STUDY OF THE PIEZORESISTIVE PROPERTIES OF Si, Ge AND GaAs MOSFETs USING
A NOVEL FLEXURE BASED WAFER BENDING SETUP

By

SAGAR SUTHRAM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2008
I would like to dedicate my dissertation to my mother and father. Their countless sacrifices to provide good education to me have made me what I am today.
ACKNOWLEDGMENTS

First of all I would like to express my sincere gratitude and appreciation to my advisor Dr Scott E Thompson. His constant encouragement and excellent mentorship have been very valuable in the course of my research. I would like to convey my special thanks to Dr Nishida for his guidance throughout the course of my research right from my first day in graduate school. I would also like to thank my other committee members, Dr Fossum, Dr Bosman and Dr Hershfield for their comments, suggestions and help from time to time. I would also like to thank Dr Rusty Harris and Dr Prashant Majhi, my managers at Sematech, from whom I have learnt a lot.

I would like thank all members of our research group, Kehuey Wu, Guangyu Sun, Nirav Shah, Ji-Song Lim, Younsung Choi, Yongke Sun, Andy Koehler, Uma Maheshwari Aghoram, Srivatsan Parthasarthy, Min Chu, Xiaodong Yang, Xiaoliang Lu, Toshi Numata, Hyunwoo Park, Nidhi Mohta and Tony Acosta. Constant encouragement, exciting discussions and pot-luck get-togethers with them have made doing research a lot of fun.

I dedicate this work to my parents. Their innumerable sacrifices to provide good education to me have made me what I am now. I dedicate all my achievements to them. I also want to thank my sister for her understanding and support. Most of all I would like to thank my girlfriend Krishna, for standing by me the last three years and being the pillar that I could lean on when I needed to.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................................................................................. 4

LIST OF TABLES ................................................................................................................................................ 7

LIST OF FIGURES .............................................................................................................................................. 8

ABSTRACT ......................................................................................................................................................... 12

CHAPTER

1 INTRODUCTION ............................................................................................................................................ 14

1.1 Overview .................................................................................................................................................. 14
1.2 Past Work and Motivation ....................................................................................................................... 15
1.3 Organization ............................................................................................................................................. 20

2 FLEXURE-BASED DESIGN AND HIGH STRESS .................................................................................... 22

2.1 Introduction to Flexures .......................................................................................................................... 22
2.2 Design Considerations for a Flexure-Based Setup .............................................................................. 23
  2.2.1 Basic Requirements for Four-Point Bending .................................................................................. 23
  2.2.2 Drawbacks of Previous Four-Point Bending Setups ................................................................... 24
  2.2.3 Flexure System Concept ................................................................................................................. 24
  2.2.4 Notch Design .................................................................................................................................. 26
  2.2.5 Material Considerations .................................................................................................................. 28
  2.2.6 Design Simulations on Ansys ....................................................................................................... 28
2.3 How is the Flexure-Based Setup Better? ............................................................................................... 30
2.4 Flexure Setup Operation ....................................................................................................................... 30
2.5 Sample Preparation ................................................................................................................................ 31
2.6 Stress Calibration ..................................................................................................................................... 32
  2.6.1 Load Cell Measurements ............................................................................................................... 32
  2.6.2 Strain Gauge Measurements ......................................................................................................... 33
  2.6.3 Optical Wafer Curvature Measurements ...................................................................................... 34
2.7 High Stress Measurements ................................................................................................................... 35
2.8 Wire Bonding to Measure Short Channel Devices .............................................................................. 37

3 SILICON CHANNEL MOSFET PIEZORESISTANCE .............................................................................. 39

3.1 Planar MOSFETs ....................................................................................................................................... 39
  3.1.1 Electron Mobility Enhancement ................................................................................................. 39
  3.1.2 Hole Mobility Enhancement ....................................................................................................... 46
  3.1.3 Effect of Stress on n and p – MOSFET I_g-sat ............................................................................... 51
3.2 Double-Gate FinFETs ............................................................................................................................... 59
  3.2.1 Mobility Enhancement .................................................................................................................... 59
3.2.2  FinFET Reliability ........................................................................................................65
  3.2.2.1 Bias instability ........................................................................................................65
  3.2.2.2 Hot carrier injection ...............................................................................................67
  3.2.3 Ring Oscillator Delay Performance .........................................................................69

4  GERMANIUM CHANNEL MOSFET PIEZORESISTANCE ........................................70
  4.1 Introduction ....................................................................................................................70
  4.2 Device Fabrication and Performance ............................................................................70
  4.3 Uniaxial Stress Enhancement .......................................................................................73

5  GaAs CHANNEL MOSFET PIEZORESISTANCE ............................................................79
  5.1 Introduction and Motivation .........................................................................................79
  5.2 Experimental Approach ..............................................................................................81
  5.3 Results and Discussion ...............................................................................................82
    5.3.1 Bulk GaAs ...............................................................................................................82
    5.3.2 GaAs nMOSFETs .................................................................................................83
    5.3.2 GaAs pMOSFETs .................................................................................................89

6  STRESS EFFECTS ON TANOS FLASH MEMORY DEVICES ........................................92
  6.1 Introduction and Motivation .......................................................................................92
  6.2 Experiment and Results .............................................................................................93

7  SUMMARY AND FUTURE WORK .................................................................................97
  7.1 Summary ......................................................................................................................97
  7.2 Future Work .................................................................................................................98

LIST OF REFERENCES ........................................................................................................100

BIOGRAPHICAL SKETCH .................................................................................................108
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>62</td>
</tr>
</tbody>
</table>

3-1  Longitudinal piezoresistance coefficients for n and p-FinFETs along with previous published data for planar (110) MOSFETs and bulk Si (110) is shown.................................62
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Schematic diagrams of different wafer bending techniques</td>
<td>18</td>
</tr>
<tr>
<td>1-2</td>
<td>Photograph of the uniaxial four-point wafer bending setup used by Wu et al.</td>
<td>19</td>
</tr>
<tr>
<td>2-1</td>
<td>Schematic diagrams of different flexure elements</td>
<td>23</td>
</tr>
<tr>
<td>2-2</td>
<td>A schematic diagram of a circular notch type flexure element operation concept is shown</td>
<td>25</td>
</tr>
<tr>
<td>2-3</td>
<td>A schematic diagram of eight identical flexure elements joined to form a system which provides uniaxial displacement for a four-point bending setup</td>
<td>25</td>
</tr>
<tr>
<td>2-4</td>
<td>Schematic of the analogy between a car jack and the flexure system is shown. Each arm of a car jack is equivalent to a couple of flexure beams</td>
<td>25</td>
</tr>
<tr>
<td>2-5</td>
<td>Schematic diagram of a flexure element is shown</td>
<td>26</td>
</tr>
<tr>
<td>2-6</td>
<td>Final design drawings of the flexure setup obtained after performing Ansys simulations</td>
<td>29</td>
</tr>
<tr>
<td>2-7</td>
<td>Photograph of the final fabricated flexure based wafer bending setup</td>
<td>29</td>
</tr>
<tr>
<td>2-8</td>
<td>A photograph of a Sip wafer edge formed by laser cutting</td>
<td>31</td>
</tr>
<tr>
<td>2-9</td>
<td>A photograph of a Si wafer edge formed by cleaving taken through a microscope</td>
<td>32</td>
</tr>
<tr>
<td>2-10</td>
<td>A load cell is placed under the top mount in order to measure the force being applied on the rods and hence calibrate the stress in the wafer</td>
<td>32</td>
</tr>
<tr>
<td>2-11</td>
<td>Stress estimation from the load cell measurements also show that the flexure setup applies uniaxial stress based on the equation even up to ~1GPa</td>
<td>33</td>
</tr>
<tr>
<td>2-12</td>
<td>A strain gauge is glued to a bare silicon sample before performing the stress calibration measurement</td>
<td>34</td>
</tr>
<tr>
<td>2-13</td>
<td>Strain gauge measurements reveal that the actual stress on the surface of the wafer is within 5% of the calculated stress from the equation</td>
<td>34</td>
</tr>
<tr>
<td>2-14</td>
<td>A schematic diagram of the optical wafer curvature measurement technique is shown</td>
<td>35</td>
</tr>
<tr>
<td>2-15</td>
<td>A photograph of the complete flexure based wafer bending setup</td>
<td>36</td>
</tr>
<tr>
<td>2-16</td>
<td>A photograph showing a silicon wafer bent under high stress</td>
<td>36</td>
</tr>
</tbody>
</table>
2-17  Drive current enhancement data after wire bonding to a device.  ..................................................38
3-1   NMOSFET drive current enhancement and channel piezoresistance plotted as a function of wafer bending stress. .................................................................................................................41
3-2   Extracted effective mobility enhancement ($\mu_{\text{eff}}$) vs. stress plotted for long and short devices with $R_{s/d}$ correction versus stress at $E_{\text{eff}} \approx 0.7\text{MV/cm}$. ..................................................42
3-3   Schematic diagram of the physics of electron mobility enhancement in Si is shown.  .................43
3-4   Plot of electron mobility extracted from drive current data as a function of vertical field for different uniaxial tensile stresses. ........................................................................................................44
3-5   Dependence of $R_{\text{tot}}$ on $L_{\text{mask}}$ for different wafer bending stresses. .......................................48
3-6   Hole mobility enhancement as a function of stress along with other published work and calculation results using the k-p method is shown. ..........................................................49
3-7   Sub-band splitting and optical phonon scattering rates plotted as a function of stress......50
3-8   Hole mobility of short devices with a built-in stress of $\sim 1\text{ GPa}$ as a function of gate over-drive for different mechanical stresses. ..........................................................51
3-9   Saturation drive current enhancement as a function of stress for a 60nm nMOSFET device compared with long channel electron mobility enhancement trend..........................57
3-10  Prediction of percentage $I_{\text{d-sat}}$ change with percentage mobility change using both models as compared with experimental results is plotted.................................................59
3-11  Drive current enhancement data for wide and narrow FinFETs with uniaxial tensile stress ...............................................................................................................................................61
3-12  Comparison of CESL enhancement for long and short devices is shown. .................................63
3-13  CESL stress transfer mechanism. ..................................................................................................64
3-14  The layout of a gated diode structure used for performing the charge pumping measurements is shown...................................................................................................................66
3-15  Application of both uniaxial tension and compression shows negligible change in the charge pumping current. ..........................................................66
3-16  Increased FinFET hot carrier degradation is seen for both tensile and compressive uniaxial stresses. ..................................................................................................................67
3-17  Gm degradation does not show any change with uniaxial tensile stresses suggesting that increased HCI degradation is due to bulk traps and not interface state generation. ...68
3-18 Both tensile and compressive stresses cause bandgap narrowing at the drain-body junction resulting in increased impact-ionization.

3-19 FinFET ring-oscillator delay/stage measured as a function of bias voltage showing reduced delay with increasing supply voltage as expected.

3-20 Opposing effects on n and p-FinFETs due to strain result in marginal delay reduction.

4-1 Schematic of gate-first process flow for QW pMOSFET used in this work.

4-2 Biaxially strained SiGe heterostructure pMOSFET is shown.

4-3 Extracted hole mobility of the SiGe QW as compared with control Si.

4-4 Si, Ge and SiGe all show similar enhancements for stress less than 3 GPa.

4-5 Topmost valence band structure change with 1.0 GPa of uniaxial compressive stress in Si and Ge.

4-6 Enhancement from mass change is larger for Ge than Si.

4-7 2D-density of states change for the top valence band of Ge and Si with uniaxial stress is plotted.

4-8 Linear drive current enhancement for different Ge concentrations as a function of uniaxial compressive stress.

4-9 Linear drive current enhancement for different epi-Ge layer thicknesses as a function of uniaxial compression stress.

4-10 Dependence of drive current enhancement on epi-Ge thickness for a given stress is shown.

4-11 Drive current enhancement with negligible change in off-state current with stress. Inset shows 30% drive current enhancement with stress plotted in linear scale.

4-12 Negligible change in threshold voltage for both SiGe and Ge devices with uniaxial compressive stress is observed.

5-1 Stress calibration for GaAs piezoresistance measurements.

5-2 Percentage resistance change plotted as a function of applied wafer bending stress on a bulk resistor device.

5-3 Temperature dependence of mobility for Si and GaAs nMOSFETs is shown.

5-4 Drive current enhancement of GaAs nMOSFET as a function of applied uniaxial stress.
5-5 Theoretical estimation of the effect of uniaxial stress on electrons in GaAs is shown. ....86
5-6 Gate leakage current change with uniaxial stress is shown. ..............................................87
5-7 GaAs nMOSFET channel mobility and mobility enhancement as a function of inversion carrier density..............................................................................................................87
5-8 Drive current enhancement of GaAs nMOSFET as a function of applied biaxial stress..................................................................................................................................88
5-9 Theoretical estimation of the effect of biaxial stress on electrons in GaAs is shown. ......89
5-10 2D-Constant Energy Contours of the topmost valence band in GaAs...............................90
5-11 GaAs hole mobility calculation using k·p method as a function of uniaxial stress. ........91
5-12 Drive current enhancement of (100) GaAs p-MOSFETs with uniaxial stress compared with (100) Si..............................................................................................................91
6-1 Schematic shows the Si-SiO2 barrier and the dominant parameters controlling the gate tunneling current. .......................................................................................................93
6-2 Id-Vg plot showing the change in the threshold voltage after application of a 1 sec program pulse..................................................................................................................................94
6-3 Change in threshold voltage with different program pulse widths and different applied mechanical stresses is shown. .....................................................................................95
6-4 Id-Vg plot showing the change of threshold voltage with application of a 1 sec erase pulse. ..................................................................................................................................95
6-5 Change of threshold voltage with different erase pulse widths and different applied mechanical stresses is shown. .....................................................................................96
6-6 Change in threshold voltage with time under different applied mechanical stresses is shown. No change in retention time observed. ........................................................................96
For more than three decades, transistor density in microprocessor chips has doubled every two years keeping with Moore’s law by following the traditional scaling methods. But at the sub-100nm regime, this traditional scaling gave diminishing results and performance boosters like strain engineering were incorporated. Beyond the 90nm technology node, larger and larger amounts of channel strain are being incorporated. Even though the bulk and low stress piezoresistive behavior of Si MOSFETs (metal-oxide-semiconductor field-effect transistors) are well reported, there are no reports of the piezoresistive behavior of Si MOSFETs at high stress levels which are of more technological importance. Hence in this work, we systematically study the drawbacks of wafer bending setups used in previous works and design a novel flexure based wafer bending setup capable of applying stresses greater than 1 GPa on Si wafers. This bending setup is then used to study the high stress piezoresistive properties of Si n and p-MOSFETs. Physical insight into the trends observed is given, and the high stress behavior is compared with that of low stress and bulk Si. A discussion and analysis on the effect of stress on saturation drive current is also given which is a more important parameter for logic devices.

