapter 1  Introduction

1.1  Motivation for Strained Si and SiGe Quantum Devices

Silicon has been dominating the semiconductor industry for over forty years because of the widely used Si metal-oxide-semiconductor field-effect transistor (MOSFET), the basic unit of a logic device. Si MOSFETs are probably the most fabricated electronic devices; for example, in 2011, there were more than $10^8$ computers sold [1]. Assuming there were $10^9$ transistors in a CPU such as Intel Core i7 [2] for each computer, at least $10^{17}$ Si MOSFETs were fabricated in a year.

The great success of Si MOSFETs may be attributed to Moore’s law, which has successfully predicted the pace of transistor development over recent decades. However, to follow Moore’s law further, several critical issues such as gate leakage, junction leakage, low transconductance, interconnect capacitance, and subthreshold conduction [3] must be solved. Furthermore, as a device is further scaled down, quantum mechanics must be used for nano-devices, and novel processing techniques have to be employed to create atom-sized structures in such small devices.

In the quantum regime, several interesting proposals for the next generation of computing devices have been proposed, such as tunneling transistors [4] and quantum computing [5]. A tunneling field-effect transistor (TFET) was first proposed by Baba [6] in 1992. The major benefit of this device is the low leakage current in the OFF state due to the sharp subthreshold slope, which is crucial for low-power applications.
Furthermore, because its structure is similar to a MOSFET and its fabrication steps are compatible with those of MOSFETs [7], remarkable TFET efforts have been made in the past decade.

On the other hand, for better system performance, quantum computing has been suggested to out-perform the classical computation by Shor [8]. He proposed a quantum algorithm with a much faster computation speed, as high as $10^{200}$ bits/second, by harnessing the power of quantum superposition in a quantum system. There are several candidates for the implementation of quantum computing, such as nuclear magnetic resonance [9], superconducting devices [10], and double quantum dots in semiconductors [11]. Quantum dot devices have drawn increasing attention from physicists and material scientists recently since the successful demonstrations of coherence control of quantum bits in GaAs [12] and Si [13]. The scalability and compatibility with semiconductor technology make quantum dots more feasible and promising than other proposed devices.

### 1.2 Thesis Organization

In this thesis, we first report band-to-band tunneling in strained SiGe for TFET applications. Then we present the electron transport properties in strained Si two-dimensional electron gas for quantum dot applications.

In chapter 2, band-to-band tunneling in strained $p^+\text{-SiGe}/n^+\text{-Si}$ heterojunctions and $p^+\text{-SiGe}/n^+\text{-SiGe}$ homojunctions, prepared by chemical vapor deposition (CVD), is investigated, and the experimental results are presented with a model comparison. This experimental verification of the model allows the device designers to predict the
dependence of band-to-band tunneling on germanium fraction and can be used in device simulators for the prediction of TFET performance and the relevant devices.

In chapter 3, the surface segregation of phosphorus in relaxed SiGe films grown by CVD is investigated. The effect of surface hydrogen on phosphorus segregation is explored and a phenomenological model including surface hydrogen is proposed. The model is in good agreement with the experimental results. A record sharp turn-off slope for phosphorus of 6 nm/dec is achieved.

A two-dimensional electron gas (2DEG) in strained Si is studied in chapter 4 for quantum dot applications. First, the basic physics of a 2DEG is reviewed. Efforts towards achieving high electron mobility in strained Si are described. The experimental results and the effects of layer structure on electron mobility are also presented.

For quantum computing applications, spin decoherence must be reduced for better device performance. $^{29}$Si induces spin decoherence due to its nuclear spin via hyperfine interactions. In chapter 5, we present our work on the growth of 2DEGs in isotopically enriched $^{28}$Si quantum wells. Furthermore, we report extremely high electron mobility in an inverted modulation-doped Si 2DEG grown by low-temperature epitaxy.

Finally in chapter 6, a brief summary is presented to conclude this thesis, and is followed by a few suggestions for future work.
Chapter 2  Band-to-Band Tunneling in Strained $p^+\text{-SiGe}/n^+\text{-Si}$ Heterojunctions and $p^+\text{-SiGe}/n^+\text{-SiGe}$ Homojunctions

2.1  Motivation

Tunneling is a fundamental quantum mechanical process. While quantum mechanics was developed in the 1920’s, the first tunneling device was demonstrated in 1958 by Esaki in a heavily doped p-n junction of Ge [14]. Since then, semiconductor quantum devices, such as resonant tunneling diodes [15], quantum cascade lasers (QCLs) [16], and tunneling field-effect transistors (TFETs) [17], have been of great interest to scientists and device engineers. In the past five years particularly, extensive studies on TFETs have been conducted because of their potential for low-power applications [4].

Unlike a conventional metal-oxide-semiconductor FET (MOSFET), a TFET is operated by switching its tunneling junction on and off between the source and channel.

![Schematics of (a) n-type MOSFET and (b) n-type TFET in the ON state.](image)

Fig. 2.1  Schematics of (a) n-type MOSFET and (b) n-type TFET in the ON state.
regions. The device structures of an n-type MOSFET and an n-type TFET are illustrated in Fig. 2.1. A p⁺-source is used in an n-type TFET, instead of an n⁺-source in an n-type MOSFET. If the gate voltage \(V_g\) is below the threshold voltage \(V_{th}\), there is ideally no available state for electrons below the Fermi level, in the valence band of the p⁺-source, to tunnel into the bandgap of the channel region for current conduction (Fig. 2.2 (a)). Furthermore, the distance for electrons in the source region to directly tunnel to the conduction band of the drain region is fairly large, leading to a very small leakage current at OFF state. On the other hand, when \(V_g\) is larger than \(V_{th}\), the tunneling barrier between the source and channel regions is greatly reduced, such that electrons in the valence band of p⁺-source can tunnel for current conduction (Fig. 2.2(b)). Because of this tunneling nature, the subthreshold slope can be smaller than the thermal limit of 60 mV/decade in a Si MOSFET at room temperature [18].

Si TFETs demonstrate great promise for low-power applications because of its sharp subthreshold slope. However, the ON state current in a Si TFET is fairly low.
because of the large tunneling barrier resulting from the large bandgap of Si, limiting the potential for device performance (i.e. switching speed). SiGe has been considered a candidate because of its smaller bandgap energy and the compatibility of Si technology [19]. Furthermore, theoretical work has predicted that the subthreshold slope can be particularly sharp [20]. Two types of SiGe TFETs were proposed: (i) a heterojunction of p^+-SiGe source with a Si channel [20] and (ii) a homojunction of p^+-SiGe source with a SiGe channel [21]. In the following sections, we present the work on band-to-band tunneling in both p^+-SiGe/n^+-Si heterojunctions and p^+-SiGe/n^+-SiGe homojunctions, which serves as a baseline for TFET device modeling.

2.2 Fundamentals of Band-to-Band Tunneling

2.2.1 Negative Differential Resistance

The most important feature of band-to-band tunneling in a heavily doped p-n junction is the presence of negative differential resistance (NDR) in forward bias [22]. To observe NDR, two requirements have to be met. Firstly, a small tunneling barrier resulting from a large electric field (usually > 10^8 V/m) in the p-n region is required for tunneling to occur. Secondly, the Fermi level \( (E_F) \) must be located within the valence band in the p-type region and in the conduction band in the n-type region (Fig. 2.3); otherwise, a monotonic increase of tunneling current with applied voltage would be observed, such as Zener tunneling in reverse bias. Usually, in a degenerate doped p-n junction, both requirements are satisfied. For example, if the doping levels are \( 5 \times 10^{19} \) cm\(^{-3} \), the electric field can be as large as \( 3 \times 10^8 \) V/m by a device simulator [23],
Fig. 2.3 A degenerately doped p-n junction in Si at a small forward bias. $eV_n$ and $eV_p$ represent the energy difference between the Fermi level in the n-type region and the conduction band edge, and the Fermi level in the p-type region and the valence band edge, respectively.

leading to a short tunneling distance of ~ 10 nm. At this doping level, $E_{fp}$ is 88 meV below the edge of the valence band in the p-type region, and $E_{fn}$ is 94 meV above the conduction band edge in the n-type region, at small forward biases.

In Fig. 2.4, the band diagrams (top) and the associated I-V curves (bottom) of a degenerately doped tunneling diode are plotted. Initially, in reverse bias, Zener tunneling occurs with the electrons in the valence band of the p-type region tunneling to the conduction band of the n-type region (Fig. 2.4(a)). As the negative bias is increased, the electric field in the p-n junction increases. Therefore, the tunneling barrier for electrons becomes smaller, leading to a higher current.

At forward bias, the tunneling direction reverses. The electrons in the conduction band of the n-type region now tunnel to the available states in the valence band of the p-type region at the same energy level. At a certain bias ($V_2$), a band of electron energy in the n-type conduction band, aligned with a band of the unoccupied states in the p-type valence band, leads to a peak current in forward bias (Fig. 2.4(b)). As the applied voltage is increased slightly, some electrons at the highest energy levels in
the n-type conduction band cannot tunnel to the p-type valence band for current conduction, since there is no available state at the same energy levels in the bandgap of the p-type region (red arrow in the top part of Fig. 2.4(c)). As a result, the current starts to drop (Fig. 2.4(c)). As the entire band filled with electrons in the n-type region is completely “uncrossed” with all available states in the valence band of the p-type region, tunneling ceases and the current becomes zero (Fig. 2.4(d)). When the applied bias is further increased, more electrons (holes) in the conduction (valence) band of n-type (p-type) region have enough energy to move towards the junction, with normal diode characteristics (Fig. 2.4(e)).

Typically, the WKB approximation (Wentzel-Kramers-Brillouin method) is used to estimate the tunneling probability $T_t$ [24]:

$$T_t \approx \exp \left[ -2 \int_{r_p}^{r_e} |k(x)| \, dx \right],$$

(2.1)

where $|k(x)|$ is the absolute value of the wavevector of the electron in the tunneling
barrier, and $x_p$ and $x_n$ are the positions of the valence band edge and conduction band edge in the p-type and n-type regions, respectively. The wavevector is

$$k(x) = \sqrt{\frac{2m^*}{\hbar^2}} (PE - E),$$

(2.2)

where $m^*$ is the effective mass of the tunneling electron, $PE$ is the potential energy, and $E$ is the electron energy (top of Fig. 2.4 (b)). The tunneling probability is

$$T_t \approx \exp \left(-\frac{\pi \sqrt{\frac{m^* E_g^3}{2\sqrt{2}e\hbar F_{field}}}}{2 E_{\perp}} \right) \exp \left(-\frac{2E_{\perp}}{E_{eff}} \right),$$

(2.3)

where $E_g$ is the bandgap energy, $F_{field}$ is the peak electric field in the p-n junction, $E_{\perp}$ is the energy associated with momentum perpendicular to the tunneling direction, and $E_{eff}$ is the measure of transverse momentum as given by [24]

$$E_{eff} = \frac{4\sqrt{2}e\hbar F_{field}}{3\pi \sqrt{m^* E_g}}.$$  

(2.4)

In thermal equilibrium, the net tunneling current $I_{net}$ in a tunneling diode from the p-type region to the n-type region is the difference between tunneling current of $I_{p\rightarrow n}$ and $I_{n\rightarrow p}$, as given by

$$I_{p\rightarrow n} = B \int_{E_{C}}^{E_T} F_C(E)n_T(E)T_t[1-F_C(E)]n_C(E)dE,$$

(2.5)

$$I_{n\rightarrow p} = B \int_{E_{C}}^{E_T} F_T(E)n_C(E)T_t[1-F_T(E)]n_T(E)dE,$$

(2.6)

where $B$ is a pre-factor, $F_C(E)$ and $F_T(E)$ are the Fermi-Dirac distribution functions, $T_t$ is the tunneling probability and is assumed to be equal for both directions, and $n_C(E)$ and $n_T(E)$ are the density of states in the conduction band and valence band respectively.
Thus, the peak tunneling current can be calculated as [25]

$$I_{\text{peak}} = A \frac{e m^*}{4\pi^2\hbar^2} \exp \left( -\frac{\pi \sqrt{m^* E_g^3}}{2\sqrt{2e}\hbar E_{\text{field}}} \right) E_{\text{eff}} \cdot D,$$

(2.7)

$$D \equiv \int [F_c(E) - F_v(E)] \cdot \left[ 1 - \exp \left( -\frac{2E_{\text{eff}}}{E_{\text{eff}}} \right) \right] dE,$$

(2.8)

where $A$ is the device area, $D$ is the effective density of states, $E_S$ is the smaller of $E_{fp}$ and $E_{fn}$, which are the energies of the hole and electron, measured from the edges of the p-type valence band and n-type conduction band respectively.

To enhance tunneling current for TFET applications, a large electric field and a small bandgap energy are preferred. By increasing the p-type doping level in the source region of an n-type Si TFET, a higher ON state current is expected. Moreover, by replacing Si with SiGe in the source region, the tunneling barrier between the source-channel junction becomes smaller due to the smaller bandgap energy of SiGe than Si. Therefore, the tunneling can be boosted further [7, 21]. Consequently, in this chapter we focus on the dependence of band-to-band tunneling (BTBT) on the p-type doping level and Ge fraction, in both forward and reverse biases. Experimental results and model comparisons will be presented in the following sections.

### 2.2.2 Defect-Assisted Tunneling

In the preceding section, the fundamentals of tunneling were introduced qualitatively and quantitatively, based on an assumption that the p-n junction is free of defects. The derivation of the tunneling model was proceeded by assuming an ideal p-n
junction without any defects. However, in practice, defects exist in the junction and affect the tunneling characteristics. For all reported tunneling junctions made by degenerately doping, an “excess” current at biases between negative differential resistance and the onset of the thermal current was observed [22], and is defined as defect-assisted tunneling (DAT) current in this thesis. Chynoweth et al. suggested that this excess current results from an additional tunneling mechanism via junction defects [26] at biases, where band-to-band tunneling is prohibited due to the complete misalignment of energy bands of the n-type and p-type regions.

By intentionally introducing defects into the junction, Sah [27] and Weaver [28] confirmed the existence of defect-assisted tunneling, and showed that NDR disappeared when this excess tunneling current swamps the BTBT current. Thus, for an accurate measure of BTBT in a p-n junction, the defect-assisted tunneling component has to be eliminated (or heavily reduced) so that NDR can be observed.

In Fig. 2.5, a tunneling junction with defect states in the bandgap is illustrated [26]. At this bias, the electron band in the n-type region is completely uncrossed with the valence band in the p-type region, and is still far away from the normal diode operation;
thus, there is no tunneling. With the presence of defects in the junction, electrons in the conduction band of the n-type region can actually move to the valence band of the p-type region via several paths, such as (i) $C \to D_p \to V$, (ii) $C \to D_n \to V$, or (iii) $C \to V$ directly. For (i), electrons first tunnel to the defect state $D_p$ and recombine with holes in the valence band. Conversely, for (ii) electrons move to the defect state $D_n$ first, and then tunnel to the valence band. The last route for defect-assisted tunneling involves the existence of continuous defect levels between the conduction band and the valence band, where an electron loses its energy by these continuous transitions.

For simplicity, Chynoweth et al. assumed that path (ii) would dominate and that the limiting step is the tunneling from $D_n$ to $V$, but not the recombination of $C$ to $D_n$ [26]. The defect-assisted tunneling (DAT) current is given by

$$I_{DAT} = K D_x \exp \left\{ -\alpha_x \left[ E_g - eV + 0.6 (E_n + E_p) \right] \right\},$$

(2.9)

where $K$ is a pre-factor, $D_x$ is the defect density, $\alpha_x$ is a constant, and $V$ is the applied

![Fig. 2.6 Three main current components in a tunneling diode in practice: BTBT current (blue), DAT current (red), and thermal current (green), with their sum also shown (black). There is no NDR, because DAT dominates over BTBT due to high defect density in the p-n junction.](image-url)
voltage. The main factors of DAT current are the defect density, bandgap energy, and the applied voltage (electric field). With a smaller bandgap or a larger bias (stronger electric field), the tunneling probability from D to V becomes higher, with the result that $I_{DAT}$ increases. Between the peak voltage and the onset of the thermal current, the current never falls to zero because of the presence of defect-assisted tunneling. Fig. 2.6 shows I-V curves of a tunneling diode in practice with high defect density, where DAT swamps BTBT at small biases so that NDR disappears.

### 2.3 Band-to-Band Tunnelling in $p^+-SiGe/n^+Si$ Heterojunctions

#### 2.3.1 Device Fabrication

For tunneling FET applications, SiGe has been suggested to replace Si for higher tunneling currents due to its smaller bandgap energy [20]. In this section we focus on band-to-band tunneling in $p^+-SiGe/n^+Si$ heterojunctions. Since the level of n-type dopant is limited to the level of $10^{19}$ cm$^{-3}$ due to the low incorporation [29], for the heavily doped n$^+$-Si layer, we used ion implantation of phosphorus with multiple steps to achieve a high doping level of $2 \times 10^{20}$ cm$^{-3}$ (Table 2.1). Due to the high-dose implantation, Si substrates became amorphous and defective. Thus, an annealing step is required to re-crystallize the Si substrates and to remove the defects by implantation.

<table>
<thead>
<tr>
<th>Species: Phosphorus</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy (keV)</td>
<td>15</td>
<td>40</td>
<td>80</td>
<td>120</td>
</tr>
<tr>
<td>Dose (cm$^2$)</td>
<td>$5 \times 10^{14}$</td>
<td>$7 \times 10^{14}$</td>
<td>$1 \times 10^{15}$</td>
<td>$2 \times 10^{15}$</td>
</tr>
</tbody>
</table>
Prior to being annealed, the implanted Si (100) wafers of phosphorus level of \( \sim 10^{20} \text{ cm}^{-3} \) were cleaned by \( \text{H}_2\text{SO}_4: \text{H}_2\text{O}_2 \) (2.5:1) for 15 minutes followed by diluted HF (1:100) in deionized (DI) water for 2 minutes to remove the residual oxide. Then the wafers were loaded into the CVD reactor for high temperature annealing (700 \( \sim \) 1050 °C) for 5 minutes, followed by \( \text{p}^{+}-\text{SiGe} \) epitaxial growth. The growth pressure was 6 torr. The temperatures for SiGe growth were 625 °C for Ge fractions of 14, 21, and 27%, and 575 °C for 35 and 39%. The gas precursors were dichlorosilane (\( \text{SiH}_2\text{Cl}_2 \)) and diluted germane (0.8% \( \text{GeH}_4 \) in hydrogen) for SiGe growth. Diluted diborane (100 ppm in hydrogen) was used for in-situ p-type doping. The thicknesses of the SiGe films were kept below the critical thickness for each Ge fraction to avoid strain relaxation, which could induce dislocation defects in the junction. After the epitaxial growth, the wafers were mesa-etched by reactive-ion etching (RIE) to isolate the p-n junction and to define the device area of 25 \( \mu \text{m} \times 25 \mu \text{m} \). The etching gases were \( \text{CF}_4 \) of 50 sccm and \( \text{O}_2 \) of

**Fig. 2.7** Fabrication steps of \( \text{p}^{+}-\text{SiGe}/n^{+}-\text{Si} \) heterojunction tunneling diodes.
10 sccm. The chamber pressure was 100 mtorr and the RF power was 100 W. The etching rate of SiGe films was \(\sim 100\) nm/min. Lastly, Ti/Al was deposited on the top and bottom surfaces for electrical contacts. The details of the fabrication steps are illustrated in Fig. 2.7. I-V measurements were performed at room temperature.

