MOSFET Channel Engineering Using Strained Si and Strained Ge Grown on SiGe Virtual Substrates

by

Minjoo Lawrence Lee

Sc.B. with Honors in Materials Engineering
Brown University, 1998

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

Doctor of Philosophy in Electronic Materials

at the
Massachusetts Institute of Technology

June 2003

Signature of Author: _____
Department of Materials Science and Engineering
May 5, 2003

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

Accepted by: _____
Harry L. Tuller
Professor of Ceramics and Electronic Materials
Chair, Departmental Committee on Graduate Students

MOSFET Channel Engineering Using Strained Si and Strained Ge Grown on SiGe Virtual Substrates

by

Minjoo Lawrence Lee

Submitted to the Department of Materials Science and Engineering
on May 9, 2003 in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Electronic Materials

ABSTRACT

As the economic and technological benefits of scaling in very-large-scale-integrated (VLSI) circuits decreases, the use of alternative channel materials such as germanium and strained silicon (ε-Si) is increasingly being considered as a method for improving the performance of MOSFETs. While ε-Si grown on relaxed $Si_{1-x}Ge_x$ (i.e. single-channel heterostructure) is drawing closer to widespread commercialization, it is currently believed that almost all of the performance benefit in CMOS implementations will derive from the *n*-MOSFET. ε-Si *p*-MOSFETs demonstrate enhanced hole mobility, but the enhancement has been shown to degrade at high vertical fields for reasons that are still poorly understood. Dual-channel heterostructures, where a compressively-strained, Ge-rich layer is grown between the ε-Si cap and relaxed $Si_{1-x}Ge_x$ virtual substrate have been shown to offer much larger hole mobility enhancements. One of the primary goals of this thesis is to understand and improve the performance of both single- and dual-channel heterostructure *p*-MOSFETs. The approach taken was to grow novel heterostructures and then fabricate MOSFETs using a short process flow. Cross-sectional TEM was constantly employed as a way to connect microstructure with mobility characteristics. In this way, constant and rapid feedback between device results and the design of improved layer structures was achieved, and the map of available mobility enhancements in $Si_{1-x}Ge_x$-based heterostructures was greatly extended.

The step preceding all of the device work was to optimize the growth of highly strained layers via ultrahigh-vacuum chemical vapor deposition (UHVCVD). The deposition of highly strained layers in compression and tension creates problems that are not encountered in the growth of ε-Si on Si-rich $Si_{1-x}Ge_x$ virtual substrates. Through a combination of low-temperature growth and a novel two-step passivation-and-heating sequence, a wide variety of fully planar single- and dual-channel heterostructures can now be achieved in UHVCVD.

While ε-Si *p*-MOSFETs tend to lose much of their mobility enhancement at large vertical fields, previous work shows that the situation improves as x in the $Si_{1-x}Ge_x$ virtual substrate is increased to 0.5. The work presented here demonstrates that enhancements continue to improve for even higher Ge content. At $x = 0.7$, hole mobility enhancements

of 2.9 times were observed with no degradation at very large inversion densities (i.e. $>10^{13}$ cm^{-2}). Also, for the first time, a *p*-MOSFET with mobility enhancements that are independent of inversion density has been demonstrated through the use of a digital-alloy heterostructure. In general, it is shown that engineering the layer structure allows great control over the slope of hole mobility versus gate overdrive and that hole mobility enhancements that increase or remain constant with respect to inversion density can be attained.

While the first demonstration of high hole mobility in strained Ge (ε-Ge) was published nearly 10 years ago, little or no work on enhancement mode *p*-MOSFETs utilizing ε-Ge had been published prior to this thesis. In the work presented here, a thin, epitaxial Si capping layer both protects the ε-Ge from standard cleaning steps and serves as the interface with a standard SiO_2 gate dielectric. For the first time, the effects of all growth variables (including virtual substrate Ge content, ε-Ge thickness, and ε-Si thickness) on hole mobility are presented, and a *p*-MOSFET with mobility enhancement exceeding 10 times is demonstrated.

Having fabricated high mobility *p*-MOSFETs using Ge-rich single- and dual-channel heterostructures, the challenges of fabricating *n*-MOSFETs on the same wafers are next addressed. Experimental results presented here show that the highest electron mobility can only be attained when the wave function is completely confined in the ε-Si layer. Despite Ge's high electron mobility, simultaneous electron transport through both ε-Si and ε-Ge is shown to lead to depressed *n*-channel mobilities. Rampant intervalley phonon scattering from Ge's *L*-valley conduction band minima to Si's *X*-valley minima is stipulated as the mobility-limiting mechanism, and experimental evidence supporting this hypothesis is presented. Taking the best device results together, it is shown that symmetric-mobility *n*- and *p*-MOSFETs can be realized on a single heterostructure grown upon $Si_{0.5}Ge_{0.5}$.

Thesis Supervisor: Eugene A. Fitzgerald
Title: Professor of Materials Science and Engineering

Table of Contents

List of Figures

Acknowledgements

There are a lot of people whom I need to acknowledge here, but the first has to be my advisor, Gene Fitzgerald. He handed me a really interesting project and simply let me go with it. He was always patient with the setbacks that I encountered and enthusiastic with my successes. I don't think that we'll soon forget some of the "eureka" moments that we had while discussing ideas in front of his whiteboard.

Dimitri Antoniadis has always been a supportive force for me over in Course 6. He is one of those rare people who can be an expert in his field, yet never treat the most basic questions condescendingly. I am very thankful that he has always taken time to look over my device results and that he made the resources of his measurement lab available to me. I've been honored to have Professor Sam Allen on my committee. He has always been friendly and encouraging during my years here.

Every member of the Fitzgerald group has been a huge part of my time here... *Steve Ting* was an early member who graduated soon after I arrived, but he never hesitated to take time to teach me about TEM or anything else. I remember going to eat at Goosebeary's food truck with *Mayank Bulsara* way back in the day. Mayank may sometimes be on a mental plane of his own, but he is still a very down-to-earth guy. I'm looking forward to our next Hell Night adventure. I didn't have much time to get to know *Andy Kim*, but to this day, I still use the notes I took way back in '99 when he taught me double- and triple-axis XRD. *Matt Currie* was one of my initiators into the world of SiGe. He's a great writer and scientist but never takes himself so seriously that he can't enjoy a good practical joke. *Tom Langdo* is never satisfied with a research problem until all angles have been thoroughly attacked. His natural skepticism combined with virtuosic technical ability are tremendous assets. I can recount dozens of hours spent with him training me on sample prep and imaging those pesky Ge undulations. I won't forget that he really took the time and energy to work with me and train me as a first year. The next guy out of the pool was *Chris Leitz*. His focus and discipline are legendary around here. He really taught me everything about MOSFET processing and UHVCVD maintenance. His thesis was the model that I based much of my work on, and his excellent writing style (and clear lab-books) made it easy for me to pick up what he'd done. How he graduated in under 4 years, managed a half dozen collaborations, and the reactor itself, all while running 40+ miles per week will probably never be known. *Vicky Yang's* patience and ability to take everything in stride in this often frustrating environment were amazing. I wish her and Matt years of health and happiness together! *Mike Groenert* was a good friend around here, and his research accomplishments are quite an inspiration to us all. He always had his own take on things, and I really enjoyed all of our conversations here in the computer room. Mike is an amazingly resourceful individual, and even though he's been in Germany for a while now, all anyone talks about are his "lasers." I've also enjoyed getting to know *Gianni Taraschi*. He is an excellent writer and speaker and a consummate scientist. He was always enthusiastic about discussing any problem of scientific (or economic) interest. *Lisa McGill* has been extremely helpful in editing my papers and providing delicious sweets to the group. She has such a warm personality that I sometimes forget that she is, in fact, a mighty blackbelt. *Arthur Pitera* has been a close friend and ally for years now. We spent countless hours planning growths, fixing the reactor, and discussing our research together, in addition to all the time passed going out for coffee and lunch. He's got a no-nonsense approach to solving just about any problem and was one of the few people who would get sushi with me. *Nate Quitoriano* has always had a positive outlook on things, even when the upgraded reactor deposited pure Arsenic on one side of the tube and a puddle of Gallium on the other. It's been really fun working alongside him and having another music minor in the group. I've enjoyed discussing everything from microbatteries to world politics with *Nava Ariel*. She's now one of the senior students in the group, and I'm sure she'll be successful at being a pioneer in her

nascent field. *Juwell Wu* is another person who is forging her own path in the group. Her creativity and good spirits will serve her well as she finds new applications for visible lasers. It's been a pleasure getting to train *Saurabh Gupta* on everything from TEM to MOS measurements. I'm still thankful that he took the time to look after my apartment and cat over Christmas break, and I look forward to being his friend and collaborator for years to come. *Dave Isaacson* and I only intersected briefly, but I thank him for taking the time to proofread a lot of my thesis. Lastly, I really enjoyed watching *Carl Dohrman* and *Mike Mori* dig into the upgrades to the MOCVD during IAP, and I wish them the best of luck.

Thank you to the many people who made working around here possible: Vicky Diadiuk and the MTL staff, Mike Frongillo, and Isaac Lauer all maintained facilities that were critical to me. Thank you also to the many great friends that I made over in Course 6: Jongwan Jung's expertise in processing was tremendously helpful, Tonya Drake really helped make those last three transistor runs more fun, Charles Cheng helped me out a lot with measurements, Andy Ritenour has been a cool guy to talk to and never complained about having to do TEM sample prep or CMP, Scott Yu taught me a lot about MOS-physics, and Professor Judy Hoyt also provided lots of friendly advice. It's really been a privilege to get to be a collaborator with all of these talented people. Thanks are also due to Bill Cutter, for being a teacher and a friend over the years.

I have really loved spending hundreds of hours driving in the van, carrying equipment, messing up takes, and playing for crowded and empty rooms with the guys in Words for Snow. Thanks also to the many other amazing musicians around town who I've met over the years. Shout-outs to all my friends from high school and college who helped to keep me sane either by visiting or just calling: Joel, Ed, Pete, Jon, Scott, Mindy, Catherine, Dave, Steve, and Mike.

Kathy Kew has been with me from the day I applied to MIT through the morning of my final defense. It wouldn't be possible for me to describe all the ways that she has helped and supported me for these past five years, and I'm looking forward to many more years together. She read every single word of this thesis, for which I cannot thank her enough. And even though I am not yet officially a member of the family, the Kews have always treated me like one of their own. Thank you all so much for everything.

My parents have always supported me and taught me the value of truly committing to things and giving my best. This thesis is a testament to their love for me and constant commitment to my always getting the best education possible. Ever since mom and dad moved back to Korea, I really feel I've gotten even closer with my two big brothers (and my new sister in-law!). Whether skiing, eating live sea animals, or cooking Thanksgiving dinner, time with them has always been exciting, and I hope that HJ and TJ know that they're two of my best friends in the world.

Finally, I would like to dedicate this thesis to all four of my grandparents, particularly to the memory of my two grandfathers who passed away before they could see this day. I hope that this thesis does honor to them both. Both of my halmuni's were two of my biggest fans throughout my time at MIT, and I'm so happy to finally get to share this occasion with them.

Chapter 1. Introduction

1.1. Motivation

The practice of scaling in microelectronics has allowed constant improvements in the cost, performance, and functionality of Si-based integrated circuits. However, in the current era of multi-billion dollar Si fabrication facilities and deep sub-micron metal-oxide-semiconductor field-effect transistors (MOSFETs), the practical benefit of scaling is declining as physical and economic limits are approached. The incorporation of new materials, such as Cu and silicon-on-insulator wafers (SOI), is emerging as an important way to continue to improve the performance of very large scale integrated (VLSI) circuits.[1] Another sign of industry's willingness to utilize new materials is the growing prominence of strained silicon (ε-Si). Just over ten years after the first laboratory demonstration of high mobility ε-Si grown on relaxed $Si_{1-x}Ge_x$ from Bell Labs, all of the major manufacturers of high-performance microprocessors have publicly expressed their intention of utilizing strained Si in future technology nodes.[2] Currently, it is widely believed that much of the benefit of incorporating strained Si will originate from the improved mobility of the *n*-MOSFET,[3] with little or no benefit from the *p*-MOSFET. While strained Si *p*-MOSFETs with enhanced mobility have been demonstrated by numerous research groups,[4-7] the enhancement has been shown to degrade at high vertical fields for reasons that are still poorly understood.[8]

1.2. Scope and Organization of Thesis

One of the primary goals of this thesis is to understand and improve the performance of $Si_{1-x}Ge_x$-based *p*-MOSFETs. The approach taken was to grow and process simple MOSFETs in rapid succession, allowing constant feedback between

device results and the design of new heterostructures. Chapters 2 and 3 review key aspects of $Si_{1-x}Ge_x$ epitaxial growth and the physics of $Si_{1-x}Ge_x$-based MOSFETs, and Chapter 4 describes the problems encountered and solutions devised for depositing compressive and tensile films on relaxed $Si_{1-x}Ge_x$. In Chapter 5, wave function penetration into the oxide is examined as a possible mechanism responsible for the rapid drop in hole mobility observed in ε-Si *p*-MOSFETs at large inversion densities. The use of highly-strained Si films on Ge-rich $Si_{1-x}Ge_x$ is proposed as a method for preserving the enhancement at high field, and experimental results proving the validity of this approach are presented. A novel digital-alloy heterostructure *p*-MOSFET exhibiting mobility enhancements that are independent of inversion strength is also demonstrated. Chapter 6 investigates the use of a compressively strained Ge (ε-Ge) layer just below the ε-Si cap (dual-channel heterostructure) for attaining even greater hole mobility enhancements. All channel engineering variables are investigated experimentally, and a *p*-MOSFET with mobility enhancement exceeding 10 times is demonstrated. Having fabricated high mobility *p*-MOSFETs using ε-Si / ε-Ge dual-channel heterostructures, Chapter 7 addresses the integration of *n*-MOSFETs on the same wafers. Experimental results show that the highest electron mobility can only be attained when the wave function is completely confined in the ε-Si layer. The consequences of electron transport through both the ε-Si and ε-Ge are examined. Taking the results of Chapter 6 and 7 together, it is shown that symmetric-mobility *n*- and *p*-MOSFETs can be realized on a single heterostructure.

Chapter 2. SiGe virtual substrates - Growth and Applications

In this chapter a motivation for achieving fully-relaxed, low defect density $Si_{1-x}Ge_x$ alloys on bulk Si wafers is presented. The theory and practice of growing $Si_{1-x}Ge_x$ graded buffers is reviewed, along with a description of the growth system that was used in this research. The main characterization techniques which were used to study the physical and microstructural properties of $Si_{1-x}Ge_x$ films are outlined, and finally, the installation and qualification of a disilane source in the Fitzgerald Group reactor is described.

2.1. Introduction to the SiGe Integration Platform

Since silicon and germanium form a completely miscible alloy system (Figure 2.1), $Si_{1-x}Ge_x$ films of any composition can be grown with a lattice constant that varies linearly between 5.43 Å - 5.66 Å according to Vegard's law:[9]

(2.1) $$a_{SiGe} = a_{Si}(1-x) + a_{Ge}(x)$$

The ability to access such a wide range of lattice constants without the possibility of phase segregation makes $Si_{1-x}Ge_x$ alloys extremely useful. For example, $Si_{1-x}Ge_x$ can be used as a template for depositing Si-rich layers in a state of biaxial tensile strain and Ge-rich layers in a state of biaxial compressive strain. Applied strain causes changes in the band structure that can be exploited in a variety of electronic devices, and discontinuities in the band alignment allow the formation of quantum wells (high-performance $Si_{1-x}Ge_x$ heterostructure devices are reviewed in Chapter 3). However the $Si_{1-x}Ge_x$ substrates needed to enable bandgap-engineered $Si_{1-x}Ge_x$ heterostructures cannot be grown as bulk single crystals; Czochralski-grown $Si_{1-x}Ge_x$ single crystals particularly suffer from non-uniform Ge-content.[9]

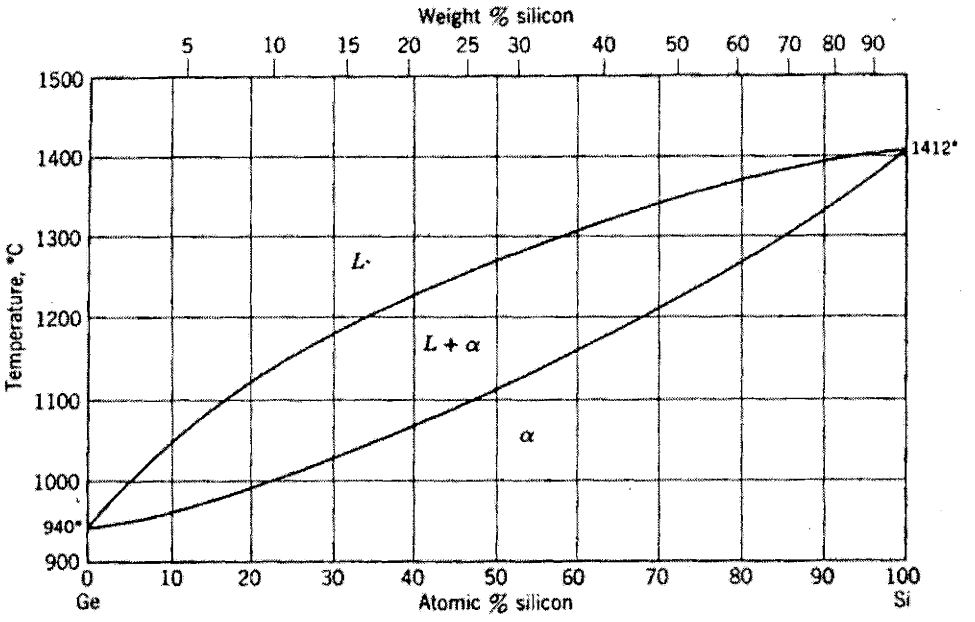
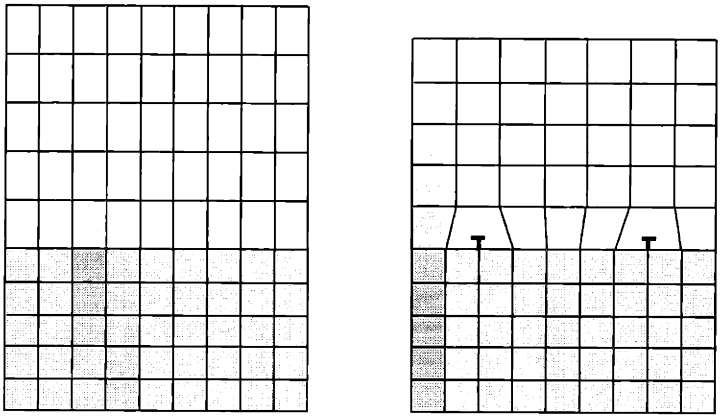