Going forward, there is much interest in the semiconductor industry in novel device architectures and novel channel materials for the sub-22nm technology nodes. Since strain
engineering has been successful for planar MOSFETs, there is an obvious interest in knowing how these novel systems respond to strain and gaining physical insight into the mechanisms for enhancement which can guide transistor design for future nodes. The flexure setup is used to first study double-gate (DG) FinFET devices and compare their piezoresistive behavior against planar (110) MOSFETs and bulk Si (110). Novel channel materials such as Ge and GaAs are also investigated. Their piezoresistive properties are compared with simulation predictions of the same and deviations are explained. Temperature dependent measurements are also done to determine dominant scattering mechanisms in these novel systems. For n-GaAs, mass change is explained to be the main mechanism for enhancement in the bulk case, while repopulation between different valleys is the reason for the enhancement in the MOSFET case. With growing interest in applying strain in memory devices and not much reported on the same, in this work we also study stress effects on TANOS flash memory transistor cells and report the effect of stress on TANOS program, erase and retention times.
CHAPTER 1
INTRODUCTION

1.1 Overview

The piezoresistive effect is described as the change in resistance of a material due to an applied mechanical stress. This effect for semiconductors such as Si and Ge was first discovered and studied by C S Smith in 1954 [1]. His experiments involved applying uniaxial stresses of under 10MPa on bulk Si and Ge samples and extracting the piezoresistance coefficients. The coefficient, \( \pi \), used to characterize the piezoresistance is defined as [1],

\[
\pi = \frac{1}{\sigma} \left( \frac{R_\sigma - R_0}{R_0} \right) = \frac{1}{\sigma} \frac{\Delta R}{R_0},
\]  

(1-1)

where \( R_0 = V_0/I \) and \( R_\sigma = V_\sigma/I \) are unstressed and stressed resistances respectively, \( V_0 \) and \( V_\sigma \) are unstressed and stressed voltages respectively, \( I \) is the electric current, and \( \sigma \) is the uniaxial stress. This discovery along with seminal work by Bardeen and Shockley on electron-phonon coupling based on their deformation potential theory [2] is considered the starting point for our current understanding of strain effects on semiconductors. With the advent of CMOSFET technology, the piezoresistance effect in Si was found to enhance MOSFET performance while researching growth of epi-Si layers on SiGe buffers in the 1980s [3, 4]. But all of this work was on biaxial stress induced due to the lattice mismatch between Si and SiGe. It is only recently, almost three decades after the advent of the MOSFET technology when channel lengths have been scaled to the sub-100nm regime and further geometric scaling gives diminishing returns, that there has been renewed interest in the Si piezoresistive behavior under uniaxial stress to boost the performance of MOSFETs. Successful integration of uniaxial process induced stress into the mainstream MOSFET process flow was demonstrated by Thompson et al. [5] and it has now become an integral part of most 90, 65 and 45nm technologies. The discovery of the
piezoresistive effect in semiconductors is also considered a starting point in the field of semiconductor sensors and actuators. With advanced semiconductor processing capabilities available now, this field has spawned into a vast area of research of its own. In this work, we focus only on the MOSFET applications.

### 1.2 Past Work and Motivation

Uniaxial process induced strained Si has become the preferred method to enhance transistor performance for sub-100 nm technologies [6-8]. The magnitude of channel stress in second generation strained Si technologies (65nm technology node) [9, 10] is approaching 1GPa and even higher levels of channel stress will likely be used for the 45 nm and beyond technology nodes. Modeling of the highly strained Si devices has largely used piezoresistance coefficients, due to the difficulty in predictively modeling the stress enhanced electron and hole mobility [11, 12]. However, nearly all the experimentally determined coefficients are used outside the stress range in which they were extracted and non-linearity in the piezoresistance should be expected [13]. For example, to model the mobility enhancement in 90 and 65nm technologies [5, 10], Smith $\pi$ coefficients [1] are used which were extracted at stresses up to only 10MPa. Recently, there have been studies [14-20] measuring MOSFET piezoresistance coefficient for slightly larger stresses in the range of a few 100 MPa with one exception being Wang et al. [19] with nMOSFET data up to $\sim$580MPa for long channel and 300MPa for short channel MOSFETs, all of this work still being below the high stress levels already in production [5, 10]. So there is a need for measuring and studying the piezoresistance behavior of Si MOSFETs in the high stress regime. Also going forward into the sub-22nm regime there is much interest in novel channel materials for MOSFETs such as Ge and III-Vs, mainly because of their ultra high mobility and narrow band-gap. Since strain has been successful for Si, it is interesting to see if it can be used in conjunction
with these novel materials also. Hence in this work, we also study the piezoresistance behavior of Ge and GaAs MOSFETs. In addition to logic devices, there is rising interest in strain induced performance enhancement of memory devices. In order to answer that question, we also study the effect of stress on TANOS flash memory transistors.

Piezoresistance measurements are done by applying external mechanical uniaxial stress on wafers and measuring the change in resistance of a device due to the application of stress. There are basically four methods used to achieve this in the literature, namely:

**Hanging weights to a slab of semiconductor:** This method was used by Smith [1] in the first piezoresistance measurement, and the setup is as shown in Figure 1-1A. As can be seen, large slabs of semiconductor are required to perform this experiment and hence this method is useful only for bulk measurements. The highest stress reported using this method in Smith’s work is just 10MPa.

**Cantilever bending:** In this method a rectangular strip of semiconductor of length $l$, thickness $t$ and width $w$ is either cleaved or diced and fixed to a rigid support on one end and weights are hung on the free end to apply the force $F$ as shown in the schematic in Figure 1-1B. The device to be measured is at a fixed distance $x$ from the rigid end and hence the stress experienced by the device can be calculated as follows [17]:

$$\sigma(x) = \frac{6Fl}{wt^3} \left[ \left( 1 - \frac{u^2}{15} \right) \frac{x}{l} \right]$$  \hspace{1cm} (1-2)

where

$$u = \frac{12Fl^2}{Ywt^3}$$  \hspace{1cm} (1-3)
and $Y$ is the young’s modulus of the semiconductor wafer being used. The maximum stress applied using this method as reported in Ref [17] is ~150MPa. It is evident that this method is not useful to perform measurements at high stress.

**Three-point bending:** In this method, a strip of semiconductor of a given thickness is fixed at both ends, and stress is applied by bending the wafer piece by applying a force at the center of the sample as shown in the schematic in Figure 1-1C. Uniaxial stress is induced along the length of the wafer piece because of the bending induced by the center point. The magnitude of stress is estimated by measuring the force being applied and then using the Equation $^{(21)}$,

$$\sigma = \frac{3aF}{wt^2} \quad \text{ (1-4)}$$

where $a$ is as indicated in the schematic, $F$ is the force applied, $w$ is the width of the sample, and $t$ is the thickness of the sample. This method is not accurate because the equation gives the correct value of the stress only at the center of the wafer. Any slight deviation from the center will lead to an inaccurate estimation of the stress. The highest stress achieved by this method is slightly higher than that achieved by the cantilever bending method and is ~300MPa $^{(21)}$.

**Four-point bending:** The schematic for a four-point wafer bending setup is shown in Figure 1-1D. As can be seen from the schematic, the two top points are held rigidly fixed while stress is applied by gradually moving the bottom points upward. The stress on the top surface of the wafer is estimated by using the Equation $^{(22)}$,

$$\sigma = \frac{Ety_{x=a}}{2a\left(\frac{L}{2} - \frac{2a}{3}\right)} \quad \text{ (1-5)}$$

where $E$ is the young’s modulus, $t$ is the thickness, $y$ is the vertical displacement at the point of contact of the inner rods, $a$ and $L$ are as shown in the schematic Figure 1-1D. Another alternate Equation for the same is $^{(15)}$,
where $y$ is the vertical displacement of the center of the wafer piece being used.

\[
\sigma = \frac{12Ey_0}{(3L^2 - 4a^2)}
\]  

(1-6)

Figure 1-1. Schematic diagrams of different wafer bending techniques, A) weights hanging to a slab of semiconductor, B) cantilever bending, C) three-point bending and D) four-point bending.

This method is more favorable for measuring piezoresistance of MOSFETs because in four-point bending, the bending moment induced on the top surface of the wafer is uniform.
between the center rods and hence the stress is uniform between the center rods. So this method does not suffer the same problem as in the three-point bending case of inaccurate estimation of stress. Any slight variations in the position of the device which is usual during measurements will not affect the accuracy of the results Figure 1-2 shows an image of one such four-point wafer bending setup taken from Wu et al. [23]. The maximum stress achieved using this setup is ~500MPa, much higher than that achieved in works before this, but there is a need to be able to bend wafers to even higher stresses (> 1GPa).

One of the key contributions of this work is to develop a novel flexure based four point wafer bending setup capable of applying greater than 1 GPa uniaxial stresses on semiconductor
wafers. This novel bending setup is used to study the Si n and p-MOSFET piezoresistance trends over a 0-2 GPa stress range which has never been reported before and physical insight into the mechanisms for mobility and drive current enhancements at these high stress levels is given. A similar study for narrow Double Gate (DG) – FinFET devices is performed and differences from planar MOSFETs and bulk Si are explained. Experimental investigation of the enhancement trends expected for Ge and GaAs channel MOSFETs is also done, which is of great importance to industry in determining the choice of channel materials for future technology nodes. The key mechanisms for enhancement in MOSFETs are observed and compared to the bulk case. Temperature dependent measurements are also performed to determine the dominant scattering mechanisms in the devices. Stress effects on TANOS flash devices is investigated and key mechanisms for the observed results are explained, which is again a key contribution for guiding future process technology development for memory devices, given the interest in incorporating strain to improve their performance.

1.3 Organization

In this work, we study the piezoresistive properties of Si, Ge, and GaAs MOSFETs by designing and developing a flexure based high stress wafer bending setup. The organization of the chapters is as follows:

Chapter 2 introduces flexures and discusses the design issues of a flexure based setup. It then covers the drawbacks of bending setups used in past works and shows how using this flexure setup and some novel sample preparation methods, we can achieve very high wafer bending stresses.

Chapter 3 shows the results obtained for piezoresistance measurements on Si n-and p-MOSFETs and Si DG-FinFETs using the flexure based wafer bending setup. It also includes a discussion on effect of stress on I_d-sat and velocity saturation and overshoot. A qualitative model
based on physics is used to explain the results. This chapter also covers some additional measurements done to study FinFET reliability and ring oscillator delay.

Chapter 4 discusses results obtained from piezoresistance measurements on Si/Si$_{1-x}$Ge$_x$/Si heterostructure MOSFETs. A physics based understanding of the results is developed and explained. Chapter 5 covers the study of the piezoresistance behavior of GaAs MOSFETs. Both k·p calculation results and multi-valley conduction band models are used to get a physical insight into the physics from theory and explain the differences between the bulk and the MOSFET case. Temperature dependence of mobility is also measured to determine the dominant scattering mechanism in the MOSFETs.

The final topic on the effect of stress on TANOS Flash cell program, erase and retention times is covered in chapter 6. Chapter 7 presents a summary and some suggestions for future work.
CHAPTER 2  
FLEXURE-BASED DESIGN AND HIGH STRESS

2.1 Introduction to Flexures

A flexure is usually considered to be a mechanism consisting of a series of rigid bodies connected by compliant elements that is designed to produce a geometrically well-defined motion upon application of force \cite{note24}. Flexure based systems offer many advantages such as, they are simple and inexpensive to manufacture, virtually wear-free unless caused by fatigue cracks, provide smooth and continuous displacements and complete mechanisms can be produced from a single monolith. There exists a linear relationship between applied force and displacement for small distortions, but accurate prediction of force-displacement characteristics requires accurate knowledge of the elastic modulus and geometry. Even tight manufacturing tolerances can produce relatively large uncertainty between predicted and actual performance. Flexures are restricted in the length of translation for a given size and stiffness and also cannot tolerate large loads. This poses a design challenge while choosing the material for our application because the maximum applicable stress would depend on the stiffness of the material for given size limitations of the setup.

Most flexure systems can be divided into two broad categories, notch and leaf type hinges. A notch hinge can be made by making two holes with a small separation between them to form a circular notch, or web. It is this thin web which serves as the flexure element as shown in the schematic Figure 2-1A. Leaf hinges typically consist of a slender member connected at each end by two rigid bodies to provide a compliant coupling as shown in the schematic in Figure 2-1B. The notch is one of the popular flexure elements used in designing systems, and it will be used to design the flexure-based wafer bending setup in this work too.
2.2 Design Considerations for a Flexure-Based Setup

To set design considerations for a flexure-based wafer bending setup, first we need to know two things: (1) The basic requirements for four-point bending and (2) Drawbacks of previous four-point wafer bending setups.

2.2.1 Basic Requirements for Four-Point Bending

- Two free points of contact should have uniaxial upward displacement to induce uniaxial stress on the wafer surface and remain parallel to the other two controls.

Figure 2-1. Schematic diagrams of different flexure elements, A) a circular notch hinge and B) a leaf hinge.
• Two points (either top or bottom) of contact should remain rigidly fixed.

• Rods used to provide the four points of contact in the setup should have a point-contact with the wafer, remain fixed in place and should not undergo any form of elastic or inelastic deformation.

• The maximum vertical displacement of the free points should not be much larger than the thickness of the sample in order to remain within the linear regime of operation of four-point bending.