<table>
<thead>
<tr>
<th>Ge Fraction (%)</th>
<th>Gas flow rates (sccm) (DCS/GeH(_4)/B(_2)H(_6))</th>
<th>Growth Rate (nm/min)</th>
<th>Thickness of p(^{+})-SiGe (nm)</th>
<th>Doping level (cm(^{-3}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 (625(^\circ)C)</td>
<td>26/50/100</td>
<td>3</td>
<td>25</td>
<td>(1.0 \times 10^{20})</td>
</tr>
<tr>
<td>21 (625(^\circ)C)</td>
<td>26/100/125</td>
<td>6.5</td>
<td>44</td>
<td>(1.3 \times 10^{20})</td>
</tr>
<tr>
<td>27 (625(^\circ)C)</td>
<td>26/200/167</td>
<td>10</td>
<td>20</td>
<td>(1.8 \times 10^{20})</td>
</tr>
<tr>
<td>35 (575(^\circ)C)</td>
<td>26/300/250</td>
<td>6.8</td>
<td>16</td>
<td>(2.5 \times 10^{20})</td>
</tr>
<tr>
<td>38 (575(^\circ)C)</td>
<td>26/400/350</td>
<td>5.7</td>
<td>14</td>
<td>(0.8 \times 10^{20})</td>
</tr>
</tbody>
</table>

Fig. 2.8 SIMS analysis of a p\(^{+}\)-Si\(_{0.73}\)Ge\(_{0.27}\)/n\(^{+}\)-Si heterojunction tunneling diode (sample #5150). The slopes of B trailing edge and P leading edge are 13 and 14 nm/decade respectively.

The information of Ge fraction, gas flow rates, growth rates, layer thicknesses, and doping levels were collected by secondary ion mass spectrometry (SIMS) and listed
in Table 2.2. A typical SIMS profile is shown in Fig. 2.8. For this device, the annealing temperature was 950°C, and the phosphorus level after annealing is $2 \times 10^{20} \text{ cm}^{-3}$ and flat. The boron level is $\sim 1.5 \times 10^{20} \text{ cm}^{-3}$. The slopes of the boron trailing edge and phosphorus leading edge are 13 nm/decade and 14 nm/decade respectively. The former is believed to be an artifact of SIMS due to the knock-on effect [30], and the later is caused by surface segregation during epitaxy [31], which will be discussed in chapter 3.

### 2.3.2 Effects of Annealing Temperature on Negative Differential Resistance

We now study the effects of annealing temperature on the presence of negative differential resistance in the implanted tunneling diodes. Several $p^+\text{Si}_{0.73}\text{Ge}_{0.27}/n^+\text{Si}$ tunneling diodes with different annealing temperatures to remove the junction defects were fabricated and were measured at room temperature, with their J-V curves shown in Fig. 2.9. The boron and phosphorus levels are $1.5 \times 10^{20} \text{ cm}^{-3}$ and $2 \times 10^{20} \text{ cm}^{-3}$ respectively. With 700 °C annealing, NDR was not observed because its defect-assisted tunneling current is higher than the band-to-band tunneling current at small forward biases, due to the lack of defect annealing. To observe NDR, a higher annealing temperature is required to remove the defects. For the device annealed at 800 ~ 950°C, NDR was observed. For 1050°C annealing there was no NDR, most likely because strong dopant diffusion at such high temperatures reduces the junction abruptness and the electric field, leading to a greatly reduced tunneling current despite the lower defect density. As a figure of merit, the peak-to-valley current ratio (PVCR) at forward bias is usually used to indicate the junction quality in a tunneling diode [22]. As shown in Fig. 2.9, a best PVCR of 2 was achieved in the device annealed at 900°C.
Fig. 2.9 J-V curves of $p^{+}$-Si$_{0.73}$Ge$_{0.27}$/n$^{+}$-Si heterojunction tunneling diodes at different annealing temperatures.

Note that for some devices, a bump or hump of current in the region of NDR was observed (e.g., diodes annealed at 850 °C and 900 °C in Fig. 2.9). We confirmed this was due to the oscillations in the measurement circuit [32], which includes an Agilent 4155C, electrical cables and probe station, and the tunneling devices. When the oscillations occur, a dc I-V measurement contains ac components resulting from the resonance of the entire circuit. As the tunneling diode enters the NDR region biased at a voltage slightly larger than the peak voltage, ac components of current are induced by the ac voltage swings. Near the peak voltage, any ac voltage swing results in the reduction of the tunneling because of the fewer available states in the p-type valence band at the same energy levels of the electrons in the n-type conduction band due to the ac voltage deviations from the peak voltage. Thus, a dc I-V measurement gives a sharp drop in current near the peak voltage, and no other physics is involved.

To further investigate the effects of annealing temperature on BTBT and DAT
Fig. 2.10 \( J_{\text{peak}} \) and \( J_{\text{DAT}} \) (at 0.3 V) of \( p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si} \) heterojunction tunneling diodes vs. annealing temperature.

For the annealing temperature of 800 ~ 900°C, \( J_{\text{DAT}} \) decreases with annealing temperature, so \( J_{\text{peak}} \) decreases as well, assuming to first order the BTBT current is constant. At 1050°C, despite the high efficiency of defect removal, the abruptness of doping profiles is reduced even more due to strong dopant diffusion, leading to a smaller electric field and less tunneling. Thus, \( J_{\text{peak}} \) drops further until NDR is gone.

### 2.3.3 Effects of Electric Field and Bandgap Energy on Band-to-Band Tunneling

To observe NDR in implanted \( p^+\text{-SiGe}/n^+\text{-Si} \) heterojunctions, a post-annealing
Fig. 2.11 J-V curves of $p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si}$ heterojunction tunneling diodes of three different boron concentrations.

step at certain temperatures is required. In the last section, we found that the best annealing temperature was 900 °C which gave the highest PVCR. Thus, for all heterojunctions devices of different boron levels and Ge fractions investigated in this section, 900°C annealing was applied.

Firstly, $p^+\text{-Si}_{0.73}\text{Ge}_{0.27}/n^+\text{-Si}$ diodes of a fixed n-type doping (phosphorus) of $2 \times 10^{20}$ cm$^{-3}$, and three different p-type doping levels (boron) at 0.75, 1.8, and $3.6 \times 10^{20}$ cm$^{-3}$, were fabricated. The J-V curves are shown in Fig. 2.11. For each device, NDR was clearly seen. As the boron level increases, the resulting higher electric field reduces the tunneling barrier, so the peak tunneling current increases from 0.2 to 1 kA/cm$^2$. While BTBT current increases with the boron level, DAT current also increases because of its tunneling nature.

Next, we present the results of the dependence of tunneling on Ge fraction in $p^+\text{-Si}_{1-x}\text{Ge}_x/n^+\text{-Si}$ diodes with various Ge fractions of 14, 21, 27, 35, and 38%. The
Fig. 2.12 J-V curves of p⁺-Si₁₋ₓGeₓ/n⁺-Si tunneling diodes of different Ge fractions.

Fig. 2.13 $J_{\text{peak}}$ vs. $\Delta V_{\text{peak}}$ for different Ge fractions. The linear dash line represents the spreading resistance of n⁺-Si substrate.

annealing temperature for these devices was 900°C, and J-V curves are shown in Fig. 2.12. As the Ge fraction increases, the tunneling barrier reduces due to the smaller bandgap energy, leading to stronger tunneling. As the Ge fraction increases from 0.14 to
0.39, the peak tunneling current density increases by a factor of $10^4$ from 0.01 to 9 kA/cm$^2$, which is believed to be the highest reported among all Si-based tunneling diodes by CVD. The peak voltage shifts to a larger value with Ge fraction because of the series resistance. In those devices, the series resistance is dominated by the spreading resistance in the Si substrate, so it is approximately constant for each device. Thus, the shift of the peak voltage ($\Delta V_{peak}$) is proportional to the peak current density and increases with the Ge fraction (Fig. 2.13).

### 2.4 Band-to-Band Tunneling in p$^+$-SiGe/n$^+$-SiGe Homojunctions

#### 2.4.1 Device Fabrication

In this section, we discuss band-to-band tunneling in p$^+$-SiGe/n$^+$-SiGe homojunctions. Similar to the fabrication steps of p$^+$-SiGe/n$^+$-Si heterojunction tunneling diodes, the Si substrates were prepared by the same ex-situ cleaning steps, before being loaded into the CVD reactor. Unlike for the heterojunction devices, both p$^+$ and n$^+$ layers were epitaxially grown by in-situ doping CVD. Without implant steps, radiation damage could be eliminated completely. Furthermore, the abruptness of the doping profiles could be preserved, since post-annealing was no longer required before epitaxial growth. Before growth, high temperature baking at 850°C for 5 minutes was performed to remove the residual oxide on the Si substrates. SiH$_2$Cl$_2$ and diluted GeH$_4$ (0.8% in H$_2$) were the precursors for SiGe growth. Diluted phosphine (PH$_3$) and diborane (B$_2$H$_6$) (both 100 ppm in H$_2$) were used for n-type and p-type in-situ doping.
Table 2.3 Growth parameters of p⁺-SiGe and n⁺-SiGe layers

<table>
<thead>
<tr>
<th>Ge Fraction (%)</th>
<th>Gas flow rates (scm) (DCS/GeH₄/B₂H₆/PH₃)</th>
<th>GR (nm/min)</th>
<th>Thickness of p⁺-n⁺-SiGe (nm)</th>
<th>Critical thickness (nm)</th>
<th>Doping levels of p⁺-n⁺-SiGe (cm⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 (625°C)</td>
<td>26/50/100/100</td>
<td>3</td>
<td>32/18</td>
<td>450</td>
<td>1.3/0.7×10²⁰</td>
</tr>
<tr>
<td>21 (625°C)</td>
<td>26/100/125/125</td>
<td>6.5</td>
<td>44/22</td>
<td>180</td>
<td>1.3/1.1×10²⁰</td>
</tr>
<tr>
<td>27 (625°C)</td>
<td>26/200/167/167</td>
<td>10</td>
<td>25/16</td>
<td>80</td>
<td>1.8/1.2×10²⁰</td>
</tr>
<tr>
<td>35 (575°C)</td>
<td>26/300/250/250</td>
<td>6.8</td>
<td>15/12</td>
<td>35</td>
<td>3.5/2.4×10²⁰</td>
</tr>
</tbody>
</table>

After baking, an n⁺-SiGe layer was immediately grown, followed by the deposition of a p⁺-SiGe layer by fast switching between the doping gases. The diodes with Ge fractions of 14, 21, and 27% were grown at 625°C with thicknesses of 50, 66, and 40 nm, and the diode with a Ge fraction of 35% was grown at 575°C with a thickness of 30 nm. The layer thicknesses, Ge fraction, and doping concentrations were determined by SIMS (Table 2.3). Metastable critical thicknesses of strained SiGe layers on Si substrates for Ge fraction of 14, 21, 27, and 35% are 450, 180, 80, and 35 nm, respectively [33]. Therefore, we expected the layers to be biaxially compressively strained, pseudomorphic to the Si substrates. Lastly, square mesas were dry-etched with an area of 25μm × 25 μm. Ti/Al was evaporated on the bottom of the wafers for the n-side ohmic contact. For the p-side ohmic contact, Ti/Al was patterned on top of the mesa by a combination of photolithography and lift-off with a contact area of 23 μm × 23 μm between Ti/Al and the mesa.

2.4.2 Forward-Biased Band-to-Band Tunneling

First, we studied the effect of p-type doping levels on BTBT by varying the boron concentration ($N_A$: 1.7 to 5.1 × 10²⁰ cm⁻³) with a fixed phosphorus level of $N_D$ ~
Fig. 2.14 J-V curves of $p^+\text{Si}_{0.73}\text{Ge}_{0.27}/n^+\text{Si}_{0.73}\text{Ge}_{0.27}$ homojunction tunneling diodes of three different boron levels.

$1.0 \times 10^{20}$ cm$^{-3}$ in $\text{Si}_{0.73}\text{Ge}_{0.27}$ tunneling diodes. NDR is clearly seen from the J-V curves of these three devices (Fig. 2.14). The peak tunneling current density in forward bias ($J_{\text{peak,FB}}$) increases with the boron concentration from 1.7 to 8 kA/cm$^2$, due to the reduced tunneling barrier. The peak voltage is shifted to a larger value because of the series resistance. A best PVCR of 3.6 is achieved with a peak tunneling current density of 8.2 kA/cm$^2$, an indication of the high quality of the tunneling devices by in-situ doping CVD.

Next, to study the effect of Ge fraction, strained $p^+\text{SiGe}/n^+\text{SiGe}$ homojunction tunneling diodes, with four different Ge fractions of 14, 21, 27, and 35%, were grown and fabricated. J-V curves are shown with NDR clearly seen in Fig. 2.15. $J_{\text{peak,FB}}$ increases from 0.03 A/cm$^2$ to 8.2 kA/cm$^2$ as the Ge fraction increases from 0.14 to 0.35 because of the smaller bandgap energy. While the increasing $J_{\text{peak}}$ at forward bias is evidence of the effect of Ge fraction on tunneling, the effect of different p-type doping levels must be isolated to accurately extract the dependence of tunneling on Ge fraction.
To understand how the forward-biased BTBT current is affected by SiGe bandgap energy, we used an empirical model of bandgaps by Robbins et al. based on their photoluminescence measurements at 4 K [34], with a subtraction of 50 meV for the difference between their measurements at 4 K and ours at room temperature [35], to convert the measured Ge fractions into bandgap energies:

\[ E_g(300K) = 1.17 - 0.896x + 0.396x^2 - 0.05 \text{ (eV)}, \quad (2.10) \]

where \( x \) is the germanium fraction in Si\(_{1-x}\)Ge\(_x\) alloys. We then used a correction procedure to separate out the effect of doping profiles (shown schematically in Fig. 2.16), which were not the same for all Ge fractions. (E.g. \( N_A \) varied from 1.2 to 3.6 \( \times 10^{20} \text{ cm}^{-3} \) and \( N_D \) varied from 0.7 to 2.5 \( \times 10^{20} \text{ cm}^{-3} \)). In summary, we used Eqs. (2.4), (2.7), and (2.8) to predict how the peak tunneling current density scales with the electric field, and then adjusted the data points to reflect a single dopant profile at different Ge fractions. First we calculate the peak voltage (\( V_{\text{peak}} \)) and its peak electric field (\( F_{\text{field}} \), by
assuming perfectly abrupt doping profiles and no series resistance, with a single set of fixed doping levels of \( N_A = 1.2 \times 10^{20} \text{ cm}^{-3} \) and \( N_D = 0.7 \times 10^{20} \text{ cm}^{-3} \) for each Ge fraction. We calculate the peak current density \((J_{\text{peak, abrupt}})\) using Eqs. (2.4), (2.7), and (2.8). Next, we use the actual doping profiles measured by SIMS to calculate \( V_{\text{peak}} \) and \( F_{\text{field}} \) using a device simulator. We then calculate the peak current density \((J_{\text{peak, SIMS}})\) at that field. The ratio of these two current densities gives a correction factor to be applied to the experimental data for the adjustment of doping levels for all Ge fractions, so a comparison of data with a single set of doping profiles \((N_A = 1.2 \times 10^{20} \text{ cm}^{-3} \text{ and } N_D = 0.7 \times 10^{20} \text{ cm}^{-3})\) can be made. After measuring the I-V curves of the devices to obtain \( J_{\text{peak, measured}} \), we can calculate the corrected peak current density \((J_{\text{peak, corrected}})\) using the relationship of

\[
J_{\text{peak, corrected}} = J_{\text{peak, measured}} \times \frac{J_{\text{peak, abrupt}}}{J_{\text{peak, SIMS}}}. \tag{2.11}
\]
Fig. 2.17 $J_{\text{peak}}$ vs. Ge fraction. Blue squares are the experimental data, red squares represent corrected data to a single set of fixed doping levels ($N_A = 1.2 \times 10^{20}$ cm$^{-3}$ and $N_D = 0.7 \times 10^{20}$ cm$^{-3}$), and the solid line is the model prediction based on Eqs. (2.4), (2.7), and (2.8).

Fig. 2.17 shows the experimental ($J_{\text{peak,measured}}$) and corrected peak current density ($J_{\text{peak,corrected}}$) as a function of Ge fraction (blue and red squares), and also a theoretical prediction for a single set of doping levels assuming abrupt profiles (solid line). For a Ge fraction of 0.14, $J_{\text{peak,corrected}}$ is larger than $J_{\text{peak,measured}}$ because the actual doping profiles were not abrupt, resulting in a smaller electric field. For higher Ge fractions however, $J_{\text{peak,measured}}$ was larger than $J_{\text{peak,corrected}}$ because of the higher doping levels for those devices. No adjustable parameters were used in the correction process. Significantly, there is a good agreement between the slope of the theoretical calculation of peak tunneling current and data corrected to a single set of doping levels versus bandgap. This confirms that Eqs. (2.4), (2.7), and (2.8) can be used to predict the band-to-band tunneling of p$^+$-SiGe/n$^+$-SiGe homojunctions at current density up to $\sim 10$ kA/cm$^2$ in forward bias.
2.4.3 Reverse-Biased Band-to-Band Tunneling

2.4.3.1 Effects of Defect-Assisted Tunneling

The operation of TFETs relies on BTBT under reverse bias, also known as Zener tunneling [36]. In reverse bias, there is no simple clear feature such as NDR in forward bias which can be used to confirm that the observed current is due to BTBT. For example, defect states in the bandgap at the junction can lead to defect-assisted tunneling (DAT) [26], in which an electron first tunnels from the valence band of the p-type region to a defect state in the bandgap of the p-n junction, and then tunnels from the defect state to the conduction band of the n-type region (Fig. 2.18). Because each step of this process has a much lower tunneling barrier than direct BTBT, the two-step DAT process can easily swamp the direct BTBT. To investigate the effects of the DAT process, we examined tunneling in both forward and reverse biases in $p^+\text{SiGe}/n^+\text{Si}$ heterojunction tunneling diodes. The preparation steps of the heterojunction tunneling diodes were introduced in section 2.3.1.