Figure 2.1 The Ge-Si binary phase diagram (adapted from Gandhi[10])

2.2. Mismatched heteroepitaxy and SiGe virtual substrates

Given the difficulty in creating $Si_{1-x}Ge_x$ single crystal substrates, the most feasible method for accessing the wide range of $Si_{1-x}Ge_x$ lattice constants is through epitaxial growth of $Si_{1-x}Ge_x$ buffers on Si wafers. In this sense, the wafer can be transformed into a $Si_{1-x}Ge_x$ "virtual substrate"; while the bulk of the wafer is Si, the lattice constant of the top surface is that of a relaxed $Si_{1-x}Ge_x$ alloy.[11] Of course, another reason to start with bulk Si is that the most mature processing technology and tools in existence are designed to accommodate large diameter Si substrates. Therefore, semiconductor materials that are heterogeneously integrated onto Si wafers maintain basic compatibility with state-of-the-art, multibillion-dollar Si processing infrastructure. In the initial stages of mismatched growth, the in-plane lattice constant of the SiGe epitaxial layer is constrained by the Si substrate, resulting in the accumulation of strain in the growing film [Figure

2.2(a)]. This stage of mismatched epitaxy is known as 'pseudomorphic growth' and is characterized by a fully coherent interface between film and substrate.[12]



(a) (b)

Figure 2.2 Schematic diagram of (a) pseudomorphic growth and (b) dislocation introduction to relieve lattice strain

At a certain critical thickness, it becomes energetically favorable for dislocations to form at the film-substrate interface in order to accommodate the lattice strain in the epitaxial layer [Figure 2.2(b)]. The critical thickness h_c is given by:

$$h_c = \frac{D(1-\nu\cos^2\alpha)(b/b_{eff})\left[\ln\left(h_c/b\right)+1\right]}{2Yf} \tag{2.2}$$

where D is the average shear modulus for an interfacial dislocation, ν is the Poisson ratio, α is the angle between the dislocation line and the Burgers vector b, b_{eff} is the component of the Burgers vector resolved to the interface, Y is the Young's modulus of the film, and f is the film-substrate mismatch; in low-mismatched heteroepitaxy of diamond cubic semiconductor materials, $\alpha = 60°$ and $b_{eff} = a/2$ for the majority of dislocations.[13] Figure 2.3 plots the critical thickness of $Si_{1-x}Ge_x$ grown on Si as a function of Ge content.[14] Note that Equation 2.2 is an equilibrium equation, and that fully strained $Si_{1-x}Ge_x$ films can be grown well past their critical thickness if very low growth temperatures are used. Such

films are considered to be metastable, because subsequent high temperature steps will cause dislocation nucleation and relaxation.

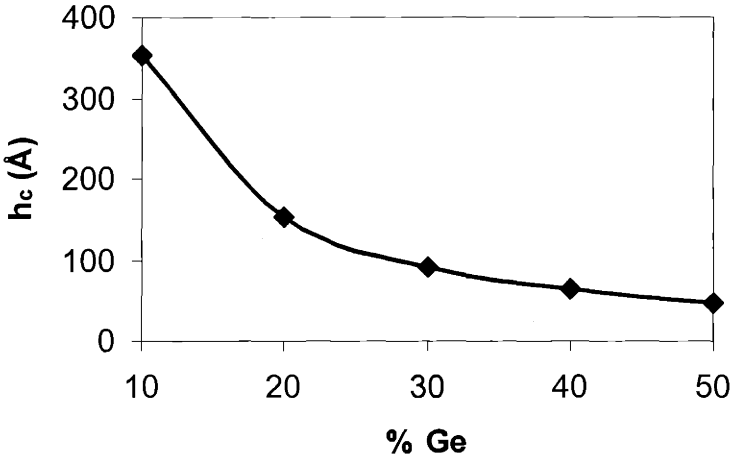


Figure 2.3 Critical thickness (h_c) vs Ge content for $Si_{1-x}Ge_x$ grown on Si (001). For high Ge content films, the critical thickness becomes very low.

The misfit dislocations that form at the film-substrate interface to relieve the lattice strain cannot terminate within the crystal. Since extremely long line lengths would be needed to cause all of the misfits to glide to the edge of the wafer, the majority of misfit dislocations thread upwards and terminate at the film surface.[12] The dislocation morphology of highly mismatched (i.e. $f > 1.5\%$) epilayers is qualitatively very different from that of low mismatched epilayers. In high-mismatch heteroepitaxy, the misfit dislocations are closely spaced and have short line lengths, resulting in high densities of threading dislocations; for fully-relaxed epitaxial GaAs on Si ($f = 4\%$), the threading dislocation density is typically 10^8 to 10^9 cm^{-2}. The lifetime and reliability of optoelectronic devices are severely limited when they are fabricated from such highly defective starting material.[15] The defect morphology in low-mismatch heteroepitaxy is characterized by low dislocation nucleation rates, long misfit line-lengths, and relatively

few threading dislocations intersecting the growth surface. In theory, both small and large amounts of film-substrate mismatch can be accommodated by forming numerous low-mismatch heterointerfaces. As practiced, this strategy is known as compositional grading.[11]

Growth of low defect density, fully-relaxed SiGe graded buffers relies upon two intertwined factors: minimizing dislocation nucleation rates and maximizing dislocation glide velocity. The dependence of final threading dislocation density on various growth parameters can be summarized as follows:[16]

$$\rho = \frac{2R_g R_{gr} \exp\left(\frac{E_{glide}}{kT}\right)}{bBY^m \varepsilon_{eff}^m} \tag{2.3}$$

where R_g is the growth rate, R_{gr} is strain grading rate, B is a constant, Y is the Young's modulus of the film, and ε_{eff} is the effective strain driving dislocation glide. In practice, the requirement for low dislocation nucleation rates is met by slowly increasing the Ge content over some thickness of film, typically on the order of microns. Dislocation glide is thermally activated, and thus the highest velocities and lowest defect densities are attained by growing at high temperatures. Furthermore, once threads are nucleated, subsequent layers can be relaxed by the glide of pre-existing threads (Figure 2.4). Since the strain in the growing film is applied slowly and the velocity of dislocations is high, a low-mismatch defect morphology can, in theory, be maintained throughout growth, regardless of final Ge content.[11]

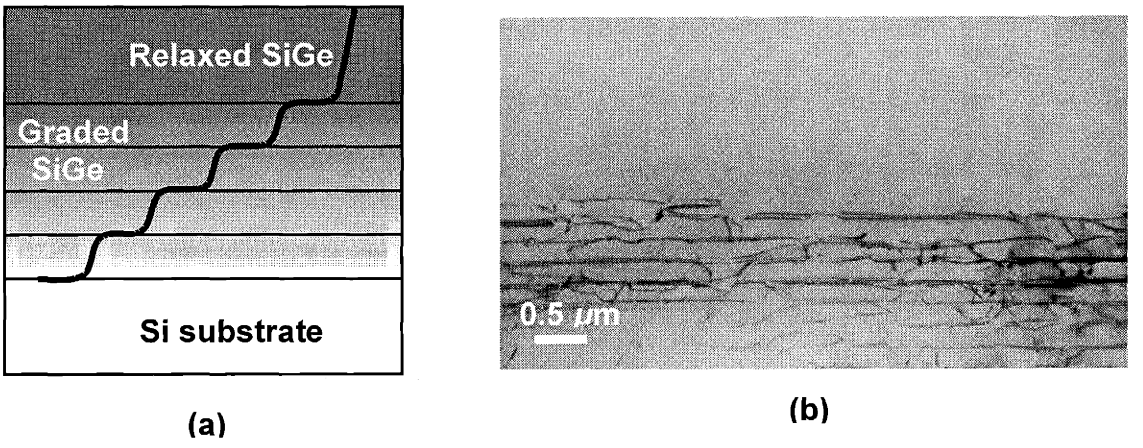


Figure 2.4 Schematic diagram and XTEM of step-graded buffer defect morphology. Each interface is grown at a low strain rate, allowing the nucleation rate of dislocations to be minimized. Pre-existing threads can be used to relax the strain in graded layers, preventing new dislocations from nucleating.

However, Figure 2.5 shows that dislocation densities in graded buffers grown at the same temperature (750°C) and grading rate (10% Ge/μm) vary by over an order of magnitude depending on final Ge content.[16-18] In fact, for pure Ge, the threading dislocation density climbs to nearly 10^8 cm^{-2}, making it unsuitable for subsequent optoelectronic integration. The deviation between experiment and theory is caused by the fact that Equation 2.3 fails to take kinetic barriers to dislocation glide into consideration. In SiGe epitaxy, the primary impediment to the glide of threads is dislocation pileups.[18] Dislocation pileups are believed to form as a result of the crosshatch surface roughness that naturally arises in mismatched epitaxy. Since threads trapped in a pileup can no longer contribute to strain relief, new dislocations must nucleate in order to relax successive graded layers, and the overall defect density in the final film increases.

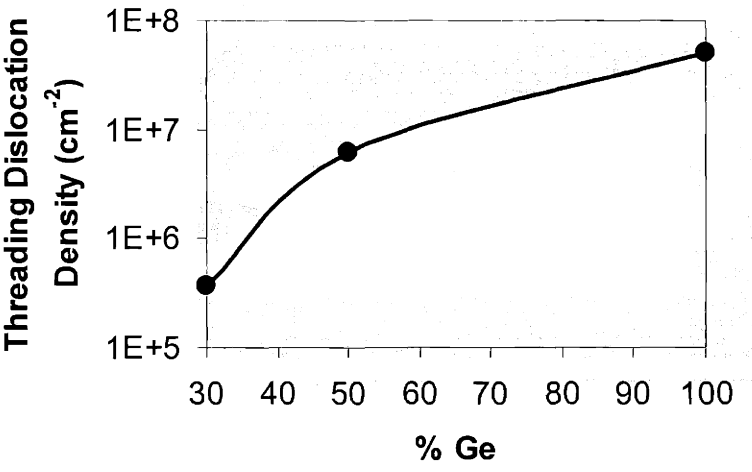


Figure 2.5 Defect density of $Si_{1-x}Ge_x$ virtual substrates grown at a grading rate of 10% Ge/μm at 750°C. Kinetic barriers to dislocation glide cause a direct dependence of threading dislocation density on Ge content.

Currie *et al.* have found that a chemical-mechanical planarization step (CMP) can be used to remove crosshatch surface roughness in $Si_{1-x}Ge_x$ graded buffers and greatly reduce the density of dislocation pileups.[18] Upon further compositional grading on the polished surfaces, total dislocation densities were actually found to decrease, because threads, once trapped in pileups, were free to glide and undergo annihilation events. Incorporating intermediate CMP steps into the compositional grading process has allowed growth of SiGe graded buffers of any composition, including pure Ge, with threading dislocation densities lower than 2×10^6 cm^{-2}. GaAs deposited onto these "virtual Ge" substrates displays minority carrier lifetimes close to those of bulk GaAs, suggesting that such low threading dislocation densities have little effect on minority carriers. The most powerful proof of the efficacy of the graded buffer technique has been the demonstration of working AlGaAs/GaAs lasers on virtual Ge wafers.[19] The successful integration of high quality III-V material[20] on Si wafers holds great promise for future circuit technologies where optical devices may be monolithically-integrated[21] in high densities with the most advanced VLSI logic technology.

2.3. Ultrahigh-vacuum chemical-vapor-deposition

While low-defect density $Si_{1-x}Ge_x$ graded buffers have been grown by a variety of methods, including molecular beam epitaxy [22](MBE) and low-pressure rapid thermal chemical vapor deposition (LPRTCVD),[11] all of the material for this thesis was grown via ultrahigh-vacuum chemical vapor deposition (UHVCVD). The Fitzgerald group UHVCVD reactor is a hot-walled, load-locked system equipped with a quartz growth tube (Figure 2.6). Both the growth chamber and loadlock are evacuated by turbomolecular pumps that are, in turn, backed by a roots-trivac combination pump. The seal between the quartz tube and the stainless steel chamber is accomplished by concentric o-rings. The space between the o-rings is pumped by a separate, smaller turbomolecular pump so that each o-ring supports a vacuum of about 10^{-6} Torr. This scheme, called differential o-ring pumping, allows base pressures of 1×10^{-9} Torr to be attained at 900°C. Growth rates of around 10Å $second^{-1}$ are typical for temperatures from 800°C to 900°C and up to 10 × 6" wafers can be loaded simultaneously. A manual gate valve can be used to vary the growth pressure in the chamber from 5 to 25 mTorr.

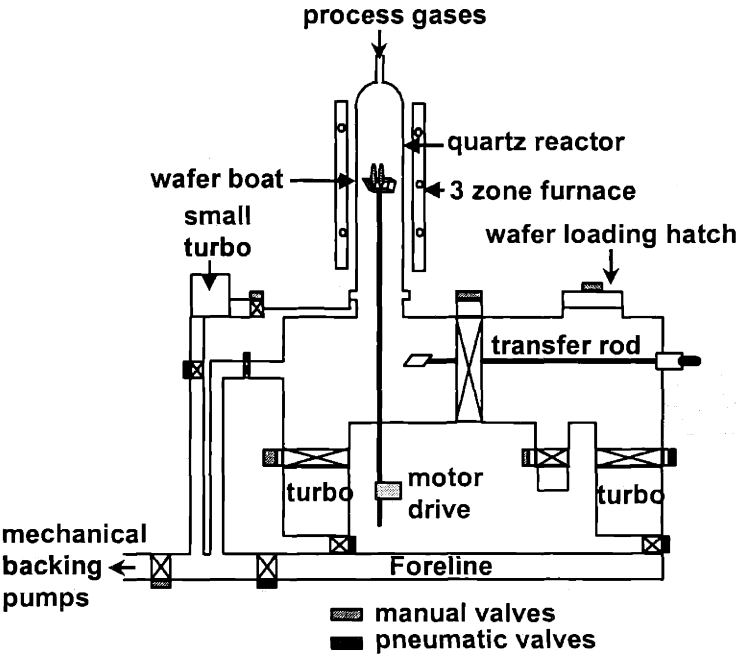


Figure 2.6 Schematic of the Fitzgerald Group UHVCVD system

Prior to growth, wafers are cleaned by a two-step procedure consisting of a 10 minute clean in 3:1 H_2SO_4:H_2O_2 (piranha clean) followed by deionized water (DI) rinsing and a 1 minute dip in 50:1 H_2O:HF. The dilute hydrofluoric acid step serves to strip the chemical oxide that grows during piranha cleaning as well as terminating the surface with H atoms to prevent formation of native oxide in atmosphere. Wafers are typically pumped down in the loadlock chamber to a pressure of $\sim 1 \times 10^{-7}$ Torr before introduction into the growth chamber. The stepper-controlled pedestal shown in Figure 2.6 is then used to raise the wafers into an area just below the furnace. This step is intended to serve as a low-temperature desorption to encourage volatilization of water molecules and other atmospheric contaminants. Finally, just before introduction of source gases, the wafers are held at 900°C for several minutes in order to volatilize any native oxide that may have formed after cleaning.

While the growth of microns-thick graded buffers is typically carried out at high temperatures, much lower growth temperatures are needed to grow the nanometer-scale heterostructures studied in this thesis. In conventional atmospheric pressure Si homoepitaxy, very high growth temperatures (1000-1100°C) and growth rates are used to prevent formation of native oxide on the surface.[10] In high vacuum techniques such as MBE and UHVCVD, considerably lower growth rates and growth temperatures can be used for epitaxial growth, because the background levels of impurities are orders of magnitude lower.[23] The ability of the Fitzgerald Group UHVCVD to deposit high quality films at extremely low growth temperatures has been indispensable to this research, since low temperatures help to stabilize surface morphology in strained films. The technique of growing highly strained quantum wells is described in detail in Chapter 4.

2.4. Materials Characterization Techniques

While the reactor has been calibrated for the growth of relaxed $Si_{1-x}Ge_x$ buffers spanning the entire composition range, little previous work required growth temperatures below 550°C. Since considerably lower growth temperatures are needed for the growth of fully planar strained layers investigated in this thesis, the reactor had to be calibrated for growth at temperatures as low as 350°C. The composition and thickness of strained layers are key variables in determining the mobility characteristics of heterostructure MOSFETs. Therefore, thorough materials characterization is a requirement for understanding the effect of growth parameters on carrier transport.

2.4.1. Transmission Electron Microscopy

In transmission electron microscopy (TEM), samples are manually ground and polished to a thickness of approximately 10 μm and then Ar-ion milled to render them electron-transparent. The usefulness of TEM lies in its ability to directly image crystalline structure at very high magnifications and resolution. Magnifications on the order of 10,000 times are ideal for imaging micrometer-scale features, such as graded buffers and MOSFET gate stacks. For strained layers, dimensions of approximately 1-10 nm are typical, requiring imaging at magnifications on the order of 10^5 times. Since biaxial strain causes a disruption to the periodicity of the diamond cubic lattice, dark/light contrast can be seen between layers under different strain states if proper tilt conditions are used. For example, tensile Si grown on relaxed (100) $Si_{1-x}Ge_x$ grows in a tetragonal form where the *c*-axis is shorter than the in-plane axes and therefore does not possess perfect cubic symmetry. As a result, tilting the sample to cause preferential excitation of a plane lying in the <001> direction causes the tensile Si layer to exhibit slightly enhanced brightness compared to the relaxed $Si_{1-x}Ge_x$. Compressively strained layers likewise appear dark.