2.2.2 Drawbacks of Previous Four-Point Bending Setups

Out of the four basic criteria mentioned above, the last three are met by most four-bending setups, but the critical one is the first one. Most four-point wafer bending setups are like the one shown in Figure 1-2. They have fours rods, two at the bottom and two on the top of the wafer piece fixed in place by the v-groves on the steel plates. The bottom rods remain rigidly fixed, while the top rods can be moved manually with the help of two screw gauges. This manual control with two screw gauges is what causes error in providing uniaxial vertical displacement for the top rods. Even a slight error will cause the wafer to crack at relatively low stresses of \(~400-500\text{MPa}\) (highest that has been achieved using this type of four-point bending setup). Hence the main design consideration for the flexure based setup is to eliminate manual error using two screw gauges and provide uniaxial vertical displacement.

2.2.3 Flexure System Concept

The notch type flexure element can be represented schematically as shown in Figure 2-2. The notch on one end of the beam suppresses two of the rotational degrees of freedom for the beam. So a design based on such a structure can be used to provide a fixed angular displacement and in turn a fixed vertical displacement required in a four-point bending setup. A system of eight such symmetric notch type flexure elements also referred to as a “double compound linear spring” is known to provide displacement with no deviation from rectilinear motion.\[^{[23]}\] A schematic of such a system is shown in Figure 2-3.
Figure 2-2. A schematic diagram of a circular notch type flexure element operation concept is shown. The notch suppresses two rotational degrees of freedom (DOF) for the beam and makes it into a system with only one DOF.

Figure 2-3. A schematic diagram of eight identical flexure elements joined to form a system which provides uniaxial displacement for a four-point bending setup.

Figure 2-4. Schematic of the analogy between a car jack and the flexure system is shown. Each arm of a car jack is equivalent to a couple of flexure beams.

To understand how this system works, we can draw an analogy to a car jack. The eight flexure elements can be thought of as four-pairs, each pair being equivalent to each arm of the car jack.
as shown in Figure 2-4. When the bottom screw is turned, each flexure element rotates by exactly the same angle (since they are identical) and similar to the car jack moves the top mount (C) vertically upwards. Since the side frames (A and B in Figure 2-3) of the system always remain parallel, the top mount (C) also remains parallel to the base (D) of the setup, and we achieve uniaxial vertical displacement.

2.2.4 Notch Design

From the above conceptual understanding, we see that each of the eight flexure elements is identical, and the design of the flexure element is the most crucial part of the whole setup. So the three design elements to be determined are the thickness \( t \) of the notch, the radius \( R \) of the notch, and the length \( L \) of the beam (Figure 2-5).

![Diagram of a flexure element](image)

Figure 2-5. Schematic diagram of a flexure element is shown. The three design parameters for a notch type flexure element are (1) the radius of the circular notch (2) the thickness of the thinnest portion of the beam and (3) the length of the beam. The length of the beam is limited by the size constraints of the setup, which in turn affects the thickness \( t \) in order to achieve the maximum possible vertical displacement.

From simple bending theory, the angular stiffness (K) of such a circular notch hinge is given by \(^{[24]}\),

\[
K = \frac{M}{\theta},
\]  

(2-1)
where M is the bending moment and θ is the angular displacement. For semicircular notches this expression is given by \(^{[24]}\):

$$\frac{1}{K} = \frac{3}{2EwR^2} \left[ \frac{1}{2\beta + \beta^2} \right] \left[ \frac{3 + 4\beta + 2\beta^2}{(1 + \beta)(2\beta + \beta^2)} \right] + \frac{6(1 + \beta)}{(2\beta + \beta^2)^{3/2}} \tan^{-1}\left( \frac{2 + \beta}{\beta} \right) \tag{2-2}$$

$$\equiv \frac{1}{K} = \frac{3}{2EwR^2} f(\beta) \tag{2-3}$$

where E is the Young’s modulus of the material, w is the width of the beam, and \(\beta = t/2R\) is a dimensionless quantity representing the hinge geometry. For small values of \(\beta\), Eq. 2-7 can be simplified to produce \(^{[24]}\):

$$K = \frac{2Ewt^{5/2}}{9\pi R^{3/2}} \tag{2-4}$$

The stress \(\sigma\) experienced by the notch is given by \(^{[24]}\)

$$\sigma = K_r \frac{6M}{t^2w} \tag{2-5}$$

where \(K_r\) is the stress concentration factor. For a wide range of \(\beta\) values, the stress concentration factor, to within better than 2%, is given by \(^{[24]}\)

$$K_r = (1 + \beta)^{0.20} \quad 0 < \beta < 2.3 \tag{2-6}$$

In the current application, we require a fixed angular displacement from the notches and hence the applied bending moment M is an unknown quantity. So the stress in the notch for a given angular displacement \(\theta\) is given by \(^{[24]}\)

$$\sigma = \frac{E(1 + \beta)^{0.20}}{\beta^2 f(\beta)} \theta. \tag{2-7}$$

At the onset of yielding, the maximum thickness of the notch for a given deflection of the hinge can be determined from
\[ t^2 = \frac{4ER^2}{f(\beta)} (1 + \beta)^{9/20} \theta \sigma_y \]  \hspace{1cm} (2-8)

or, from \[^{24}\]

\[ t = \left( \frac{\sigma_y}{\theta} \right)^2 \frac{9\pi^2 R}{16K^2 E^2} \]  \hspace{1cm} (2-9)

The fixed angular displacement can be calculated based on the size limitations which determine the length of the beam \( L \).

### 2.2.5 Material Considerations

From Eq. 2-9, it is evident that if we want large angular displacement along with a manufacturable notch thickness, the material should have large yield (fracture) strength. Generally materials that have large yield strengths are difficult to machine and pose a fabrication challenge to make a complicated flexure system with thin notches as in this case. There exists a trade-off between the two. The material chosen for fabricating the flexure system is 7075 aerospace grade Al, which has a yield strength of 500MPa and is also very easy to machine.

### 2.2.6 Design Simulations on Ansys

After determining the material, simulations were performed on a finite element software called “Ansys” to estimate the exact amount of stress induced on the notch for different angular displacements. Based on these simulations, the notch geometry was fixed, in order to achieve +/-5mm vertical displacement which in turn corresponds to \textasciitilde +/-2GPa range of stresses. The final design dimensions of the setup after the simulations are shown in Figure 2-6 and a photograph of the fabricated setup is shown in Figure 2-7. A single screw is also designed, using which we can provide vertical displacement to the four-point bending setup. This screw has a tapping of 20 threads per inch. The underside of the setup has 24 markings along the circumference and hence splitting one full rotation of the screw into 24 graduations. So rotation of the screw by one
graduation provides a vertical displacement of ~50μm or 2mils. Steel rods of 4mm diameter are used to provide the point contact. A separate mount with v-groves is designed to provide a platform to load the bottom rods and the wafer piece. The top frame of the setup also has v-groves to fix the position of the top rods.

Figure 2-6. Final design drawings of the flexure setup obtained after performing Ansys simulations.

Figure 2-7. Photograph of the final fabricated flexure based wafer bending setup. Since the entire flexure system is equivalent to a spring, holding bars and pins are used to hold the system fixed when not in use and prevent damage to the flexure elements due to vibrations.
2.3 How is the Flexure-Based Setup Better?

The flexure-based wafer bending setup is better than previous setups because it provides uniaxial vertical displacement, involves the use of only one screw and hence eliminates human error involved with rotating two screws, is very easy to use, capable of achieving high stresses in a short period of time and is easy to maintain as it is fabricated from a monolith and does not have multiple independent moving parts.

2.4 Flexure Setup Operation

The following are the steps to apply stress on a wafer piece using the flexure based wafer bending setup:

- Gently remove the holding pins and bars to release the flexure arms. Be careful not to shake the flexure arms.
- Lower the screw so that there is enough room to load the wafer and the rods.
- First place the bottom rods on the mount equidistant from the centre.
- Place the wafer on the bottom rods.
- Slowly place the top rods one by one while holding the wafer in place. Lower the screw as and when required to have adequate room to push the rods in place.
- Center the wafer on the mount so that the device under test is at the centre of the opening on the top plate and can be reached easily while probing.
- Gently raise the screw till you feel some resistance. Push the rods to see if they are taut. If they move raise the screw some more, till all the rods are rigidly in place and the marking on the screw is aligned with any marking on the underside of the setup. This ensures that we have the correct zero stress point.
- Now the wafer is ready for the bending measurement.
- Use a wrench to rotate the screw for applying stress. Do so slowly. Sudden motion of the screw can cause the wafer to break easily and also might damage the setup.
- Once measurement is done, slowly lower the screw so that the rods and the wafer can be pulled out.
- Now put the holding pins and bars in place, to avoid any damage to the flexure arms.
2.5 Sample Preparation

Preliminary wafer bending experiments using the flexure setup on bare wafer pieces revealed that the sample preparation also plays a vital component in the maximum stress achieved via bending. Generally samples which have poor edge quality tend to crack very easily, because they have many defects along the edges which can propagate at relatively low stress levels. There are many ways of preparing a sample wafer piece from the whole wafer, such as dicing, laser cutting, cleaving etc. Of these, dicing yields the poorest results with the most number of edge-defects and hence is not suitable for high stress bending experiments. Laser cutting seemed a good choice, but a closer look at the edge quality under a microscope showed that it too had a large number of edge defects (Figure 2-8). Cleaving the wafer into smaller pieces using a diamond tip scribe yields the best quality edges of all the methods as revealed in the photograph taken through a microscope in Figure 2-9. Hence in all the experiments in this work, the wafers are cleaved using a micropositioner controlled diamond tip scribe to get optical quality edges.

Figure 2-8. A photograph of a Sip wafer edge formed by laser cutting. From the photograph it is evident that there are many defects, and the edge-quality is poor.
Figure 2-9. A photograph of a Si wafer edge formed by cleaving taken through a microscope. We can see no defects visibly even through a microscope, and the edge quality is very good.

2.6 Stress Calibration

Before we start carrying out the high stress measurements, first we need to calibrate the setup to check if it indeed does apply the stress that it is designed to do. The stress calibration was performed using three techniques:

2.6.1 Load Cell Measurements

The first method used was by utilizing a load cell under the top mount to directly measure the force experienced by the bottom rods. Figure 2-10 shows the setup with the load cell in place to do the calibration measurements.

Figure 2-10. A load cell is placed under the top mount in order to measure the force being applied on the rods and hence calibrate the stress in the wafer.
The stress measured from this calibration measurement is plotted in Figure 2-11 along with that calculated from Eq. 1-2. It can be seen that the stress actually applied by the setup is very accurate even at stresses as high as 1GPa.

![Graph showing stress measurements](image)

Figure 2-11. Stress estimation from the load cell measurements also show that the flexure setup applies uniaxial stress based on the equation even up to ~1GPa.

### 2.6.2 Strain Gauge Measurements

The second method used for stress calibration is to epoxy a strain gauge on the top surface of the wafer and directly measure the strain experienced by the top surface of the wafer. This method gives the most accurate estimate of the stress on the surface of the wafer, but care has to be taken to prepare the sample properly before the measurement. Figure 2-12 shows a photograph of a wafer sample with a strain gauge epoxied on the surface just before the calibration measurement was started. If the sample is not prepared properly (epoxy is not cured properly or there is too much epoxy between the strain gauge and the wafer piece), the calibration measurement will give erroneous results. The result of the calibration measurement using the strain gauge is shown in Figure 2-13. Again we see that the actual stress experienced by the sample is within 5% of what is expected from the equation.
Figure 2-12. A strain gauge is glued to a bare silicon sample before performing the stress calibration measurement.

Figure 2-13. Strain gauge measurements reveal that the actual stress on the surface of the wafer is within 5% of the calculated stress from the equation.

2.6.3 Optical Wafer Curvature Measurements

In this method, a laser is pointed on the top surface of the wafer and scanned across its length to measure its curvature. The stress on the top surface is then estimated from the curvature measurements. Figure 2-14 shows the schematic of this measurement technique. This method is
difficult, time consuming and prone to error because it is very dependant on the quality of the laser equipment and accuracy of the curvature measurement. The stress estimated by this technique also reveals that the stress on the wafer surface is close to what is expected.

![Schematic diagram of optical wafer curvature measurement technique](image)

Figure 2-14. A schematic diagram of the optical wafer curvature measurement technique is shown.

### 2.7 High Stress Measurements

After the calibration, the setup is ready to perform the high stress wafer bending experiments. A Keithley 4200 Semiconductor Characterization System is used to perform the electrical measurements on the devices. Standard coaxial cables, micropositioner probe arms, and tungsten probe tips are used. Figure 2-15 shows a photograph of the complete setup. Figure 2-16 shows a close-up image of the Si wafer piece bent under high stress. Despite the large curvature of the wafer at high stress, it is observed that landing the probe tips repeatedly to measure devices does not affect the stress and does not cause the wafer to crack. This confirms our understanding that the wafer cracking occurs due to the edge defects which propagate at higher stresses and samples with optical quality edges are essential to achieve high stress.
Figure 2-15. A photograph of the complete flexure based wafer bending setup. Not shown in the photograph is the Keithley 4200 semiconductor parameter analyzer to which the coaxial cables from the probe arms are connected.

Figure 2-16. A photograph showing a silicon wafer bent under high stress. The large curvature of the wafer is evident from the picture.
2.8 Wire Bonding to Measure Short Channel Devices

It is known that as the channel length of a device reduces, the external resistance contributes a significant fraction to the total resistance and true extraction of the channel resistance change or the piezoresistance coefficients is very challenging. Stress measurements with probe tips require raising and lowering of the tips on the device pads for each stress point measured. This can lead to significant variation in the total resistance measured based on the exact point where the tip is landed each time. In order to eliminate this variation and keep the external resistance fixed, a scheme of wire bonding directly to the device pads was devised. There are three ways to wire bond to a device pad:

**Epoxy bonding:** The most direct method is to epoxy a thin gold wire to the bond pad. But this method failed, for the device pads for most technologies at the M1 level are 50μm x 50μm (2mil x 2mil) and are too small and have too fine a pitch to epoxy bond without shorting the four pads of a transistor together. Hence this method was abandoned.