![Fig. 2.18 Schematic of a band energy diagram for a $p^+\text{Si}_{0.73}\text{Ge}_{0.27}/n^+\text{Si}$ heterojunction to show the processes of band-to-band tunneling and defect-assisted tunneling in reverse bias.](image-url)
Fig. 2.19 J-V curves of $p^+\text{Si}_{0.73}\text{Ge}_{0.27}/n^+\text{Si}$ heterojunction tunneling diodes with three annealing temperatures, showing the importance of the presence of NDR in forward bias on the BTBT current in reverse-bias.

For a relatively low annealing temperature (700 °C), a fairly high current with ohmic characteristics was observed at both forward and reverse biases, with no evidence of NDR (Fig. 2.19). With 900°C annealing, a lower current with NDR at forward bias was observed, with a lower current in reverse bias as well. We hypothesize that for the device with 700 °C annealing, the current in forward and reverse biases was dominated by the DAT process due to the incomplete annealing of the implanted damages, which swamped the true BTBT current. The defect density (and thus the DAT process) was reduced by 900 °C annealing, so that NDR at forward bias and the true BTBT current density could be observed. In the sample annealed at 1050°C, NDR disappeared with a much lower current density at both forward and reverse biases, because the diffusion of dopants at 1050°C reduces the junction abruptness and the electric field. The main message is that if NDR was not observed (e.g. at 700°C), the observed current in both forward and reverse biases has a large DAT component, and cannot be used as a true
measure of BTBT current. Consequently, we strongly suggest that to use reverse-biased tunneling data for the calibration of a BTBT model, a demonstration of NDR at forward bias is necessary to exclude the possibility of a dominant contribution of a defect-assisted tunneling current.

2.4.3.2 Effects of Series Resistance

Zener tunneling (reverse-biased BTBT) in silicon-based devices was in the past usually characterized at moderate doping levels ($\sim 10^{18}$ cm$^{-3}$). At such doping levels, the current density is very low ($\sim 10^{-4}$ kA/cm$^2$ [36, 37]), so the effect of series resistance could be ignored. However, for SiGe TFETs, operation at electric fields $> 10^7$ V/m is desired, leading to a high current level. Thus, the series resistance effect such as current crowding must be considered for a precise calibration of BTBT in SiGe tunneling diodes. Guo et al. suggested that by scaling down the mesa width ($W$) of the diodes, current

![J-V curves of Si$_{0.73}$Ge$_{0.27}$ homojunction tunneling diodes with different mesa widths.](image)

Fig. 2.20 J-V curves of Si$_{0.73}$Ge$_{0.27}$ homojunction tunneling diodes with different mesa widths.
crowding can be eliminated [38]. Therefore, we fabricated several Si$_{0.73}$Ge$_{0.27}$ homojunction p$^+$/n$^+$ diodes with mesa widths ranging from 50 μm to 0.35 μm using a combination of photolithography and electron-beam lithography. The growth parameters and layer structures are listed in Table 2.3. The current density at forward bias is shown in Fig. 2.20, with clear NDR for each device. The average of $J_{peak, FB}$ is 1.73 kA/cm$^2$ and the deviations are within 5% for all mesa sizes, confirming a negligible contribution of the leakage current via the mesa edges [39]. In reverse bias, the current density approaches a constant level as the mesa width decreases (Fig. 2.21(a) and (b)). Because (i) NDR is clearly seen at forward bias and (ii) the effect of series resistance was eliminated by scaling down the mesa width, we believe the plateaus of the current density Fig. 2.21(b) represent the true BTBT current density in reverse bias.

2.4.3.3 Comparison of Experimental Results and Models

We now seek to present BTBT in reverse bias as a function of the Ge fraction and electric field for device designers to model TFETs and related devices. Similar to
the peak current density at forward bias, the reverse-biased BTBT current \(J_{BTBT,RB}\) also strongly depends on the electric field and bandgap energy. An analytical form of \(J_{BTBT,RB}\) in reverse bias based on Fair’s model [36] is

\[
J_{BTBT,RB} = \frac{\sqrt{2m^*e^*}E_{field}V}{4\pi\hbar^2\sqrt{E_g}} \exp\left(-\frac{4\sqrt{2m^*E^*_g}}{3e\hbar F_{field}}\right),
\]

where \(V\) is the applied voltage. The tunneling probability of electrons depends on the tunneling distance. In a semiconductor p-n junction, the tunneling distance is inversely proportional to the peak electric field [24], so \(\log(J_{BTBT,RB})\) in Eq. (2.12) depends on the inverse of the electric field. The parameter of \(J_{BTBT,RB}/V\) is plotted versus the electric field \(F_{field}\) for different Ge fractions in Fig. 2.22, along with theoretical calculations based on Eq. (2.12). For our devices, the electric field was varied by adjusting the reverse bias from 0.1 to 0.9 V. As in forward bias, the measured SIMS doping profiles

![Graph showing the relationship between electric field and reverse-biased BTBT current density per volt.](image)

Fig. 2.22 Reverse-biased BTBT current density per volt \(J_{BTBT,RB}/V\) vs. electric field in SiGe BTBT. Symbols are the experimental data, and multiple lines represent the model predictions based on Eqs. (2.10) and (2.12). An inset provides an enlarged view for high electric fields of 2 to \(3 \times 10^8\) V/m.
and the device simulator were used to calculate the electric field for all devices. The horizontal error bars in the electric field, which resulted from an estimated uncertainty of ±15% in the doping levels, are also presented in Fig. 2.22.

At small electric fields ($< 1 \times 10^8$ V/m), Eq. (2.12) predicts that $J_{BTBT,RB}/V$ rises sharply with $F_{field}$. On the other hand, at large electric fields, $J_{BTBT,RB}/V$ increases much more slowly. For example, at low fields ($F_{field} = 5 \times 10^7$ V/m), $J_{BTBT,RB}/V$ of Si$_{0.73}$Ge$_{0.27}$ (red line in Fig. 2.22) increases by a factor of $2 \times 10^5$ as $F_{field}$ increases by 60%. However, at $F_{field} = 2 \times 10^8$ V/m, it only increases by a factor of 30. Our data of reverse-biased tunneling conductance density (current density/voltage) at electric fields of $> 2 \times 10^8$ V/m, which are between 3 – 1000 kA/cm$^2$-V, are in close agreement with the model predictions (the inset in Fig. 2.22). This is fortuitous as our data points of $J_{BTBT,RB}/V$ are three to five orders of magnitude higher than those for Si BTBT at low electric fields ($< 1 \times 10^8$ V/m) [36, 37], where Eq. (2.12) has been applied.

To isolate the effect of bandgap energy, $J_{BTBT,RB}/V$ vs. Ge fraction was plotted at $F_{field} = 2 \times 10^8$ in Fig. 2.23 by interpolating between points in Fig. 2.22, along with the model of Eqs. (2.10) and (2.12). The horizontal error bars of ±6% variations in the peak electric field of Fig. 2.22 were used to estimate the resulting errors in the BTBT current. Within the uncertainty in current introduced by the variation in electric field, the results show that Eqs. (2.10) and (2.12) can be used to model the dependence of $J_{BTBT,RB}$ on Ge fraction. More complete modeling might include the effect of heavy doping on bandgap energy, and the effects of strains on the effective density of states and the effective mass of electrons, which are not considered in this work.
Fig. 2.23 $J_{BTBT,RB}$ vs. Ge fraction at $F_{field}$ of $2 \times 10^8$ V/m. Points are the experimental data with error bars, representing the effects of the variations in electric field (the horizontal error bars in Fig. 2.22).

2.5 Summary

SiGe-based TFETs were proposed as replacements for Si TFETs because higher tunneling current is expected for the smaller bandgap. We studied two types of SiGe tunneling junction: p$^+$-SiGe/n$^+$-Si heterojunctions and p$^+$-SiGe/n$^+$-SiGe homojunctions, to establish the relationship between the experiment results and models. To observe NDR in implanted p$^+$-SiGe/n$^+$-Si heterojunction tunneling diodes, annealing at high temperature is required to reduce defect-assisted-tunneling (DAT) via the junction defects introduced during the implant process for n$^+$-Si layer. For low-temperature annealing, DAT swamps band-to-band tunneling, so no NDR was observed. For high-temperature annealing, there was also no NDR observed because of the reduced electric field resulting from strong dopant diffusion. For a Ge fraction of
0.35, a peak tunneling current density of 8.2 kA/cm² in a homojunction tunneling diode was reported, which is the highest for all Si-based tunneling diodes grown by CVD. For p⁺-SiGe/n⁺-SiGe homojunction tunneling diodes by in-situ doping CVD, defect-assisted tunneling can be greatly reduced with an improved peak-to-valley current ratio of 3.6.

For both types of tunneling junctions, the effects of the electric field and Ge fraction on BTBT at forward bias and bias bias were investigated. Increasing the doping levels and Ge fraction enhances tunneling. We compared our experimental results in forward bias with Kane’s model [25] and the reverse-biased experimental data with Fair’s model [36], both in good agreement. This suggested the adequacy of Kane’s model for predicting BTBT in strained SiGe junctions, at least up to a current density level of $10^4$ A/cm² and $10^6$ A/cm² in forward and reverse biases, respectively.

We qualitatively demonstrated the importance of the presence of NDR in reverse biased BTBT. With NDR observed in each device and the elimination of series resistance by scaling down the device, the true measure of reverse-biased BTBT in SiGe p⁺/n⁺ homojunction was identified for the first time.
Chapter 3  Surface Segregation of Phosphorus in Relaxed Si$_{0.7}$Ge$_{0.3}$ Layers Epitaxially Grown by Chemical Vapor Deposition

3.1  Introduction

As device dimensions are scaled down, a sharp profile of dopants is becoming a key factor to realize nano-scale semiconductor devices such as tunneling diodes [40], tunneling field-effect transistors [20], and two-dimensional electron gases (2DEGs) in a modulation-doped Si/SiGe heterostructure [41]. 2DEGs are of particular interest for quantum dot (QD) applications. A QD is usually fabricated on a 2DEG, with top metal Schottky depletion gates used to isolate a single electron in the underlying 2D EG layer. However, the strong surface segregation of n-type dopants in a relaxed SiGe epitaxial film can cause a high dopant concentration at the surface, resulting in high gate leakage current and ineffective gating. Therefore, a sharp turn-off slope of n-type dopants is necessary.

Although a turn-off slope of 2-3 nm/dec for antimony was reported in Si epitaxial films grown by molecular beam epitaxy (MBE) [42], it has been difficult to obtain such abrupt profiles for phosphorus and arsenic, the most common n-type dopants in chemical vapor deposition (CVD) systems [43]. Several works were reported to reduce phosphorus segregation in Si by ex-situ cleaning (13 nm/dec) [44] or by introducing substitutional carbon atoms into Si epitaxial films (11 nm/dec) [45].
However, the former approach requires a growth interruption, which may introduce contaminants into the growth interface. For the latter, the control over carbon atoms into substitutional sites is critical since the interstitial carbon could degrade device performance due to their midgap energy states [46].

In this chapter, we report an extremely sharp phosphorus turn-off slope of 6 nm/dec in relaxed $\text{Si}_{0.7}\text{Ge}_{0.3}$ films without any ex-situ cleaning step or introduction of carbon into the epitaxial films. We found that the hydrogen coverage on the surface during the growth plays an important role in the suppression of phosphorus segregation in the CVD process at low temperatures (500 $\sim$ 600 °C). Finally, a phenomenological model is proposed to explain the effect of surface hydrogen on phosphorus segregation.

### 3.2 Two State Model

In our work, a matrix of Si and Ge atoms and surface hydrogen complicate the analysis of phosphorus segregation. In a simpler case of phosphorus in Si (100), a segregation energy of 0.64 eV, which was defined as the energy difference of phosphorus in the surface and bulk layers, was firstly reported by Nützel et al. [47] from SIMS results. This work used a so-called two-state model (TSM) [48] of atoms moving between the surface and sub-surface layers in Si. Later, by temperature programmed desorption (TPD), Cho et al. [49] reported a segregation enthalpy of 0.86 eV, which is essentially the same as the segregation energy defined in the TSM. The P coverage in Nützel’s work [47] and Cho’s work [49] was larger than 0.1 monolayer (ML), and the major Si surface structures with P surface coverage > 0.1 ML was previously by Yu et al. [50] as a mixture of Si–Si, Si–P, and P–P dimers. Sen et al. [51] used density functional
theory to predict the different favorable sites for surface phosphorus atoms at coverage above and below 0.13 ML. Thus, the work in [50, 51] might not be directly relevant to our experimental results since the integrated phosphorus doses in our samples are at most $5 \times 10^{12}$ cm$^{-2}$ ($\sim 0.01$ ML). Other works on Sb [52], As [53], and Ge [54, 55] surface segregation in Si have also been modeled by using a TSM. In addition, those works all ignored any temperature dependence of attempt frequency and used the segregation energy to reflect all temperature effects. Thus, in this chapter, we also use a modified TSM to investigate phosphorus segregation in a more complicated structure of relaxed Si$_{0.7}$Ge$_{0.3}$ layers.

A two-state model (TSM) describes the dopant segregation as an exchange process of P atoms and host atoms (Si or Ge in this study) between the surface layer and the sub-surface layer (Fig. 3.1). The rate equations governing this exchange process between those two layers are

$$\frac{dn_1}{dt} = -r_{12}n_1(1-n_2) + r_{21}n_2(1-n_1), \quad (3.1)$$

$$\frac{dn_2}{dt} = -r_{21}n_2(1-n_1) + r_{12}n_1(1-n_2), \quad (3.2)$$

$$r_{12} = e^{-\frac{E_1}{kT}}, \quad (3.3)$$

$$r_{21} = e^{-\frac{E_1+\Delta E_{surf}}{kT}}, \quad (3.4)$$

where $n_1$ and $n_2$ are the normalized concentrations of phosphorus in the sub-surface layer (layer 1) and the surface layer (layer 2), $r_{12}$ and $r_{21}$ are the jumping rates of phosphorus from the sub-surface layer to the surface layer, and vice versa. $E_1$ is the
activation energy barrier facing phosphorus in the sub-surface layer, $\Delta E_{\text{surf}}$ is the segregation energy, which represents the difference of activation barriers of layer 1 and layer 2, and $v$ is the attempt frequency (Fig. 3.2). We assume a single attempt frequency independent of temperature as previous reports suggested for phosphorus segregation in Si [47, 48, 49, 53, 54, 55]. It is assumed that P atoms below the sub-surface layer are trapped and cannot diffuse during the time of the growth. Assuming $n_1$ and $n_2 << 1$, the differential rate equations can be solved and an analytical form of phosphorus turn-off slope $x_0$ (nm/decade) is given by [56]

$$
x_0 = a_0 \ln 10 \frac{1}{4} \frac{1}{\ln \left(1 + e^{-\Delta E_{\text{surf}}/kT}\right) - \ln \left[1 - e^{-\left(n_2 + n_1\right)/4GR}\right]}, \tag{3.5}
$$

where $a_0$ is lattice constant and $GR$ is the growth rate.

According to the TSM, the surface segregation occurs because P atoms at the sub-surface layer tend to stay in the surface layer due to the lower energy level in the surface layer (Fig. 3.2). At thermal equilibrium (low growth rates), the surface segregation is determined by the ratio of the phosphorus concentrations at the surface
Fig. 3.2  Schematic of phosphorus energy near the surface during the epitaxial growth in a two-state model. Layer 2 represents the surface layer and layer 1 (sub-surface layer) represents the next layer below the surface.

and sub-surface layers \( n_2/n_1 \), which only relies on the temperature and the segregation energy \( (\Delta E_{\text{surf}}) \) through the first term in the denominator of Eq. (3.5). At lower temperatures, if still in equilibrium, more P atoms are trapped in the surface layer than in the sub-surface layer due to the lower energy state of the former, so the segregation is stronger. On the other hand, in the kinetic-limited regime of high growth rates, P atoms in the sub-surface layer cannot reach the equilibrium with those in the surface layer. Therefore, phosphorus will be trapped in the sub-surface layer and the limiting factor is its activation energy barrier \( E_1 \). As the temperature is reduced, the probability of phosphorus jumping across the barrier to the surface layer is smaller because of the lower kinetic energy of phosphorus. As a result, phosphorus segregation is reduced. This physical limit has been applied to reduce phosphorus segregation in Si (100) grown by MBE at temperature below 500 °C [47], and the best turn-off slope (4 nm/dec) for P in
Si grown by a combination of scanning tunneling microscopy and MBE was reported by room-temperature growth [57].

3.3 Failure of Two-State Model

3.3.1 Experimental Results

In this work, Si (100) substrates (for strained SiGe) and polished relaxed Si$_{0.7}$Ge$_{0.3}$ virtual substrates with a graded Si$_{1-x}$Ge$_x$ ($0 < x < 0.3$) buffer layer grown on Si (100) substrates (for relaxed SiGe) were used to study phosphorus segregation. Prior to being placed into the reactor, substrates were cleaned by the following steps: 5 min in diluted HF (1%), 15 min in H$_2$SO$_4$: H$_2$O$_2$ (2.5:1), followed by 2 min in diluted HF (1%). Then the samples were heated to 850 °C in hydrogen gas at 6 torr to remove the residual oxide before the epitaxial growth starts. The gas precursors were diluted silane (10 % in argon) and GeH$_4$ (0.8 % in hydrogen) for Si and SiGe growth, and a diluted phosphine (100 ppm in hydrogen) was the doping gas. The test structure for phosphorus segregation is as follows: first, a 20-nm undoped Si$_{0.7}$Ge$_{0.3}$ buffer layer was grown followed by a 10-nm n-type Si$_{0.7}$Ge$_{0.3}$ layer doped with phosphorus of peak level between $10^{18}$ and $10^{19}$ cm$^{-3}$. Both layers were grown at 575 °C and the growth rate was 5 nm/min. Then an undoped Si$_{0.7}$Ge$_{0.3}$ cap layer was grown at 500 °C ~ 600 °C to study the effect of growth temperature on phosphorus segregation, with its thickness between 30 to 150 nm. To investigate the effect of growth rate, we varied the growth rate of the Si$_{0.7}$Ge$_{0.3}$ cap layers between 0.1 to 30 nm/min by adjusting the partial pressures of silane and germane. The Ge fraction in the SiGe cap layer was between 0.28 to 0.30.
Last, a thin Si cap layer of 4 nm was grown at 625 °C, with the growth rate of 2.5 nm/min. The films were subsequently characterized by SIMS to determine the phosphorus profiles and the growth rates.

Most prior works of phosphorus segregation in SiGe were done in compressively strained SiGe layers [53], not in the relaxed SiGe layers required for a modulation-doped 2DEG, i.e. a higher conduction band edge in the SiGe layer than in the Si layer. Thus, we compared the phosphorus profiles in strained and relaxed SiGe films first. We found that the segregation is much worse in relaxed SiGe films than in strained SiGe films (Fig. 3.3). For a growth temperature of 575 °C and growth rate of 5 nm/min, the turn-off slopes of phosphorus in strained and relaxed films are 27 and 41 nm/decade, respectively [56]. The fundamental reasons for this difference are unknown and a further study is required. We then focused on phosphorus segregation in the
Fig. 3.4  Phosphorus profiles in relaxed Si$_{0.7}$Ge$_{0.3}$ layers grown at 500, 550, and 600 °C. Phosphorus supply was turned off at the depth of 45 nm. P turn-off slopes were 127, 40, and 9 nm/decade for 600, 550, and 500 °C, respectively [56].