2.4.2. X-ray Diffraction

Triple- and double-axis X-ray diffraction (XRD) are techniques for determining both strain state and composition in epitaxial films. The generated x-ray beam is collimated through slits and a Ge-crystal (for triple-axis scans) in order to produce a highly focused, monochromatic beam; the use of the Ge-crystal decreases the signal intensity, but greatly increases resolution. Figure 2.7 is a $\theta/2\theta$ scan taken on a fully

relaxed $Si_{1-x}Ge_x$ graded buffer. Given the intensity of the x-ray generator employed, layer thicknesses of around 200-300 nm are typically required in order to create suitable diffraction intensity. The strain in MOSFET channel layers is directly determined by the Ge content of the relaxed buffer. Therefore, accurate knowledge of the lattice constant of the virtual substrate is a necessary precondition to understanding the impact of strain on MOSFET characteristics.

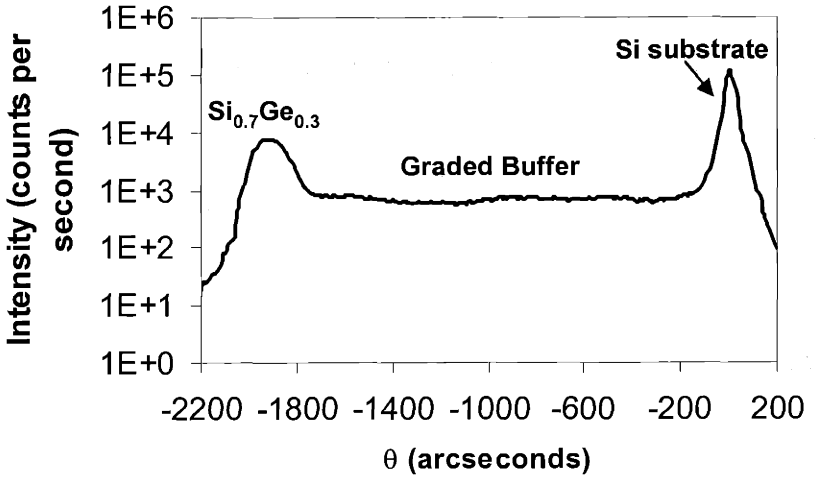


Figure 2.7 (004) $\theta/2\theta$ double-axis rocking curve of a $Si_{1-x}Ge_x$ graded buffer with a final composition of $x = 0.3$

2.4.3. Secondary Ion Mass Spectroscopy

In secondary ion mass spectroscopy (SIMS), the surface of a sample is slowly milled while simultaneously collecting the resulting ion flux. Extremely low growth temperatures and growth rates were necessary in many cases in order to grow the heterostructures studied in this thesis. With a depression in growth rate comes increased risk of residual impurity contamination. However, SIMS done on samples grown at temperatures as low as 450°C revealed no increase in impurity incorporation. The ability to grow extremely high purity films at growth rates as low as 1×10^{-3} Å sec^{-1} is testament

to the exceptionally clean background in the Fitzgerald group UHVCVD. Information on dopant incorporation was also routinely obtained using SIMS.

2.5. Disilane Calibration and Qualification

Previous work in the Fitzgerald group UHVCVD focused primarily on silane (SiH_4) and germane (GeH_4) source gases for SiGe epitaxy. In order to expand the low temperature capabilities of the reactor, a disilane (Si_2H_6) tank was installed. The lower activation energy of Si_2H_6 compared to SiH_4 makes it useful in industry as a source gas for low temperature growth of amorphous-Si films. In the reactor used here, it was found that Si growth with disilane at temperatures below 650°C could be fit to an activation energy of 2.18 eV, while silane's activation energy was found to be 2.34 eV. As Figure 2.8 shows, growth rates of pure Si with disilane were increased by a factor of ~3-4 over silane at low temperatures. Note that the activation energy for GeH_4 decomposition is much lower, allowing relatively high growth rates for pure Ge even at very low temperatures.

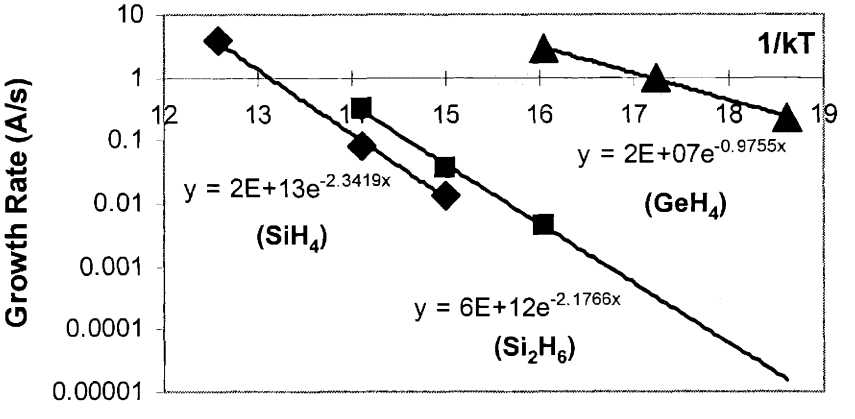


Figure 2.8 Arrhenius Plot of growth rate vs temperature for source gases employed in the Fitzgerald Group UHVCVD

Up to this point, all source gases used in the Fitzgerald group UHVCVD were passed through a PALL brand GasKleen purifier in order to reduce moisture contamination. However, PALL corporation has stated that their standard hydride purifiers, which rely upon reactive metal resins to remove moisture, are not compatible with disilane. Therefore, Si_2H_6 had to be piped into the reactor with no in-line purification or filtration. Since Si_2H_6 in research quantities is not available in purity levels as high as that which is available for SiH_4 and GeH_4, there was concern that films grown with Si_2H_6 would suffer from increased contamination levels. In order to test the purity of the disilane source, we deposited a 0.5 μm Si film on a $Si_{0.7}Ge_{0.3}$ relaxed buffer that had been grown with our standard silane and germane sources. SIMS analysis revealed no significant increase in incorporation of oxygen. A significant increase in nitrogen and carbon content was observed, but both were maintained at levels below 1×10^{17} cm^{-3}.

As a final test, we grew a single-channel ε-Si heterostructure on $Si_{0.8}Ge_{0.2}$ using Si_2H_6 only for the channel region and processed MOSFETs. Figure 2.9 shows that the extracted electron mobility for the disilane-grown device was nearly identical to a co-processed ε-Si *n*-MOSFET grown with silane. The mechanism for mobility enhancement in ε-Si *n*-MOSFETs is explained in Chapter 3.

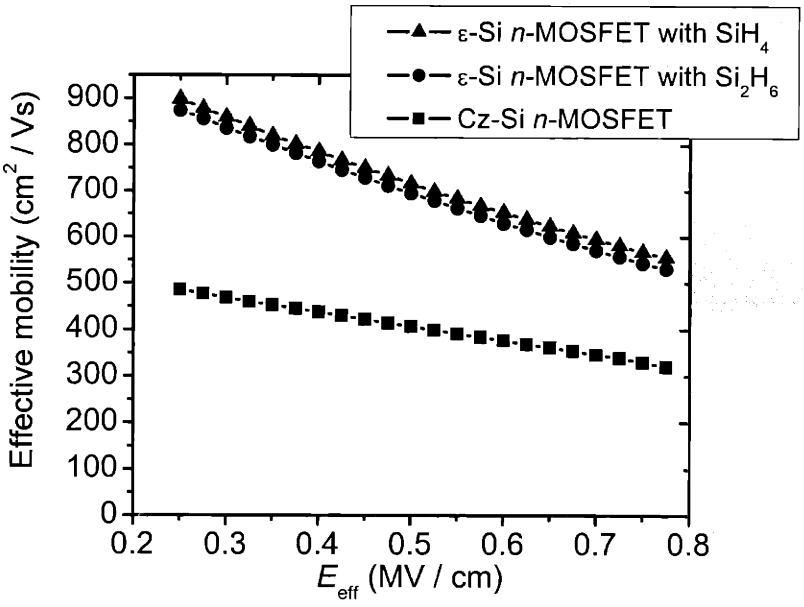


Figure 2.9 Effective electron mobility versus E_{eff} for strained Si *n*-MOSFETs grown with SiH_4 and Si_2H_6.

Chapter 3. $Si_{1-x}Ge_x$-based field-effect transistors

The past twenty years have brought huge improvements in the performance of Si-based very large-scale integrated (VLSI) circuits. This chapter reviews the fundamentals of MOSFET operation and also describes how scaling has been used to improve MOSFET performance. A motivation for the use of $Si_{1-x}Ge_x$-based heterostructures to improve MOSFET performance is presented, and several characteristic device designs and their associated performance benefits are described.

3.1. Review of MOSFET fundamentals

Viewed from the outside, a MOSFET is a 3-terminal semiconductor device consisting of electrodes for the source, drain, and gate (Figure 3.1).[24] Physically, the source and drain are heavily-doped conductive regions near the semiconductor surface that are connected to the outside world by ohmic metal contacts and separated from each other by the channel region. The gate, on the other hand, is a capacitive element comprised of a dielectric layer sandwiched between a metal contact and the semiconductor substrate. In present-day Si CMOS technology, the gate dielectric is typically formed by thermally oxidizing the substrate wafer, and the gate metal is heavily doped polycrystalline silicon (polysilicon). The source and drain are made conductive by implantation of ionized dopant atoms followed by high temperature annealing. Annealing causes the implanted atoms to move into lattice sites, thus becoming electrically active, as well as removing crystalline damage caused by the high-energy ion beam. Several excellent reviews of MOS processing science and technology are available, including one by Gandhi.[10]

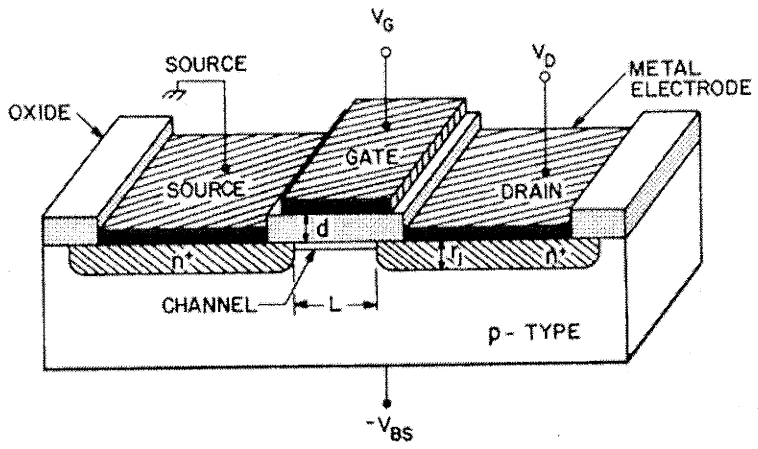


Figure 3.1 Schematic diagram of an enhancement-mode *n*-MOSFET.[24] The source-drain region of the semiconductor is implanted *n*+, and the channel region is *p*-type. At zero gate bias and positive drain voltage a negligible amount of current flows. Once the gate voltage exceeds the threshold voltage, electrons can flow from source to drain.

Biasing the gate electrode (V_{GS}) modulates the density and type of charge carriers located in the channel region, and biasing the drain-source (V_{DS}) modulates the potential for current flow in the channel. The doping of the source-drain regions determines whether the channel conducts positive or negative charge (*p*- and *n*-MOSFETs, respectively), and the doping in the channel determines the sign and magnitude of the gate voltage required to turn the device on and off. Beyond distinguishing between *n*- and *p*-type, MOSFETs are further classified as either enhancement-mode or depletion-mode. In a depletion-mode MOSFET, carriers are free to flow from source to drain with $V_{GS} = 0$; the device is "on" with no applied gate bias. To turn such a device "off," a gate voltage must be applied in order to deplete the carriers in the channel region. In an enhancement-mode MOSFET, the source and drain are doped the opposite type of the bulk of the semiconductor, forming two diode junctions. In operation, the drain is reverse biased, and for $V_{GS} = 0$, the only current that can flow from source to drain is the reverse

bias junction leakage current; the device is normally "off." The threshold voltage V_T for the device to turn "on" can be defined as the gate bias that provides sufficient band bending to create an inversion in the carrier population at the oxide-semiconductor interface. Many factors influence the absolute value of V_T for a given device, but in enhancement-mode devices, the value is typically positive for *n*-MOS and negative for *p*-MOS.

3.2. The CMOS Inverter

The complementary-MOS (CMOS) inverter, which has become the basis of modern Si-based digital electronic systems, is a circuit that employs two enhancement mode devices: one *n*-MOSFET and one *p*-MOSFET. The drains of the two devices are tied together to serve as the output, and a common connection to the two gates serves as the input. As Figure 3.2 shows, for $V_{in} = 0$, $V_{GS} = 0$ for the *n*-MOSFET and $V_{GS} = -V$ for the *p*-MOSFET. Therefore, the *p*-MOSFET is on (assuming that $|-V| > |V_{T,pMOS}|$), the *n*-MOSFET is off, and $V_{out} = V$. Conversely, for $V_{in} > 0$ (a binary "1"), the *n*-MOSFET is on while the *p*-MOSFET is off, leading to $V_{out} = 0$. Since the drains of both devices are connected in series, no drain current can flow aside from the charging current that is drawn during the switching process, and the power consumption is very low. The amount of time delay needed to accomplish the switching action, and hence, the speed of the inverter, is determined by the drain current I_D of the MOSFETs.[25]

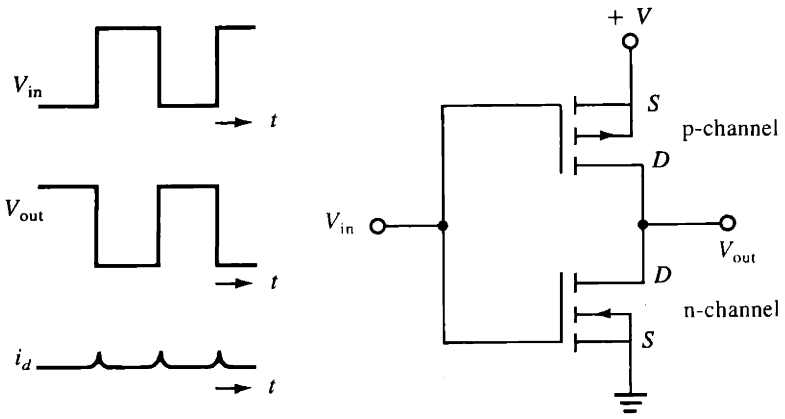


Figure 3.2 (a) Input and output characteristics and (b) schematic diagram of the CMOS inverter (from Streetman[25])

The on state for a MOSFET can be divided into two regimes of operation: the linear regime, where $|V_{DS}| < |V_{GS} - V_T|$, and the saturation regime, where the slope of I_D versus V_{DS} becomes close to zero. The current-voltage relationship for a long-channel MOSFET in the linear regime is given by:

$$I_D = \frac{W}{L}\mu_{eff}C_{ox}(V_G - V_T)V_D \tag{3.1}$$

where μ_{eff} is the effective mobility of the inversion carriers, W is the width of the transistor, L is the length of the channel, and C_{ox} is the capacitance of the oxide. The oxide capacitance is given by:

$$C_{ox} = \varepsilon_{ox} A / t_{ox} \tag{3.2}$$

where ε_{ox} is the dielectric constant of the oxide, A is the area of the capacitor, and t_{ox} is the thickness of the oxide. The relationship given by Eq 3.1 changes in the saturation regime, and additional adjustments are required for accurately predicting the current-

voltage relationships in short-channel devices.[24] However, the key dependencies on device geometry and carrier mobility are still preserved in all cases, and device designers sometimes refer to the pre-factors of Eq 3.1 collectively as the gain factor of the transistor:[26]

$$GainFactor = \frac{W}{L}\mu_{eff}C_{ox} \tag{3.3}$$

One immediate insight to draw from Eq 3.1 is that for a specified set of dimensions and voltages, I_D will be larger for an *n*-MOSFET than a *p*-MOSFET due to the fact that the electron mobility in Si is ~3× higher than the hole mobility. In order to allow for symmetric drive currents and, thus, symmetric switching times, the *p*-MOSFET is typically laid out ~3× wider than the *n*-MOSFET ($W_p/W_n = 3$) as in Figure 3.3.[27]

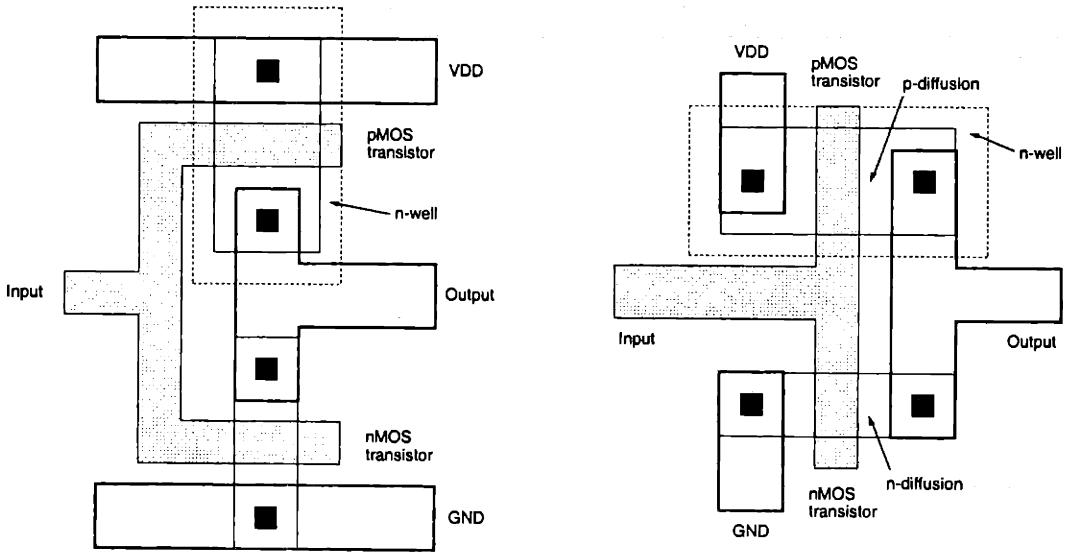


Figure 3.3 Planar view of typical CMOS inverter layouts. The *p*-MOSFET is ~3× wider than the *n*-MOSFET in order to provide symmetric drive current and switching times.[27]