**Wedge bonding:** The next method is to wedge bond using a wedge bonder on the device pad. This method had two problems. The first was that most wedge bonders use gold wires, while most device pads are made of copper. It is a well known fact that copper and gold do not bond unless in an clean forming gas environment. So an alternate material Al was tried on devices which had aluminum pads. This too did not work, because the wedge bonder works on the principle of wedging a piece of wire on the pad and this method damaged the devices. Even using the lowest force setting did not yield any working devices after the bonding.

**Ball bonding:** The final method available was ball bonding. For this method a Kulike & Soffa 4124 Ball Bonding machine was used. This instrument is also generally used to bond gold balls on to gold pads using a combination of force, ultrasonic energy and temperature. Use of
high temperature was not an option because it forms a native oxide on both copper and aluminum and prevents the formation of a good bond. Aluminum wire was used to bond on to devices with Al pads. Using this method which is very time-consuming and challenging, a couple of devices were successfully bonded and Figure 2-17 shows the plot of stress measurements after the bonding process. Despite the small success achieved using this method, it was not used for the rest of the high stress measurements because of the poor yield. It took 200 attempts before we could get one working device. Also most of the times, even the ball bonding method damages the device electrically even though there is no visible physical damage.

![Graph](image.png)

Figure 2-17. Drive current enhancement data after wire bonding to a device. We see that there is considerable deviation from published work.
CHAPTER 3
SILICON CHANNEL MOSFET PIEZORESISTANCE

3.1 Planar MOSFETs

3.1.1 Electron Mobility Enhancement

Modeling of the highly strained Si devices has largely used piezoresistance coefficients, particularly for nMOSFETs, due to the difficulty in predictively modeling the stress enhanced electron mobility.\(^{[11, 12]}\) However, nearly all the experimentally determined coefficients are used outside the stress range in which they were extracted and non-linearity in the piezoresistance should be expected \(^{[13]}\). For example, to model the mobility enhancement in 90 and 65nm technologies \(^{[6, 9]}\), Smith \(\pi\) coefficients \(^{[1]}\) are used which were extracted at stresses up to only 10 MPa. Recently, there have been studies \(^{[14-20]}\) measuring MOSFET piezoresistance coefficient for slightly larger stresses in the range of a few 100 MPa with one exception being Wang et. al. \(^{[19]}\) with nMOSFET data up to \(~580\) MPa for long channel and 300MPa for short channel MOSFETs, all of this work still being below the high stress levels already in production \(^{[5, 9]}\). In this work, we extend the stress range and measure piezoresistance coefficients and mobility enhancement at both low and high stresses up to \(~1.5\) GPa.

The flexure based four-point mechanical wafer bending setup is used to apply large uniaxial tensile stress (up to 1.2 GPa) on industrial nMOSFETs with 0 to \(~700\) MPa of process induced stress in order to study the electron mobility enhancement trend at high stresses. The samples used for this work consist of nMOSFETs with \((100)\) surface, \(<110>\) channel, \(~1.2\-1.4\) nm nitrided \(\text{SiO}_2\) gate oxide, phosphorus doped n+ poly-silicon gates and body doping density \((N_A) ~10^{18}/\text{cm}^3\). The devices either had an unstressed or stressed nitride capping layer creating samples with various amounts of process induced stress ranging from 300 MPa to 700 MPa \((\pm100\text{MPa})\) for the short devices. The effective process induced built in stress is the sum of
longitudinal and out of plane stresses as determined using the relation $\sigma_{\text{eff}} = \sigma_{<110>} + \frac{\pi_{\text{op}}}{\pi_{110}} \sigma_{<\text{op}>}$.[1,25] (“op” stands for out of plane) and 2D-FLOOPS stress simulations[26]. The measured devices had the following width/ gate lengths: 10/10, 10/0.1, 1/0.14, and 1/0.05 (in microns). Linear drain current was measured with a Keithley 4200 semiconductor parameter analyzer at $V_{ds} = 50\text{mV}$ and $V_{gs}$ being swept from 0 to 1.2V. All drain current and mobility enhancement extractions were done at $E_{\text{eff}} \approx 0.7\text{MV/cm}$[27].

The stress altered drain current versus total uniaxial longitudinal stress for long and short devices is shown in Figure 3-1A. Observed is a linear drain current enhancement for the long devices for longitudinal tensile stresses up to 820 MPa and hence an approximately constant $\pi$-coefficient with a value of -315 (extracted and plotted in Figure. 3-1B). Stress altered short channel data are also shown for stresses greater than 1 GPa (The stress due to the bending setup is assumed to add linearly to the built-in process stress) in Figures 3-1A and 3-1B. For the short devices, a non-linear slope in the drain current and a considerable degradation in the $\pi$-coefficient are observed (with respect to the long devices). The intrinsic channel $\pi$-coefficient values for the short devices ($L_g = 0.14\mu\text{m}, 0.1\mu\text{m}$ and $0.05\mu\text{m}$) can be estimated by correcting for the source and drain external resistance, $R_{s/d}$, from the stress altered channel mobility. A simple $R_{\text{tot}}$ slope-based effective channel mobility ($\mu_{\text{eff}}$) extraction method outlined in [28] is used to correct for the parasitic $R_{s/d}$ dependence of the short devices. This method offers a simple way of finding $\mu_{\text{eff}}$ by plotting the $R_{\text{tot}}$ for both long and short channel devices at a given vertical field vs. channel length with the intercept of this plot giving us the $R_{s/d}$ (50-60% of $R_{\text{tot}}$ in this work) contribution to the $R_{\text{tot}}$. The total resistance is given by $R_{\text{tot}} = \frac{V_{ds}}{I_{d\text{-lin}}}$. 


Figure 3-1. NMOSFET drive current enhancement and channel piezoresistance plotted as a function of wafer bending stress. A) $\Delta I_d/I_{d0}$ enhancement vs. uniaxial longitudinal tensile stress for 10 and 0.1 micron devices along with other published work. B) Extracted channel $\pi$-coefficient values for 10 and 0.1 micrometer devices (before and after $R_{sil}$ correction).
For the long devices ($L_g = 10\mu m$), $\mu_{\text{eff}}$ is defined in the linear region of operation by \cite{27} 

$$\mu_{\text{eff}} = \frac{I_{ds} L_g}{C_{ox} W \left(V_{gs} - V_t\right) V_{ds}}$$

where $I_d$ is the drain current, $L_g$ is the channel length, $C_{ox} = \varepsilon_{ox}/t_{ox}$ is the gate oxide capacitance per unit area, $W$ is the width of the device, and $V_{gs}$, $V_{ds}$ and $V_t$ are the gate bias, drain bias, and the threshold voltage respectively. The effective mobility enhancement for long and short devices obtained by the above two methods is plotted in Figure 3-2 along with more samples with various built in process stresses. It can be seen that after $R_{s/d}$ correction, the intrinsic channel $\pi$-coefficient values (extracted and plotted in Figure 3-1B) and mobility enhancement (in Figure 3-2) are independent of channel length and are linear even up to a stress of 1.5 GPa.

![Figure 3-2](image)

Figure 3-2. Extracted effective mobility enhancement ($\mu_{\text{eff}}$) vs. stress plotted for long and short devices with $R_{s/d}$ correction versus stress at $E_{\text{eff}} \approx 0.7$ MV/cm.

The underlying physics for stress induced enhancement of electron mobility in Si has been widely reported \cite{11, 12, and 23}. We know that the Si conduction band is six-fold degenerate and has
the band-minimum at the \( \Delta \)-point. The dominant scattering mechanism at room temperature is phonon scattering (both inter-valley and intra-valley). The application of either a uniaxial or biaxial stress splits the degeneracy of the six \( \Delta \)-valleys into \( \Delta_2 \) and \( \Delta_4 \), with the \( \Delta_2 \) valleys being lower in energy for the tensile stress case. This results in a repopulation of carriers from the higher energy \( \Delta_4 \) to the lower energy \( \Delta_2 \) valleys. Since the \( \Delta_2 \) valleys have a low mass along the \( <110> \) direction which is usually the channel direction, as compared with the \( \Delta_4 \) valleys, we see an enhancement in mobility. Also the band splitting due to stress causes a reduction in inter-valley phonon scattering because of the reduction in the density of states in the \( \Delta_4 \) valleys, which also contributes to an enhanced mobility. A schematic diagram showing the above conceptual picture is shown in Figure 3-3.

![Schematic diagram](image)

Figure 3-3. Schematic diagram of the physics of electron mobility enhancement in Si is shown.

But the above simple picture is valid only for the bulk or 3D case. In a Si-MOSFET with vertical field confinement and high inversion charge densities, it is no longer a 3D picture, but the carriers now have to be treated as a 2DEG (2D electron gas). It is known that vertical field
confinement also splits the degeneracy of the conduction band valleys in a similar fashion as stress. It can also be shown from simple band-calculations that for high inversion charge densities or high vertical fields, the band splitting due to confinement is larger than that due to stress [29]. This implies that the repopulation of carriers and suppression of phonon scattering has already occurred in the unstressed case alone and enhancement due to stress should be smaller at high fields. But the stress induced drive current enhancement data at high fields in Figure 3-1 and the field dependence of enhancement shown in Figure 3-4 suggest otherwise.

![Figure 3-4](image_url)

Figure 3-4. Plot of electron mobility extracted from drive current data as a function of vertical field for different uniaxial tensile stresses. We can see that the enhancement is retained at high vertical fields.

This large magnitude of electron mobility enhancement is surprisingly not well understood and cannot be explained by mass change and phonon scattering reduction alone [11]. It is known that at high vertical fields, the majority of the carriers populate the ground sub-band which is
strongly confined and is hence strongly influenced by the roughness of the surface. If the field is strong enough the mobility could be limited by surface roughness scattering as opposed to the usual phonon scattering. In this case, the reduction in phonon scattering is expected to have an even smaller effect on mobility enhancement. Given this scenario for the state of the art devices, the large mobility enhancement seen in experiments has been explained only by ad hoc assumptions of reduced surface roughness scattering with stress or the dominance of f vs. g type scattering \cite{11,12}. The linearity and enhancement up to \(~1.6\) times seen for the uniaxial stress case is similar to what has been observed for biaxial stress \cite{30} even though the maximum mobility enhancement should not be expected to be universal and depends on channel doping (smaller enhancement at higher doping because of ionized impurity scattering). It has recently been suggested that uniaxial stress along \(<110>\) provides a significantly larger enhancement over biaxial tensile stress \cite{20} because of band warping leading to reduced conductivity mass but not all calculations support this conclusion \cite{31}. Since under high vertical confinement fields, there is not much repopulation of carriers, strain induced band warping for longitudinal \(<110>\) stress may play some role in the enhanced mobility, but it appears the effect is not large (\(<20\%\)) and cannot explain the entire enhancement.

The model based on reduced surface roughness scattering with tensile stress has recently gained more popularity and acceptance with some experimental and theoretical works which confirm the same \cite{32,33}. In Ref \cite{32}, an atomic scale model is developed for the MOSFET surface roughness and its changes with application of external stress are monitored. The model predictions match quite well with experimental results for the biaxial stress case. A similar analysis on evaluating the effect of stress on surface roughness scattering using an experimental approach is being done by researchers in our group \cite{34}. They carried out temperature dependent
measurements of electron mobility enhancement and observed greater enhancements at lower temperatures. This clearly suggests that stress reduces surface roughness scattering, since at low temperatures, it is the dominant scattering mechanism and there is also not much mass change effect at high vertical fields.

After correcting for external resistance, the nMOSFET intrinsic piezoresistance coefficients are found to be approximately constant and the electron mobility increases linearly with stress (up to ~1.5GPa), respectively. The highest level of channel stress in production logic technologies consist of ~1GPa for pMOSFETs \[^9\] but only ~500MPa \[^9\] for n-MOSFETs. Thus, the piezoresistance and mobility data suggest stress enhanced electron mobility is not saturated and additional performance enhancement via higher levels of channel stress are possible up to 1.6 times even for high ~10\(^{18}\) cm\(^{-3}\) channel doping. This large enhancement at high gate fields is probably due to the reduction in surface roughness scattering due to stress and some small mass change effects. It is quite surprising that despite the different mechanisms coming into play at low fields and high fields, the magnitude of the piezoresistance is the same all the way up to 1 GPa of uniaxial stress. However, to realize the additional electron mobility enhancement at high stress on short devices, the external resistance needs to be reduced. Furthermore, this work highlights that for nano-scale MOSFETs in logic applications, the real limiter will likely be external resistance and not the magnitude of channel mobility.

3.1.2 Hole Mobility Enhancement

Applying uniaxial stress to the Si-channel to enhance hole mobility is standard in sub-90nm logic technologies \[^6\], \[^9\], \[^10\]. Increased Ge concentration in the SiGe, closer proximity of the SiGe to the channel, and high stress compressive contact etch-stop layer has resulted in channel stress greater than 1GPa and mobility enhancement of ~200\% \[^35\]-\[^37\]. To date, the uniaxial stress (> 1GPa) data is on p-channel metal-oxide-semiconductor field-effect transistor
(p-MOSFET) samples with large process-induced stress [6, 9-10, 35-38] where uncertainty in the stress, non-zero stress components in all 3 directions [36, 37], and changing external resistance with the Si$_x$Ge$_{1-x}$ complicates the stress enhanced mobility extraction. Despite this, the experimental data shows a non-linear trend for the mobility enhancement with stress calling for a more accurate study at high stresses. In this work, the super-linear hole mobility enhancement is studied more closely through high stress wafer bending experiments with the results compared to $k \cdot p$ calculations to gain more physical insight.