Phosphorus profiles measured by SIMS for relaxed SiGe layers grown at 500, 550, and 600 °C are shown in Fig. 3.4. At a depth of 45 nm, the phosphine supply for the doped layer was turned off and growth was continued without interruption. As the growth temperature is reduced, the surface segregation is reduced, with the phosphorus turn-off slope declining from 127 nm/dec at 600 °C to 9 nm/dec at 500 °C. By adjusting the gas flow rates of silane and germane at 500 °C, an extremely sharp slope of 6 nm/min was obtained with the growth rate of 0.08 nm/min, which we believe is the sharpest reported turn-off slope of phosphorus in relaxed SiGe films.

Our experimental data (points) of phosphorus turn-off slope vs. growth rate at different temperatures (500 °C to 600 °C) and theoretical curves based on the TSM are shown in Fig. 3.5. At 600 °C the segregation is near the transition between thermal
Fig. 3.5 Phosphorus turn-off slope vs. growth rate for different temperatures. Experimental results (points) and the theoretical prediction (lines) are presented for comparison [56].

equilibrium and kinetic-limited regime. Thus, $\Delta E_{surf} = 0.47 \text{ eV}$ can be fitted by assuming the experimental results were in the regime of thermal equilibrium. Furthermore, as Nützel et al. [58] suggested that the attempt frequency would be between $10^{11}$ to $10^{13} \text{ Hz}$, we selected $\nu = 1 \times 10^{12} \text{ Hz}$ to fit the data at 600 °C and found $E_1 = 1.84 \text{ eV}$. Despite a good match between the experimental data and the theoretical curve at 600 °C, there is a large discrepancy between them at 500 °C to 575 °C. The low dependence of the segregation on growth rate suggests the data at low temperatures were in the equilibrium regime, not kinetically limited. However, the phosphorus slopes are much sharper at lower temperatures, in contrast with what would be expected from the TSM in equilibrium. This discrepancy cannot be resolved by simply adjusting the fitting parameters.
3.3.2 Effects of Hydrogen on Phosphorus Segregation

In the CVD process, hydrogen is used as a carrier gas and it is well established that hydrogen could cover the surface by forming Si–H [59] or Ge–H bonds [60]. At high temperatures, most of those bonds break easily and hydrogen desorbs, so the surface coverage of hydrogen is nearly zero. At low temperatures, however, the thermal energy is too low to break the Si–H or Ge–H bonds efficiently, so hydrogen will cover most of the surface layer. Our data show that the phosphorus turn-off slope is nearly constant with growth rate at a fixed temperature, suggesting that it is in the regime of thermal equilibrium, with the segregation then depending on $\Delta E_{\text{surf}}$.

At the heart of our model, we assume phenomenologically that the presence of surface hydrogen changes the relative energy of P atoms in the surface and sub-surface layers such that the segregation energy $\Delta E_{\text{surf}}$ is reduced. Thus, at lower temperatures, the segregation will be suppressed due to higher hydrogen coverage on the surface. Because of two types of surface sites (with or without H), in principle we could model the problem with two segregation energies and a fraction of phosphorus segregation to each site. However, the two energies would probably depend on the local numbers of Si or Ge atoms, leading to too many parameters. Prior works by MBE and TPD used an effective segregation energy [54, 55] to investigated the effect of hydrogen on Ge segregation into Si (100), which we follow in this study. Thus, we treat $\Delta E_{\text{surf}}$ as a single effective parameter which varies with hydrogen coverage as the temperature changes.

By the introduction of the ideas of an effect of hydrogen on phosphorus segregation, a fit between the experimental data and the model is obtained by using $\Delta E_{\text{surf}}$ as a fitting parameter at different temperatures or hydrogen pressures (Fig. 3.6),
Fig. 3.6  Comparison of experimental data and modified TSM of phosphorus turn-off slope vs. growth rate by including the effect of surface hydrogen on the segregation energy ($E_{surf}$) [56]. The hydrogen pressure was 6 torr.

with $\Delta E_{surf}$ plotted vs. temperature in Fig. 3.7. At hydrogen pressure of 6 torr, the reduced segregation at lower temperatures originates from a smaller segregation energy $\Delta E_{surf}$. The decrease of $\Delta E_{surf}$ with decreasing temperature will reduce the ratio of phosphorus populations in the surface vs sub-surface layers ($n_2/n_1$), resulting in a reduction of the segregation. On the other hand, $\Delta E_{surf}$ increases with temperature because of less hydrogen coverage at higher temperatures. The reduced segregation energy of P in Si$_{0.7}$Ge$_{0.3}$ with more surface hydrogen coverage follows the trend of Ge segregation in Si [54, 55, 61].

To confirm that the main effect of segregation reduction comes from a change of hydrogen coverage rather than a change of attempt frequency or other effects with temperature, we adjusted hydrogen pressure to vary its surface coverage [62] at 575 °C (Fig. 3.7). As expected, with higher hydrogen pressure (23 torr), the segregation is
suppressed, leading to a lower effective segregation energy. The opposite trend is shown with a lower hydrogen pressure of 2 torr. The effective segregation energy obtained in this work ranges between 0.37 ~ 0.52 eV, which is lower than those published for P in Si of 0.67 ~ 0.86 eV [47, 49] without hydrogen coverage. The difference could be explained by the presence of surface hydrogen on Si$_{0.7}$Ge$_{0.3}$ surface in this work, and the SiGe matrix in our work vs. Si matrix in other work.

### 3.4 Summary

We studied the surface segregation of phosphorus in Si$_{0.7}$Ge$_{0.3}$ films grown by CVD. The segregation in relaxed layers is stronger than in strained layers. Furthermore, the segregation in relaxed SiGe layers is reduced as the temperature is decreased due to
the higher surface coverage of hydrogen. We proposed a phenomenological model to explain the effect of surface hydrogen on phosphorus segregation, where surface hydrogen changed the bonding structure of host atoms near the surface and reduced the segregation energy $\Delta E_{\text{surf}}$ as the growth temperature is decreased. Thus, the segregation is suppressed at lower temperatures. An extremely sharp phosphorus turn-off slope of 6 nm/dec in relaxed Si$_{0.7}$Ge$_{0.3}$ layers grown at 500 °C was also reported, enabling effective Schottky gating on a modulation-doped Si two-dimensional electron gas (chapter 4).
Chapter 4  High Mobility in Modulation-Doped Si

Two-Dimensional Electron Gases

4.1  Introduction

The two-dimensional electron gas (2DEG) in semiconductors is a fundamental low-dimensional system [63] for condensed matter physics. Important physical phenomena such as the integral and the fractional quantum Hall effects were discovered in high-quality 2DEGs [64, 65]. Furthermore, practical applications such as modulation-doped field-effect transistors (MODFETs) [66] or quantum computation [12] were also realized by utilizing a 2DEG. The most important feature of a 2DEG is its high electron mobility using the so-called modulation-doping technique to separate the supply impurities and electrons [67]. A record high mobility of $3.5 \times 10^7 \text{ cm}^2/\text{V-s}$ was demonstrated in a GaAs 2DEG at $T < 1 \text{ K}$ [68]. In silicon, the electron confinement could be achieved in the inversion layer of Si MOS structure [64]. However, strong scattering from the impurity charges of the Si/SiO$_2$ interface has restricted the electron mobility to the level of $2 \times 10^4 \text{ cm}^2/\text{V-s}$ [69].

In 1985, Abstreiter et al. showed that electrons can be confined in a strained Si layer sandwiched between the relaxed SiGe layers [70]. In their Si/SiGe heterostructure, the mobility was fairly low ($\sim 2000 \text{ cm}^2/\text{V-s}$) due to the high density of dislocation defects by a large lattice mismatch between Si and SiGe layers. With a thick graded Si$_{1-x}$Ge$_x$ layer ($0 < x < 0.3$) of several microns grown on a Si substrate followed by a
relaxed Si$_{0.7}$Ge$_{0.3}$ layer, Ismail et al. demonstrated a much improved mobility of 535,000 cm$^2$/V-s [71]. A yet higher electron mobility of 800,000 cm$^2$/V-s was reported by Sugii et al. with a combination of MBE and solid phase epitaxy to provide an atomically flat interface, reducing the interface roughness scattering [72]. On the other hand, Huang et al. recently demonstrated a record high mobility of $2 \times 10^6$ cm/V-s in their enhancement-mode device of undoped Si 2DEG by top gating (no modulation doping), and suggested the background impurity scattering as the limiting factor of electron transport [73].

To summarize, the dominant scattering source in Si 2DEGs appeared very different for various structures and there is no systematic work on the effects of the layer structure on electron transport properties in a Si 2DEG so far. In this chapter, the basic properties of a 2DEG will first be briefly introduced, followed by our experimental results of modulation-doped Si 2DEGs of different layer structures. We also demonstrate effective Schottky gating on a modulation-doped Si 2DEG in order to manipulate the electron density and mobility so as to identify the dominant scattering mechanisms.

### 4.2 Characteristics of a Si 2DEG

#### 4.2.1 Band Offset between Strained Si and Relaxed SiGe layers

For a strained SiGe layer grown pseudomorphically on a relaxed Si substrate, such as the heterojunction tunneling devices in chapter 2, the entire band offset is almost in the valence band, providing hole confinement [74]. On the other hand, for electron confinement, a tensile strained Si layer must be sandwiched between relaxed SiGe
Fig. 4.1 (a) Band diagram of a relaxed-SiGe/strained-Si/relaxed-SiGe heterostructure. The energy levels in the conduction band of strained Si are split into two states: $\Delta_2$ and $\Delta_4$; (b) under tensile strain, electrons reside in the two states of $\Delta_2$ along (001) in Si with an in-plane effective mass of $m^* = 0.19\ m_0$ [75].

layers (Fig. 4.1 (a)). A conduction band offset ($\Delta E_c$) exists between a strained Si layer and a relaxed SiGe layer because the six degenerate valleys in the conduction band of the strained Si layer split into two sub-sets of $\Delta_4$ and $\Delta_2$ (Fig. 4.1(b)) [75]. Due to the lower energy level of $\Delta_2$ states than that of $\Delta_4$ states, the electrons are in $\Delta_2$ states. In relaxed SiGe layers, the degeneracy of six is preserved. The conduction band offset $\Delta E_c$ between a strained Si layer and a relaxed SiGe layer represents the energy difference between the conduction band edge in SiGe and the energy level of $\Delta_2$ in Si, which depends on the Ge fraction in the SiGe layer.

4.2.2 Layer Structure and Epitaxial Growth

A typical layer structure of a Si 2DEG and the associated band energy diagram are shown in Fig. 4.2. From the substrate to the surface, there are several epitaxial layers:
Fig. 4.2  (a) Layer structure of a Si 2DEG and (b) the associated band diagram with a Ge fraction of 0.27 in a Si/SiGe heterostructure without gate bias.

(i)  $\text{Si}_{1-x}\text{Ge}_x$ graded buffer ($0 < x < 0.27$ for this work) + $\text{Si}_{0.73}\text{Ge}_{0.27}$ relaxed buffer,

(ii) strained Si quantum well (2DEG),

(iii) relaxed $\text{Si}_{0.73}\text{Ge}_{0.27}$ setback,

(iv) relaxed $\text{Si}_{0.73}\text{Ge}_{0.27}$ supply of electrons,

(v) relaxed $\text{Si}_{0.73}\text{Ge}_{0.27}$ cap,

(vi) strained Si cap.

The basic functions of these layers in a Si 2DEG are described as follows: first, the $\text{Si}_{1-x}\text{Ge}_x$ graded layer ($0 < x < 0.27$) provides a buffer between the growth interface and the Si 2DEG layer on top. Due to the large lattice mismatch between Si and SiGe, the misfit dislocations exist at the growth interface and the threading dislocations punch through the 2DEG layer [19]. With a slow grading rate of 10% Ge/μm, a misfit dislocation as low as $5 \times 10^5$ cm$^{-2}$ was be achieved in a $\text{Si}_{1-x}\text{Ge}_x$ graded buffer ($0 < x < 0.3$) [76]. Next, a thick relaxed SiGe layer of constant Ge fraction is grown on top of the $\text{Si}_{1-x}\text{Ge}_x$ graded layer. This layer offers a buffer between the buried dislocations and the
2DEG layer on top to reduce the electron scattering by those dislocations.

For (ii), a strained Si layer - where the 2DEG resides - is grown between two relaxed \( \text{Si}_{0.73}\text{Ge}_{0.27} \) layers, forming a quantum well for electron confinement. The thickness of this layer is limited by the critical thickness of a strained Si layer on a relaxed SiGe layer, so as to prevent any further induced dislocations. A relaxed \( \text{Si}_{0.73}\text{Ge}_{0.27} \) setback layer (iii) not only provides the required energy confinement for electrons in the Si QW layer, but also separates the supply layer of ionized impurities from the 2DEG, leading to less impurity scattering. With the remote supply layer of n-type doping (iv), electrons can be “supplied” to the Si QW layer. Due to the separation of 2DEG and those ionized impurities in the supply layer, electron scattering can be significantly reduced, resulting in very high electron mobility.

Due to the surface segregation of the n-type dopants in the SiGe layers (chapter 3), a thick SiGe cap layer (v) has to be grown to avoid a high doping level at the surface which induces high gate leakage current through a Schottky barrier on the surface. Finally, a Si cap layer offers a much more stable surface than a SiGe layer, which is crucial for the subsequent processing steps.

In our work, the relaxed SiGe buffer layers were provided by Amberwave Inc. with a graded \( \text{Si}_{1-x}\text{Ge}_x \) layer \( (0 < x < 0.27) \) of a grading rate of 10% Ge/\( \mu \text{m} \) grown on top of Si substrates followed by a \( \text{Si}_{0.73}\text{Ge}_{0.27} \) buffer layer of 2 \( \mu \text{m} \). Chemical mechanical polishing (CMP) was applied to reduce the surface roughness prior to the epitaxial growth. The wafers were cleaned by the following steps: 5 min in diluted HF (1%), 15 min in \( \text{H}_2\text{SO}_4: \text{H}_2\text{O}_2 \) (2.5:1), followed by 2 min in diluted HF (1%). Next, the wafers were heated to 850 °C for 5 min in hydrogen gas at 6 torr to remove the residual oxide.
Fig. 4.3  SIMS analysis of a typical Si 2DEG (sample #5737). Ge, P, B, C, and O were measured from the surface to below the growth interface at a depth of 255 nm.

before the epitaxial growth began. The flow rate of hydrogen was 3 slm. Diluted SiH₄ (10 % in argon) of 50 sccm and GeH₄ (0.8 % in hydrogen) of 76 sccm were used for the growth of the Si and SiGe epitaxial layers. The growth rates of Si at 625 °C and SiGe at 575 °C were 0.5 nm/min and 5 nm/min, respectively. Diluted phosphine (100 ppm in hydrogen) of 2 sccm was used as a doping gas for the growth of an n-type SiGe supply layer. After baking, a SiGe buffer layer of 100 ~ 150 nm was grown at 575 °C, followed by a strained-Si layer (2DEG layer) of 6 ~ 30 nm at 625 °C, a SiGe setback layer of 10 ~ 70 nm at 575 °C, a n-type SiGe supply layer of 10 nm at 575 °C with a doping level between 10¹⁸ to 10¹⁹ cm⁻³, a SiGe cap layer of 20 ~ 50 nm at 575 °C, and a Si cap layer of 2 nm at 625 °C.

The layer thicknesses, doping level and Ge fraction were measured by SIMS.
Ge, P, B, C, and O in a typical modulation-doped Si 2DEG (sample # 5737) are plotted versus depth in Fig. 4.3. At a depth of 96 nm, a strained Si QW layer of 12 nm is sandwiched between the relaxed SiGe layers. The thickness of the SiGe setback layer is 50 nm between the upper SiGe/Si interface and the position of the peak phosphorus level. The surface segregation of phosphorus was suppressed with a turn-off slope of 13 nm/decade in the SiGe cap layer, which was grown at 525 °C. At the re-growth interface of depth 255 nm, a spike of C and O indicates the incomplete removal of contaminants. Although baking at higher temperatures can be used for the better cleaning of the growth interface, dislocations can be created via the strain relaxation of the underlying graded Si$_{1-x}$Ge$_x$ layer. Thus, other approaches such as in-situ cleaning using HCl or HF [77], or a thick SiGe buffer layer which increases the distance of remote scattering at the growth interface to 2DEG [78], were suggested. In this work, the latter approach was used and the experimental results are presented in section 4.4.1. The C levels in the SiGe and Si layers are $1.8 \times 10^{17}$ and $8 \times 10^{16}$ cm$^{-3}$, respectively. The O levels in the SiGe and Si layers are $7.3 \times 10^{18}$, and $1.5 \times 10^{17}$ cm$^{-3}$, respectively.
After the growth, the wafers were mesa-etched to the Hall bar geometry, followed by AuSb deposition (1% Sb) and annealing at 450 °C over 10 minutes for electrical contacts. For the gating experiment (in section 4.5), palladium was thermally evaporated in order to cover the entire Hall bar for Schottky gating on single 2DEG devices. The typical Hall bar device is shown in Fig. 4.4. For low-temperature measurements, samples were first wire bonded and dipped into a liquid helium dewar at 4 K or 0.3 K. The Hall measurements were performed by lock-in technique to measure the longitudinal ($R_{xx}$) and transverse ($R_{xy}$) magneto-resistances for the analysis of electron density and mobility. The typical drive current is 100 nA.

4.2.3 Electrostatics

In this section, we focus on the electrostatics of a Si 2DEG. With a simulator solving Poisson’s and Schrödinger’s equations at the same time [23], the energy diagram of the conduction band in a Si 2DEG at 4 K is plotted in Fig. 4.5. Two boundary conditions were assumed for this band diagram. First, the Fermi level ($E_F$) is pinned at the midgap of the Si surface layer due to the presence of many surface defects. The second boundary condition is that $E_F$ is also pinned at the donor level ($E_d$) in the supply layer, since the devices were measured at low temperatures. The electron density of the 2DEG can be solved by the following equation [79]:

$$\Delta E_C = E_{edge} + E_d + \frac{e^2 N_d L^2}{2\varepsilon \varepsilon_0} + \frac{e^2 n_{2D} d}{\varepsilon \varepsilon_0},$$

(4.1)

where $E_d$ is the energy difference between the donor level and the conduction band edge in the SiGe layer, $E_{edge}$ is the energy difference between $E_F$ and the conduction band
Fig. 4.5  Conduction band diagram of a Si 2DEG at 4 K with two boundary conditions for the Fermi level ($E_F$) to be pinned at (i) the mid-gap of Si surface and (ii) the donor level.

edge at the upper SiGe/Si interface, and the last two terms represent the energy drop across the ionized impurity in the charged supply layer of length $L$ and the neutral setback layer of length $d$. We assume zero doping in the Si QW and in the structure below the QW, so the electric field below the quantum well is zero.