Another important consideration in device design is that MOSFETs with different channel doping levels display widely differing μ_{eff}. The actual value of μ_{eff} in a given device can be obtained by performing measurements of the linear-regime I_D-V_{GS} characteristics (known as the transfer characteristic), and the general trend for both *p*- and *n*-channel devices is that higher channel dopings lead to lower mobilities. However, plotting μ_{eff} with respect to the vertical effective field, E_{eff}, reveals that devices with different doping levels all fall onto a single "universal mobility curve."[28] E_{eff} is defined as:

$$E_{eff} = \frac{Q_b + \eta Q_{inv}}{\varepsilon_s} \tag{3.4}$$

where Q_b is the bulk charge per unit area in the depletion region, Q_{inv} is the charge per unit area in the inversion layer, η is an empirical fitting-factor (0.5 for electrons, 0.33 for holes) and ε_s is the permittivity of the semiconductor.[28] A schematic plot of μ_{eff} vs E_{eff} is shown in Figure 3.4. According to the universal mobility model, carrier mobility at low inversion densities (i.e. the lefthand portion of Figure 3.4) is limited by Coulomb scattering events. In channels with high amounts of ionized impurities, inversion carriers will undergo more Coulomb-type scattering than in an intrinsic channel. Hence the severity of Coulomb-scattering increases with higher channel dopings. As E_{eff} is increased, Coulomb-scattering diminishes because the increasing density of inversion charge screens out any charged impurities, and the mobility displays a positive exponential dependence on E_{eff}. As the gate overdrive is increased, phonon scattering becomes the primary mobility-limiting mechanism, and a dependence on $E_{eff}^{-0.3}$ is typically observed. At large gate overdrives, the universal mobility model stipulates that surface-roughness scattering events are the dominant factor in limiting carrier mobility,

with an E_{eff}^{-2} dependence for electrons and an E_{eff}^{-1} dependence for holes. Due to the presence of multiple scattering mechanisms, the overall inversion layer mobility is typically calculated using Mathiesen's Rule:[29]

$$\frac{1}{\mu} = \frac{1}{\mu_{coulomb}} + \frac{1}{\mu_{phonon}} + \frac{1}{\mu_{surface-roughness}} \tag{3.5}$$

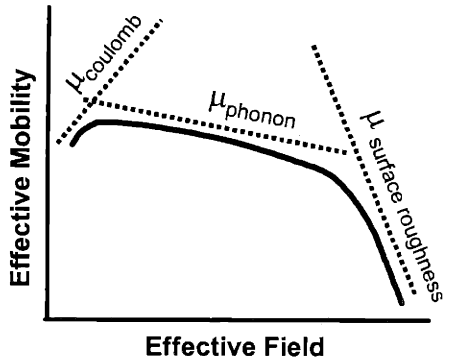


Figure 3.4 Universal mobility curve for Cz-Si MOSFETs.

Besides the gain factor, another important performance metric for MOSFETs is the subthreshold swing, S^{-1}. The subthreshold slope S is the gate voltage required to increase I_D by one order of magnitude and is given by:

$$S = nkT/q \ln(10) \tag{3.6}$$

where n is an ideality factor that is determined by the gate-stack characteristics. If $n = 1$, then $S^{-1} = 60$ mV/dec at room temperature; S^{-1} is typically ~ 70 to 90 mV/dec in well-engineered devices. The subthreshold characteristics are extremely important to the overall circuit performance, because the off current I_{off} is exponentially dependent on n. Reducing t_{ox} is one way to decrease the value of n, because, qualitatively speaking, the

gate is in better "communication" with the carriers in the channel. Good subthreshold behavior has become progressively more important as MOSFET packing densities have increased to tens of millions per chip, because small increases in I_{off} can lead to huge increases in DC power dissipation.[30]

3.3. Improving CMOS performance by scaling – a simple view

The practice of scaling in microfabrication has led to decades of steady improvement in the cost, performance, and functionality of VLSI circuits. The chief variables that define a scaling generation are the minimum feature size that can be patterned and the gate oxide thickness. Thus, a simplistic example of the performance enhancement gained by scaling is that as the resolution of photolithography has improved, the gate length L has shrunk, resulting in larger MOSFET drive currents per unit width. Likewise, improvements in the ability to grow thin, highly uniform thermal oxides have allowed devices with larger C_{ox} (Eq 3.2), also resulting in improved MOSFET gain. However, as L is scaled into the deep sub-micron regime, a number of short-channel effects become dominant, including V_T reduction (rolloff), drain-induced barrier lowering, and punch-through.[30]

Among the many problems associated with shrinking device dimensions are those related to extremely thin gate oxides. Gate leakage currents can become substantial for oxides that are only 1-1.5 nm thick, leading to high static power dissipation. Furthermore, the risk of dopant atoms implanted into the poly-Si gate diffusing through the gate oxide and into the channel increases. Since unintentional doping in the channel must be avoided in order to prevent V_T shift, the portion of the poly-Si adjacent to the

gate oxide is typically left with lower doping density. In operation, the gate bias can actually cause the poly-Si adjacent to the gate oxide to deplete (a phenomenon known as poly-depletion), causing a reduction in overall gate capacitance and an increase in S^{-1}. In order to address these problems, researchers have devoted considerable attention to the engineering of the gate stack itself. For example, a metallic gate could be used instead of poly-Si, thus avoiding the poly-depletion problem. Another complementary approach to problems associated with thin SiO_2 gates is the use of high-κ gate dielectrics such as ZrO_2 and HfO_2. With a higher dielectric constant, the thickness of the insulating film can be increased without sacrificing C_{ox}, drastically reducing the probability of carrier tunneling. To date, however, relatively immature process technology has limited the performance of MOSFETs with high-κ gate dielectrics.[31] While orders of magnitude reduction in leakage current have been attained, degraded carrier mobilities and higher interface state densities are among the problems that researchers have encountered. Some believe that no matter what improvements may come in the engineering of high-κ dielectrics the inversion layer mobility will always be degraded due to carriers strongly coupling with phonons in the high-κ material.[31]

Clearly, it is becoming progressively more difficult to improve MOSFET performance via scaling. Returning to Eq. 3.3, we see that μ_{eff} is the final parameter that can be used to increase the MOSFET gain factor. In the next section, we explore how SiGe heterostructures can be used to increase the effective mobility, and thus the performance, of both *n*-MOSFETs and *p*-MOSFETs.

3.4. High mobility SiGe heterostructures

Mobility determines carrier velocity in an applied electric field. While traditionally viewed as a fixed parameter, μ originates from the band structure of the material, and thus changes to the band structure can significantly change μ. Drift mobility is given by:[29]

$$\mu = \frac{e\tau}{m^*} \tag{3.7}$$

where e is the elementary charge, τ is the scattering time, and m^* is the carrier effective mass. The low-field electron and hole mobilities in lightly doped bulk Si and Ge are shown in Table 3.1.[25] Modern epitaxial growth technique allows both τ and m^* to be altered via the modulation of strain and composition over short length-scales. In this section, I will show how the use of $Si_{1-x}Ge_x$ alloys allows the design of devices with mobilities that are greatly enhanced over bulk Si.

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

Table 3.1 Electron and hole mobilities in bulk Si and Ge[24]

Figure 3.5 plots the variation in bandgap, E_g, for $Si_{1-x}Ge_x$ alloys versus x. As the Ge content in the alloy is increased, E_g drops slowly, remaining substantially Si-like (X-valley minima) until $x \sim 0.8$, after which the conduction band minima crosses over to the L-valley. For $x > 0.8$, E_g drops rapidly to the value of bulk Ge ($E_{g,Ge} = 0.66$ eV).[32] In this section, several devices that take advantage of the band offsets available in the $Si_{1-x}Ge_x$ alloy system as well as the effects of biaxial strain are examined.

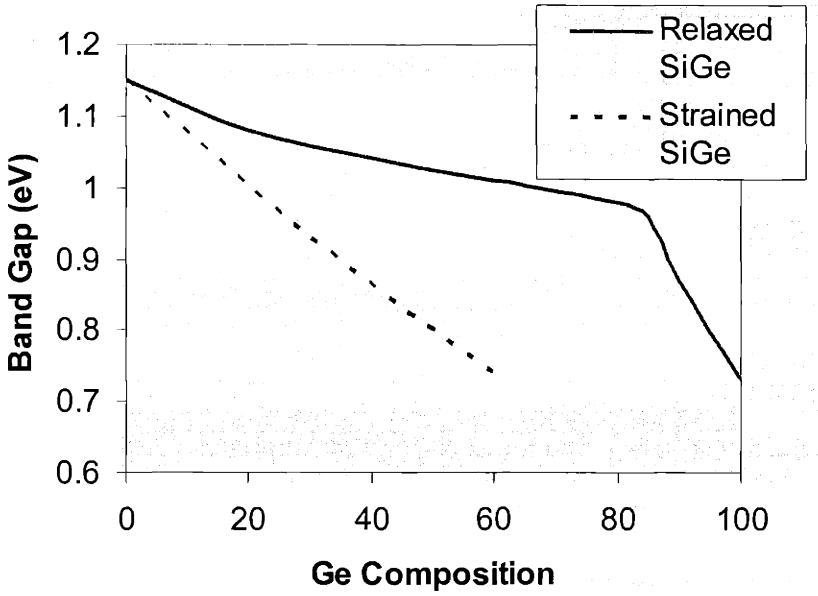


Figure 3.5 Bandgap variation in $Si_{1-x}Ge_x$ alloys. The conduction band minima crosses over from *X* to L at $x \sim 0.8$

3.4.1. Pseudomorphic SiGe-channel *p*-MOSFET

One of the first devices conceived by researchers to make use of the advantageous properties of $Si_{1-x}Ge_x$ was the pseudomorphic SiGe-channel *p*-MOSFET. When $Si_{1-x}Ge_x$ is grown epitaxially upon a Si substrate, a type-I band offset results (Figure 3.6), forming a shallow quantum well for holes in the $Si_{1-x}Ge_x$.[33] Since SiGe-oxides have poor electrical properties,[34] a Si cap layer must then be deposited to allow the formation of an SiO_2 gate. In inversion, the type-I band offset confines holes in the $Si_{1-x}Ge_x$, and surface-roughness scattering at the Si/SiO_2 interface is reduced.[35] The compressive strain in the $Si_{1-x}Ge_x$ breaks the light-hole/heavy-hole degeneracy of the valence band, reducing the rate of intervalley scattering for holes.[36] The lower rate of both surface-roughness and intervalley scattering increases τ for holes and leads to higher μ. Strain also distorts the shape of the hole valleys, causing a reduction in the hole effective mass.[36] However there

exist many limitations to the performance of pseudomorphic SiGe *p*-MOSFETs. First, the valence band offsets for $Si_{1-x}Ge_x$ alloys do not become very large until $x \sim 0.8$. Most pseudomorphic SiGe devices utilize alloys with $x < 0.5$ due to the extremely low h_c of Ge-rich films on Si, and therefore the valence band offset for the hole well is relatively small.[37] Second, with sufficiently large vertical fields, holes can be pulled into the Si cap layer, and much of the performance benefit associated with the strained $Si_{1-x}Ge_x$ layer is lost. Holes in the Si cap are free to undergo surface-roughness scattering events, and they also screen the gate electrode from the buried layer, making it difficult to significantly increase the carrier density in the $Si_{1-x}Ge_x$ layer.[35,37] Also, the increased hole mobility in the buried layer does not translate directly into enhancements in drive current, because the Si cap decreases the gate-to-channel capacitance. Pseudomorphic $Si_{1-x}Ge_x$ *p*-MOSFETs also suffer from many of the same problems seen in traditional Si-based buried channel devices (where the buried channel is defined by variations in doping density perpendicular to the channel), including higher subthreshold swing and worse susceptibility to short-channel effects. While this technology is relatively simple to implement from the standpoint of wafer supply and epitaxy steps, the performance enhancements reported to date have been modest.[35,37] Additionally, this heterostructure provides no benefit for electron mobility.

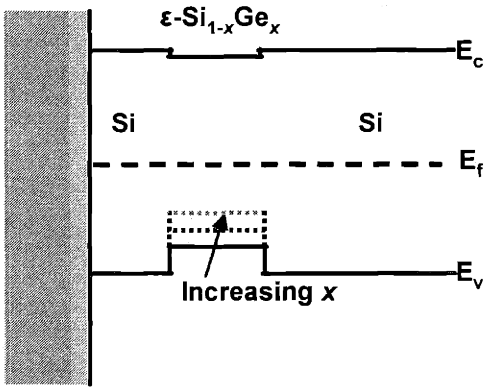


Figure 3.6 Band alignment of the pseudomorphic SiGe-channel *p*-MOSFET

3.4.2. Tensile strained Si *n*- and *p*-MOSFETs

When Si is grown onto a (100) relaxed $Si_{1-x}Ge_x$ virtual substrate, a type-II band alignment results, allowing the formation of a two-dimensional electron gas (Figure 3.7). The magnitude of the conduction band offset in eV is approximately given as:[38]

(3.4) $$\Delta E_c = 0.6x$$

where x is the Ge-fraction in the buffer (note that this equation only applies to $x < 0.6$).

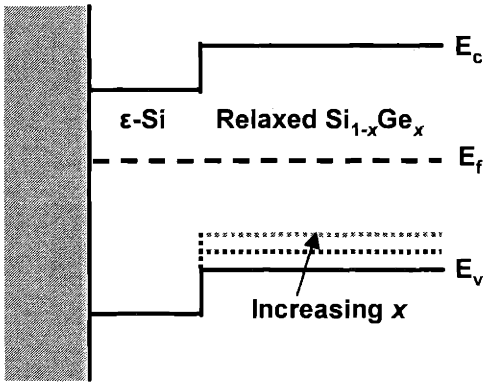


Figure 3.7 Band alignment of strained Si grown on relaxed $Si_{1-x}Ge_x$

Biaxial tensile strain causes significant changes to both the valence band and conduction band. In bulk Si, the conduction band is comprised of 6 degenerate valleys in the <100> directions [Figure 3.8(a)]. The effective mass of each cigar-shaped lobe is highly anisotropic, with the transverse mass (m_t=0.19m_o) being much lower than the longitudinal mass (m_l = 0.98m_o).[25] The contributions of the six degenerate valleys can be added to result in a single expression known as the conductivity effective mass of electrons, m^*_c:[29]

$$m^* = \left[\frac{1}{3}\left(\frac{1}{m_l}\right) + \frac{2}{m_t}\right]^{-1} \tag{3.5}$$

However, under biaxial tension, the degeneracy of the valleys is broken, and the energy of the two valleys perpendicular to the growth-plane is lowered with respect to the four in-plane valleys [Figure 3.8(b)]. Therefore electrons only experience the low transverse effective mass m_t for in-plane transport. Another important consequence of the broken degeneracy is that the effective mass in the out-of-plane direction is m_l, meaning that the electrons can be effectively confined even for thin ε-Si layers.[39] The combination of reduced intervalley scattering and decreased effective mass leads to greatly enhanced electron mobility.

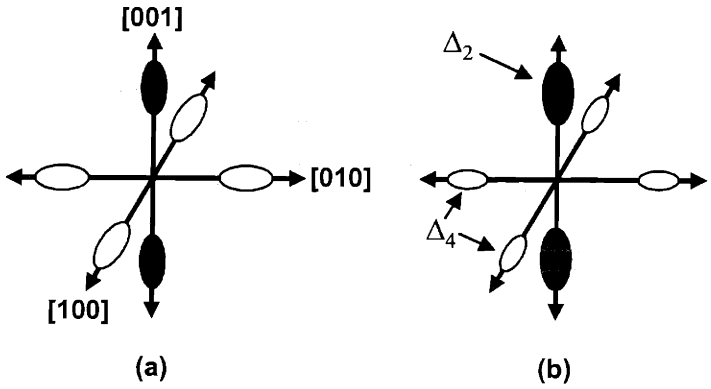


Figure 3.8 Conduction band structure of (a) bulk and (b) strained Si. Under biaxial tensile strain, the 6-fold degeneracy is broken, with the two out-of-plane valleys lowered in energy. The reduction in degeneracy decreases intervalley scattering and increases electron mobility.

High quality epitaxial growth has been the key enabling factor in the successful implementation of $Si_{1-x}Ge_x$ band-engineered devices. In early attempts, a thick uniform $Si_{1-x}Ge_x$ buffer layer was first deposited directly onto Si wafers to serve as the template for subsequent deposition of ε-Si.[40,41] These buffers were not fully relaxed and contained threading dislocation densities on the order of 10^8 cm^{-2}, and the electron mobility of the ε-Si layer was in some cases found to be lower than that of bulk Si. Not until the advent of fully-relaxed $Si_{1-x}Ge_x$ graded buffers with a 100-fold reduction in defect density[11] was the predicted enhancement in electron mobility observed (Figure 3.9).[42-44] The need for high quality $Si_{1-x}Ge_x$ virtual substrates makes the implementation of ε-Si devices much more complicated than pseudomorphic SiGe-channel *p*-MOSFETs. However, the performance benefits associated with ε-Si have proven to be much greater.