Wafer bending experiments are performed using the flexure-based wafer bending setup on Si wafers cleaved using a diamond tip scribe to provide edges suitable to achieve high stress. The wafer samples consist of industrial p-MOSFETs with (100) surface, <110> channel, ~1.2 nm nitrided SiO$_2$ gate oxide, phosphorus doped p+ poly-silicon gates, and ~10$^{18}$/cm$^3$ body doping density ($N_D$). The measured devices have the following width/gate lengths: 1/0.05, 1/0.055, 1/0.06, and 1/0.065 (in microns). The devices have a compressive stressed nitride capping layer and SiGe source/drain creating samples with ~1GPa (±100MPa) built-in stress for the shortest devices. External uniaxial compressive and tensile stresses are applied on these samples with 1 GPa built-in stress to obtain data with a net stress over the entire 0-2 GPa stress range. Linear drain current is measured at $V_{ds} = -10$ to -50mV and $V_{gs}$ being swept from 0 to -1.2V. The short channel mobility is extracted by correcting for the parasitic source/drain resistance using the $dR/dL$ method as described by [28] and [39] at a gate overdrive of -0.5V as shown in Figure 3-5. for three cases: 1 GPa built-in stress and the same with 500 MPa external uniaxial tensile and compressive stresses. Since the mobility is inversely proportional to the slope, $dR/dL$, the change in mobility with stress is apparent from the change in slope in Figure 3-5. $Q_{inv}$ is estimated in this
work by \( C_{\text{ox}} (V_g-V_{\text{on}}) \) where \( C_{\text{ox}} = \varepsilon_{\text{ox}}/t_{\text{ox}} \). \( V_{\text{on}} \) is determined by the linear interpolation of \( I_d \) vs. \( V_g \) at the maximum transconductance point.

The hole mobility enhancement extracted from the wafer-bending experiments in this work along with other published work \(^{36-37}\) is plotted in Figure 3-6 up to \( \sim 2 \) GPa of stress. This is the widest range of wafer bending p-MOSFET mobility enhancement data reported till date. A super-linear trend for stress greater than \( \sim 700\text{MPa} \) is observed which is consistent with other works \(^{35-36}\). The data also matches previous reported wafer bending results for stress less than \( 500\text{MPa} \).

![Figure 3-5. Dependence of \( R_{\text{tot}} \) on \( L_{\text{mask}} \) for different wafer bending stresses. Mobility is extracted using the dR/dL method.](image)

In order to explain the super-linear behavior of the hole mobility enhancement, six-band \( k\cdot p \) calculation of the Si valence bands for a (100)/\(<110>\) pMOSFET inversion layer using the finite difference formalism \(^{41}\) is employed. Both phonon and surface roughness scattering
(dominant scattering mechanisms) \cite{13} are taken into account using scattering rate parameters and deformation potentials in \cite{13}. The calculated mobility enhancement is plotted in Figure 3-6.

Figure 3-6. Hole mobility enhancement as a function of stress along with other published work and calculation results using the k·p method is shown. The dotted linear line shows the low stress intrinsic piezoresistance coefficient.

Good agreement is observed at stresses below ~500 MPa \cite{40}, but not at higher stresses despite the calculations also showing a super-linear trend. The super-linear characteristics can be understood from the sub-band splitting and optical phonon scattering rate change as a function of stress which is plotted in Figure 3-7. For stress less than 500 MPa, the optical phonon scattering rate change due to 2D- density of states change is small. Also for stress less than 1 GPa, the sub-band splitting is less than the optical phonon energy (~60meV), hence the enhanced mobility results primarily from mass change due to band warping and repopulation. For stress greater than ~700 MPa, the sub-band splitting becomes comparable to the optical phonon energy. Thus along with a reduced density of states resulting from the band warping, a significant reduction in
phonon scattering occurs, which gives added mobility enhancement resulting in a super-linear enhancement. A straight line indicating the low stress piezoresistance coefficient which matches the bulk $\pi$-coefficient value \cite{1} is shown in Figure 3-6 clearly emphasizing the deviation from the linear trend. Although it has been suggested for stress enhanced electron mobility that simulation can match experimental results by making *adhoc* assumptions of reduced surface roughness scattering at higher tensile stresses, there is no physical reasoning for this \cite{13}. There is uncertainty in the scattering rate calculations for surface roughness \cite{13} and coupling of $f$ vs. $g$ type inter-valley phonon scattering \cite{12}. Similar uncertainty exists for the hole mobility case. But a recent study taking into account atomic scale surface roughness suggests that strain indeed has a large impact on the surface roughness scattering and hence the mobility \cite{32}. The $k\cdot p$ calculation in this work does not take into account the detailed atomic scale surface roughness effects as done in Ref. [32] and hence this is suggested as the cause for the deviation from experimental results.

![Figure 3-7. Sub-band splitting and optical phonon scattering rates plotted as a function of stress.](image)
The vertical field dependence of the hole mobility enhancement is shown in Figure 3-8. The band-splitting due to confinement and uniaxial compressive stress is additive for the Si valence band. Therefore the enhancement due to stress is expected to be retained at both low and high vertical fields\(^{[23]}\) as observed. Nearly constant mobility enhancement is seen in Fig. 3-8 as a function of gate overdrive. This is unlike the biaxial stress case where the confinement splitting and stress splitting oppose each other. Hence under strong vertical confinement fields, we see that biaxial tensile stress provides negligible mobility enhancement\(^{[23]}\). It is only beyond stresses of 2 GPa that we actually start to see enhancement for biaxial stress.

![Figure 3-8: Hole mobility of short devices with a built-in stress of ~1 GPa as a function of gate overdrive for different mechanical stresses. Data for a long channel device (4\(\mu\)m) is also plotted as a reference.](image)

**3.1.3 Effect of Stress on n and p – MOSFET \(I_{d-sat}\)**

The effect of stress on electron and hole mobility is well studied and understood, but not much has been reported on the effect of stress on the saturation drive current, which probably is a
more important parameter for logic devices. For long devices, we know that the saturation drive
current is directly proportional to the mobility of the carriers in the channel and hence will show
the same trend as that of the long channel mobility. But as we scale the channel lengths to the
regime where velocity saturation and velocity overshoot start occurring (true for sub 90nm nodes
where stress engineering is incorporated), the picture is quite complicated and the mechanism for
enhancement of the saturation drive current with stress is not well understood. The picture is
even more complicated when quasi-ballisticity in ultra-short channel length devices is
considered.

The above problem can be tackled by two different approaches. The first is by considering
the velocity saturation model as the starting point, where the saturation drive current is controlled
by the drain-end of the channel. We know that when the channel length of a MOSFET is scaled,
due to the high lateral fields, the carrier velocity saturates causing the drain current of the carriers
to saturate at values much lower than expected from the traditional long-channel models. This
saturation occurs at a critical field which is again dependent on carrier mobility as follows [27]:

\[ E_c = \frac{v_{sat}}{\mu_{eff}} \]  

(3-1)

where \( \mu_{eff} \) is the effective channel mobility, and \( v_{sat} \) is the saturation velocity. From experimental
measurements the velocity-field relationship takes the following empirical form [27]:

\[ V = \frac{\mu_{eff} E}{\left[1 + \left(\frac{E}{E_c}\right)^n\right]^{1/n}} \]  

(3-2)

where \( n = 2 \) for electrons and \( n = 1 \) for holes. \( n > 1 \) is a measure of how rapidly the carriers
approach saturation. For simplicity assuming \( n = 1 \), the saturation drive current has the following
dependence on mobility and \( v_{sat} \) [27]:

52
\[ I_{dsat} = W \nu_{sat} C_{ox} \left( V_g - V_t \right) \sqrt{\frac{1 + 2 \mu_{eff} \left( V_g - V_t \right) / \left( m \nu_{sat} L \right) - 1}{1 + 2 \mu_{eff} \left( V_g - V_t \right) / \left( m \nu_{sat} L \right) + 1}} \] 

(3-3)

This expression is equivalent to

\[ I_{dsat} = W \nu_{sat} Q_{inv-D} \] 

(3-4)

or

\[ I_{dsat} = W \nu_{sat} Q_{inv-S} \sqrt{\frac{1 + 2 \mu_{eff} \left( V_g - V_t \right) / \left( m \nu_{sat} L \right) - 1}{1 + 2 \mu_{eff} \left( V_g - V_t \right) / \left( m \nu_{sat} L \right) + 1}} \] 

(3-5)

where \( Q_{inv-S} \) is the inversion charge density at the source, which is just \( C_{ox}(V_g - V_t) \) and \( m \) is the body effect coefficient as defined in Ref [27]. We know that even though the carrier velocity saturates, unlike the pinch-off condition for long channel devices, there is a finite amount of charge at the drain end which contributes to the saturation current. From Equation 3-5, it is evident that not all the carriers injected from the source make it to the drain with the saturation velocity, some of them are injected back into the source due to scattering events in the channel region. Hence, we see that the fractional term in Equation 3-5 has a dependence on mobility which determines what fraction of the carriers that enter the channel region from the source make it to the drain. So application of either external or process induced mechanical stress, which reduces scattering and carrier effective mass and hence increases the channel mobility will increase this fractional term and hence the saturation drain current of the MOSFET. It is also argued whether stress improves the magnitude of the saturation velocity itself. Though this idea seems very appealing, it is not plausible. The energy of the carriers when they attain the saturation velocity is very large and hence they are no longer located at the band minimum. Mechanical stress warps the bands and reduces, the transport effective mass only close to the
conduction or valence band minimum. When we move away from the band minimum to higher energy k-points in the reciprocal space which is where the carriers with saturation velocity are located, there is little or no effect due to stress. Hence the saturation velocity of the carriers is unlikely to change much, and the change in the saturation drive current probably results solely from the dependence on mobility. The velocity saturation model though adequate to explain drive-current characteristics for sub-micron devices is inadequate for sub-100nm devices where significant velocity overshoot is observed. So to understand the stress effects in such ultra-short channel devices from a different perspective we look at the scattering based model.

The second approach involves the scattering based model described in Ref [42-43]. This scattering theory relates the steady-state current to transmission and reflection coefficients. Traditionally in MOSFET analysis, the transport of carriers across the source-channel barrier which exists even beyond threshold has been ignored because the transport across the channel has been the limiting factor. But as the channel lengths are scaled to the sub-100nm regime where there is significant velocity overshoot, transport across the source-channel barrier becomes more and more important, and the saturation drain current is limited by the transport at the source-end of the channel instead of the drain end. The scattering theory assumes the source of the MOSFET to be a reservoir of carriers which injects a flux \( a_s \) into the source-channel barrier, of which a fraction \( t_s \) transmits across the barrier into the channel. A fraction \( r_c \) backscatters and reenters the source, while the remaining exits the drain contributing to the saturation drain current. Given this formulation, the saturation drain current can be written as \(^{42} \):

\[
I_{Dsat} = C_{ox} W \nu_T \left( \frac{1 - r_c}{1 + r_c} \right) (V_{GS} - V_T) \tag{3-6}
\]
Here $C_{ox}$, $V_{gs}$, $V_t$ and $W$ have their usual definition, and $v_T$ is the thermal injection velocity at the source-end. A backscattering coefficient of 0 signifies the ballistic limit in which case all carriers entering the channel across the source-channel barrier are collected at the drain before encountering even a single scattering event. Under this limit, the saturation drain current is limited by the source thermal injection velocity. This source thermal injection velocity is a function of both the effective mass of the carriers and the carrier density. The carrier density determines the amount of degeneracy and also the populations in the different sub-bands which have different effective masses. A closed form expression assuming a single-band occupation and degenerate Fermi-Dirac carrier distribution has been derived in Ref [43]:

$$v_T = \sqrt{\frac{2\kappa_B T_L}{\pi m^*}} \left\{ \frac{F_{1/2}(\eta)}{\ln(1 + e^{\eta})} \right\}$$

(3-7)

where $\eta = (E_F - \varepsilon_1)/k_B T$, and $F_{1/2}(\cdot)$ is the Fermi integral of the order $1/2$. Equation 3-7 suggests that application of stress which reduces the effective mass of the carriers due to both repopulation between sub-bands and band-warping will improve the saturation drain current even in the ballistic limit. The current technology nodes are not yet at the ballistic limit but close to it due to the significant velocity overshoot observed. There have been recent works on estimating how close to the ballistic limit current state-of-the-art MOSFETs are, and have found that they are no more than 40% ballistic [44-45] by looking at the effective carrier velocity using empirical techniques.

From this theory, it is evident that the saturation drain current is directly a function of the channel backscattering coefficient. For a given channel electric field $E$, the backscattering coefficient can be given as follows [43].
\[ r_{cE} \approx \frac{l}{l + \lambda} \tag{3-5} \]

where \( \lambda \) is the mean free path of the carriers and \( l \) is the distance traveled by the carriers at the source-end of the channel before they drop a potential \( k_B T/q \). The mean-free path of the carriers under a field is proportional to their mobility. So the above expression can be rewritten as \[^{[43]}\]:

\[ r_{cE} \approx \frac{1}{1 + 2 \mu_n^0 E/\nu_T} \tag{3-6} \]