Since the typical 2DEG density and the doping level are on the order of $10^{11}$ cm$^{-2}$ and $10^{18}$ cm$^{-3}$, respectively, the required width $L$ can be calculated by $N_d L = n_{2D} \sim 1$ nm, which is much smaller than the thickness of the setback layer (20 ~ 70 nm). Therefore, the third term in Eq. (4.1) is usually ignored. Furthermore, because $E_{edge}$ of ~ hundreds μeV is fairly small compared to $\Delta E_c$ (~ 180 meV for Si$_{0.7}$Ge$_{0.3}$ [80]), to first order only, the second and last terms are considered in the analysis. The electron density $n_{2D}$ in the Si QW layer can be solved as:
\[ n_{2D} \approx \frac{e\varepsilon_0}{e^2} \frac{\Delta E_C - E_d}{d}. \]  

This shows that the density at equilibrium (without external gating) can be manipulated by adjusting the conduction band offset (i.e., changing the Ge fraction in the SiGe barrier) or the thickness of the setback SiGe layer. In this work, the Ge fraction is fixed at 0.27 and we focus on the effect of the SiGe setback layer thickness on 2DEG characteristics (section 4.4.3).

### 4.2.4 Mobility Model

To investigate the scattering mechanisms in a Si 2DEG, we followed the procedures derived by Davies [79]. For quantum dot applications, achieving a single electron would require a low density in a 2DEG or a small dot area by a simple relation of \(1 = n_{2D} \times \text{Area}.\) With a QD of 30 nm \(\times30\) nm (typical of advanced lithography), \(n_{2D}\) must be at most \(\sim 1 \times 10^{11}\) cm\(^{-2}\). Thus, we focus on the transport properties of the 2DEGs at densities below \(4 \times 10^{11}\) cm\(^{-2}\), where remote or background impurity scattering has been suggested as the dominant scattering mechanism [81, 82].

The strongest scattering at low temperatures in many 2DEG systems arises from ionized impurities such as n-type donors in the supply layer or throughout the entire material. The former is usually called ‘remote impurity scattering’ and the latter is usually referred to as ‘background impurity scattering’ [79, 82]. For remote impurity scattering, we assumed for simplicity that the supply layer is delta-doped with a two-dimensional density of \(N_{imp}\) ionized impurity, and located at a distance of \(d\) from the doping plane to the edge of the Si QW layer. The resulting electron transport time based
on Fermi-Golden’s rule can be expressed as [79]

\[
\frac{1}{\tau_{\text{r}}} = \frac{\hbar N_{\text{imp}}}{32m^*d^3 \sqrt{\pi n_{2D}^3}},
\]

(4.3)

where \( m^* \) is the effective mass of electron in the Si QW layer and the effects of the degeneracy of the two valley states and two spin states were already included. Next, the mobility \( (\mu = \frac{e\tau_{\text{r}}}{m}) \) is given by

\[
\mu_{\text{remote}} = \frac{32e d^3 \sqrt{\pi n_{2D}^3}}{\hbar N_{\text{imp}}},
\]

(4.4)

When \( n_{2D} \) increases, there is stronger electron screening [79] and so the mobility increases. As the distance between the plane of ionized impurities and the 2DEG increases, electron scattering is weaker and the mobility is higher. On the other hand, as the number of ionized impurities is reduced, the remote scattering becomes weaker and the mobility increases.

Ideally, in a 2DEG structure, the only doped region is the supply layer. However, in practice - in other layers a minimum level of certain types of impurities exists, depending upon the reactor history and growth conditions. The ionized background impurities across the entire sample also scatter the 2DEG with a predicted mobility as:

\[
\mu_{\text{background}} = \frac{2e \sqrt{n_{2D}}}{\sqrt{\pi \hbar N_{\text{background}}}}.
\]

(4.5)

In a similar manner to remote impurity scattering, as the electron density increases, the background impurity scattering is reduced due to stronger electron screening. When the
number of ionized background impurities is reduced, the scattering will be weaker.

The total mobility related to these two types of impurities can be written as:

\[
\frac{1}{\mu_{\text{total}}} = \frac{1}{\mu_{\text{remote}}} + \frac{1}{\mu_{\text{background}}}. \tag{4.6}
\]

In the next section, we will show the importance of background scattering in a Si 2DEG that has been grown. Afterwards, we present a much improved mobility by reducing the background impurity level in the CVD system at Princeton University.

The above model of 2DEG mobility was proposed with several assumptions, such as the delta doping in the supply layer. In practice, broadened n-type doping profiles in the SiGe layer are caused by strong surface segregation (chapter 3), which might offset the predicted mobility by the above model. Furthermore, the abruptness of the SiGe/Si/SiGe heterostructure may not be perfectly sharp, resulting in a perturbation on the quantum states in the Si QW layer. During the derivation for electron screening, many numerical approximations were made [79]. This made the derivation easier, but is not of spectacular accuracy. As a result, in this work, we use this theoretical model only as a rough guide to help us to understand the electron transport properties, but not for any precise numerical fitting or model comparison with our data.

4.3 Efforts Toward High Mobility in Si 2DEGs

4.3.1 Effects of Phosphorus Background Impurity on 2DEG Mobility

In CVD systems, the background level of n-type dopants such as phosphorus or arsenic may be very high due to the memory effect [83]. In Fig. 4.6, the SIMS data of a typical Si 2DEG (sample # 5144) is illustrated. The gas precursors for the epitaxial
growth of this sample were SiH$_2$Cl$_2$ and diluted GeH$_4$ with diluted PH$_3$ as the n-type doping gas. At a depth of 40 nm, a phosphorus doped layer of $1.5 \times 10^{18}$ cm$^{-3}$ was grown. The phosphorus background levels in the Si and SiGe layers are $6 \times 10^{17}$ cm$^{-3}$ and $2 \times 10^{17}$ cm$^{-3}$, respectively. Note that below the growth interface at 250 nm, the phosphorus level in the relaxed buffer grown by Amberwave Inc. is below $10^{15}$ cm$^{-3}$, showing that the phosphorus levels in the epitaxial layers grown by our CVD are true and not artifacts or detection limits from the SIMS measurements. By assuming $n_{2D} \sim 5 \times 10^{11}$ cm$^2$ and ignoring the remote impurity scattering, the electron mobility at a background impurity of $5 \times 10^{17}$ cm$^{-3}$ can be estimated by Eq. (4.5) as 5,000 cm$^2$/V-s, which is close to our experimental results of 3,000 ~ 8,000 cm$^2$/V-s. Thus, the first task in improving the 2DEG mobility is to lower the background impurity level in the epitaxial layers.
4.3.2 The Reduction of Phosphorus Background Impurity Levels

A low phosphorus background level has been suggested for a 2DEG of high mobility. In this section, we will introduce the means by which we identified the sources for high background level of phosphorus in our CVD system, followed by an approach to reduce the phosphorus level and our experimental results.

First, we evaluate the impact of the potential sources for a high phosphorus level, such as the quartz tube and the quartz wafer holder. Since the Si wafers were often baked or annealed at high temperatures (~ 1000 °C) in our CVD system, phosphorus adsorbed on the quartz walls might desorb and incorporate into the epitaxial films. To remove the coated phosphorus on the quartz wall and the wafer stand, we used HF:HNO₃:H₂O (1:4: 6) to etch the deposited films on the quartz surface, followed by a DI water rinse. However, in a test 2DEG structure grown after the wet cleaning steps, the phosphorus background level was rather high (~ 10^{17} \text{ cm}^{-3}) and the mobility was still low (5,000 ~ 10,000 \text{ cm}^2/\text{V-s}). Even with a brand new tube and wafer stand used for the growth of 2DEG devices, identical results of the high background level of phosphorus and low mobility were obtained. This suggests that the main source of phosphorus contamination was not the tubes or wafer stands.

In addition to the quartz tube and stand, the manifold for gas mixing in the gas supply system could be another source of phosphorus contamination. To overcome this problem, we designed a gas delivery system which isolates PH₃ and other gases (Fig. 4.7). The distance from the mixing valve of PH₃ and other process gases to the chamber is 90 cm. The phosphorus level is expected to be lower than that in the old panel.
Fig. 4.7  New gas supply system with separation of process gases (H₂, SiH₄, GeH₄, etc.) from PH₃ to reduce the memory effect of phosphorus.

Fig. 4.8  B, P, Ge profiles of a test structure of multiple Si and SiGe layers grown. Both the B and P levels are down to the SIMS detection limits of $5 \times 10^{15}$ and $3 \times 10^{14} \text{ cm}^{-3}$, respectively, except for a phosphorus doped layer grown at the depth of 360 nm.

because there is only a small proportion of tubing which could have a memory effect to contaminate the other process gases. The SIMS results of a test structure of an undoped 2DEG structure (sample #5503) grown by this new gas panel supported our concept (Fig.
Fig. 4.9  P and B profiles in Si layers grown at different temperatures (sample #5823).

4.8). The boron level and the phosphorus level are as low as the SIMS detection limits of $5 \times 10^{15}$ cm$^{-3}$ and below $10^{15}$ cm$^{-3}$, respectively. In this sample, a phosphorus doped layer was grown at the depth of 360 nm. The phosphorus levels before PH$_3$ turn-on and after PH$_3$ turn-off are identical, showing that the memory effect is insignificant using the new gas panel equipped with the gas separation system.

A test sample was grown after more than 100 runs of 10 μm deposition in the reactor to confirm the negligible memory effect of phosphorus in the entire system. In this sample, only Si layers were grown at different temperatures, with very thin SiGe layers as indicators for the SIMS analysis (Fig. 4.9). For phosphorus, in the Si layer grown below 750 °C, its level is as low as the detection limit of $1.5 \times 10^{14}$ cm$^{-2}$, which is three orders of magnitude lower than the level by using the old gas panel. Between 260 to 350 nm, the phosphorus level is slightly higher. Probably because the layer was
grown immediately after high temperature baking, the desorption of the absorbed phosphorus on the quartz wall was stronger, leading to higher incorporation into the films. For boron, although it is as low as $3 \times 10^{14}$ cm$^{-3}$, it is surprising that its level is slightly higher, since a boron precursor has never been used in the new gas panel. The low boron level in the system could be a SIMS effect or the auto-doping from the lightly boron-doped SiGe relaxed buffer substrates grown by Amberwave Inc.

### 4.3.3 Transport Properties of Si 2DEGs with Low Phosphorus Background Impurity Levels

To estimate the effect of a low background level of mid $10^{14}$ cm$^{-3}$ on 2DEG mobility, we assumed that the 2DEG mobility is limited by the background impurity scattering. Using Eq. (4.5), we estimated a mobility of $3 \times 10^{6}$ cm$^2$/V-s, which is almost 10 times higher than the highest mobility reported in this work. Therefore, it is suggested that the mobility in 2DEGs grown by this new gas system is not limited by background impurity scattering, but rather that it is limited by remote impurity scattering. Further experimental results in the following sections confirm our conclusions.

Si 2DEGs of low background impurity were epitaxially grown with the new gas panel and fabricated into Hall bar devices. The low-temperature (4 K and 0.3 K) electron density and mobility of Si 2DEGs with different layer structures (i.e. varying the SiGe setback layer thickness, the doping level, SiGe cap layer thickness, etc.) were measured (Fig. 4.10). The highest mobility was 522,000 cm$^2$/V-s [84], which was ~ 100
Fig. 4.10 Hall mobility vs. electron density at low temperatures (4 K or 0.3 K). A highest mobility of 522,000 cm$^2$/V·s was measured at 0.3 K [84]. Note that each data point was taken from different samples with different layer structures.

Fig. 4.11 Longitudinal ($R_{xx}$) and transverse ($R_{xy}$) magneto-resistances vs. magnetic field at 0.3 K [84].

times higher than the level of the samples grown by the old gas panel. The longitudinal ($R_{xx}$) and transverse ($R_{xy}$) magneto-resistances of sample #5446 were also measured at
0.3 K by Prof. Rokhinson at Purdue University (Fig. 4.11). Clear Shubnikov-de Haas oscillations in $R_{xx}$ and quantum Hall plateaus in $R_{xy}$ were observed, showing the high quality of this sample. The spin splitting and valley splitting occurred at 0.8 T and 3.5 T, respectively. Overshoots of $R_{xy}$ at odd filling factors were observed. Several suggestions have been made to explain such effects, such as a rapid decoupling of overlapping spin-split states [85] and a mixing of overlapping edge states and the intrinsic spin-orbit interaction of 2DEGs [86]. These effects are beyond the scope of this thesis.

4.4 Effects of Layer Structure on 2DEG Mobility

4.4.1 SiGe Buffer Layer

Typically, a Si 2DEG structure is epitaxially grown on a Si (100) substrate, either by MBE or CVD. In Princeton, we grew the epitaxial layers of a Si 2DEG on a virtual substrate which consisted of a relaxed SiGe buffer layer on top of a graded SiGe layer on a Si substrate (provided by Amberwave Inc.). Prior to the growth, chemical mechanical polish was applied to reduce the surface roughness, followed by a wet cleaning to remove the residual contaminants and native oxide on the polished relaxed SiGe surface. Next, SiGe and Si epitaxial layers were grown using the new gas supply system, and Hall bar devices were fabricated for electron transport measurement at low temperatures.

As depicted in Fig. 4.3, there were spikes of C and O at the substrate interface due to inefficient cleaning before epitaxy. Paul et al. have suggested that with a thick SiGe buffer layer on the SiGe virtual substrate, electron scattering from the bottom
growth interface was much reduced [78]. Therefore, we grew test samples (#5858, #5838, and #5862) with three different thicknesses of the SiGe buffer layer (50, 155, and 310 nm), while the rest of the layer structures in those three samples were identical to each other (Table 4.1).

Table 4.1 Layer structures of sample #5858, #5838, and #5862 for study of the effects of SiGe buffer layer thickness on the 2DEG mobility

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Thickness (nm)</th>
<th>Growth Temp (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si cap</td>
<td>4</td>
<td>625</td>
</tr>
<tr>
<td>SiGe cap</td>
<td>50 ~ 70, 50 ~ 70</td>
<td>550</td>
</tr>
<tr>
<td>SiGe supply</td>
<td>10</td>
<td>575</td>
</tr>
<tr>
<td>P level (cm⁻³)</td>
<td>1 ~ 4 × 10¹⁸</td>
<td>575</td>
</tr>
<tr>
<td>SiGe setback</td>
<td>30</td>
<td>575</td>
</tr>
<tr>
<td>Si QW</td>
<td>10</td>
<td>625</td>
</tr>
<tr>
<td>SiGe buffer</td>
<td>50</td>
<td>575</td>
</tr>
</tbody>
</table>

The electron mobility measured at a 4 K vs. SiGe buffer layer thickness is shown in Fig. 4.1. The electron mobility was only slightly reduced for the device with the 75-nm SiGe buffer layer, which followed the experimental results in [78]. Although the density of C and O are high, the degree to which they affect the electron transport (e.g., the ionization ratio and charge type) is not yet known. Since the SiGe buffer layer separates the growth interface and the 2DEG, the remote impurity scattering at this interface could be much reduced with a thicker SiGe buffer layer. With 2DEG mobility saturated as the SiGe buffer layer thickness > 150 nm, we think that 2DEG transport is not limited by the remote impurity scattering from the growth interface, and that the mobility does not depend on the levels of C and O at the growth interface.
4.4.2 Si Quantum Well Layer

Since the 2DEG is in a strained Si layer, several devices with different thicknesses of Si layer grown at 575 °C and 625 °C were made in order to investigate the effects of strained relaxation on 2DEG mobility (Table 4.2). Electron mobility vs. Si

Table 4.2 Layer structures of test samples for study of the effects of Si QW layer thickness on the 2DEG mobility

<table>
<thead>
<tr>
<th>Layer Type</th>
<th>Growth Temp (°C)</th>
<th>Thickness</th>
<th>Growth Temp (°C)</th>
<th>Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si cap</td>
<td>625</td>
<td>4</td>
<td>625</td>
<td>4</td>
</tr>
<tr>
<td>SiGe cap</td>
<td>550</td>
<td>50 ~ 70</td>
<td>550</td>
<td>50 ~ 70</td>
</tr>
<tr>
<td>SiGe supply</td>
<td>575</td>
<td>10</td>
<td>575</td>
<td>10</td>
</tr>
<tr>
<td>P level (cm⁻³)</td>
<td>1 ~ 4 × 10¹⁸</td>
<td>1 ~ 4 × 10¹⁸</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SiGe setback</td>
<td>575</td>
<td>30</td>
<td>575</td>
<td>30</td>
</tr>
<tr>
<td>Si QW</td>
<td><strong>575</strong></td>
<td><strong>12, 21, 30</strong></td>
<td><strong>625</strong></td>
<td><strong>6, 10, 12, 18, 26</strong></td>
</tr>
<tr>
<td>SiGe buffer</td>
<td>575</td>
<td>150</td>
<td>575</td>
<td>150</td>
</tr>
</tbody>
</table>
Fig. 4.13 Mobility at a 4 K vs. Si QW layer thickness grown at 575 °C and 625 °C. The arrow indicates the predicted critical thickness of strained Si on the relaxed Siₐ₀.₇₃Ge₀.₂₇ [87].

QW thickness is shown in Fig. 4.13. A peak near 10 nm is observed. For the 625 °C data, the Si QW is thicker than 12 nm, and the mobility drops significantly as the thickness increases. This may due to dislocations created by the strain relaxation of the thick Si QW layer. The thickness at which mobility drops significantly is close to the reported critical thickness (~ 12 nm) of a strained Si layer on a relaxed Siₐ₀.₇₃Ge₀.₂₇ layer [87]. For 575 °C, a similar trend was observed.

On the other hand, when the Si QW is under 10 nm (e.g. 6 nm), the 2DEG mobility drops slightly when the well is thinner, which could be attributed to the stronger scattering from the roughness at the upper SiGe/Si heterointerface [88]. In a narrow QW, the spreading of electron wavefunction into the SiGe barrier layers becomes stronger, which might also lead to stronger scattering from the remote impurity in the supply layer or alloy scattering in the SiGe barrier [82]. Thus, when we discuss
the ionized impurity scattering in the next section, the thickness of the strained Si QW layer is controlled between 9 to 12 nm and to more easily isolate the effects of different scattering mechanisms on 2DEG mobility.