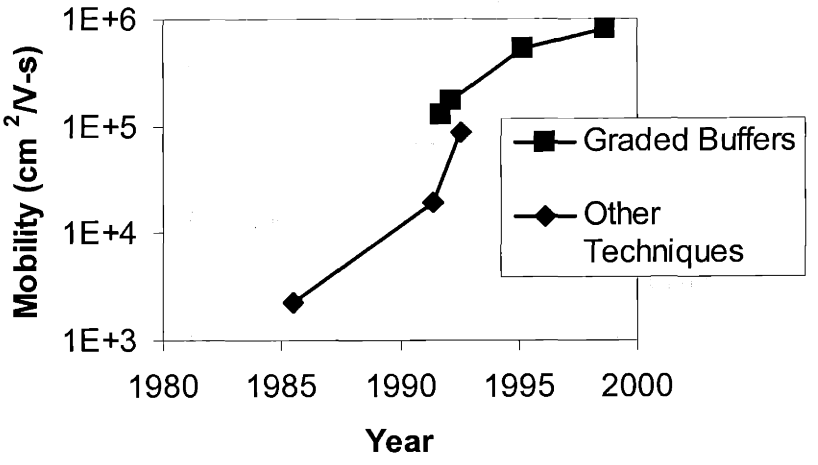


Figure 3.9 Evolution of record low temperature electron mobility in ε-Si modulation-doped heterostructures.[43]

ε-Si n-MOSFETs

Numerous researchers have fabricated ε-Si *n*-MOSFETs and observed mobility enhancements over bulk Si of 1.7 to 2.0.[7,8,45,46] A strain of 0.8%, corresponding to Si grown on relaxed $Si_{0.8}Ge_{0.2}$, results in sufficient conduction band splitting to completely suppress intervalley scattering, and little improvement in electron mobility is gained by further increasing the strain.[7,47] The large effective mass of electrons in the out-of-plane direction is convenient from the perspective of growth, because ε-Si layers as thin as 6 nm can completely confine electrons away from the relaxed $Si_{1-x}Ge_x$ buffer.[7] For *n*-MOSFETs with thinner ε-Si layers, some electron mobility enhancement tends to be lost. Currie *et al.* examined ε-Si *n*-MOSFETs with Si caps as thin as 2 nm and found their mobility to be degraded below that of bulk Si, presumably due to electrons undergoing alloy scattering in the relatively low-mobility relaxed $Si_{1-x}Ge_x$.[7] Two simple design criteria for the growth and fabrication of high mobility ε-Si *n*-MOSFETs can therefore be stated: First, the strain in the Si layer, ε_{Si}, must be at least 0.8%, and second,

the thickness of the ε-Si layer h_{Si} must be at least 5-6 nm after device processing. If both of these conditions are met, the electron mobility enhancement has been shown to be quite robust.

Just as in bulk Si, electrons in ε-Si *n*-MOSFETs demonstrate a "universal" behavior where electron mobilities plotted with respect to E_{eff} fall onto a single curve. In deeply scaled devices, the channel doping tends to be quite high ($\geq 10^{18}$ cm^{-2}) in order to ensure proper V_T control, leading to very high vertical effective fields. However, even in heavily doped channels with fields of 2.0 MV/cm, electron mobility enhancements of 1.5 to1.7 times have been observed.[48] The lateral fields in submicron devices are also extremely high, and there has been considerable debate over the effect of velocity saturation in ε-Si *n*-MOSFETs.[49] Since the saturation velocity of electrons in ε-Si is not significantly higher than that in bulk Si, some believed that the enhanced electron mobility would not translate into higher drive currents. However ε-Si *n*-MOSFETs with sub-100 nm gate lengths have been shown to exhibit significant drive current enhancement over bulk Si counterparts.[8]

ε-Si p-MOSFETs

Biaxial tensile strain also splits the light-hole/heavy-hole degeneracy, though the rate of subband splitting in the valence band is known to be lower than for the conduction band.[5] Unlike electrons, both the out-of-plane and in-plane effective mass of holes is reduced by strain (Figure 3.10). Recent experimental work by Leitz *et al.*[6] demonstrated that hole mobility enhancements in ε-Si *p*-MOSFETs increase with higher Ge content in the virtual substrate up to 40% Ge, in good agreement with earlier theoretical calculations where a complete suppression of intervalley scattering was predicted for strain of 1.6%.[50]

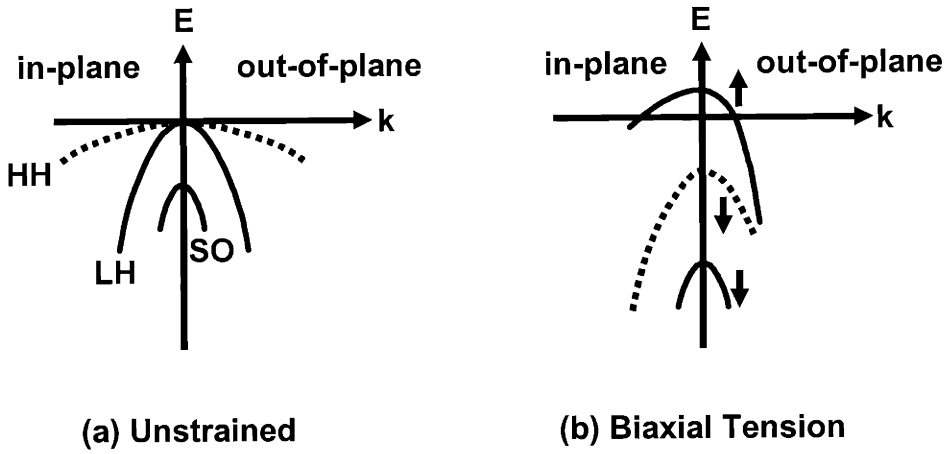


Figure 3.10 Valence band of (a) bulk Si and (b) Si under biaxial tension.

Unlike the *n*-MOS case, hole mobility enhancements vary as a function of E_{eff} (Figure 3.11). At low vertical fields, the hole wave function tends to be weighted below the surface[51] due to the type II band alignment between the relaxed $Si_{1-x}Ge_x$ and the ε-Si (Figure 3.7). The magnitude of the valence band offset (ΔE_v) increases with x according to:[38]

$$\Delta E_v = (0.74 - 0.53x)x \tag{3.7}$$

As gate overdrive is increased, the hole wave function can overcome the band offset and shift toward the surface. The gradual increase of inversion density in the ε-Si results in increasing mobility enhancement with E_{eff} up to ~0.35 to 0.5 MV/cm, where a peak is typically observed. With higher Ge-content in the buffer, the barrier to hole occupation in the ε-Si cap increases, and a stronger vertical field is needed to shift the hole wave function up to the surface. Therefore, the value of E_{eff} at peak mobility enhancement also increases with the Ge-content,[14] as shown in Figure 3.11. Beyond 0.5 MV/cm, the mobility enhancement in ε-Si *p*-MOSFETs, regardless of buffer composition, slowly

decreases for reasons that are still poorly understood.[6-8,51-53] Recent experimental results show that deeply scaled ε-Si *p*-MOSFETs grown on $Si_{0.72}Ge_{0.28}$ lose all mobility enhancement at E_{eff} = 1.0 MV/cm,[8] though no comparable study has been published for higher strain levels. As stated in the Introduction, ε-Si is drawing nearer to widespread commercialization due to its ability to increase circuit performance without scaling. However, it is currently believed that the performance benefit of implementing ε-Si will derive almost solely from the enhanced drive current of the *n*-MOSFET.

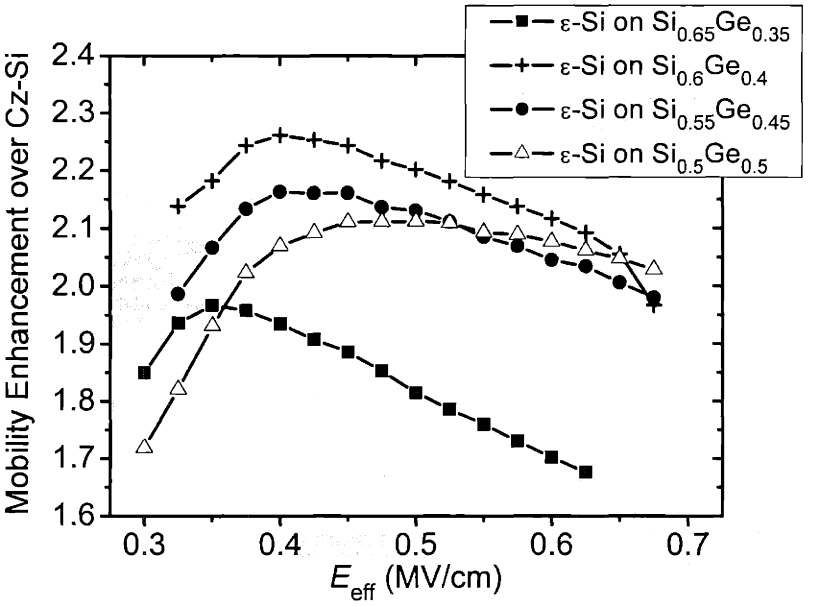


Figure 3.11 Hole mobility enhancement vs E_{eff} for ε-Si *p*-MOSFETs grown upon $Si_{1-x}Ge_x$ buffers with x=0.35 to 0.5. Adapted from Leitz *et al.*[14] The point of peak enhancement is pushed to higher E_{eff} with higher x.

While a fairly simple set of "design rules" seems to exist for the growth of high mobility ε-Si *n*-MOSFETs (i.e. h_{si} > 5-6 nm and $\varepsilon_{Si} \geq 0.8\%$), no such rules seem applicable to ε-Si *p*-MOSFETs. For example, Leitz *et al.* investigated ε-Si *p*-MOSFETs on $Si_{0.7}Ge_{0.3}$ with the Si cap thickness ($h_{Si\ cap}$) ranging from 4 and 25 nm and found that while μ_{eff} differed substantially at low vertical fields, for E_{eff} > 0.5 MV/cm all devices had

nearly identical μ_{eff} regardless of $h_{Si\ cap}$.[6] Currie *et al.* observed an analogous result in analyzing the thermal budget of ε-Si *p*-MOSFETs on $Si_{0.8}Ge_{0.2}$. In Currie's study, a 1000°C RTA was performed on device wafers for varying amounts of time ranging from 1-30 seconds.[7] TEM performed after processing showed that $h_{Si\ cap}$ was reduced by interdiffusion at the ε-Si/$Si_{0.8}Ge_{0.2}$ interface. In the extreme case of a 30 second RTA, an as-grown Si cap of 10 nm was no longer visible in TEM. But just as Leitz *et al.* found, all of the *p*-MOSFETs had nearly identical μ_{eff} for $E_{eff} > 0.5$ MV/cm. In ε-Si *n*-MOSFETs subjected to the same treatment, a significant drop in effective electron mobility was observed for RTA times longer than 10 seconds. Therefore, while mobility enhancements in ε-Si *p*-MOSFETs are more sensitive to variations in E_{eff} and strain than their *n*-channel counterparts, the *p*-FET seems to be less sensitive to process variations.

3.4.3. Dual-channel heterostructure *p*-MOSFETs

The incorporation of compressively strained layers beneath the surface has been investigated to address the challenges in controlling hole mobilities in surface ε-Si devices.[54-56] In these devices, termed dual-channel heterostructures, a compressively strained $Si_{1-y}Ge_y$ layer is grown upon a relaxed $Si_{1-x}Ge_x$ buffer ($y>x$) and capped with tensile ε-Si. In many ways, the dual-channel heterostructure is an improved version of the pseudomorphic strained-SiGe *p*-MOSFET discussed above. Since the dual-channel heterostructure is grown on a relaxed $Si_{1-x}Ge_x$ buffer rather than a Si substrate, a compressively strained layer with considerably higher Ge content can be used. Furthermore, the capping Si layer is strained, so that holes pulled upward to the SiO_2/Si interface are still in a material with enhanced hole mobility. Large hole mobility enhancements are predicted in dual-channel heterostructures for the following reasons:

first, compressive strain in the $Si_{1-y}Ge_y$ splits the valence band degeneracy and reduces the in-plane and out-of-plane effective masses.[36,56] Second, the type I band alignment between $Si_{1-y}Ge_y$ and $Si_{1-x}Ge_x$ ($y>x$) causes the formation of a deep quantum well for holes in the ε-$Si_{1-y}Ge_y$ (Figure 3.12).[40] Third, Ge-rich $Si_{1-y}Ge_y$ alloys can possess intrinsically high hole mobilities due to their Ge-like band structures.

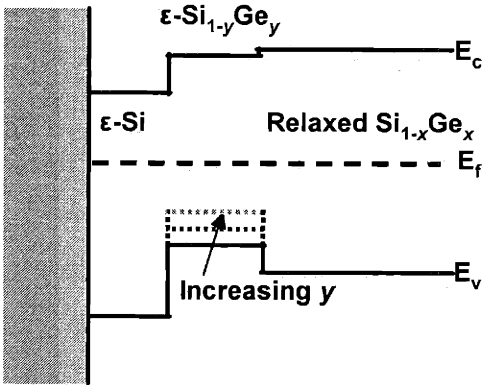


Figure 3.12 Schematic band alignment of a dual-channel heterostructure. Higher *y* in the compressive layer leads to a deeper well for holes. In *p*-MOSFETs based on these structures, large gate overdrives force holes towards the SiO_2/Si interface.

While some debate over the impact of alloy scattering in $Si_{1-y}Ge_y$ alloys has persisted,[57,58] experimental results have repeatedly shown that Ge-rich alloys possess large hole mobilities.[56,59-61] Earlier work in modulation-doped structures showed room temperature mobilities of 700[56] and 1050 cm^2/Vs[62] in compressively strained $Si_{0.3}Ge_{0.7}$ and $Si_{0.2}Ge_{0.8}$ quantum wells. Leitz *et al*. examined a wide range of compressively strained $Si_{1-y}Ge_y$ alloys with $0.5 < y < 0.8$ in dual-channel heterostructure MOSFETs and found all compositions to display significant mobility enhancements over Cz-Si (Figure 3.13).[54] Furthermore, high mobilities were observed even in the samples with $y = 0.6$, where alloy scattering is expected to be severe. However, the trend in Figure 3.13

indicates that compressively strained pure Ge should demonstrate the highest hole mobility.

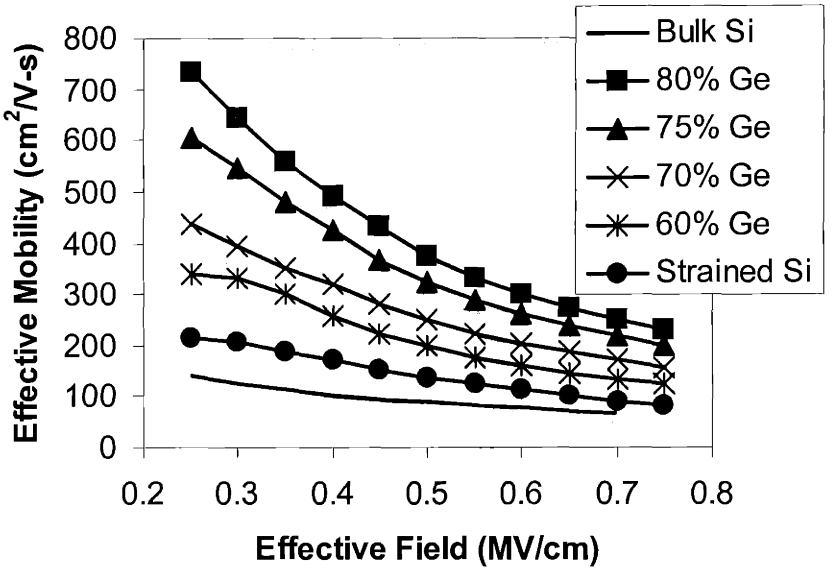


Figure 3.13 Effective hole mobilities of dual-channel heterostructure *p*-MOSFETs (from Leitz *et al.*[54]). The difference in Ge content between the virtual substrate and compressive layer, y-x, is held constant at 0.3.

While a single-channel heterostructure can be completely described by the Ge content in the virtual substrate and the thickness of the Si surface channel, a dual-channel heterostructure must be further specified by the composition and thickness of the buried layer. The design of the layer structure in dual-channel heterostructures is therefore considerably more complicated than that of single-channel heterostructures. Since holes are expected to populate both the buried and surface channel, the average inversion mobility μ_{avg} is a number-weighted average of inversion carriers in both layers. Leitz *et al.* experimentally found that a thin Si cap and a thick buried channel lead to high μ_{avg}, simply because more of the inversion charge would be expected to reside in the higher

mobility compressive layer.[54] Thinner Si caps also benefit subthreshold characteristics since the gate is in better electrostatic communication with holes in the $Si_{1-y}Ge_y$.

Dual-channel heterostructures may become relevant as the "next-generation" technology after the deployment of ε-Si (single-channel) CMOS. The situation for electrons in dual-channel heterostructures does not differ significantly from the single-channel case described in Section 3.4.2; as long as the ε-Si cap is thicker than 5-6 nm, the large vertical effective mass of electrons and type I band offset will prevent electrons from ever "seeing" the material below the surface channel. ε-Si *n*-MOSFETs exhibiting the familiar 1.7-2 times electron mobility enhancement can thus be fabricated on dual-channel heterostructures.[63] Moreover, while the mobility enhancements in ε-Si *p*-MOSFETs grown on $Si_{0.72}Ge_{0.28}$ have been shown to be completely lost at $E_{eff} \sim 1.0$ MV/cm,[8] dual-channel heterostructure *p*-MOSFETs with enhancements of ~2.2 at high field have recently been reported.[64] Therefore, CMOS technology based on dual-channel heterostructures could have symmetric enhancement for both *n*- and *p*-MOSFETs. While a significant amount of study will undoubtedly be needed to find optimum tradeoffs in subthreshold characteristics and current drive, the added design and process complexity that comes with dual-channel heterostructures may be justified by the improvement in performance.