So we see that the channel backscattering coefficient is inversely proportional to low-field carrier mobility (field is low at the source-end even though it is very high at the drain-end). When we apply mechanical stress and improve the low field carrier mobility, we reduce the channel backscattering coefficient and hence improve the saturation drain current. So it can be said that low field mobility continues to be an important factor even in the sub-100nm regime where there is significant velocity overshoot. There have been studies to quantify this dependence between low-field carrier mobility and saturation drain current by looking at the change in effective source velocity with mechanical stress for low stress levels and have reported a value of \(~0.5\) for electrons \[^{[46]}\]. So if the electron mobility is improved by a certain amount, the saturation drain current is improved by half that value. A similar experiment was conducted in this work for the stress range of 0-1 GPa for 50-60nm devices, and the drain current enhancement as a function of stress is plotted in Figure 3-9. It is seen that the same fraction of 0.5 is maintained between drain current enhancement and mobility enhancement. A recent work extracted similar data for holes and observed the ratio of 0.5 between mobility and saturation drain current enhancement \[^{[37]}\]. All this data hints that the ratio of \(~0.5\) is universal.
In order to gain further insight into this dependence, a simple qualitative assessment of the two models described before is made. Both models were first fitted to the experimentally measured values of drain current under the no stress condition, and then the mobility was changed to see how much change in drain current each model predicted. For this analysis, the same 60nm nMOSFET device data as in Figure 3-9 was used. First looking at the velocity saturation model, the values for W, Vg-Vt, \( \mu_{\text{eff}} \) and L are those measured experimentally, while those for \( C_{\text{ox}} \), m and \( v_{\text{sat}} \) are varied about their mean value to match the measured \( I_{d,\text{sat}} \). We know that the value for \( v_{\text{sat}} \) in MOSFETs is smaller than that for bulk and for electrons it is around 7-8 x10^6 cm/s \[^{[27]}\]. Even after using a value of 8x10^6 cm/s for \( v_{\text{sat}} \), we can match the observed drain current only by assuming a very large value for \( C_{\text{ox}} \) or in essence an inversion charge density of \( \sim 3\times10^{13}/\text{cm}^2 \). But the inversion charge density at the measured gate overdrive condition is just \( \sim 1\times10^{13}/\text{cm}^2 \). This clearly implies that the velocity at the drain is significantly larger than the
saturation velocity, and there is significant velocity overshoot. Then, the value for $v_{sat}$ was increased instead of $C_{ox}$ (EOT ~ 2nm), and the experimental drain current value was matched for a $v_{sat}$ of $2.5 \times 10^7$ cm/s. Given this starting point, the mobility value was allowed to increase, and the change in the $I_{d-sat}$ was tracked. The percentage change in $I_{d-sat}$ using this technique as a function of percentage mobility change is plotted in Figure 3-10. Next we look at the scattering based model; the values for $W$, $C_{ox}$ and $\mu_{eff}$ are measured experimentally while those for $E_0$ (lateral field at the source-end of the channel) are fitted to match the data. From Ref [42] we know that if we assume carrier degeneracy and single band model (high vertical fields and hence all carriers populate only the lowest sub-band and no repopulation effects), the source thermal injection velocity tends to a value of $2 \times 10^7$ cm/s. This is reasonable because it is lower than the value set for the $v_{sat}$ in the previous model. Again the mobility is allowed to change and the changes in the saturation drive current are tracked and plotted in Figure 3-10 along with the experimentally observed values. We see that the prediction from the velocity saturation model is closer to the experimental value than the scattering based model. Even though the deviation of the scattering based model is not much, we should keep in mind that we have not taken into account the increase in the source injection velocity with mass change which could cause further deviation from the measured values. Detailed Poisson-Schrödinger simulations are required to accurately predict the thermal injection velocity change and hence the saturation drain current change with stress. This analysis shows that both the velocity saturation model and the scattering model demonstrate the importance of mobility even in the saturation regime. To first order, the velocity saturation model if modified to take into account the overshoot effects, is a better tool to predict stress induced enhancement of the saturation drive current in these 60nm channel length nMOSFETs.
Figure 3-10. Prediction of percentage $I_{d\text{-sat}}$ change with percentage mobility change using both models as compared with experimental results is plotted.

3.2 Double-Gate FinFETs

3.2.1 Mobility Enhancement

Non-planar 3D-devices such as double-gate FinFETs are considered a possible replacement for planar devices to extend the Si roadmap at the 22nm and beyond [47, 48]. Additional performance enhancement in FinFETs via strain using different types of stressors such as SiGe [49] and SiC [50] source/drain, CESLs [48, 51, 52], global strain engineering [53, 54] and various combinations [47, 55, 56] have been recently investigated. Among the various strain techniques, CESLs are the simplest and most direct way to induce stress into the FinFET channels [52]. Performance enhancement from CESL in planar devices is well understood, but there lacks a quantitative understanding of the same for the FinFET case mainly because most existing strain enhancement models rely on bulk $\pi$-coefficients [48, 51-53, 55, 57]. The first externally applied mechanical stress effect on MuGFETs was reported in [57], but this work was limited to biaxial stress which does not yield the channel $\pi$-coefficients. In this work, we measure the
longitudinal $\pi$-coefficients for both n and p-FinFETs using uniaxial wafer bending over a 1GPa stress range and compare with bulk. This understanding is further used to explain the stress transfer mechanism in FinFETs from the CESLs.

FinFETs are fabricated using SOI wafers with 80nm silicon top layer and 145nm buried oxide. Fins of width down to 20nm and gate lengths ranging from 100nm to 10$\mu$m are patterned with 193nm lithography and reactive-ion-etching (RIE). A 2nm high-k gate dielectric followed by TiN metal gate stack is deposited by atomic layer deposition (ALD). This is followed by a 200nm poly-Si gate deposition using chemical vapor deposition (CVD). The gates are patterned and resist etched to form devices with (110) sidewall surfaces and $<110>$ channel direction. Implanted source/drain dopants are activated by spike anneal, while the fins are left undoped. Then 50nm tensile and compressive-CESLs with built-in stress of $\sim$1GPa and $\sim$1.4GPa are deposited on n and p-FinFET, respectively. The control devices have a very low-stress (<60MPa) c-CESL. The presence of a SiN hard mask on the top of the fin makes it a DG instead of a tri-gate FinFET. Tungsten vias and AlSi metal at M1 complete the device processing.

Since CESL induces stresses in three directions in a FinFET, modeling requires three $\pi$-coefficients; $\pi$-longitudinal (x), $\pi$-in-plane-transverse (z) and $\pi$-out-of-plane-transverse (y). A schematic diagram of a FinFET with the stress application direction is shown in the inset in Figure 3-11A. Because of the 3-dimensional nature of the fins and the different aspect ratio of the fin with respect to the Si substrate, there is a lot of uncertainty in estimating the mechanical stress induced in the y and z directions. Uniaxial stress application along the channel direction (x) does not suffer from this problem.
Figure 3-11. Drive current enhancement data for wide and narrow FinFETs with uniaxial tensile stress. A) 1μm long, 200nm wide n-FinFETs compared with published data on planar (110)/<110> nMOSFETs. Inset shows schematic of a fin and the stress application direction. B) The same for 20nm wide n and p FinFETs.
In order to verify the accuracy of the stress applied by the wafer bending setup, the linear drive current enhancement of a long (1μm) channel length 200nm wide n-FinFET is measured versus externally applied stress and compared with (110)/<110> planar nMOSFET. From Figure 3-11A, we see the expected result that the wide FinFET device which has negligible physical confinement of the carriers behaves similar to a planar MOSFET and that the longitudinal $\pi$-value extracted is accurate. With this knowledge, linear and saturation drive current enhancement are measured on 1μm channel length, 20nm wide n and p-FinFET devices for over a 1GPa stress range and plotted in Figure 3-11B. The extracted longitudinal $\pi$-coefficients from the wafer bending data along with those for planar MOSFETs $^{[58, 59]}$ and bulk values $^{[1, 60]}$ for both n and p-type devices are shown in Table 3.1. The bulk $\pi$-values differ from those of FinFETs as should be expected due to strong physical and electrical confinement induced band-splitting.

Table 3-1. Longitudinal piezoresistance coefficients for n and p-FinFETs along with previous published data for planar (110) MOSFETs and bulk Si (110) is shown.

<table>
<thead>
<tr>
<th>Device Type</th>
<th>Longitudinal Piezoresistance Coefficients (110)/&lt;110&gt; Units (x10^-11Pa^-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N-type</td>
</tr>
<tr>
<td>FinFET (This Work)</td>
<td>51.4</td>
</tr>
<tr>
<td>Planar MOSFET</td>
<td>40 [58]</td>
</tr>
<tr>
<td>Bulk Si (110)</td>
<td>31.2 [1, 60]</td>
</tr>
</tbody>
</table>

Figure 3-12A shows the saturation drive current enhancement of 100nm channel length, 20nm wide n and p-FinFETs with tensile and compressive CESLs respectively as compared with the control wafer at a gate overdrive of 0.85V. The data is averaged over 9 sites on the wafer. Both n and p-FinFET devices show ~30% enhancement which is ~2x larger than the enhancement of ~15% on planar devices with the same CESL $^{[61]}$. Using the piezoresistance data in Table 3-1 and enhancement data in Figure 3-12A, a qualitative estimate of an effective
channel stress (along x) can be made but it is not accurate since the actual stress is in all the three
directions. Figure 3-12B shows similar data for 10μm channel length, 20nm wide FinFETs,
which show no enhancement compared to the control devices. This result implies that the stress
induced by the CESL has a strong channel length dependence which is explained in the next
section.

![Figure 3-12: Comparison of CESL enhancement for long and short devices](image)

Figure 3-12. Comparison of CESL enhancement for long and short devices is shown. A)
Saturation Id-Vg of 100nm channel length and 50nm wide n & p-FinFETs with
tensile and compressive CESLs respectively along with control at a gate overdrive of
0.85V. B) Same for 10μm devices. The data is averaged over 9 sites on the entire
wafer.

A TEM cross-section of the fins, gate-stack and the CESL is shown in Figure 3.13A. As
observed in the cross section, the CESL is not wrapped around the fin along the gate and hence
we can expect little stress coupling to take place into the fins through the gate. However the
tensile-CESL is wrapped directly around the fin at the fin-extensions and induces a compressive
stress along the y and z directions as seen in the schematic diagram in Figure 3.13B. From wafer
bending results in Ref [58] on planar devices, it is known that compressive stress along y and z
directions enhances n-MOSFET performance. Since the above stress coupling takes place along
the fin-extension, the proximity of the channel to the source/drain fin-extensions will be a determining factor for the amount of stress transfer into the channel region under the gate. This is evident from the strong channel length dependence observed in the performance enhancement due to the CESL where the shortest devices (100nm) show the maximum performance enhancement of ~30% while the long devices (10μm) show little or no difference from the control wafers.

Figure 3-13. CESL stress transfer mechanism. A) TEM cross-section view of a DG-FinFET with a CESL along the gate. B) Schematic diagram of a FinFET with a CESL along the fin-extension. The arrows indicate the direction of stress in the CESL, which is the same as that in the fin.
Similar arguments hold for a compressive CESL for p-FinFET where a tensile stress is induced along the y and z directions which enhance the p-FinFET performance [51]. On the other hand, if the gate is scaled in height to be more conformal, similar arguments would hold true as there would be an additive effect from both, the stress coupling from the fin-extensions and directly through the gate. Both these factors induce stress into the channel in a similar fashion; though, the stress-coupling directly through the gate will be channel length independent.

3.2.2 FinFET Reliability

3.2.2.1 Bias instability

Strain effects on FinFET bias instability is studied by performing charge pumping measurements. This is a challenging task because a FinFET device unlike planar MOSFETs does not have a bulk contact and a traditional charge pumping measurement cannot be performed. No previous work on measurements of FinFET bias instability exist due to this reason except for a recent work by Ref [62], where a gated diode structure as shown in the schematic Figure 3-14 is used to do charge pumping. Here the anode behaves like the bulk contact to the slightly p-doped body of the fin. A similar gated diode structure with the same gate stack as in the devices mentioned in the previous section is used in this work too, to study the strain effects on the charge pumping current and hence the effect on interface state generation rate with stress. The result of such a measurement is shown in Figure 3-15. As can be seen, there is negligible impact of strain on the interface state generation rate. This is probably due to the large physical thickness of the high-k dielectric, which results in a very small magnitude of gate leakage current and hence an even smaller (negligible) change with stress.
Figure 3-14. The layout of a gated diode structure used for performing the charge pumping measurements is shown. The anode serves as the body contact.

Figure 3-15. Application of both uniaxial tension and compression shows negligible change in the charge pumping current. This implies that there is negligible change in the interface state density and interface state generation with stress.
3.2.2.2 Hot carrier injection

Strain effects on FinFET hot carrier injection induced degradation is studied by looking at the threshold voltage and transconductance change with mechanical stress as function of time when biased under hot carrier stress. Wafer bending experiments show increased Vt shifts with both tension and compression under hot carrier stress of Vg = Vd = 2.5V on 20nm wide and 90nm long n-FinFETs (Figure 3-16). Negligible change in Gm degradation is seen (Figure 3-17) suggesting that there is negligible interface state generation change with mechanical stress. Increased impact ionization at the drain-body junction due to bandgap narrowing with mechanical stress is attributed as the cause for the increased hot carrier induced Vt shift (Figure 3-18), though a more clear understanding of the mechanism is required for, tension and compression show different results.

![Figure 3-16. Increased FinFET hot carrier degradation is seen for both tensile and compressive uniaxial stresses.](image)
Figure 3-17. Gm degradation does not show any change with uniaxial tensile stresses suggesting that increased HCI degradation is due to bulk traps and not interface state generation.

Figure 3-18. Both tensile and compressive stresses cause bandgap narrowing at the drain-body junction resulting in increased impact-ionization. This in turn could lead to increased HCI degradation.
3.2.3 Ring Oscillator Delay Performance

Wafer bending experiments are also done to study strain effects on FinFET RO delay performance. Figure 3-19 shows good FinFET RO performance. Figure 3-20 shows that RO delay change with stress is small and also saturates for less than 100MPa stress. This is expected due to the opposing effects of wafer bending stress on n and p-FinFETs. Tensile stress shows a slightly larger change due to the larger n-FinFET $\pi$-value than that of the p-FinFET.

![Figure 3-19. FinFET ring-oscillator delay/stage measured as a function of bias voltage showing reduced delay with increasing supply voltage as expected.](image1)

![Figure 3-20. Opposing effects on n and p-FinFETs due to strain result in marginal delay reduction. Larger enhancement is seen at low stress levels for the n-FinFET case due to their larger p-coefficient than p-FinFETs.](image2)
CHAPTER 4
GERMANIUM CHANNEL MOSFET PIEZORESISTANCE

4.1 Introduction

Low-power and high-performance requirements at the 22nm and beyond technology generations may not be met with a Si channel MOSFET even with the highly successful performance boosters like process induced uniaxial stress\[63\]. This has led to alternate high-mobility channel materials (group IV and III-V) being investigated extensively. However unstrained hole mobility in these materials is unlikely to be competitive with the 3-400% uniaxial stress enhanced mobility in strained Si\[59\]. Selective epitaxy to obtain defect-free high mobility strained Si$_{1-x}$Ge$_x$ channels have been recently investigated but only for small (<30%) Ge concentrations\[64-66\]. Whether uniaxial stress provides additional hole mobility enhancement for SiGe channel devices has been investigated via wafer bending on the (110) surface orientation and for low stresses (<150MPa)\[67\]. In this work the additivity of strain for (100) surface oriented Si$_x$Ge$_{1-x}$ and Ge channel devices are addressed. In this work, we show that Si/ Si$_{1-x}$Ge$_x$/Si quantum wells (QW) devices have significant additional mobility enhancements under compressive uniaxial stress as shown using $k\cdot p$ calculations and wafer bending experiments.