4.4.3 SiGe Setback Layer

According to the analysis of electrostatics in section 4.2.3, the electron density in the 2DEG layer depends upon the thickness of the SiGe setback layer wherever enough carriers are provided from the supply layer. By adjusting the thickness of the SiGe setback layer, the density can be modulated following Eq. (4.2). The experimental results of electron density vs. the SiGe setback layer thickness is shown in Fig. 4.14(a) with a theoretical calculation of Eq. (4.2). The growth parameters for those devices are listed in Table 4.3. For quantum dot applications, a low electron density is preferred and so a thick setback layer is required. However, for a Si 2DEG of a thick setback layer, the control of patterned top metal gates over the underlying 2DEG will be weaker. Thus, an optimized thickness of a SiGe setback layer is usually used for quantum dot fabrication.

Table 4.3 Layer structures of test samples for study of the effects of SiGe setback layer thickness on the 2DEG mobility.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Growth Temp (°C)</th>
<th>Thickness (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si cap</td>
<td>625</td>
<td>4</td>
</tr>
<tr>
<td>SiGe cap</td>
<td>550</td>
<td>40 ~ 70</td>
</tr>
<tr>
<td>SiGe supply</td>
<td>575</td>
<td>10</td>
</tr>
<tr>
<td>P doping level (cm⁻³)</td>
<td>1 ~ 4 × 10¹⁸ cm⁻³</td>
<td></td>
</tr>
<tr>
<td><strong>SiGe setback</strong></td>
<td><strong>575</strong></td>
<td><strong>12, 20, 30, 40, 50, 70</strong></td>
</tr>
<tr>
<td>Si QW</td>
<td>625</td>
<td>10</td>
</tr>
<tr>
<td>SiGe buffer</td>
<td>575</td>
<td>150</td>
</tr>
</tbody>
</table>
Fig. 4.14 (a) Electron density and (b) mobility vs. SiGe setback layer thickness at 4 K.

In Fig. 4.14(b), electron mobility is also plotted versus the setback layer thickness with a peak mobility at 30 nm. This could be explained as follows: as the setback layer is thicker, the remote impurity scattering from the supply layer becomes weaker – thus, initially the mobility increases. When the thickness of the setback layer is further increased, the electron screening becomes weaker due to lower electron density in the Si QW layer. As such, mobility drops even though the remote scattering itself is reduced with a longer setback distance.

4.5 Extremely Low Electron Density by Effective Schottky Gating in Single 2DEG Devices

In the previous sections, the effects of different layers on 2DEG characteristics were investigated by comparing different 2DEG samples. It is well known that the scattering mechanisms in a 2DEG can be identified and interpreted more accurately in a single device by gating [63]. This allows mobility vs. density to be studied. For a modulation-doped Si 2DEG device, the surface segregation of n-type
Fig. 4.15 Top view of the Hall bar devices. The contacts were made by AuSb alloying (1 % Sb) and top Pd Schottky gate covers the Hall bar.

dopants leads to a high surface phosphorus level and high gate leakage current. By low-temperature epitaxy [56], we could suppress the segregation significantly (see chapter 3), and effective gating without leakage was enabled [84]. In this section, we present 2DEG characteristics at different densities in a single device obtained by palladium Schottky barrier gating.

Two wafers (#5414 and # 5613) were grown for the leakage test of top Pd Schottky gates covering the entire the Hall bar (Fig. 4.15). P and Ge profiles of these two wafers are illustrated in Fig. 4.16. The layer structures and growth parameters of these two samples are listed in Table 4.4. The growth temperature of the SiGe cap layer for #5414 and #5613 are 575 and 525 °C, respectively. The resulting P turn-off slopes of #5414 and #5613 are 70 nm/dec and 13 nm/dec, respectively. Hall bar devices were fabricated and the leakage current via the Pd Schottky gate were measured at 4 K (Fig. 4.17). For #5414 with a slow P turn-off of 70 nm/dec, the P surface level is above $10^{18}$ cm$^{-3}$, leading to a very leaky Hall bar device through the top gate. At low growth temperature of SiGe cap layer (525 °C) for #5613, a P turn-off slope of 13 nm/dec was
Fig. 4.16 SIMS results of Ge and P profiles of (a) sample #5414 and (b) sample #5613. The phosphorus level at the surface are \(2 \times 10^{18}\) and \(2 \times 10^{16}\) cm\(^{-3}\) for sample #5414 and #5613, respectively.

Fig. 4.17 Gate leakage current density at 4 K vs. gate voltage for sample #5414 and #5613, respectively. Pd Schottky gate covered the entire Hall bar shown in Fig. 4.15.

achieved with a much reduced P surface level of \(2 \times 10^{16}\) cm\(^{-3}\). This enabled extremely low gate leakage current at negative applied voltage until -7 V. At forward bias, a normal Schottky diode operation was observed at 0.2 and 0.7 V for #5414 and #5613, respectively.

Two wafers (#5737 and #5747) with different setback layer thicknesses (Fig.
Fig. 4.18 Layer structures of (a) sample #5737 and (b) sample #5747. The phosphorus doping levels are $1 \times 10^{18}$ and $3 \times 10^{18}$ cm$^{-3}$ for (a) and (b), respectively.

Fig. 4.19 SIMIS results of Ge and P profiles for (a) #5737 and (b) #5747. The bump of P in the Si QW layer is thought the artifact of SIMS measurement.

4.18) were grown and Hall bar devices were fabricated with Pd Schottky gates on top, covering the entire Hall bar. The SIMS results of these two wafers confirm the low levels of surface phosphorus ($< 10^{16}$ cm$^{-3}$) by low-temperature epitaxy at 525 °C (Fig. 4.19). Thus, effective Schottky gating was enabled with very low gate leakage current of 0.1 nA compared to the drive current of 100 nA (Fig. 4.20(a)).

The electron density vs. gate voltage of #5737 by Pd Schottky gating is shown
Fig. 4.20 (a) Gate leakage current ($I_g$) and (b) electron density vs. gate voltage ($V_g$) for sample #5737 at 4 K. At $V_g = 0.5$ V, a normal forward-bias Schottky diode operation was observed. The drive current for the Hall measurement was 100 nA.

in Fig. 4.20(b). The effective capacitance between the surface and the 2DEG layer can be extracted from the slope of $n_{2D}$ vs. $V_g$ as $1.2 \times 10^{-7}$ F/cm², which is close the calculated value of $1.3 \times 10^{-7}$ F/cm² based on a parallel capacitor model. By extrapolating the line to $n_{2D} = 0$, $V_g = -0.06$, V was obtained for full depletion of electrons in the 2DEG layer. However, below $V_g = 0.02$ V, there was no conduction in this sample because of metal-insulator transition (MIT). MIT occurs because of the potential fluctuations from the remote impurities of the supply layer, which creates numerous barriers blocking electron transport at low densities [90]. Thus, conduction ceases.

The mobility vs. density of sample #5737 (setback: 50 nm) and #5747 (setback: 20 nm) are shown in Fig. 4.21. At the same density, the mobility of #5737 is higher than that of #5747 due to the thicker setback layer that separates the ionized impurity and 2DEG. As the density increases, the mobility of both samples increases because of the stronger electron screening. The slopes of mobility vs. density for #5737
Fig. 4.21 Hall mobility vs. density at 4 K for sample #5737 (setback: 50 nm) and #5747 (setback: 20 nm).

and #5747 were extracted as $\mu \propto n^{1.4}$ and $\mu \propto n^{3.7}$, respectively. It is suggested in the mobility models [63, 79, 82] that if the mobility is limited by the remote impurity scattering, the exponent of $\mu \propto n^\alpha$ will be close to 1.5. On the other hand, if the background impurity scattering dominates, the exponent is close to 1. This has been experimentally verified with exponents of $0.5 \sim 1.5$ for modulation-doped GaAs 2DEGs [91] and undoped Si 2DEGs [73]. For sample #5737, the mobility is probably limited by the remote impurity scattering, as the exponent is close to 1.5. For sample #5747, the exponent is 3.7, which is much higher than any theoretical numbers. Prior work on Si 2DEGs has also reported an exponent of 2.4, and suggested the inadequacy of the Thomas-Fermi approximation on electron screening as the electrons move closer to the MIT regime [92].
Note that the MIT occurs at $6.5 \times 10^{10}$ and $3.2 \times 10^{11}$ cm$^2$, with mobilities of 30,000 and 7,000 cm$^2$/V-s for #5737 and #5747, respectively. For sample #5737, the mobility at $n_{2D} = 1 \times 10^{11}$ cm$^2$ is 45,000 cm$^2$/V-s. To our knowledge, this is the highest mobility reported for modulated-doped Si 2DEGs at a density of $10^{11}$ cm$^{-2}$ - an indication of the high quality of our 2DEG samples. The higher critical density for MIT to occur in #5747 is attributed to its shorter setback layer, leading to stronger effects of potential fluctuations on 2DEG so that the metallic conduction ceases at a higher electron density [89].

There is a saturation in mobility at $n_{2D} = 3.6 \times 10^{11}$ and $7.9 \times 10^{11}$ cm$^2$ for #5737 and #5747. Several reasons could be attributed to the maximum mobility, such as parallel conduction in the supply layer [93, 94], or the second subband occupancy in the Si QW layer [95], or the interface roughness scattering [88] at the upper SiGe/Si interface coming from a higher electric field pushing the 2DEG to the interface with a high electron density. In chapter 5, a preliminary analysis will show that, in our 2DEG devices, the electron occupancy of the second subband in the Si QW layer is the main reason for the presence of peak mobility.

Finally, concerning the amazingly high exponent of $\mu \propto n^\alpha$ in sample #5747, we note that when we derived the 2DEG mobility in section 4.2.4, electron tunneling between the Si QW layer and the SiGe supply layer was ignored. However, for a 2DEG with a thin setback layer, such as sample #5747 (20 nm), electron tunneling could occur. When tunneling happens, there might be a parallel conduction channel in the remote supply layer [93], even if the electron density in the supply layer is below the critical density as MIT occurs. Thus, the measured Hall density and mobility do not represent
the true 2DEG characteristics and an elaborate analysis is required [94]. Lu et al. reported an upper limit of 2DEG density in an undoped Si 2DEG and suggested that it could be related to the tunneling between the 2DEG and the Si surface layer [96]. However, no further experimental result was presented to support their concept of electron tunneling in a 2DEG structure. Furthermore, exactly how this tunneling would affect electron screening and mobility is not yet known. Thus, a quantitative interpretation of our data of $\mu$ vs. $n$, and the high exponent cannot be made at this time. More work is required in order to understand the effect of tunneling on 2DEG properties.

### 4.6 Summary

In this chapter, we first reviewed the electrostatics and a mobility model of a Si 2DEG. Then, we demonstrated much improved mobility in Si 2DEGs grown by piping PH$_3$ separately from the other gases, which reduces the phosphorus background level to below $10^{15}$ cm$^{-2}$ in epitaxial films grown by our CVD system. A high mobility of 522,000 cm$^2$/V-s without external gating was achieved, which is the highest reported among all un-gated modulation-doped Si 2DEGs grown by CVD.

We then investigated the effects of different layers on the 2DEG properties. The feasibility of using a SiGe virtual substrate for high-quality Si 2DEGs was demonstrated by increasing the thickness of the SiGe buffer layer. Furthermore, the thickness of the Si QW layer was shown to be crucial for the 2DEG mobility due to the detrimental strained relaxation of the Si layer as it is over the critical thickness. A Si QW layer of 9 to 12 nm was suggested for high-mobility Si 2DEGs.
By low-temperature epitaxy, effective Schottky gating on single 2DEG devices was enabled, and an extremely low electron density of $6.5 \times 10^{10}$ cm$^{-2}$ was obtained by gating with a mobility of 45,000 cm$^2$/V-s at the important density of $1 \times 10^{11}$ cm$^{-2}$. The relation between modulated mobility and density was used to identify the dominant scattering mechanism. It was expected that the remote impurity scattering would be the dominant mechanism at densities below $4 \times 10^{11}$ cm$^{-2}$. However, the mobility model could not completely explain our experimental results, particularly the exact numerical dependence of $\mu$ on $n$. A new model to include possible tunneling between 2DEG and the remote supply layer is required in order to accurately explain the 2DEG characteristics in a device with short setback distance.
Chapter 5  Isotopically Enriched $^{28}$Si 2DEGs and Inverted Modulation-Doped Si 2DEGs

5.1  Introduction

In 2005, Petta et al. demonstrated the first spin-based quantum computing in a double quantum dot device operation in a GaAs two-dimensional electron gas (2DEG) [12]. However, strong spin decoherence exists in a GaAs-based QD because of the hyperfine interactions of electron spins and nuclear spins. Alternatively, for Si QDs, greatly reduced spin decoherence was demonstrated with a longer dephasing time ($T_2^*$) [13] because of the lower fraction of the only nuclear-spin-carrying isotope of $^{29}$Si (4.7 %) [97]. Moreover, $^{29}$Si can be reduced by $^{28}$Si enrichment [98] and reduced spin decoherence is predicted [99]. Thus, in the first part of this chapter, we present the work concerning the electron transport properties of enriched $^{28}$Si 2DEGs.

The second part of this chapter is the demonstration of inverted modulation-doped 2DEGs in silicon with natural isotope abundance, with record high mobility of 470,000 cm$^2$/V-s. This inverted structure is crucial to the realization of a bilayer device of two adjacent 2DEGs with a thin tunneling barrier in between [100]. While a GaAs-based bilayer device has been fabricated successfully by the delta-doping technique [101], there is no experiment reported on a Si bilayer device yet. The major obstacle for the demonstration of a Si bilayer device is the severe surface segregation of n-type dopants, which would reduce the mobility in the bottom channel of Si 2DEG due
to stronger impurity scattering from the n-type dopants in the setback layer. In the second part of this chapter, we investigate the effects of phosphorus segregation on electron transport in an inverted modulation-doped Si 2DEG. By low-temperature epitaxy to suppress phosphorus segregation, an inverted modulation-doped Si 2DEG of high mobility of 470,000 cm²/V-s is presented.

5.2 Isotopically Enriched $^{28}\text{Si}$ 2DEGs

Quantum dots (QDs) containing single electrons are very promising for the realization of spin-based quantum computing in solid-state systems because of their scalability [5] and the mature nature of semiconductor technology. The coherent control over the spin states of two single electrons in a double quantum dot was demonstrated in GaAs for the first time [12]. However, its short dephasing time $T_2^*$ of ~7 ns of electron spins, due to the severe hyperfine interactions with the host nuclei [11], imposes a lower limit on the speed of gate switching to preserve the quantum phase information before a gate switching operation is completed. To increase the dephasing time of electron spins, silicon has been suggested as a replacement for GaAs because of the very low spin decoherence resulting from its only spin-carrying isotope $^{29}\text{Si}$ of 4.7 % [97]. A much longer $T_2^*$ of ~360 ns was recently demonstrated in a Si double QDs [13]. With the reduction of $^{29}\text{Si}$ to below 0.01 %, a very long dephasing time has been predicted because of the greatly suppressed hyperfine interactions between electron spins and nuclear spins of $^{29}\text{Si}$ [99].

A 2DEG in an isotopically-enriched $^{28}\text{Si}$ quantum well (QW) was
demonstrated by molecular beam epitaxy (MBE), with the mobility of $55,000 \text{ cm}^2/\text{V-s}$ at density of $3 \times 10^{11} \text{ cm}^{-2}$ \cite{98}. The reason of its relatively low electron mobility compared to Si 2DEGs of natural abundance and the limiting factor of electron scattering are still unknown. Moreover, there is no report yet on an enriched $^{28}\text{Si}$ 2DEG prepared by CVD. Thus, we study the transport properties of enriched $^{28}\text{Si}$ 2DEGs by CVD and estimate the spin decoherence in this section.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Growth Temp. (°C)</th>
<th>Thickness (nm)</th>
<th>Modulation-Doped (Depletion-mode)</th>
<th>Undoped (Enhancement-mode)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si cap</td>
<td>625</td>
<td></td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>SiGe cap</td>
<td>575</td>
<td></td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>SiGe supply</td>
<td>575</td>
<td></td>
<td>10</td>
<td>(no doping)</td>
</tr>
<tr>
<td>(doping level)</td>
<td>575</td>
<td>$(4 \times 10^{18} \text{ cm}^{-3})$</td>
<td>60 (#5854) or 150 (#5853)</td>
<td></td>
</tr>
<tr>
<td>SiGe setback</td>
<td>575</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Si quantum well</td>
<td>625</td>
<td>16</td>
<td>9*</td>
<td></td>
</tr>
<tr>
<td>SiGe buffer</td>
<td>575</td>
<td>110</td>
<td>150</td>
<td></td>
</tr>
</tbody>
</table>

* $^{28}\text{Si}$ was only enriched in the Si quantum well.

The preparation of enriched $^{28}\text{Si}$ 2DEG samples was similar to the steps described in chapter 4, except that a specialized silane of enriched $^{28}\text{Si}$ (provided by Voltaix Inc.) was used as a gas precursor. After the cleaning steps, relaxed SiGe substrates were loaded into the CVD reactor for epitaxial growth. A SiGe buffer layer of 100 ~ 150 nm was first grown at 575 °C, followed by a strained-Si layer (2DEG layer)
at 625 °C, a SiGe setback layer at 575 °C, a n-type SiGe supply layer at 575 °C, a SiGe cap layer at 525 °C, and a Si cap layer at 625 °C (Table 5.1). There are two sets of samples in this section. First, a depletion-mode device of modulation-doped 2DEG with $^{28}$Si enriched throughout the entire epitaxial growth (including Si and SiGe epitaxial layers) was fabricated. For enhancement-mode samples, the supply layer was grown without n-type doping and $^{28}$Si was only enriched in the Si QW layer.

### 5.2.1 Reduction of Spin-Carrying Isotope $^{29}$Si by $^{28}$Si Enrichment

The concentrations of three isotopes, $^{28}$Si, $^{29}$Si, and $^{30}$Si, and Ge vs. depth in a modulation-doped enriched $^{28}$Si 2DEG device are shown in Fig. 5.1. Below the growth interface at 185 nm depth, the fractions of $^{28}$Si, $^{29}$Si, and $^{30}$Si are 92 %, 4.7 %, and 3.3 %, respectively, which are identical to the compositions of the natural isotopic

![Fig. 5.1](image.png)  
Fig. 5.1 Three isotopes $^{28}$Si, $^{29}$Si, and $^{30}$Si vs. depth in 2DEG structure with Ge as an indicator by SIMS measurements. The growth began at a depth of 185 nm and Si QW is at a depth of 75 nm.
Fig. 5.2  Dephasing time vs. $^{29}$Si fraction. Solid line is the model prediction, the solid square is the experimental result in Si QD [13], and the vertical dash line represents the fraction of $^{29}$Si in our enriched $^{28}$Si 2DEGs. Dephasing time of GaAs QDs [12] is also presented for comparison.

abundance [97]. For Si and SiGe epitaxial layers grown with silane of enriched $^{28}$Si, the fractions for those three isotopes became 99.72 %, 0.08 %, and 0.002 %, respectively. The enrichment factors, defined as the ratios of $^{28}$Si to $^{29}$Si and $^{28}$Si to $^{30}$Si, were increased from 20 to 1250 (i.e., 60 times enhancement), and 27 to 50,000 (i.e., 2000 times enhancement), respectively.