3.4.4. Germanium as a channel material

Historically, MOS technology never advanced for Ge due to the poor chemical and electrical properties of GeO_2; GeO_2 dissolves in water, making device processing nearly impossible, and also absorbs water from the air, causing rapid electrical degradation upon exposure to atmosphere.[65,66] Still, there has been considerable interest

in implementing devices with Ge as the channel material due to its extremely high carrier mobilities (Table 3.1).[67] Several researchers have explored the use of Ge oxynitride (Ge_2N_2O) as a gate dielectric for Ge MOSFETs.[68-71] In the late 80's, it was found that subjecting thermally-grown GeO_2 to a treatment in pure NH_3 greatly improved the oxide's chemical and electrical properties by conversion to Ge_2N_2O.[65] Interface state densities ranging from 3×10^{10} cm^{-2} to 4×10^{11} cm^{-2}-eV and *p*-MOSFET mobilities of 325[71]-770 cm^2/Vs[69] have been reported. At this time the wide variation of hole mobilities reported for Ge MOSFETs with Ge oxynitride gates is unclear. More recently, other researchers have reported the use of high-κ gate dielectrics including ZrO_2[72,73] and HfO_2[74] on Ge.

As described in Section 3.2.3, biaxial compressive strain in a pure Ge layer grown on relaxed $Si_{1-x}Ge_x$ should provide the highest hole mobilities. The first high mobility ε-Ge heterostructures grown on $Si_{1-x}Ge_x$ with $x = 0.6$ to 0.7 were reported by Xie *et al.* in 1993.[75] Since then, numerous researchers have reported extremely high room-temperature Hall mobilities in modulation-doped ε-Ge quantum wells, ranging from 1300 to 2940 cm^2/Vs,[43,61,75-79] as well as the fabrication of Schottky-gated *p*-channel FETs.[80-83] In these devices, modulation doping was employed to provide holes to the Ge well while preventing ionized impurity scattering. The high hole mobility in the ε-Ge channel leads to *p*-channel transconductances as high as 622 mS/mm (measured at 77K); this value is considered to be the highest transconductance ever measured in a *p*-FET.[84] Also, since these FETs are buried-channel devices, the holes never undergo any surface-roughness scattering, allowing excellent low-noise behavior. However, as is the case for all Schottky-gated devices, relatively small gate voltages have to be used in order to prevent

large gate leakage currents. Furthermore, while carrier densities of ~10^{12} cm^{-2} can readily be provided by modulation doping, the carrier density in the inversion layer of an enhancement-mode MOSFET typically reaches 1 to 2 × 10^{13} cm^{-2}. Thus, the drive current that Schottky-gated FETs can provide is quite low compared to MOSFETs. Still ε-Ge MODFETs remain an excellent candidate for the integration of high-speed analog devices (e.g. low-noise amplifiers) on $Si_{1-x}Ge_x$/Si virtual substrates.

As described in Chapter 2, compressively strained Ge layers can easily form undulations during growth with a wavelength on the order of 100 nm. Xie *et al.* have found that these undulations can cause considerable scattering of holes, greatly reducing their mobility.[75] From a technological standpoint, devices fabricated on undulated layers with gate lengths approaching 100 nm could exhibit large die-to-die variations in mobility due to the physical differences of the underlying microstructure (e.g. a gate centered over the trough of an undulation would be expected to have lower hole mobility than one centered over a crest). For these reasons, great care must be taken in the growth of ε-Ge to avoid these undulations. The specific challenges of growing highly strained compressive layers are described in detail in Chapter 4.

3.5. Short-flow MOSFET process

Mobilities in heterostructures are often deduced from Hall-effect measurements. In these experiments, modulation doping is generally employed in order to maximize the mobility in the quantum well, and care must be taken to avoid parallel conduction through layers outside of the region of interest. Accurate interpretation of Hall measurements is often difficult, because the Hall mobility differs from the drift mobility by a constant scattering factor that is not always precisely known. Furthermore, the Hall

mobility is not always indicative of the mobilities that will be encountered in MOS devices with large carrier densities and large band-bending near the oxide-semiconductor interface. Therefore, in this thesis, MOSFETs were fabricated[85] to explore carrier transport in $Si_{1-x}Ge_x$ heterostructures. Instead of Hall-effect measurements, room-temperature electrical measurements using standard semiconductor parameter analyzers were performed. The transistors had long gate lengths (several different sizes are on each die with L ranging from 50 to 500 μm) and a ring-shaped geometry. The typical W/L ratio in Eq 3.1 is replaced with a geometry factor, G (see Appendix B for a fuller description). A schematic of the processing steps of the short-flow MOSFET is shown in Figure 3.14, and a full outline of the process flow can be found in Appendix A.

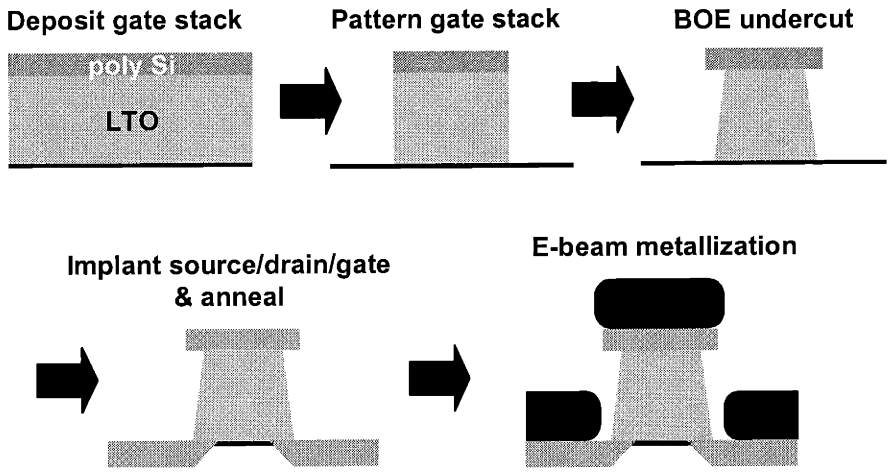


Figure 3.14 Short-flow MOSFET process flow

3.5.1. Choice of plotting method

The effective mobility of holes was extracted from the linear-regime source current-gate voltage characteristics of long-channel MOSFETs, using the relationship given in Eq 3.1 (I_D can simply be replaced with I_S).[85] While μ_{eff} is typically plotted as a

function of E_{eff} in reports on ε-Si MOSFETs in the literature, the universal mobility model does not facilitate comparison of heterostructure *p*-MOSFETs that comprise different material compositions, particularly when there may be hole conduction through the relaxed buffer or in a buried channel; the proper choice of ε_s in the equation for E_{eff} thus becomes unclear. Furthermore, the fitting-factor η is unknown for these heterostructures, and its physical origin even in bulk Si inversion layers is not fully understood.[28] Due to these uncertainties, I have chosen a conceptually simpler route and plotted μ_{eff} with respect to N_{inv}, the number of carriers in the inversion layer per unit area. In this work, N_{inv} is approximated as:

$$N_{inv} = \frac{C_{ox} \times V_{GT}}{e} \tag{3.8}$$

Where e is the elementary charge, C_{ox} is the oxide capacitance in F/cm^2 as determined by *C-V* measurements, and V_{GT} is the difference between the applied gate voltage and the threshold voltage of the device. Very large gate voltages (170-200V) are needed in order to attain large N_{inv} (i.e. $> 10^{13}$ cm^{-2}) in these devices due to the thick deposited gate oxide (300 nm) used in the short-flow process. Therefore, the error in the approximation for N_{inv} given in Eq. 3.8 is considered to be quite small. Since the material used in this study was grown nearly intrinsic (N_a, $N_d \leq 9 \times 10^{15}$ cm^{-3}), ionized impurity scattering is minimal for all devices, and the effects of Q_b are eliminated from consideration. Mobility data from several previous publications have been re-plotted with respect to N_{inv} for this thesis.

Chapter 4. Growth and characterization of strained Si and strained Ge heterostructures on relaxed $Si_{1-x}Ge_x$

The growth of highly strained heterostructures creates problems not seen in conventional single-channel ε-Si layers deposited on Si-rich $Si_{1-x}Ge_x$ virtual substrates. As was mentioned in the previous chapter, Ge-rich compressive layers have a tendency to island if they are grown at sufficiently high temperatures. Therefore understanding growth at low temperatures was necessary in order to reliably grow the novel MOSFET heterostructures described in later chapters. In this chapter, data will be presented describing the influence of both strain and growth temperature on layer morphology. Finally, an optimized process for growing fully planar compressive layers capped with strained Si is elaborated.

4.1. General procedure for growth of ε-Ge

Growing ε-Ge layers is a multi-step process, the first step of which is the growth of relaxed $Si_{1-x}Ge_x$ virtual substrates, as described in Chapter 2. After compositional grading to the desired Ge concentration, a relatively thick (0.5-1.5 μm) relaxed, uniform buffer layer is deposited. At this point, wafers may be removed for planarization via CMP in order to remove surface roughness and to reduce the dislocation pileup density. While the lowest possible defect density is absolutely necessary for optimum performance in minority carrier devices such as lasers, slightly elevated dislocation densities in the buffer are not likely to have a noticeable effect on room temperature inversion layer transport; with threading densities in the 10^5-10^6 cm^{-2} range, the dislocation spacing is so large that carrier scattering time should not be limited by scattering from dislocations. Additionally, since the MOSFET fabrication process used in this work consists of only one mask level with very large features, the crosshatch roughness does not cause any complications related to focusing the wafer surface in the

photolithography step. Thus, the CMP step was considered optional in this thesis research. If a CMP step is used, following standard cleaning and desorption steps, a thin lattice-matched buffer (typically 0.1 to 0.5 μm thick) is grown on the wafers in order to bury any residual impurities. Prior to ε-Ge growth the furnaces are cooled, but since the reactor is a hot-walled system, the cool-down step can take several hours. Once the desired growth temperature for deposition of ε-Ge is reached, a thin lattice-matched buffer layer is again deposited in order to overgrow any contaminants that may have been adsorbed onto the surface during the cool-down step. Due to the cleanliness inherent to ultra-high vacuum, the wafer surface stays relatively clean, and no ill effects were observed even when the "low-temperature buffer" preceding ε-Ge deposition was only 5 nm thick. SIMS analysis done revealed no change in the concentration of contaminants such as O, C, or N at the interface where the temperature is dropped. Occasionally, the low-temperature buffer is not perfectly lattice-matched to the virtual substrate, leading to slight strain contrast below the ε-Ge in XTEM images.

The decomposition rate of hydride precursors has been shown to depend on the composition of the substrate due to differences in surface energy.[86] While the precise reason is not clear, growth on Ge-rich surfaces tends to exhibit a slightly elevated rate compared to growth on silicon-rich surfaces. Figure 4.1 plots the variation in ε-Ge growth rate versus virtual substrate composition for thin layers (~10 nm). When the Ge film is grown thick enough, the average growth rate converges on the Ge-on-Ge homoepitaxial growth rate. However, for thin mismatched films, a single growth rate for Ge as a function temperature is inadequate due to the initial transient slow-growth period,

and great care must be exercised in order to grow heterostructures with accurate layer thicknesses.

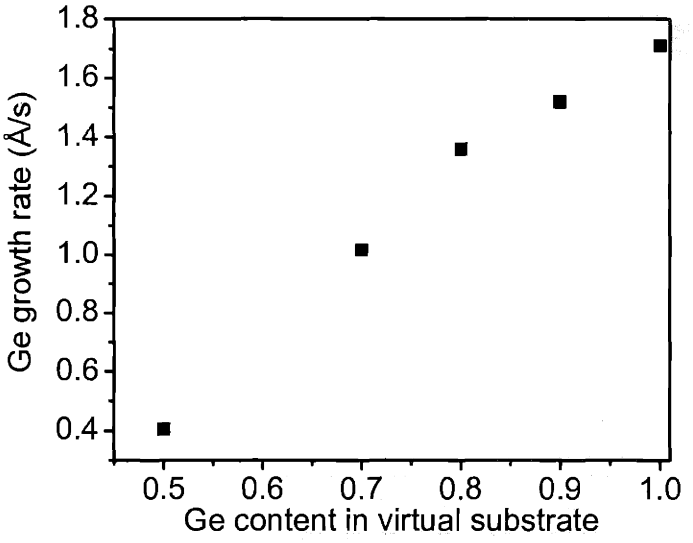


Figure 4.1 Dependence of ε-Ge growth rate on virtual substrate Ge content at 400°C.

4.2. Influence of strain and temperature on layer morphology

Compressively strained layers can relax via stress-driven surface diffusion, causing the formation of an undulated morphology. Early study of Ge on Si heteroepitaxy showed that the Ge typically grows in a layer-by-layer fashion (2D or Frank-van der Merwe growth) up to a thickness of 3 monolayers, beyond which island growth (Volmer-Weber growth) begins.[87] Thermodynamically, the increased surface energy of the undulated surface is balanced by a reduction in strain energy. Since surface diffusion is thermally activated, growth temperature is one of the primary variables in determining the morphology of a given strained layer. Figure 4.2 shows the influence of temperature on island formation in ε-Ge grown on $Si_{0.3}Ge_{0.7}$. At 550°C, island-growth

dominates due to the high surface diffusivity of Ge. At a growth temperature of 450°C, a continuous film with some undulations (Stranski-Krastanov growth mode) forms, and at 400°C, the ε-Ge grows in a fully planar fashion.

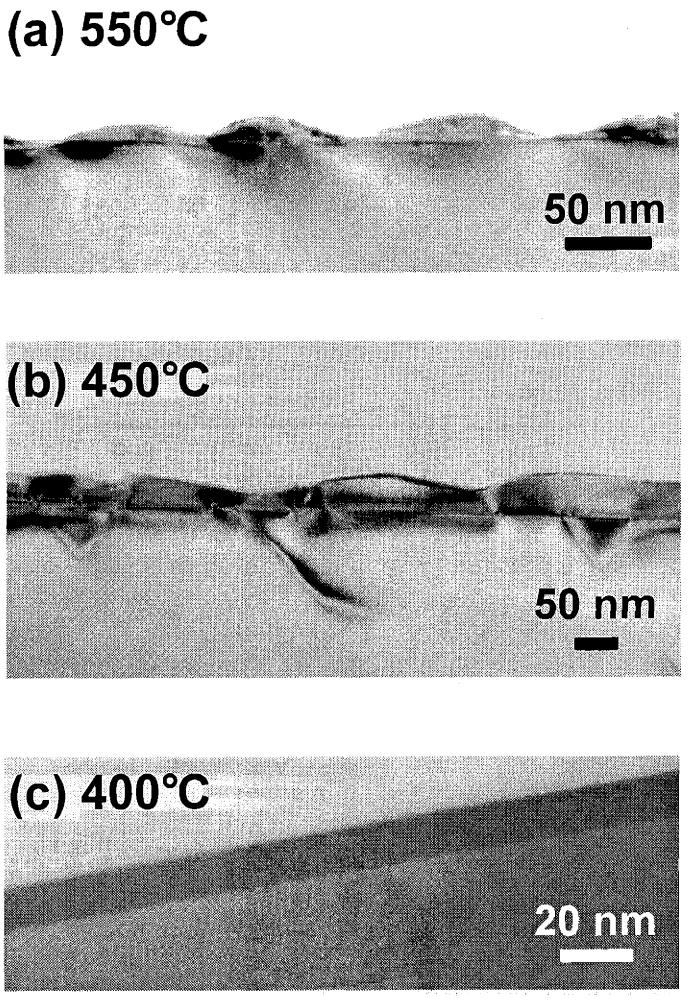


Figure 4.2 XTEM of ε-Ge grown on $Si_{0.3}Ge_{0.7}$ at (a) 550°C, (b) 450°C, and (c) 400°C showing evolution of morphology with *T* (ε-Ge is dark layer)

The experiments above show the influence of growth temperature on morphology for fixed strain and overlayer composition. However, the film-substrate mismatch is another primary variable in determining layer morphology, since the strain determines the energy balance that dictates the transition from a planar to a 3D morphology. For

example, the undulated morphology was found to return when ε-Ge was deposited on $Si_{0.5}Ge_{0.5}$ at 400°C, demonstrating the influence of strain. Reducing the growth temperature to 350°C allowed planar layers of ε-Ge to be deposited with 1.6-2% compressive strain. The low activation energy of GeH_4 decomposition is fortuitous, because it allows the growth of ε-Ge channel layers (typically ~10 nm) in a reasonable amount of time even at extremely low temperatures.

Since compressive strain relaxation via undulation is caused by surface diffusion, the melting point of the film (i.e. the growth temperature relative to the homologous temperature) is another factor that influences growth morphology. For example, while pure Ge grown at 2% compressive strain was found to island at 400°C, $Si_{0.2}Ge_{0.8}$ films grown with the same strain (i.e. on $Si_{0.7}Ge_{0.3}$) and growth temperature were planar. Recall that the melting point of $Si_{1-x}Ge_x$ alloys scales downward with Ge content ($T_{m, Si}$ = 1412°C ,$T_{m, Ge}$ = 940°C) and that the surface diffusivity of a material generally scales with its melting point. Therefore, films with lower Ge-content are less susceptible to undulation for a given growth temperature and strain.