4.2 Device Fabrication and Performance

The process flow to fabricate the SiGe channel pMOSFET is summarized in Figure 4-1. Epitaxial heterostructures of (100) surface oriented Si/ Si$_{1-x}$Ge$_x$/Si are selectively grown on the active regions isolated by shallow trench isolation. Samples with varying Ge concentration, (x values of 20%, 50%, 70% and 100%) are grown followed by a thin Si cap layer necessary to fabricate a high quality interface between HfSiO$_x$ (high-k) gate insulator and the channel. TaN is sputtered\[68\] to form the metal gate followed by a standard gate-first CMOS process with either a spike anneal at 700 – 1075 °C or a rapid thermal annealing (RTA) at 600-700 °C (targeted based
on the Ge concentration) to diffuse and activate the dopants. The device channel orientation is chosen to be $<110>$ since this direction provides the largest uniaxial stress enhanced hole mobility in group IV and III-V materials \[^{69}\].

![Schematic of gate-first process flow for QW pMOSFET used in this work. A) Standard CMOS process flow with STI. B) Selective epitaxy of Si$_{1-x}$Ge$_x$/Si heterostructure on active regions only for pMOS. C) Gate stack formation with high-k and metal gate.](image)

Due to the large lattice mismatch (4%) between Ge and Si, the Si$_{1-x}$Ge$_x$ growth follows Stranski-Krastanov kinetics (layer-by-layer followed by islands) \[^{70}\] which makes the defect-free (no island formation) growth of Si$_{1-x}$Ge$_x$ thin films directly on Si very challenging. The devices used in this work were defect free fully strained Si$_{1-x}$Ge$_x$ thin films grown directly on active Si beyond the critical thickness by extending the growth mode (onset of islands formation/strain relaxation) by selective epitaxy of SiGe/Si heterostructures on active Si bordered by STI. Figure 4-2 shows the high resolution TEM image and a schematic band diagram of the heterostructure. The quantum well structure as seen in Figure 4-2B also helps in mitigating the band-to-band
tunneling and reducing the $I_{\text{off}}$ in these devices. The use of a Si cap significantly improves the transconductance and provides a better interface with the high-k dielectric. For the large area devices, the SiGe QWs are under biaxial compressive stress due to the lattice mismatch, but the mobility enhancement mainly results from the higher intrinsic mobility of SiGe (increases with increasing Ge content) compared to Si at high vertical fields and not due to the biaxial stress [71]. A mobility enhancement of ~3x is observed (Figure 4-3) with error bars included due to uncertainty in quantifying the channel inversion charge density in the SiGe QW via CV measurements.

Figure 4-2. Biaxially strained SiGe heterostructure pMOSFET is shown. A) TEM cross-section. B) Schematic band diagram.

Figure 4-3. Extracted hole mobility of the SiGe QW as compared with control Si.
4.3 Uniaxial Stress Enhancement

To obtain significant performance advantage over strained-Si, additive strain enhanced mobility is also needed for the Si/Si$_{1-x}$Ge$_x$/Si QWs. Uniaxial stressed Si hole mobility enhancement is well understood \cite{72} but similar understanding is lacking for the Si$_{1-x}$Ge$_x$ system. Strain effects on Si-Ge channel device hole mobility enhancement are studied via $k\cdot p$ calculations and the results are plotted in Figure 4-4. Si, Ge and Si$_{1-x}$Ge$_x$ show similar hole mobility enhancement for stress less than 3 GPa. This is understood from the strain altered band structure and 2D–density of states (DOS). Plotted in Figure 4-5 is the 2D constant energy contours of the topmost valence band in Si and Ge with and without stress from which it is evident that uniaxial stress has similar impact on both systems, implying that the mechanism for mobility enhancement is somewhat similar for Si and Ge. However, in Figure 4-6, we see Ge has a larger strain induced mass change which makes the mobility enhancement in Ge slightly larger than Si. Another difference is a smaller 2D-DOS (Figure 4-7) for the top valence band in Ge as compared to Si. This requires large band splitting in Ge to fully populate the top light mass Ge band and thus increasing the stress required for maximum hole mobility enhancement (~2.5GPa for Si and ~4GPa for Ge, see Fig. 4-4).

![Figure 4-4](image)

Figure 4-4. Si, Ge and SiGe all show similar enhancements for stress less than 3 GPa.
Figure 4-5. Topmost valence band structure change with 1.0 GPa of uniaxial compressive stress in Si and Ge.

Figure 4-6. Enhancement from mass change is larger for Ge than Si.
Figure 4-7. 2D-density of states change for the top valence band of Ge and Si with uniaxial stress is plotted.

Figure 4-8 shows pMOSFET drive current enhancement measured as a function of uniaxial compressive wafer bending stress on the flexure based wafer bending setup for thick epi-SiGe on Si samples with varying Ge content. As expected, the enhancement in the different SiGe samples matched that of Si except for the 100% epi-Ge sample which has the smallest critical thickness and hence the most defects. Drive current enhancement measurements for different Ge thickness samples plotted in Figure 4-9 confirms that the enhancement has epi-Ge film thickness dependence. Epi-Ge samples thicker than the critical thickness have large number of defects/dislocations which relax the wafer bending stress and hence show smaller drive current enhancement as shown in Figure 4-10, while thinner Ge films matches with the Si case. All terminal currents in a SiGe QW pMOSFET with and without wafer bending stress are plotted in Figure 4-11 which shows that there is almost 30% drive current enhancement (see inset) with negligible change in off-state current. Gate leakage current also shows negligible change with stress as expected \cite{73}, which is further reduced because of the much larger physical thickness of
the high-k gate dielectric compared to SiO$_2$. Also Figure 4-12 shows that there is negligible threshold voltage change with uniaxial stress.

Figure 4-8. Linear drive current enhancement for different Ge concentrations as a function of uniaxial compressive stress.

Figure 4-9. Linear drive current enhancement for different epi-Ge layer thicknesses as a function of uniaxial compression stress. The thin epi-Ge sample shows similar enhancement as Si due to no strain relaxation.
Figure 4-10. Dependence of drive current enhancement on epi-Ge thickness for a given stress is shown. Below the critical thickness, strain enhancement matches that of Si.

Figure 4-11. Drive current enhancement with negligible change in off-state current with stress. Inset shows 30% drive current enhancement with stress plotted in linear scale.
Figure 4-12. Negligible change in threshold voltage for both SiGe and Ge devices with uniaxial compressive stress is observed.
CHAPTER 5
GaAs CHANNEL MOSFET PIEZORESISTANCE

5.1 Introduction and Motivation

Demand for low power and high performance makes the scaling of transistors to the next technology node very challenging. Traditionally low power was achieved by reducing leakage currents ($I_{sub}$, $I_{gate}$ and $I_{jun}$) and reducing the supply voltage, while high performance was achieved by improving the drive current by using boosters like strain and having better short channel effect control by switching to high-k dielectrics from SiO$_2$. But for the sub-22nm nodes, we need an ultra high mobility channel which also has a small bandgap so that the supply voltage can be scaled further. High mobility, narrow bandgap III-V materials are hence being considered strongly as possible replacements for Si channels. Of the different III-V materials, the high In content InGaAs quantum well field effect transistor (QWFET) is being considered as one of the most promising device candidate for heterogeneous integration on a Si substrate $^{[74, 75]}$. The quantum well structure used to mitigate the band-to-band tunneling (BTBT) which is the dominant leakage mechanism for the narrow bandgap III-V materials results in a biaxial stress in the channel layer due to the lattice mismatch between the channel and the barrier layers. This necessitates the study of the piezoresistive behavior of III-V materials to understand the effect of biaxial stress on electron mobility. Even though these III-V based materials readily demonstrate high electron mobility, they are less promising for hole transport due to lower intrinsic hole mobility and disadvantageous when compared with strained-Si $^{[76]}$. Eventual integration of III-V based semiconductors as MOSFET channels requires both high performance n- and p-MOSTETs. This necessitates the use of performance boosters such as strain if shown to be promising especially for the p-MOS devices on III-V channels. The extreme complexity involved with the processing of InGaAs heterostructure MOSFETs $^{[75]}$ and difficulty in
estimating the channel stress in such III-V devices makes the direct investigation of stress effects very challenging. The conduction bands for GaAs, InAs, and InGaAs being very similar \textsuperscript{[76]} hints that the results of a similar investigation for GaAs surface channels which is relatively easier to form, can be extended qualitatively to InGaAs channels which are of technological importance. Hence in this work, we focus on studying the stress effects on GaAs bulk and MOSFET structures.

The first piezoresistance measurements on n-GaAs were done almost 50 years ago \textsuperscript{[77-79]} in an attempt to study the energy band structure and conduction band minima of GaAs and validate theoretical predictions of the same. There are no reports as yet on the piezoresistance of p-GaAs. Uniaxial stress measurements along different directions and on different orientation crystals of n-GaAs were performed and the piezoresistance coefficients were extracted. An apparatus similar to the one used in \textsuperscript{[1]} was used in these works. All these early works found that at low stress levels, the resistance changed linearly for bulk samples and the rate of change was small due to the relatively small change in the effective mass tensor of the electrons in the lowest spherical conduction band minima at \(k = 0\) (\(\Gamma\) - valley). At higher stresses, a non-linear behavior was observed which was attributed to the repopulation of the carriers from the low mass \(\Gamma\) - valley to the heavier mass L and X valleys. But all this understanding was for bulk GaAs samples. In this work, we develop a method to perform four point wafer bending experiments on small GaAs device samples and calibrate the stress using metal strain gauges. This method is first used to measure the piezoresistance of a bulk resistor structure, and the values measured are compared with previous work. The wafer bending technique is then used to study the electron and hole mobility enhancements in ring type GaAs MOSFETs which have surface confinement effects. The observed trends are explained considering the multi-band conduction band model or \(k\cdot p\)
calculations and a comprehensive understanding of the physics of mobility enhancement for both the GaAs bulk and MOSFET case is developed. Further validation of the model is given by some measurement data on gate leakage change, field dependence, and temperature dependence of mobility for the nMOSFET case.

5.2 Experimental Approach

The detailed device description and fabrication process of the GaAs devices used in this work is given in Ref [80]. Since the GaAs samples used are very small and brittle, they are first glued on to a Si substrate as shown in Figure 5-1A using a strong non-conductive epoxy. In order to ascertain the amount of stress transfer on to the GaAs sample using the flexure based wafer bending setup, a metal strain gauge is glued on a bare GaAs sample. One such stress calibration plot from this strain gauge measurement is shown in Figure 5-1B.

Figure 5-1. Stress calibration for GaAs piezoresistance measurements. A) Photograph of bare GaAs sample glued to a Si substrate along with strain gauge for stress calibration. B) Strain gauge stress calibration plot.

This calibration measurement was repeated whenever a Si substrate of different thickness was used or when a different epoxy was used to prepare the sample. Appropriate stress calibration is used for all future device measurements. As can be seen from the calibration plot in
Figure 5-1B, the GaAs sample develops cracks at ~100 MPa, hence all future device measurements were done for stresses below 100 MPa.

5.3 Results and Discussion

5.3.1 Bulk GaAs

GaAs resistors with a nominal resistance of 7.39kΩs were measured. The resistance change with applied wafer bending uniaxial stress is plotted in Figure 5-2. The extracted $\pi$-coefficient is $\sim 3 \times 10^{-11}\text{Pa}^{-1}$ which is in agreement with that reported in Ref [77]. To understand the origin of this stress dependence, we have to look at the GaAs conduction band which consists of three predominant valleys $\Gamma$, L and X, out of which most of the carriers populate the lowest mass $\Gamma$ – valley at low doping or inversion charge densities. For small stress levels, there are no repopulation effects because the energy gap between $\Gamma$ and L-valleys is 0.29eV and the mobility change is mainly due to the effective mass change of the $\Gamma$-valley. Under uniaxial stress, the mass change due to the shear component of the stress is relatively small and hence results in the small $\pi$-coefficient for the bulk case as seen from the resistor measurements. No p-GaAs samples were available to study the bulk piezoresistance of p-GaAs.

![Figure 5-2. Percentage resistance change plotted as a function of applied wafer bending stress on a bulk resistor device.](image)
5.3.2 GaAs nMOSFETs

Undoped GaAs surface channel ring type n-MOSFETs with a width of 800μm and lengths 5, 8, 10 and 20μm are used in this work. DC I-V and split-CV measurements as a function of stress and temperature are performed on these devices to extract the piezoresistance coefficients and temperature dependence of the mobility. Gate leakage is also monitored to study its change with applied mechanical stress.

It is known from past work [81, 82] that GaAs has two additional scattering mechanisms, namely polar-optic phonon scattering and piezoelectric scattering due to its intrinsic polar nature. At room temperature, polar-optic phonon scattering is the dominant scattering mechanism in the conduction band [81]. To verify this for our devices, the temperature dependence of the mobility was measured in a liquid-nitrogen cooled low temperature setup. The result of that measurement is plotted in Figure 5-3. We see that the mobility is almost invariant with temperature which is completely opposite to the trend expected. If polar-optic phonon scattering is indeed the dominant scattering mechanism, we should see strong temperature dependence. Also the magnitude of mobility is orders of magnitude smaller than what is expected for GaAs. These two facts suggest that surface roughness scattering and not polar-optic phonon scattering is the dominant scattering mechanism. This is a reasonable argument because, the technology for fabricating surface channel III-V MOSFETs is still in very early stages, and hence the gate dielectric interface is very poor. After determining the dominant scattering mechanism, DC I-V measurements are done as a function of applied wafer bending uniaxial and biaxial stress. The drive current enhancement with applied uniaxial stress is plotted in Figure 5-4. Also plotted is the drive current enhancement trend for planar Si MOSFETs as a reference (from Figure 3-1A). We see that for uniaxial tensile stresses less than 50 MPa, the enhancement is much larger than
that for Si, but this enhancement tends to saturate at relatively small stress levels of ~100 MPa.

For uniaxial compressive stresses, there seems to be a linear degradation of drive current.

Figure 5-3. Temperature dependence of mobility for Si and GaAs nMOSFETs is shown.

![Graph showing the temperature dependence of mobility for Si and GaAs nMOSFETs.]