We now estimate the potential effects of $^{28}$Si enrichment on the spin decoherence in Si QDs. By the enrichment of $^{28}$Si, the spin decoherence of QD electrons from the nuclear spins of $^{29}$Si can be greatly suppressed. This would reduce the dephasing time ($T_2^*$), which is dominated by the hyperfine interactions between nuclear
spins and electron spins. Assuming $10^5$ nuclei in a Si QD of 100 nm $\times$ 100 nm, Assali et al. proposed a numerical model to estimate the dephasing time in Si [99]

$$T_2^* = \frac{\hbar}{4.3e\sqrt{10^3 r}} \times 10^{11}$$

(5.1)

where $e$ is the electron charge and $r$ is the fraction of $^{29}$Si. The predicted dephasing time versus $r$ was shown in Fig. 5.2 and compared with experimental results of a GaAs QD and a Si QD of natural abundance. Maune et al. reported the first dephasing time of 360 ns in Si double QDs of natural abundance [13], which is 50 times longer than that in a GaAs QD and very close to Assali’s prediction. In our sample, the fraction of $^{29}$Si was purified by a factor of 20 to 0.08% and the dephasing time is expected to be 2 μs, two orders of magnitude longer than that of a GaAs QD and 6 times longer than that of a Si QD of natural abundance. According to the model, if $^{29}$Si can be reduced to 10 ppm, the decoherence can be further reduced with $T_2^*$ longer than 10 μs.

5.2.2 Magneto-Transport Properties of a Modulation-Doped Enriched $^{28}$Si 2DEG

For the depletion-mode device (#5514), Hall electron density at 4 K was $4 \times 10^{11}$ cm$^{-2}$ with mobility of 399,000 cm$^2$/V-s. The longitudinal ($R_{xx}$) and transverse (Hall) resistances ($R_{xy}$) were also measured at 0.3 K with the magnetic field up to 8 T (Fig. 5.3). The onset of Shubnikov-de Haas (SdH) oscillations in $R_{xx}$ occurs at 0.4 T. The spin splitting, which occurs because of the associated Zeeman energy difference exceeding the Landau level broadening, occurs at 0.75 T with a filling factor of $\nu = 24$. The revelation of the two-fold degeneracy from two valleys of density of states was observed
Fig. 5.3 Magneto-resistances of depletion-mode enriched $^{28}\text{Si}$ 2DEG device measured at 0.3 K. Electron density ($4.02 \times 10^{11} \text{ cm}^{-2}$) and mobility (522,000 cm$^2$/V-s) were extracted from the periods of Shubnikov-de Haas oscillations in longitudinal resistance ($R_{xx}$) vs. (1/B) and its value at zero field.

at 1.9 T with $\nu = 9$. For Hall resistance ($R_{xy}$), the quantum Hall structures can be resolved at $B = 0.7$ T with the filling factor of $\nu = 24$ and clear plateaus were observed at $\nu = 2, 4, 8,$ etc. The two-dimensional electron densities extracted from SdH oscillations and low-field Hall resistance were $4.02$ and $4.18 \times 10^{11} \text{ cm}^{-2}$, respectively, showing that parallel conduction is insignificant. The electron mobility of this device at 0.3 K was 522,000 cm$^2$/V-s, with an associated mean free path of 6 μm, which we believe is the highest reported for any modulation-doped Si 2DEG grown by CVD, regardless of $^{28}\text{Si}$ enrichment. In previous work on isotopically enriched $^{28}\text{Si}$, the highest reported mobility was 55,000 cm$^2$/V-s in a MBE-grown 2DEG. Our results established the extremely high quality in the isotopically enriched samples.
5.2.3 Gating of Enhancement-Mode Enriched $^{28}$Si 2DEGs

For quantum dot devices, a short distance between the surface and 2DEG layer is preferred for fine gate control. Thus, we chose shallow 2DEGs with the SiGe setback layer < 150 nm thick. Two enhancement-mode devices, with undoped enriched $^{28}$Si used only in the Si QW layer, were made, without any n-type dopants, and with a SiGe setback layer of 60 (5854) and 150 nm (5853). The growth parameters and layer structures are listed in Table 5.1. Al$_2$O$_3$ of 90 nm was deposited at 300 °C over the entire device. A metal gate of Cr/Au with a Hall bar shape was deposited on top of Al$_2$O$_3$ (Fig. 5.4). In these undoped enhancement-mode devices, there cannot be parallel conduction in an n-type supply layer. Aside from this, since the Si surface layer was thin (< 3 nm), the quantum states in that layer are higher than the Fermi level, eliminating the potential parallel conduction at the surface. Furthermore, considering the intimate proximity with the Si/Al$_2$O$_3$ interface, the surface conduction is negligible because of the strong scattering from the interface impurity and roughness compared to the conduction in the 2DEG layer. Thus, the measured Hall density and mobility are considered to reflect 2DEG transport properties in this work.

With a metal gate of Cr/Au on top of 90-nm Al$_2$O$_3$ (Fig. 5.4), the electron

Fig. 5.4 (a) The layer structure of enhancement-mode devices of Si 2DEGs, and (b) the top view of the Hall bar geometry.
Fig. 5.5 Electron density vs. gate voltage by Hall measurement at 4 K for enhancement-mode devices of enriched $^{28}\text{Si}$ 2DEG with a SiGe setback layer of 60 and 150 nm. The slopes represent the effective capacitance between a Cr/Au gate and 2DEG. Density and mobility in the 2DEG layer can be modulated and the dominant scattering mechanisms could be verified by checking $\mu$ vs. $n_{2D}$ in relation to different thicknesses of the SiGe setback layer (60 and 150 nm). The electron density vs. gate voltage of those two samples was shown in Fig. 5.5. The capacitances extracted from the slopes are $5.8 \times 10^{-8}$ F/cm$^2$ and $4.1 \times 10^{-8}$ F/cm$^2$ for the setback layer of 60 and 150 nm, respectively, within 5% of the values calculated with a parallel-plate capacitor model. The lowest densities are $1.1 \times 10^{11}$ cm$^{-2}$ at $V_g = 2.2$ V, which we believed is the lowest density among all reported enriched $^{28}\text{Si}$ 2DEGs. The extrapolation of $n_{2D}$ vs. $V_g$ to zero density gives the threshold voltage ($V_T$) approximately 2 V. Below $V_g = 2.2$ V, however, there was no conduction in the 2DEG channel because of the metal-insulator transition (MIT) [90]. By assuming the threshold voltage is dominated by impurity charge ($Q_{int}$) at
Fig. 5.6 Hall mobility vs. density for enhancement-mode devices of enriched $^{28}$Si 2DEG with a SiGe setback layer of 60 and 150 nm at 4 K.

the Si/Al$_2$O$_3$ interface, a simple estimation can be made with $N_{int} = Q_{int}/e = C_{ALO} \times V_T/e$ on the order of $10^{12}$ cm$^{-2}$, where $C_{ALO}$ is the capacitance ($7.9 \times 10^{-8}$ F/cm$^2$) of 90-nm Al$_2$O$_3$ dielectric layer assuming its relative dielectric constant is 8.

In Fig. 5.6, electron mobility vs. density at 4 K for these enhancement-mode devices is illustrated. For the device with a 150-nm SiGe setback layer, the lowest density is $6.2 \times 10^{10}$ cm$^{-2}$ with mobility of 28,000 cm$^2$/V-s. Below this critical density, electron conduction ceased because MIT occurs. For both devices, the mobility scales with the density as $\mu \propto n^{1.7}$. Based on previous work [82, 91], if the 2DEG is limited by the remote impurity scattering, then $\alpha = 1.5$. Thus, we concluded that the major source of electron scattering in these enriched $^{28}$Si 2DEG devices is remote impurities at the Si/Al$_2$O$_3$ interface.
5.3 Inverted Modulation-Doped Si 2DEGs

In chapter 4, the remote impurity charges in the supply layer were suggested to be the dominant scattering sources in a modulation-doped device of Si 2DEG, with the supply layer on top of the Si QW layer. Due to a fast turn-on, the level of n-type dopants is low in the setback layer, and the effects of impurity scattering between the doped layer and 2DEG is usually ignored. For inverted structure (Fig. 5.7), with the supply layer below the 2DEG layer, due to the severe surface segregation of n-type dopants, the resulting high level would lead to stronger remote impurity scattering and a greatly reduced mobility. Furthermore, P segregates into the Si QW layer, resulting in high P background levels there (Table 5.2 or Fig. 5.8), which would reduce the 2DEG mobility further. While a sharp arsenic turn-off by ion implantation for the bottom n-type doping layer in an inverted Si 2DEG device was reported [102], the mobility was limited by the defects induced during implant process and the quality of the re-growth interface. Furthermore, high temperature annealing is required to activate the dopants.

![Fig. 5.7 Typical layer structure of an inverted modulation-doped Si 2DEGs. The SiGe supply layer of phosphorus modulation-doping is below the 2DEG layer (Si QW).](image-url)

<table>
<thead>
<tr>
<th>Layer</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si Cap (625°C)</td>
<td></td>
</tr>
<tr>
<td>SiGe Cap (525°C)</td>
<td></td>
</tr>
<tr>
<td>Si QW (625°C)</td>
<td></td>
</tr>
<tr>
<td>SiGe Setback (500 ~ 575°C)</td>
<td></td>
</tr>
<tr>
<td>SiGe Supply (575°C) P level ~ $3 \times 10^{18}$ cm$^{-3}$</td>
<td></td>
</tr>
<tr>
<td>SiGe Relaxed Buffer (575°C)</td>
<td></td>
</tr>
</tbody>
</table>
and remove the defects by implant, which would induce strain relaxation and increase the thermal budget for the subsequent processing steps. By low-temperature epitaxy (chapter 3), the phosphorus segregation can be significantly reduced with a sharp turn-off, enabling a low level in the setback layer. Thus, sample #5457, #5850, and #5630. P turn-off slopes of these three samples are 40, 14, and 8 nm/dec, respectively. The setback layer thicknesses between the phosphorus peak level and the lower Si/SiGe heterojunction are 20, 20, and 33 nm for those three in the final part of this thesis, we investigate the effect of phosphorus segregation on 2DEG characteristics and present the highest reported mobility of 470,000 cm²/V-s for all inverted modulation-doped Si 2DEGs.

5.3.1 Effects of Phosphorus Turn-Off Slope on 2DEG Mobility

To investigate the effects of phosphorus segregation on 2DEG properties in an inverted structure, three samples of different phosphorus turn-off slopes (40, 14, and 8 nm/decade) were grown at different temperatures. The details of the layer structures are listed in Table 5.2 and the associated SIMS results are illustrated in Fig. 5.8. For those three samples, the thickness of the setback layer is defined as the distance between the phosphorus peak and the lower Si/SiGe heterojunction. Hall electron density and mobility measured at 4 K are plotted versus the phosphorus turn-off slope in Fig. 5.9.

The Hall electron densities of sample #5457, #5850, and #5630 are $3.8 \times 10^{12}$ cm⁻², $7 \times 10^{11}$ cm⁻², $1.8 \times 10^{11}$ cm⁻², respectively. As P turn-off slope increases, there is more phosphorus in the SiGe setback layer, which would effectively reduce the setback layer thickness by transferring electrons at shorter distances. Furthermore, high
P background levels in the Si QW layer might contribute more electrons in the 2DEG layer. Thus, with the same setback layer thickness of 20 nm, the density of #5457 with P turn-off slope of 40 nm/dec is higher than that of #5850 with P turn-off slope of 12 nm/dec. For #5630 of a thicker setback layer, the P turn-off slope is 8 nm/dec and the P level is very low (below the detection limit), so the electron density is less than those of #5457 and #5850.

Table 5.2 Layer structures of sample #5457, #5850, and #5630 for study of the effects of P turn-off slope on 2DEG characteristics in an inverted modulation-doped structure.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>5457</th>
<th>5850</th>
<th>5630</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si cap (nm) @ 625°C</td>
<td>5</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>SiGe cap (nm) @ 575°C</td>
<td>46</td>
<td>38</td>
<td>50</td>
</tr>
<tr>
<td>Si QW (nm) @ 625°C</td>
<td>12</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>P level in Si QW</td>
<td>$0.6 \sim 2 \times 10^{18}$ cm$^{-3}$</td>
<td>$1 \sim 2 \times 10^{17}$ cm$^{-3}$</td>
<td>below $5 \times 10^{15}$ cm$^{-3}$ (SIMS limit)</td>
</tr>
<tr>
<td>SiGe setback (nm)</td>
<td>20</td>
<td>20</td>
<td>33</td>
</tr>
<tr>
<td>P turn-off slope</td>
<td>40 nm/dec</td>
<td>14 nm/dec</td>
<td>8 nm/dec</td>
</tr>
<tr>
<td>growth temperature</td>
<td>575 °C</td>
<td>525 °C</td>
<td>500 °C</td>
</tr>
<tr>
<td>P Peak level (cm$^{-3}$)</td>
<td>$4.8 \times 10^{18}$</td>
<td>$3.3 \times 10^{18}$</td>
<td>$4.6 \times 10^{18}$</td>
</tr>
<tr>
<td>SiGe buffer (nm) @ 575°C</td>
<td>160</td>
<td>125</td>
<td>170</td>
</tr>
</tbody>
</table>

Fig. 5.8 P and Ge profiles of sample #5457, #5850, and #5630. P turn-off slopes of these three samples are 40, 14, and 8 nm/dec, respectively. The setback layer thicknesses between the phosphorus peak level and the lower Si/SiGe heterojunction are 20, 20, and 33 nm for those three samples.
2DEG mobility is also affected greatly by phosphorus turn-off slope (Fig. 5.9). With a fast phosphorus turn-off slope of 8 nm/dec in sample #5630, the mobility is 60,000 cm$^2$/V-s, 20 times higher than that of sample #5457 with a slow slope of 40 nm/dec. For sample #5457, even though electron screening is expected to be stronger than sample #5630 because of the higher electron density ($3.8 \times 10^{12}$ cm$^{-2}$ >> $1.8 \times 10^{11}$ cm$^{-2}$), the mobility is actually much lower. This is attributed to the high level of phosphorus in the setback layer because of a slow phosphorus turn-off, introducing stronger remote impurity scattering for 2DEG. Furthermore, for the sample of a slow P turn-off such as #5457, the resulting high P level of $10^{18}$ cm$^{-3}$ in the Si QW layer could cause stronger background impurity scattering. Both remote and background impurity scattering affect the 2DEG mobility in the inverted modulation-doped Si 2DEGs with slow P turn-off. Thus, a fast P turn-off must be used for high-mobility 2DEGs.
5.3.2 Effects of Remote Impurity at the Si/Al₂O₃ Interface on Mobility

In this section, we study the effect of the impurity charges at the Si/Al₂O₃ interface. Unlike a top modulation-doped Si 2DEG with ionized impurities in the supply layer to screen the remote charges at the Si surface, in an inverted structure, the remote scattering effect of the Si surface charges must be considered because of the absence of an n-type doping layer between the surface and 2DEG. If the charge density at the surface is higher than the remote impurity charge in the bottom supply layer in an inverted 2DEG device, presumably, the mobility would be largely affected by those unscreened surface charges [89]. By reducing the surface charge density or the distance between the surface and 2DEG [73], the scattering is expected to be weaker.

Table 5.3 Layer structures of sample #5877 and #5630 for study of the effects of impurity charges at the Si/Al₂O₃ interface on 2DEG mobility.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Structure</th>
<th>Sample #</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>5877</td>
<td>Si cap (nm) @ 625°C</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>SiGe cap (nm) @ 575°C</td>
<td>26</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Si QW (nm) @ 625°C</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>P level in QW (cm⁻³)</td>
<td>below 5 × 10¹⁵</td>
<td>below 5 × 10¹⁵</td>
</tr>
<tr>
<td></td>
<td>P Ge setback (nm)</td>
<td>62</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>P turn-off slope</td>
<td>13 nm/dec; @ 525 °C</td>
<td>8 nm/dec; @ 500 °C</td>
</tr>
<tr>
<td></td>
<td>P Peak level (cm⁻³)</td>
<td>2.5 × 10¹⁸</td>
<td>4.6 × 10¹⁸</td>
</tr>
<tr>
<td></td>
<td>SiGe buffer (nm) @575°C</td>
<td>115</td>
<td>170</td>
</tr>
</tbody>
</table>

Hall bar devices with a Cr/Au/Al₂O₃ gate stack on top of inverted modulation-doped Si 2DEGs (Fig. 5.4) were fabricated to investigate the effect of the
surface charges on the 2DEG transport properties. Two samples of different thicknesses of top SiGe cap layer, which separates the surface charges and 2DEG were grown (Table 5.3). SIMS profiles of these two samples are plotted in Fig. 5.10. The P turn-off slopes for those two samples are different because of the different growth temperatures. The levels of phosphorus background are as low as the SIMS detection limit, of $4 \times 10^{15}$ cm$^{-3}$. Therefore, the background impurity scattering is not considered the dominant scattering mechanism. Mobility vs. density for these two devices at 4 K by gating is plotted in Fig. 5.11. The mobility of #5877 (26-nm SiGe cap) is much lower than that of #5630 (50-nm SiGe cap), which is attributed to the stronger scattering from the remote scattering of surface charges with a shorter distance for the 2DEG to the surface. Furthermore, the bottom SiGe setback layer of #5877 is two times larger than that of #5630. The remote impurity scattering at the bottom supply layer is considered much weaker than that in #5630. Thus, we believe the mobility is dominated by the remote impurity scattering from the Si/Al$_2$O$_3$ interface. In addition, the critical density for metal-insulator transition (MIT) to occur in #5877 is higher, which also implies the
Fig. 5.11 Hall mobility vs. density at 4 K for sample #5630 (50 nm from the 2DEG to the surface) and #5877 (26 nm from the 2DEG to the surface) for a comparison of the effect of upper SiGe cap layer thickness.

stronger remote scattering from the Si/Al₂O₃ interface.