Finally, film thickness is another variable that influences strained-layer morphology, since stress in a mismatched layer depends on thickness. While the 12 nm ε-Ge layer shown in Figure 4.2(c) was planar, depositing a 50 nm layer with the exact same growth conditions caused the formation of a slight ripple pattern on the surface (Figure 4.3). In this case, the buildup of compressive stress drove the transition from 2D growth to Stranski-Krastanov growth (Note that this picture was taken with the <110> axis parallel to the electron beam, and therefore little contrast appears between the ε-Ge and $Si_{0.3}Ge_{0.7}$). Clearly, a wide variety of variables influence surface diffusion, including

film thickness, melting point, mismatch with the substrate, and growth temperature. Thus, all of these variables can be factors in deciding whether a compressive layer will be planar or not.

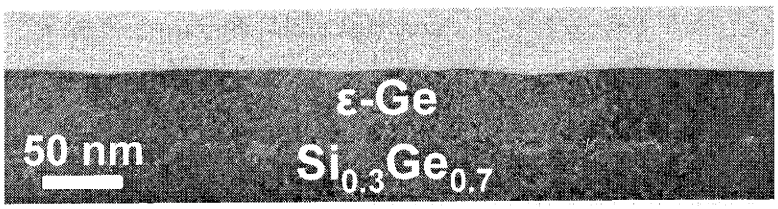


Figure 4.3 ε-Ge grown on $Si_{0.3}Ge_{0.7}$ at 400°C. While thinner Ge films at the same strain have been shown to grow planar at 400°C, the buildup of compressive stress in this sample drove the transition from 2D growth to Stranski-Krastanov growth.

Tensile layers exhibit a much reduced tendency to form islands. Xie *et al.* found that $Si_{0.5}Ge_{0.5}$ films grown at 2% compression (i.e. on pure Si substrates) undulated while those at 2% tension (i.e. on pure Ge substrates) did not.[22] Continuum mechanics approaches, where details of the surface such as reconstruction and steps are ignored, do not take the sign of the strain into consideration in modeling the surface roughness. Molecular dynamics simulations revealed that compressive strain caused a reduction in surface step energy, while tensile strain raised the surface step energy. Work in the Fitzgerald group has likewise shown that ε-Si can grow in a planar fashion at mismatches as high as 2% at 650°C. Naturally, the higher melting point of Si means that surface diffusivity will be lower than for Ge films, lending increased resistance to surface ripples. However, for a large mismatch of 2.6% (growth on $Si_{0.35}Ge_{0.65}$) at 650°C, ε-Si can grow in a nonplanar fashion. As Figure 4.4 shows, high tensile strains cause the formation of step bunches[9] that display a distinctly asymmetric, "ramp-like" morphology; the islands that were previously observed in the compressive strain case displayed a rounded, symmetric morphology. The differences in compressive undulations and tensile

undulations can be linked to the difference in surface step energies of compressive and tensile films.

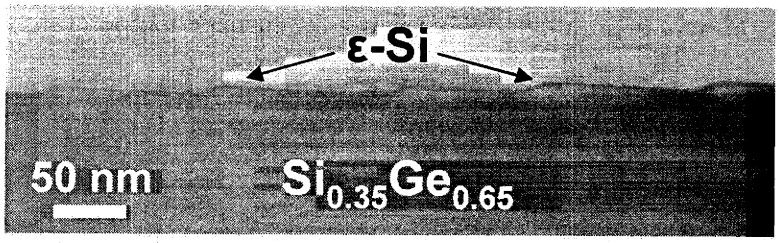


Figure 4.4 "Ramp-like" undulations in 2 nm thick ε-Si grown on $Si_{0.35}Ge_{0.65}$ at 650°C.

4.3. Growth of ε-Si / ε-Ge dual-channel heterostructures

While the highest hole mobilities reported in the literature have been in ε-Ge layers grown on relaxed $Si_{1-x}Ge_x$, there exists no mature technology for forming high quality gate dielectrics on Ge. Therefore, the most feasible way to take advantage of the high hole mobility offered by ε-Ge in a MOS device is to grow a thin Si cap to serve as the physical interface with a conventional SiO_2 gate. However, this presents a "catch-22" for the grower: Growth of planar ε-Ge requires temperatures between 350°C to 400°C, but as shown in Figure 2.8, neither SiH_4 nor Si_2H_6 can be used to grow even a thin Si layer at such low temperatures in a reasonable amount of time. Hence, the reactor temperature must be raised *after* growing the ε-Ge in order to allow the Si to deposit. An attempt was made to grow a Si capping layer by heating the furnaces to 500°C after depositing ε-Ge at 400°C. Once the furnaces stabilized at 500°C, SiH_4 was flowed to grow the Si cap. Subsequent XTEM revealed that the unprotected ε-Ge had islanded before the Si cap could be grown and demonstrated that a flat ε-Ge film is highly sensitive to post-growth annealing.

A two-step growth process was therefore adopted in order to grow the necessary ε-Si cap onto the unstable ε-Ge layer. After depositing the ε-Ge, the first step is to initiate flow of SiH_4 while simultaneously raising the furnace temperature to 450°C. Although the Si growth rate is virtually negligible at these low temperatures, the ability of hydrides to decompose more rapidly on pure Ge allows a very small amount of Si to be deposited. While this initial transient was inconvenient from the perspective of calibration as described in Section 4.1, the slightly elevated growth rate is actually advantageous in this situation since it helps to stabilize the morphology of the ε-Ge layer. The furnace is held at 450°C for some time (20 minutes was arbitrarily chosen) with SiH_4 still flowing. The second step is to heat the furnace to 550°C, all the while flowing SiH_4. Once 550°C is reached, the Si cap growth can be completed in several minutes since the SiH_4 growth rate increases to ~ 0.1 Å/s. As Figure 4.5 shows, this growth procedure yields an exceptionally flat ε-Ge channel and ε-Si cap. While the amount of Si that grows on the ε-Ge while heating the furnace to 450°C during flow of SiH_4 has not been directly determined, recall that growth of ε-Ge on $Si_{0.3}Ge_{0.7}$ at 450°C results in a slightly undulated morphology. Therefore, it can be inferred that whatever amount of Si (probably on the order of several monolayers) deposits during the initial step is adequate to prevent surface diffusion and preserve the planar morphology of the ε-Ge for some period of time.

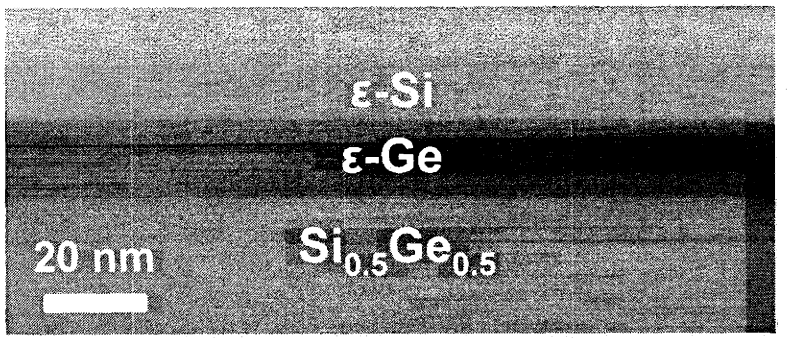


Figure 4.5 XTEM of ε-Si / ε-Ge dual-channel heterostructure grown on $Si_{0.5}Ge_{0.5}$ using the method of raising the furnace temperature while simultaneously flowing SiH_4.

The method of raising the furnace temperature while flowing SiH_4 can also be applied to the growth of highly mismatched Si caps even when there is no compressive layer. By growing the Si cap slowly, a low rate of strain is maintained over a relatively long time, and the "ramp-like" morphology seen in ε-Si (Figure 4.4) on Ge-rich relaxed virtual substrates can be avoided. Planar Si caps have been achieved on the entire range of SiGe lattice constants, including relaxed Ge (Figure 4.6). Of course, when Si is grown at such high mismatches (i.e. greater than 2.5%), extremely high dislocation densities result even for film thicknesses of only 5 nm. The dense dislocation network creates a strong repulsive force on dislocations that effectively prevents dislocation segments from looping down into the relaxed $Si_{1-x}Ge_x$, and Figure 4.6 shows that defects stay confined in the thin Si cap layer. When the mismatch is greatly reduced, as in the initial phases of graded buffer growth on bulk Si, dislocations can loop downward into the substrate before threading up to the surface due to the relatively diffuse misfit array at the film/substrate interface.[13]

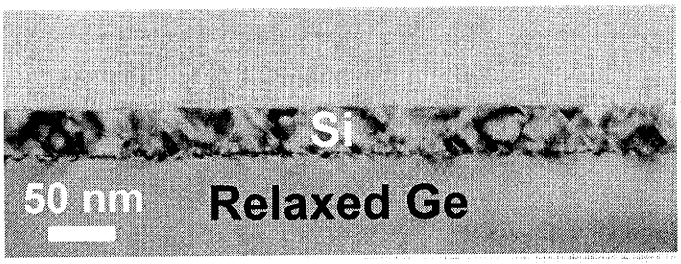


Figure 4.6 XTEM of Si deposited on relaxed Ge. The dense network of misfits causes defects to stay isolated in the cap layer. Photo taken by A.J. Pitera.

The two-step process for Si cap growth described above (hereafter referred to as Process 'A'), while very effective for attaining a wide variety of highly strained heterostructures with planar morphology, suffers notably from poor reproducibility. The growth rate of the Si cap was found to vary widely as a result of inconsistencies in furnace performance (See Figure 4.7, Process 'A', error bars correspond to standard deviation). During heating the furnace is programmed to slightly overshoot its target temperature and then to briefly cut the power to compensate; the temperature versus time curve takes on the appearance of a damped oscillator until the user-inputted temperature is reached. Since Si growth at such low temperatures is thermally activated, small changes in temperature can cause large changes in growth rate, and thus the observed uncertainty is unsurprising. Several modifications to the initial two-step growth procedure were investigated as a means to improve reproducibility.

In the first variant (Process 'B'), the first step after depositing the ε-Ge is, again, to heat the reactor to 450°C while flowing SiH_4. However, once the reactor reaches 450°C, the flow of SiH_4 is simply replaced with Si_2H_6, and the growth of the Si cap is completed without further changes in temperature. As stated above, heating the reactor to 450°C with SiH_4 allows just enough Si growth to passivate the ε-Ge layer and prevent undulation. Since the decomposition rate of SiH_4 for $T = 350\text{-}450$°C is so low, the

variation in Si cap thickness caused by the passivation step is minimal. While the Si growth rate using Si_2H_6 is considerably higher than that with SiH_4, a 3 nm cap still takes ~1.5 hours to grow at 450°C. As seen in Figure 4.7, the reduced time spent simultaneously heating and growing leads to huge improvements in repeatability. However, even layers with very little compressive strain can undulate with sufficient time. While process 'B' was found to produce excellent results with compressively strained alloy channels (e.g. ε-$Si_{0.2}Ge_{0.8}$ on $Si_{0.6}Ge_{0.4}$), when applied to a ε-Ge layer grown on $Si_{0.3}Ge_{0.7}$, the extremely long time at 450°C allowed the surface to undulate very slightly. One application where the long growth time does not pose any problems is in the growth of highly mismatched Si caps directly on Ge-rich virtual substrates (single-channel heterostructures).

In the third procedure (Process 'C') that was developed, the first step of heating the reactor to 450°C in flow of SiH_4 is preserved, but after the standard 20 minute hold period, the reactor is further heated to 500°C. Once 500°C is reached, the flow of SiH_4 is replaced with Si_2H_6, and a 3 nm Si cap can be grown in about 10 minutes. While some uncertainty returns (Figure 4.7), the reproducibility is considerably improved over Process 'A', and compared to Process 'B', the higher Si growth rate precludes the tendency for undulation of the ε-Ge with long growth times. Taken together, it is seen that while low growth rates allow greatest reproducibility, long growth times can lead to instabilities in surface morphology. Thus, the effects of time and temperature must be carefully balanced in order to create an optimum procedure for growing ε-Si / ε-Ge dual-channel heterostructures.

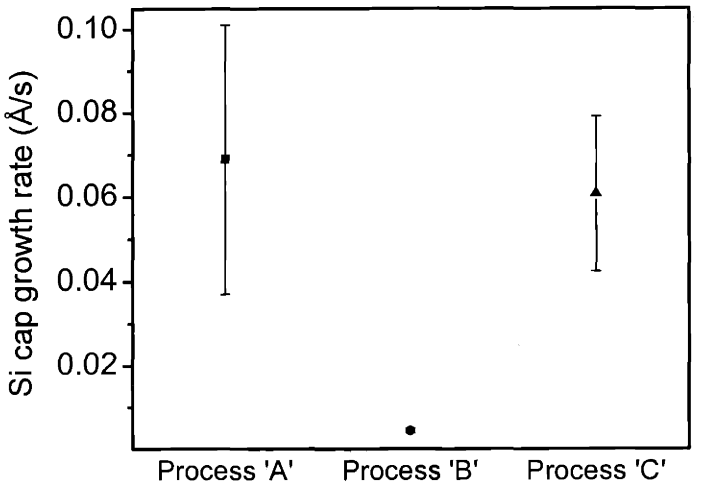


Figure 4.7 Comparison of growth rates along with associated uncertainty for different Si cap growth procedures.

4.4. Summary

The effects of temperature, strain, composition, and thickness on layer morphology have been studied. In general, compressive layers display a much greater tendency to undulate than tensile layers. However, at high enough mismatches and temperatures, ε-Si can grow with a "ramp-like" morphology due to the formation of surface-step bunches. Combining what has been learned about growth morphology with the expanded low temperature capabilities afforded by Si_2H_6, significant improvements have been made in the deposition of highly strained layers, and it is now possible to grow planar ε-Ge at mismatches as high as 2% as well as planar ε-Si on the whole range of relaxed $Si_{1-x}Ge_x$ lattice constants. The stated goal of this thesis since the outset has been to study Ge-rich heterostructures, and therefore the development of effective growth technique was the key factor that enabled the MOSFET results presented over the next several chapters.

Chapter 5. Hole mobility enhancements in nanometer-scale strained-silicon heterostructures grown on Ge-rich relaxed $Si_{1-x}Ge_x$

5.1. Introduction

Although ε-Si *p*-MOSFETs on Si-rich $Si_{1-x}Ge_x$ virtual substrates demonstrate enhanced hole mobility compared to bulk Si devices, the enhancement has widely been observed to degrade at large vertical fields. In this chapter, it is conjectured that the hole wave function in ε-Si heterostructures spreads out over distances of ~10 nm, even at large N_{inv} (i.e. > 10^{13} cm^{-2}), due to the strain-induced reduction of the out-of-plane effective mass. Relevant experimental and theoretical studies supporting this argument are presented. It is further hypothesized that by growing layers thinner than the hole wave function itself, inversion carriers can be forced to occupy and hybridize the valence bands of different materials. In this chapter, hole wave functions with mixed or hybrid character are shown to provide large *p*-channel mobility enhancements that can increase or remain constant (rather than decrease) with gate overdrive.

5.2. Evidence for wave function mixing in single- and dual-channel heterostructures

5.2.1. Single-channel heterostructures

Theoretical studies predict that when biaxial tensile strain in Si exceeds 0.8%, the out-of-plane mobility of holes becomes higher than the in-plane mobility (Figure 5.1), [58,88] resulting in a hole wave function that is substantially spread in the vertical direction. While precise calculations for holes have not been published, experimental studies show that the extent of the *electron* wave function in the vertical direction is approximately 5-6 nm, even at high E_{eff}.[7] Given the high out-of-plane effective mass of electrons in strained Si ($m_l = 0.98m_o$) and the low out-of-plane effective mass of holes ($m_\perp$) in ε-Si

(Fischetti uses $m_\perp = 0.18m_o$[58]), it can be deduced that the hole wave function must extend significantly farther than 5 nm. Thus, for a sufficiently thin Si cap, the valence band of the material *below* the ε-Si can be expected to significantly alter the effective valence band that the hole wave function "sees." Since ΔE_v between the cap and relaxed $Si_{1-x}Ge_x$ increases with x, the relative influence that the relaxed $Si_{1-x}Ge_x$ has on the hole wave function also increases with x.

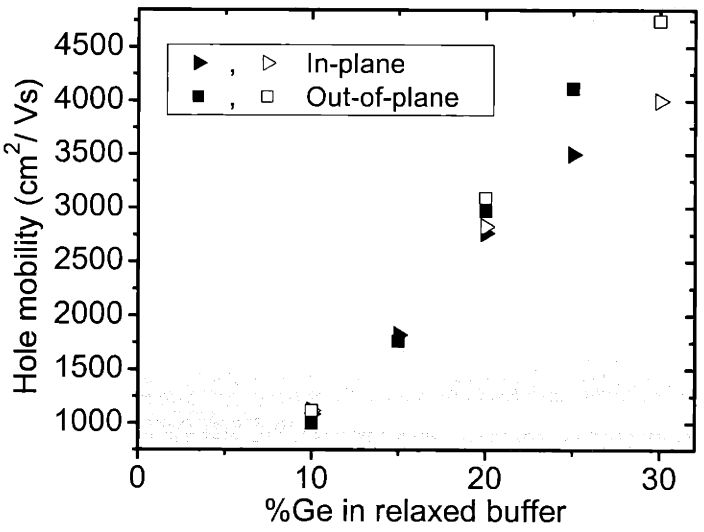


Figure 5.1 Calculated low-field hole mobility of ε-Si for both in-plane and out-of-plane transport. Solid symbols are from Nayak *et al.*[88] and open symbols are from Fischetti *et al.*[58] Both authors predict that the out-of-plane hole mobility exceeds the in-plane hole mobility at around 20% Ge in the buffer.