Figure 5-4. Drive current enhancement of GaAs nMOSFET as a function of applied uniaxial stress. Si nMOSFET data is shown as a reference. In the inset are shown the Id-Vg curves before and after application of stress.

![Graph showing drive current enhancement as a function of uniaxial stress.]
Physical understanding of the origin of these trends is obtained by considering a multi-valley conduction band model and the deformation potential values are taken from Ref [76]. We know that the GaAs conduction band consists of three predominant valleys $\Gamma$, L and X, out of which most of the carriers populate the lowest mass $\Gamma$ – valley at low inversion charge densities or low vertical field. But at high vertical fields or a large inversion carrier density (1-1.5x$10^{13}$/cm$^2$) which is the region of interest, there is a significant population of carriers in the heavy mass L and X-valleys even in the unstressed state since the density of states in the $\Gamma$-valley is very small. When we apply an uniaxial tensile stress, the hydrostatic component of the strain increases the splitting between the $\Gamma$ and L-valleys, but the shear component of the strain splits the L-valleys into two groups, out of which the energy level of one group reduces and hence the effective band gap between the $\Gamma$ and L-valleys reduces with increasing stress as shown in the schematic diagram in Figure 5-5A. This results in repopulation of carriers from L to $\Gamma$ for low stresses and then back-population of carriers from $\Gamma$ to L with increasing stresses with an inflection point at intermediate stresses (~ 300 MPa) as predicted from some carrier population calculations for an inversion charge density of $10^{13}$cm$^{-2}$ shown in Figure 5-5B. Due to the large mass difference between $\Gamma$ and L-valleys, repopulation is the dominant factor for mobility change and hence we see the non-linear trend for drive current enhancement with uniaxial stress. But we see that the saturation of enhancement occurs at a lower stress in experiment than that is expected from calculation in Figure 5-5B. This is possibly because the current model calculations depend on deformation potential constants and band-gap values taken from literature. There is much uncertainty in the literature in these values which can cause the trends in the sub-band population calculations to have some uncertainty in the stress axis and hence cause the deviation.
Another validation of the above model can be obtained by looking at the gate leakage current change with applied wafer bending stress. Since the gate leakage current just depends on the barrier height and the out of plane mass of the carriers \cite{73}, and more importantly the barrier height due to the exponential dependence, the trend of the gate leakage change will directly depend on the carrier populations in each of the conduction band valleys. The measured gate leakage change with uniaxial stress is plotted in Figure 5-6. As expected, we see a non-linear trend with tensile stress. The leakage current reduces suggesting that the barrier height increases with repopulation of carriers from L to \Gamma. The degradation tends to saturate with increasing stress implying the beginning of back-population of carriers from \Gamma to L.
Figure 5-6. Gate leakage current change with uniaxial stress is shown.

Figure 5-7. GaAs nMOSFET channel mobility and mobility enhancement as a function of inversion carrier density.
In Figure 5-7 plots the field dependence of mobility and mobility enhancement for different magnitudes of uniaxial tensile stresses. We see that for small inversion carrier densities the enhancement saturates and degrades at even smaller stress levels (< 50 MPa). This is because at small inversion charge densities most carriers are already in the Γ-valley. Small stresses show some repopulation effects between L and Γ-valleys, but at slightly larger stresses we see very little enhancement similar to the bulk case as carriers are in the Γ-valley alone. For larger inversion carrier densities, we see the same non-linear trend as seen before.

Figure 5-8. Drive current enhancement of GaAs nMOSFET as a function of applied biaxial stress.

We now look at the biaxial stress case, the drive current enhancement trend is plotted in Figure 5-8. Since there is no shear splitting of the L-valley at low stresses, there is a linear repopulation of carriers from L to Γ as shown in Figure 5-9 A and B. Hence under biaxial tensile stress, we see a linear enhancement in the drive current due to carrier repopulation from L to Γ-valleys.
Figure 5-9. Theoretical estimation of the effect of biaxial stress on electrons in GaAs is shown. A) Schematic showing the GaAs conduction band shift and splitting under biaxial stress B) Calculation showing the predicted carrier population change in the $\Gamma$, L and X valleys under biaxial stress.

5.3.2 GaAs pMOSFETs

To gain a similar understanding of hole mobility enhancement with stress in GaAs, ring type pMOSFETs were used. It is known that the GaAs valence band structure is very similar to that of Si or Ge except for three main differences. Firstly, the difference in the masses of the heavy hole (HH) and light hole (LH) bands is larger than that in Si or Ge, which results in most carriers populating only the top most valence band. Secondly, the optical-phonon energy of GaAs is 35mev which is much smaller than the 61meV for Si and, since all the carriers are already in the top most valence band, there is negligible contribution from optical-phonon scattering rate change due to stress. Finally, GaAs being a polar material, has two additional scattering mechanisms namely, polar optic-phonon scattering and piezoelectric scattering. The effect of these additional scattering mechanisms on the change of hole mobility due to stress is relatively very small. Looking at the strain effects on the constant energy contours of the top-most valence band, which is obtained by performing $k\cdot p$ simulations, we see that the band
warping due to uniaxial compressive stress in GaAs is similar to that in Si from a qualitative perspective as shown in Figure 5-10. Even calculation of actual mobility values in Figure 5-11 shows that the enhancement trend in GaAs is also similar to the Si case for stresses less than 2 GPa. But measurements on GaAs p-MOSFETs show almost 2 times the enhancement with uniaxial compressive stresses as compared with Si p-MOSFETs (Figure 5-12). This is an encouraging result, for even though GaAs has poor hole mobility, it is very sensitive to stress and hence process induced uniaxial stress can be used to achieve higher hole mobility in GaAs p-MOSFETs. But the reason for the discrepancy between simulation and experiment is not yet understood where experiment shows a much larger enhancement with uniaxial compressive stress.

Figure 5-10. 2D-Constant Energy Contours of the topmost valence band in GaAs.
Figure 5-11. GaAs hole mobility calculation using k·p method as a function of uniaxial stress.

Figure 5-12. Drive current enhancement of (100) GaAs p-MOSFETs with uniaxial stress compared with (100) Si.
CHAPTER 6
STRESS EFFECTS ON TANOS FLASH MEMORY DEVICES

6.1 Introduction and Motivation

Non-volatile memories are becoming more and more indispensable with their wide applications in mobile, music and memory applications. The dominant NVM architecture is the floating-gate (FG) type which is also being aggressively scaled along with traditional transistors in logic applications. Charge trapping memory cells like SONOS devices are now recognized to be another candidate to replace FG cells in the sub-50nm regime to overcome the interference effects in the FG devices \(^{[83]}\). But to overcome the fundamental problem of a thin tunnel oxide in SONOS devices an alternative structure, TANOS (Si-SiO\(_2\)-SiN-Al\(_2\)O\(_3\)-TaN) has been reported \(^{[84]}\). This TANOS architecture with a high-k blocking layer shows promise for future memory applications. Applying mechanical stress has been largely successful for improving the performance of devices for logic applications. Stress in addition to improving the mobility of carriers, also alters the Si-SiO\(_2\) barrier height, the out-of-plane conductivity mass and electron trap energy levels. All these hint at the fact that stress could be useful to improve the program, erase and retention times of TANOS flash transistors.

It is known that programming and erasing of a TANOS Flash cell is by the tunneling of carriers from Si through the tunnel oxide into the nitride trapping layer. So the program and erase speeds are a direct function of the tunneling current through the oxide. It is known that direct tunneling of electrons and holes is a function of their out-of-plane effective mass and barrier height as shown in the schematic Figure 6-1. When a mechanical stress is applied, the barrier height of the Si-SiO\(_2\) barrier is changed due to the splitting of the degeneracy of the carriers and the magnitude of the tunneling current changes \(^{[73]}\). Hence by applying a favorable stress, it might be possible to improve the program and erase speeds of the TANOS Flash cells.
Figure 6-1. Schematic shows the Si-SiO2 barrier and the dominant parameters controlling the gate tunneling current.

6.2 Experiment and Results

To study this effect, TANOS flash transistors with the following stack (40Å SiO2, 70Å SiN, 120Å Al2O3, and 100Å TiN) 10µm channel width and 1µm channel length were used to do wafer bending experiments. A pulse generator was used to send pulses of different widths and program the device. The threshold voltage shift after a program pulse of a maximum width of 1 sec can be seen in the I_D-V_g plot in Figure 6-2. The plot of the threshold voltage shift with pulse widths for different applied mechanical stresses is shown in Figure 6-3. We can see that there is no change in the program speed of the TANOS Flash cell with both tensile and compressive stresses. This is expected because the pulse amplitude for programming these devices is very large. At these high voltages, the field across the gate and the field in the channel is very large and splits the bands in Si by a large amount. So any further band splitting due to mechanical
stress has a negligible impact, as all the carriers are already in the bottom band. Now looking at the erase operation, Figure 6-4 shows the threshold voltage shift with a 1 sec erase pulse. Figure 6-5 plots the threshold voltage shift with pulse width for different stress conditions. Again similar to the program operation, there is no change with stress.

Even though program and erase do not show a change with stress, it is expected that retention time should vary with stress because the stress alters the electron trap energy levels. It has already been shown on NAND Flash devices in a previous work that with tensile stress the trap energy increases resulting in an improved retention performance \(^{[85]}\). So a similar retention measurement was performed on these TANOS Flash devices in this work. Plotted in Figure 6-6 is the change in the threshold voltage with time for different stress conditions. We see that there is no noticeable change in the threshold voltage with stress. But the reason for this could be the very short duration of the experiment. Ideally the retention measurement should be performed at elevated temperatures (120-150 C) for a period of 24 hrs to have accelerated degradation and see a change with stress as in Ref [85].

![Id-Vg plot showing the change in the threshold voltage after application of a 1 sec program pulse.](image)
Figure 6-3. Change in threshold voltage with different program pulse widths and different applied mechanical stresses is shown.

Figure 6-4. Id-Vg plot showing the change of threshold voltage with application of a 1 sec erase pulse.
Figure 6-5. Change of threshold voltage with different erase pulse widths and different applied mechanical stresses is shown.

Figure 6-6. Change in threshold voltage with time under different applied mechanical stresses is shown. No change in retention time observed.
CHAPTER 7
SUMMARY AND FUTURE WORK

7.1 Summary

In this dissertation, the piezoresistive properties of planar Si MOSFETs, DG-FinFETs, SiGe heterostructure MOSFETs and GaAs MOSFETs were studied over a wide range of stresses using a novel flexure based four point wafer bending setup. First, the drawbacks of existing setups were identified bringing out the need for a novel setup capable of applying high stress. Then in Chapter 2, the design and functioning of this novel flexure based setup was described and capability of applying stresses greater than 1 GPa was demonstrated.

Once this capability was achieved, Chapter 3 discussed results of the high stress piezoresistance measurements on planar n and p-MOSFETs. The n-MOSFETs show a linear enhancement trend up to 1.5 GPa with no signs of saturation. Also both long short devices show similar trends after correcting for the external resistance in short devices. This trend suggests that strain induced surface roughness scattering reduction is another important mechanism for continued enhancement at high vertical fields when most of the repopulation and phonon scattering reduction has already taken place at low stress levels. P-MOSFETs show a super-linear trend for stresses greater than 1 GPa because of added enhancement from optical phonon scattering rate reduction. An analysis of stress effect on $I_{d\text{-sat}}$ was also done, and it was found that mobility is an important metric even in the saturation regime. The velocity saturation model when modified to take into account the velocity overshoot effect was found to better predict the saturation current change than the scattering based model, implying that the state-of-the-art MOSFETs are still limited by the physics at the drain-end of the channel as opposed to the source-end of the channel. DG-FinFETs measurements show that they have slightly larger
sensitivity to strain mainly due to the physical confinement effects coming into the picture which are not there for planar devices.

Chapter 4 showed comparison of k·p calculation results and piezoresistance measurements on SiGe heterostructure pMOSFETs. As expected from theory, the enhancement trend was found to be similar to that for Si as long as the thickness of the epi-layer is less than the critical thickness for epitaxial growth without defects. This implies that Ge QW channels with stress are promising for sub-22nm nodes. Chapter 5 described the stress additivity on GaAs devices. It is found that the bulk piezoresistance of n-GaAs is small and stems from the small mass change of the Γ-valley, while in the nMOSFETs the non-linear trend stems from the repopulation of carriers between Γ and L valleys. Temperature dependent mobility measurements revealed that in these devices surface roughness as opposed to polar-optic phonon scattering is the dominant scattering mechanism. GaAs pMOSFETs showed 2 times the enhancement with uniaxial compressive stress as compared to Si. This result could not be explained by the k·p calculations.

Finally in Chapter 6 the effect of stress on TANOS flash transistor program, erase and retention times was investigated. No significant effect was observed because the vertical field splitting during the program and erase pulses is so large that the strain splitting is insignificant.

7.2 Future Work

With stress engineering becoming an essential part of mainstream VLSI fabrication technology, there is much interest in studying its effects in novel devices. As part of the future work, it is interesting to continue the work on III-V devices and investigate the piezoresistive properties of InGaAs heterostructure MOSFETs which are of more interest. Also looking at the piezoresistive behavior of 1-D devices like Si, Ge nanowires or carbon-nanotubes is suggested. Continuing investigations on stress effects other novel-memory devices is also suggested.
Up till now, stress has only been used in device fabrication to improve device performance, but there is rising interest to see if stress can be used to improve unit processes during fabrication. It has been suggested whether stressing a wafer during fabrication processes like implantation, etching or oxidation is useful. So such a study can be performed by designing a novel wafer bending setup. So there are many exciting new things that can be done in the future in this field.
LIST OF REFERENCES


BIOGRAPHICAL SKETCH

Sagar Suthram was born on the 24\textsuperscript{th} of June 1982 in Secunderabad, India. He graduated with a B. Tech in Electronics & Electrical Communication Engineering from the Indian Institute of Technology, Kharagpur in May 2004. Since then he has been pursuing a PhD in Electrical & Computer Engineering at the University of Florida with Dr Scott E. Thompson. He has also done a two semester internship with the Non-planar CMOS Extension group at Sematech, Austin and worked on many projects on FinFETs, SiGe heterostructures, TANOS Flash Memory devices and III-V devices.