For #5630, an extremely low density, as low as \(7.5 \times 10^{10}\) cm\(^{-2}\) with mobility of 33,000 cm\(^2\)/V-s, was achieved by gating. Alternatively, very high mobility of 420,000 cm\(^2\)/V-s at 4 K was achieved at density of \(5.6 \times 10^{11}\) cm\(^{-2}\). The levels of the lowest density and the highest mobility in this inverted modulation-doped 2DEG device are comparable to those of the top doped 2DEGs, an indication of the effectiveness of low-temperature epitaxy, which enables low levels of phosphorus in the setback layer and then greatly reduces the impurity scattering. The peak mobility at the high density of \(6 \times 10^{11}\) cm\(^{-2}\) for #5630 could be attributed to the interface roughness scattering [88] or parallel conduction [94, 96]. We will present our analysis in the next section.
5.3.3 Second Subband Occupancy

In a gated device, a maximum in mobility was observed at the density of mid-$10^{11}$ cm$^{-2}$ in GaAs 2DEGs because of the interface roughness scattering [88] or parallel conduction [94, 96]. For an enhancement-mode Si 2DEG device, the interface roughness scattering was suggested by Huang et al. to account for the presence of peak in the mobility [73]. While their results could be fit by a model with several adjustable parameters, it was suggested that in modulation-doped GaAs 2DEGs, the contribution of parallel conduction in other channels such as the doped layer or the second subband in the QW layer would dominate over the interface roughness scattering [63]. This was confirmed by the experimental results of Shubnikov-de Haas oscillations, showing the presence of the parallel conduction in GaAs 2DEGs [103, 104, 105]. Some of the gated devices in this work showed maximum mobility. Thus, we present the analysis of sample #5630 with a peak mobility at density of $6 \times 10^{11}$ cm$^{-2}$ (Fig. 5.11) and report the first experimental observation of second subband occupancy in a Si 2DEG.

First, Hall density ($n_{\text{Hall}}$) and mobility ($\mu_{\text{Hall}}$) vs. gate voltage at 0.3 K of sample #5630 is shown in Fig. 5.12. The device is conducting at $V_g > 1.8$ V with a threshold voltage of 1.45 V by extrapolating $n_{\text{Hall}}$ vs. $V_g$ to $n_{\text{Hall}} = 0$ (red dash line in Fig. 5.12). Between 1.45 V and 1.8 V, there was no conduction because of the metal-insulator transition. When the channel conducts, the Hall electron density increases with gate voltage. The extracted slope of the Hall density gives the effective capacitance of $6 \times 10^{-8}$ F/cm$^2$ between the Cr/Au gate and the 2DEG layer, which is close to the calculated value of $5.8 \times 10^{-8}$ F/cm$^2$, based on the parallel capacitor model. Note there is a kink in $n_{\text{Hall}}$ at $V_g = 3.4$ V, beyond which the predicted density (red dash line in Fig. 5.12 (a)) is
Fig. 5.12 (a) Hall density and (b) Hall mobility at 0.3 K vs. $V_g$ of sample #5630. The onsets of the intersubband scattering and second subband occupancy occurs at $V_g = 3.2$ V and 3.5 V, respectively. In region (i), only the first subband is populated. For region (ii), some electrons in the first subband scatter into and are trapped in the localized states, so the mobility drops. In region (iii), electrons reside in both the first and second subbands, so that the measured Hall mobility increases with electron density because of stronger screening.

actually larger than the measured Hall density (black solid line in Fig. 5.12 (a)). The kink of the Hall density was first reported in a GaAs 2DEG by Störmer et al. and
intersubband scattering between the first and the second subbands was suggested to account the kink of the measured Hall density.

A parallel conduction model including the conduction channels in the first and second subbands [79] is used here to help understand the intersubband scattering. With the onset of the occupancy in the second subband of the Si QW layer, the measured Hall density and mobility actually represent the mixed results of parallel conduction in the first and second subbands as follows:

\[
n_{\text{Hall}} = \frac{\left( n_1 \mu_1 + n_2 \mu_2 \right)^2}{n_1 \mu_1^2 + n_2 \mu_2^2}, \tag{5.1}
\]

\[
\mu_{\text{Hall}} = \frac{n_1 \mu_1^2 + n_2 \mu_2^2}{n_1 \mu_1 + n_2 \mu_2}. \tag{5.2}
\]

where \( n_1 \) and \( n_2 \) are the two-dimensional electron densities in the first and second subbands, and \( \mu_1 \) and \( \mu_2 \) represent the mobilities in the first and second subband, respectively.

For region (i) in Fig. 5.12, before the onset of the intersubband scattering (i.e. \( V_g = 3.2 \text{ V} \)), there exists only one subband. We assume \( n_2 = 0 \) in this region and \( n_{\text{Hall}} \) and \( \mu_{\text{Hall}} \) represent \( n_1 \) and \( \mu_1 \), respectively. In this region, as \( V_g \) increases, the electron density increases, leading to stronger electron screening on ionized impurity scattering and enhanced electron mobility. 2DEG mobility is plotted versus density in Fig. 5.13. The mobility scales with the Hall electron density as \( \mu \propto n^{1.4} \) for the density below \( 6 \times 10^{11} \text{ cm}^{-2} \). The exponent of 1.4 suggests that the remote impurity scattering is the dominant mechanism in this sample [79, 82]. Note that at the density below \( 2 \times 10^{11} \)
cm$^2$, the mobility drops sharply because of the MIT [106].

In region (ii) ($3.2 \, V < V_g < 3.5 \, V$), when $V_g$ increases, the mobility drops from the peak value of 470,000 cm$^2$/V-s to a minimum of 330,000 cm$^2$/V-s with a 30 % decrement (Fig. 5.12 (b)). On the other hand, the Hall density is unaffected in this region (Fig. 5.12(a)). This is surprising because the expected stronger electron screening with $V_g$ would enhance the mobility further. This discrepancy could be explained by the presence of intersubband scattering [103, 105]. When $V_g$ increases, the Fermi level ($E_F$) is raised and approaches the energy level of the second subband. Despite $E_F$ still below the level of the second subband, the levels of the localized states created by the potential fluctuations of the remote supply impurities could be below $E_F$ in practice. Thus, electrons in the first subband may scatter into and be trapped in those localized states. This additional scattering reduces the 2DEG mobility in the first subband.
To quantitatively explain this argument, we assume there are two conduction channels of electrons in region (ii): the first subband and the localized states of the second subband. Since electrons in the localized states are trapped, we assume $\mu_{\text{local}} = 0$, so $n_{\text{Hall}}$ and $\mu_{\text{Hall}}$ in Eqs. (5.1) and (5.2) represent $n_1$ and $\mu_1$, respectively. While $n_{\text{Hall}}$ (= $n_1$) increases with $V_g$, $\mu_{\text{Hall}}$ drops as $V_g$ increases (Fig. 5.12 (b)). Since $\mu_{\text{local}} = 0$, the measured Hall mobility represents the true 2DEG mobility in the first subband. With the presence of the localized states, electrons in the first subband can scatter into those localized states and be trapped, so $\mu_{\text{Hall}}$ (= $\mu_1$) drops.

In region (iii), as $E_F$ is lifted further by increasing $V_g$, the localized states are filled and mobile electrons start to populate the second subband. $n_{\text{Hall}}$ and $\mu_{\text{Hall}}$ represent the combined density and mobility of the first and second subbands following Eqs. (5.1) and (5.2). Since the FFT spectrums of $R_{xx}$ from this sample were too noisy, a quantitative measure of the population process in the first and second subbands is not possible. However, following Störmer’s arguments, we assume that when the Fermi level enters the second subband, $n_2$ increases with $V_g$ and $n_1$ is constant [103]. For $\mu_1$ and $\mu_2$, both increases with $V_g$ because of stronger screening by more populated electrons in the second subband. Since both $\mu_1$ and $\mu_2$ could be functions of $n_2$, it is not possible to solve them simultaneously without the information of the density in the second subband.

Here, we present a simplified analysis of two channel conduction in region (iii). Assuming above $V_g = 3.5$ V (Fig. 5.12 (b)), the first subband and the localized states of the second subband are fully populated. When $V_g$ increases, electrons are injected into the second subband and the relationship of $n_2$ and $V_g$ follows a simple parallel-plate capacitor model. Thus, above 3.5 V, $n_2$ increases from zero with an
Fig. 5.14 (a) Hall density and (b) Hall mobility at 0.3 K vs. $V_g$ of sample #5630. In region (i) and (ii), $n_{Hall}$ and $\mu_{Hall}$ represent $n_1$ and $\mu_1$, respectively. For region (iii), assuming $n_2$ increases from zero with $V_g$ ($V_g = 3.5$ V) by a simple parallel capacitor model, and $n_1$ is constant, $\mu_1$, and $\mu_2$ can be solved by Eqs. (5.1) and (5.2).

The effective capacitance of $6 \times 10^{-8}$ F/cm$^2$ divided by $e$ (electron charge) which is extracted from Fig. 5.12 (a). The results of $n_2$ and $n_1$ are plotted versus $V_g$ in Fig. 5.14 (a). Since now $n_1$ and $n_2$ are known, Eqs. (5.1) and (5.2) can be used to solve $\mu_1$ and $\mu_2$. 
The mobilities in the first and second subbands are plotted in Fig. 5.14 (b). Both increase with $V_g$ because of stronger screening by the increased electrons in the second subband, which follows the experimental results in [103, 105]. Note that $\mu_1$ and $\mu_{\text{Hall}}$ are very close near the onset of the second subband occupancy, suggesting that the conduction would be dominated by the electrons in the first subband. As $V_g$ increases, more electrons populated in the second subband, so the deviation between $\mu_1$ and $\mu_{\text{Hall}}$ becomes larger due to the stronger participation of the second subband.

Magneto-resistances ($R_{xx}$ and $R_{xy}$) vs. magnetic field at 0.3 K for sample #5630 at $V_g = 3$ V are shown in Fig. 5.15 (a). The onsets of Shubnikov-de Haas oscillations in $R_{xx}$ and flat quantum Hall plateaus in $R_{xy}$ occurred at 0.6 T and 0.9 T, respectively, showing the high quality of this inverted device. The spin splitting and valley splitting occurred at $B = 1.3$ T and 2.6 T. It is thought-provoking to note there is no $\nu = 3$ in $R_{xx}$, despite the presence of the Hall plateau in $R_{xy}$, which we do not understand yet. The vanishing longitudinal magneto-resistance ($R_{xx}$) indicates there is no parallel conduction from the bottom doped layer. Furthermore, a single period of $R_{xx}$ vs.
(1/B) suggests only one two-dimensional channel exists. As the gate voltage increases, the presence of the beating of the oscillations in $R_{xx}$ indicates the onset of the occupancy of the second subband (Fig. 5.15 (b)).

### 5.3.4 Comparison of 2DEG Mobility in Different Structures

In this thesis, we studied three different structures of Si 2DEGs: standard modulation-doped Si 2DEG (with a doping layer on top of the Si QW), an inverted modulation-doped Si 2DEG (with a doping layer on the bottom of the Si QW), and an enhancement-mode $^{28}$Si enriched 2DEG without any doping layer. The relationship of $\mu$ and $n$ at 4 K for those structures are plotted in Fig. 5.16 with the results of enhancement-mode 2DEG of natural Si also shown for comparison (courtesy of Chiao-Ti Huang). Clearly, at electron density below $3 \times 10^{11}$ cm$^{-2}$, the mobilities for all structures are almost identical with a slope of $\mu \propto n^{1.5}$. For those four structures, the distance from the 2DEG to the surface is 50 – 60 nm regardless of the presence of the doping layers except top-doped sample (#5737) with a distance of 100 nm from the surface to the 2DEG and a distance of 50 nm from the remote doped layer to the underlying 2DEG (Fig. 4.19(a)).

Based on the prior theoretical or experimental work [82, 91], if the exponent of $\mu \propto n^\alpha$ is close to 1.5, it is suggested the remote impurity scattering would be the dominant scattering mechanism. In Fig. 5.16, all curves of mobility vs. density show their exponents close to 1.5, which seems support our arguments on the limiting scattering mechanism. For quantum dot applications, a low electron density of $\sim 1 \times$
5.4 Summary

We first studied isotopically enriched $^{28}$Si 2DEGs, which can be used to increase the spin decoherence time for quantum computing applications. By the enrichment of $^{28}$Si, the only spin-carrying isotope of $^{29}$Si was reduced from 4.7 % to 800 \text{ppm}.
ppm with a predicted dephasing time of 2 μs. Furthermore, a high mobility of 522,000 cm²/V-s at 0.3 K in an enriched $^{28}$Si 2DEG was reported with clear Hall plateaus and Shubnikov-de Haas oscillations, showing its high material quality. By top gating, remote impurity charges at the Si/Al₂O₃ interface was suggested as the dominant scattering source, and an extremely low free electron density of $6 \times 10^{10}$ cm⁻² was also presented.

Then we present an inverted modulation-doped Si 2DEGs of high mobility ~470,000 cm²/V-s at 0.3 K by the suppression of phosphorus segregation in the lower SiGe setback layer. We showed that the sharp P turn-off (< 10 nm/decade) is a key to the success of high-quality inverted structures. An inverted modulation-doped Si 2DEG of high mobility by reducing phosphorus segregation opens the path towards the first demonstration of a bilayer device of two adjacent Si 2DEGs. Second subband occupancy and the intersubband scattering of electron between the first and second subbands were also demonstrated for the first time in a Si 2DEG.

It is interesting to note that μ vs. n, in a enhancement-mode structure Si 2DEG of isotopically enriched $^{28}$Si, in an enhancement-mode 2DEG of natural Si, in a standard modulation-doped (top-doped) Si 2DEG, and in an inverted modulation-doped (bottom-doped) Si 2DEG, all with the distance of 50 ~ 60 nm from the 2DEG layer to the remote impurity layer such as the supply layer or the Si surface layer, shows a strikingly similar form. It seems electron mobility was limited by the remote impurity scattering at the Si surface.
Chapter 6  Conclusions and Future Work

6.1  Conclusions

Strained Si and SiGe quantum devices were made from structures grown by CVD. First, we presented the work of band-to-band tunneling in strained SiGe, which is crucial for Si-based tunneling FETs for low-power applications. We investigated the effects of electric field and Ge fraction on band-to-band tunneling (BTBT) in p⁺-SiGe/n⁺-Si heterojunctions and p⁺-SiGe/n⁺-SiGe homojunctions at both forward and reverse biases. Negative differential resistance (NDR) was observed for each device, showing the high quality of tunneling devices by CVD. A peak current density of 8.2 kA/cm² in forward bias and reverse-biased BTBT current density of 10³ kA/cm² at -1 V were demonstrated, both of which we believed to be the highest among all Si-based tunneling diodes by CVD. Furthermore, we also showed the importance of the presence of NDR in forward bias to distinguish between direct BTBT and defect-assisted tunneling in reverse bias. Good agreement between experimental results and models for forward and reverse biases was achieved, which can serve as a basis for the related tunneling device modeling.

Next, we studied the surface segregation of phosphorus in relaxed SiGe epitaxial layers and investigated the effects of surface hydrogen. We experimentally observed an opposite trend of phosphorus segregation versus temperature compared to that expected from the prediction of a two-state model (TSM). We proposed a phenomenological model that accounts for the effect of surface hydrogen on the
segregation energy of phosphorus. Surface hydrogen changes the bonding structure of host atoms (Si or Ge) in the surface layer and reduces the segregation energy. An extremely sharp P turn-off slope of 6 nm/dec was also achieved, enabling effective Schottky gating on modulation-doped Si two-dimensional electron gas.

Two-dimensional electron gases (2DEGs) in a strained Si are promising for quantum dot applications. We investigated the transport properties in Si 2DEGs grown in Princeton and achieved a mobility of 522,000 cm$^2$/V-s at 0.3 K, among the best by CVD. This was attributed to the reduction of the phosphorus background level in the CVD system by a gas separation system of the doping gas of phosphorus and other process gases. The effects of different layers on ungated depletion-mode devices of modulation-doped Si 2DEG were investigated. The results showed that the remote impurity charges at the supply layer were the major source of electron scattering. The experimental results of top Schottky gated devices also supported this conclusion.

In Si of natural abundance, only the 4.7% atoms of $^{29}\text{Si}$ carry nuclear spins, leading to spin decoherence. $^{28}\text{Si}$ enrichment reduced the level of $^{29}\text{Si}$ to 800 ppm, which could extend the dephasing time in QDs to 2μs, 300 times longer than that in GaAs QDs. High mobility of 522,000 cm$^2$/V-s in an isotopically enriched $^{28}\text{Si}$ 2DEG was also reported. The remote impurity charges at the supply layer for the depletion-mode devices and at the Si/Al$_2$O$_3$ interface for the enhancement-mode devices were considered the main source for electron scattering.

Last, an inverted modulation-doped 2DEG with Si of natural abundance achieved a record high mobility of 470,000 cm$^2$/V-s at 0.3 K. High mobility in this type of structure was attributed to the lower phosphorus level by strong reduction of
phosphorus segregation, which was enabled by low-temperature epitaxy. Furthermore, we experimentally presented the occupancy of the second subband and intersubband scattering between the first and second subbands in a Si 2DEG for the first time. The fact that the initial drop and subsequent rise in the measured Hall mobility because of the intersubband scattering were observed supported our conclusions.

6.2 Future Work

In this section, some aspects of considerable interest about the topics in this thesis are given as follows:

(i) Band-to-band tunneling devices:

Based on the results in chapter 2, SiGe-based BTBT is promising for the realization of TFET of high performance. However, there is no report on any Si-based TFET with a high drive current as high as 100 μA/μm and a sharp sub-threshold slope of < 60 mV/decade at the same time, which is crucial for the replacement of conventional MOSFETs. Further work on the comparison of the TFET experiment and modeling based on our experimental results from diode structures will be required for further development of TFETs.

(ii) Phosphorus segregation:

In chapter 2, phosphorus turn-off slopes in strained and relaxed Si_{0.7}Ge_{0.3} films were different. A further analysis is required to explain this observation. Moreover, since the hydrogen coverage depends on the Ge fraction in SiGe alloys, the effect of hydrogen
on the bonding structure and segregation energy of phosphorus near the surface could also be different for various Ge fractions.

(iii) Si 2DEG:

While high mobility in Si 2DEGs was demonstrated by our CVD, the predicted upper limit of mobility is still much higher than our experimental results. Further work is required to determine the missing factor. In addition, the process of second subband occupancy is not completely understood, although some convincing evidence was presented. Last, the high-mobility inverted modulation-doped Si 2DEG enables the growth of a bilayer of parallel Si 2DEGs.
Bibliography

112


[23] http://www3.nd.edu/~gsnider/1D_Poi_PC_b8j.zip


[59] E. S. Tok, S. W. Ong, and H. Chuan Kang, “Hydrogen desorption kinetics from the Si\textsubscript{1-x}Ge\textsubscript{x}(100)–(2×1) surface,” Journal of Chemical Physics, vol. 120, pp. 5424 – 5431, 2004.


G. P. Watson, E. A. Fitzgerald, Y. -H. Xie, and D. Monroe, “Relaxed low threading defect density Si$_{0.7}$Ge$_{0.3}$ epitaxial layers grown on Si by rapid thermal chemical vapor deposition,” *Journal of Applied Physics*, vol. 75, pp. 263 – 269, 1994.


119


Appendix: Publications and Conference Presentations Resulting from this Thesis

Journal Articles


Conference Presentations