The impact of virtual substrate composition on hole transport at high vertical fields can be seen in previously published data.[6,7] Figure 5.2 is a log-log plot of μ_{eff} vs N_{inv} in devices on $Si_{1-x}Ge_x$ virtual substrates with $0.1 \leq x \leq 0.5$. For $N_{inv} > 7 \times 10^{12}$ cm^{-2}, μ_{eff} for all devices follows an exponential decay proportional to:

$$\mu_{eff} \propto N_{inv}^{-c} \tag{5.1}$$

The exponent c varies significantly with Ge content in the buffer, as seen in Figure 5.3, indicating that the mechanism that limits hole mobility at large N_{inv} is somehow dependent on the degree of strain in the Si layer, ε_{Si}, or the Ge content of the buffer, x (i.e. $c = f(\varepsilon_{Si}, x)$). For virtual substrates with $x < 0.35$, c increases with Ge content (except for the $Si_{0.9}Ge_{0.1}$ point), which is equivalent to saying that μ_{eff} at large N_{inv} is more sharply limited with higher ε_{Si}. At $Si_{0.65}Ge_{0.35}$, the decay in high-field mobility is most severe, but with still higher Ge content, c decreases and eventually returns to the same value as in Cz-Si p-MOSFETs. Observing the slower exponential decay of the mobility curve for $x > 0.35$ motivates the exploration of hole transport in single-channel heterostructures with even higher Ge content. Understanding the relationship between slope, vertical field, and composition could lead both to an explanation of why previously observed ε-Si p-MOSFETs lose their mobility enhancement at large N_{inv} and to strategies for improving their performance.

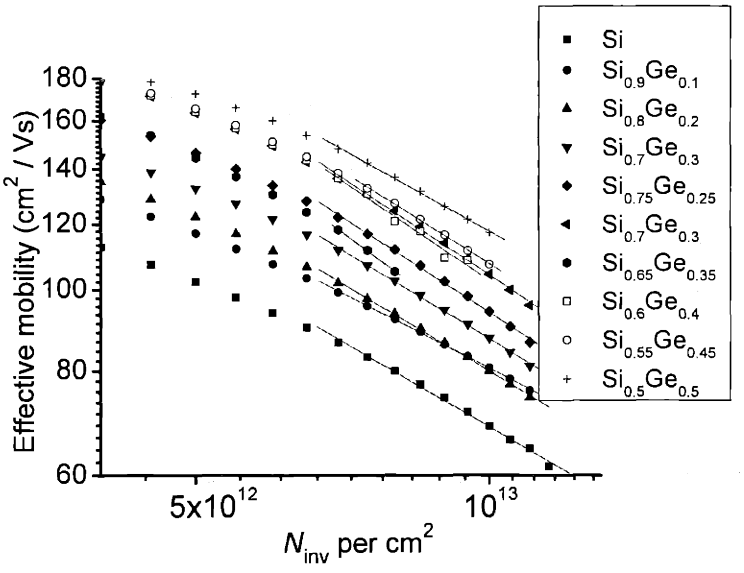


Figure 5.2 Log-log plot of hole effective mobility vs N_{inv} for ε-Si *p*-MOSFETs showing how the exponential decrease in μ_{eff} becomes weaker with higher Ge content in the buffer.

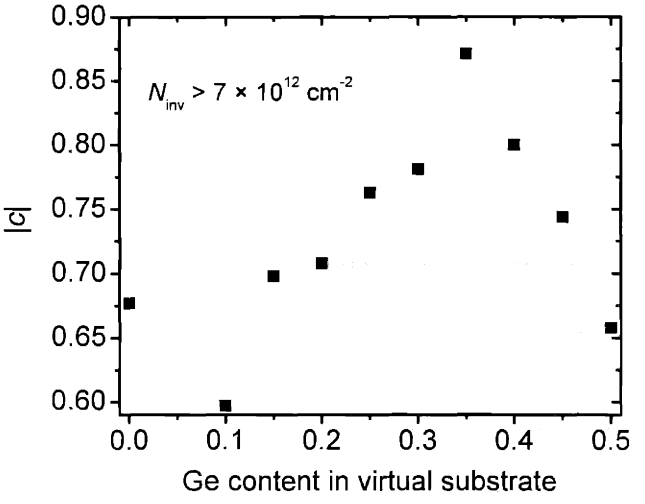


Figure 5.3 *c* vs virtual substrate Ge content, *x*. Up to $x = 0.35$, *c* increases, indicating that high-field mobility decays more quickly than in bulk Si. For $x > 0.35$, *c* starts to decrease back to the value in Cz-Si.

Plotting effective hole mobility vs N_{inv} instead of E_{eff} causes several changes when older data are re-interpreted on a different set of axes. Recall from Chapter 3 that p-MOS mobility enhancement appeared to saturate for growth on $Si_{0.6}Ge_{0.4}$.[6] However, in Figure 5.2, the highest hole mobility is seen in the sample grown on $Si_{0.5}Ge_{0.5}$. Thus the conclusion that intervalley scattering for holes in ε-Si is completely suppressed for $\varepsilon_{Si} >$ 1.6%[6,50] is called into question. The dependence of mobility enhancement on virtual substrate composition is plotted in Figure 5.4 with the enhancements computed in two different ways: in one curve, it is computed at $E_{eff} = 0.6$ MV/cm,[6] and in the other, it is computed at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2} (both calculations correspond to a large gate overdrive). Figure 5.4 indicates that computing the enhancement at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2} lowers the magnitude of the computed enhancements and changes the trend of enhancement vs Ge content in the virtual substrate.

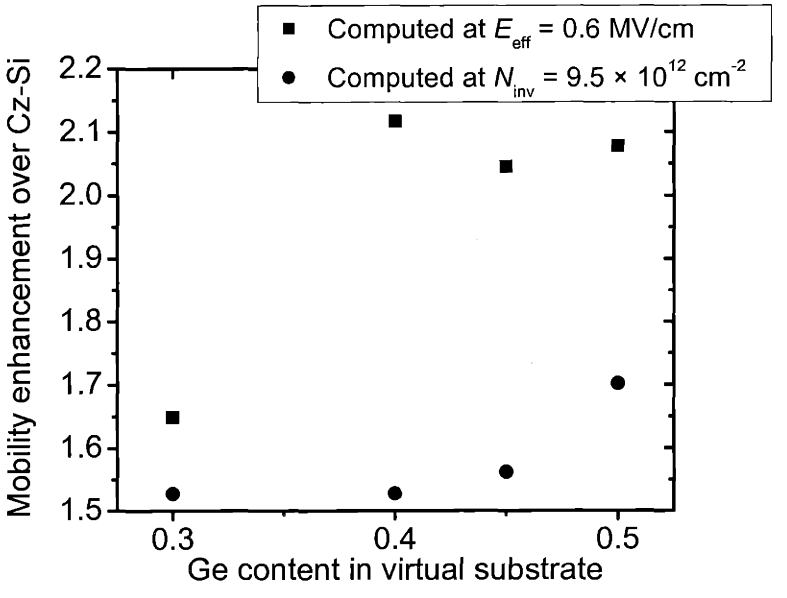


Figure 5.4 Comparison of single-channel hole mobility enhancement when computed using $E_{eff} = 0.6$ MV/cm and $N_{inv} = 9.5 \times 10^{12}$ cm^{-2}.

5.2.2. Dual-channel heterostructures

The high hole mobility of dual-channel heterostructures compared to ε-Si alone (single-channel) is evident from Figure 5.5 (adapted from Leitz *et al.*[54]). For all structures presented in Figure 5.5, the buried ε-$Si_{1-y}Ge_y$ layer has a Ge-content 30% higher than the relaxed $Si_{1-x}Ge_x$ (i.e. $y - x = 0.3$), and the ε-Si cap ranges in thickness from 7-9 nm. Thus, the compressive strain in the $Si_{1-y}Ge_y$ is held constant, but the Ge-content in the different samples varies. Because the buried compressive layer possesses a higher hole mobility than the ε-Si surface layer, mobility enhancements in dual-channel heterostructures invariably decrease as a function of gate overdrive. However, dual-channel heterostructures demonstrate higher mobility than their single-channel counterparts even at large N_{inv}, when holes are expected to populate the Si surface layer (all mobility enhancements for Figure 5.5 were computed at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2}). Although the distribution of holes shifts toward the surface with large V_{GT}, a 7-9 nm ε-Si cap does not completely isolate the hole wave function from the buried layer. For higher vertical fields, the hole wave function could theoretically be confined solely to the ε-Si cap, losing all contact with the buried layer. However, this effect could be countered by decreasing the thickness of the ε-Si cap, bringing the buried layer closer to the surface.

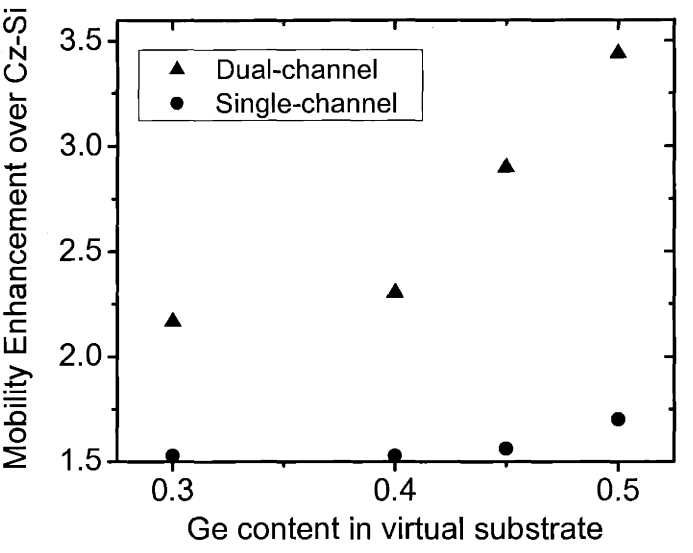


Figure 5.5 Hole mobility enhancement over Cz-Si at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2} for single-channel and dual-channel devices grown on the same relaxed buffer compositions. The buried layer Ge content in all of the dual-channel devices is 30% higher than the relaxed buffer. Even at high N_{inv}, dual-channel devices demonstrate far larger enhancements than single-channel devices. Dual-channel data adapted from Leitz *et al.*[54]

5.2.3. Motivation for using thin ε-Si layers on Ge-rich $Si_{1-x}Ge_x$

Previous work with single-channel ε-Si *p*-MOSFETs has focused on $Si_{1-x}Ge_x$ buffers with $x < 0.5$, and the relaxed $Si_{1-x}Ge_x$ was seen as a low mobility parasitic channel.[5,52,88] In most of these *p*-MOSFETs, the ε-Si channel was 15-25 nm thick, allowing the hole wave function to nearly completely occupy the surface layer at high E_{eff}. The survey of experimental and theoretical *p*-MOS results presented here shows that the hole wave function in single- and dual-channel devices can be on the order of 10 nm wide, even at large N_{inv}. In this work, rather than isolating the hole from the relaxed buffer, thin (< 5 nm) ε-Si layers were grown upon $Si_{1-x}Ge_x$ virtual substrates with $x \geq 0.6$, allowing the hole wave function to hybridize the valence band of a Ge-rich alloy with the

non-degenerate ε-Si valence band. In this chapter, single-channel ε-Si *p*-MOSFETs where the hole mobility enhancement increases with gate bias are demonstrated, reaching values as high as 2.9 at $N_{inv} = 1.35 \times 10^{13}$ cm^{-2}. A digital-alloy channel *p*-MOSFET that exhibits nearly constant mobility enhancement over a wide range of N_{inv} is also described.

5.3. Thin ε-Si grown on Ge-rich buffers- single-channel *p*-MOSFET results

Thin single channels of ε-Si were deposited on relaxed $Si_{0.4}Ge_{0.6}$ and $Si_{0.3}Ge_{0.7}$ in order to study the effects of mixed hole transport through both the cap and buffer. Due to the high Ge content of the $Si_{1-x}Ge_x$ relaxed buffers used, ε-Si growth had to be carried out at low temperature (400-550°C) in order to avoid strain-induced undulations.[89] The planarity of all heterostructures was confirmed by XTEM. For the sample on $Si_{0.4}Ge_{0.6}$, the as-grown ε-Si cap thickness was 3 nm. However, in the $Si_{0.3}Ge_{0.7}$ case, the ε-Si cap was 4.5 nm as-grown and subsequently reduced to 3 nm by dry oxidation at 600°C prior to device processing (all layer thicknesses are ±0.5 nm). Recall that cleaning steps preceding the deposition of the gate stack remove another 0.5 to 1 nm of ε-Si, leaving a final ε-Si thickness of approximately 2 to 2.5 nm for both cases.

Figure 5.6(a) plots μ_{eff} vs N_{inv} for these two devices compared with a Cz-Si control device. When the inversion layer first forms at low gate overdrive, the wave function is weighted towards the Ge-rich relaxed buffer due to the type-II valence band offset between $Si_{1-x}Ge_x$ and ε-Si. Around $N_{inv} = 4 \times 10^{12}$ cm^{-2}, both devices exhibit moderate mobility enhancements over Cz-Si devices, but the ε-Si on $Si_{0.4}Ge_{0.6}$ device exhibits a 23% lower μ_{eff} than the ε-Si on $Si_{0.3}Ge_{0.7}$ device. Since hole transport in

moderate inversion is dominated by the characteristics of the relaxed $Si_{1-x}Ge_x$, the higher μ_{eff} observed in the 70% Ge case confirms that the hole effective mass decreases significantly as x is increased towards the L-valley crossover point (i.e. the holes in the $Si_{1-x}Ge_x$ begin to resemble the holes in Ge).

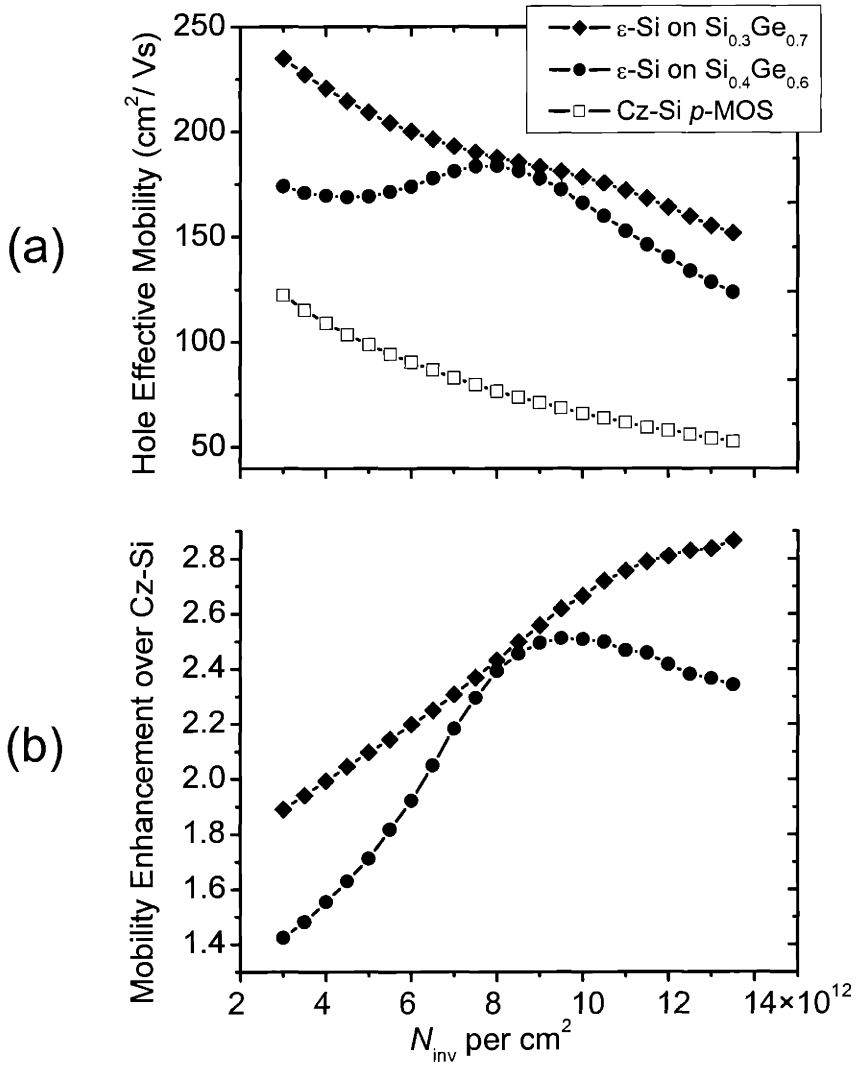


Figure 5.6 (a) Hole effective mobility vs N_{inv} of single-channel devices grown upon relaxed $Si_{0.3}Ge_{0.7}$ and $Si_{0.4}Ge_{0.6}$. (b) Mobility enhancement vs N_{inv} of the same two devices. The enhancement of the device on $Si_{0.3}Ge_{0.7}$ never degrades in the testable range of N_{inv}.

With increased gate overdrive, the hole wave function in both single-channel heterostructures is pulled closer and closer to the surface, and the hole begins to sample the valence band splitting in the ε-Si cap. Since the hole wave function occupies both the cap and the buffer, a hybrid valence band that mixes the low effective mass of the Ge-rich relaxed $Si_{1-x}Ge_x$ with the valence band splitting in the ε-Si is formed. The influence of the ε-Si causes μ_{eff} in the 60% Ge device to *increase* slightly with N_{inv} in the range of 4 to 8×10^{12} cm^{-2}. The combination of a Ge-like effective mass with strain-induced valence band splitting leads to effective hole mobilities that are considerably larger than those seen in prior single-channel *p*-MOSFETs[5-7,51] and comparable to those seen in dual-channel *p*-MOSFETs.[54] By combining the properties in the relaxed buffer and the tensile cap, the valence band that the wave function experiences in these single-channel devices resembles that of a compressed $Si_{1-y}Ge_y$ alloy. Figure 5.7 is a schematic diagram illustrating how large gate overdrives cause the holes to occupy a "virtual ε-$Si_{1-y}Ge_y$ channel." With more of the wave function in the relaxed $Si_{1-x}Ge_x$ buffer (i.e. low V_{GT}), the effective Ge content y of the virtual channel is close to x, and the amount of valence band splitting is low. Conversely, with more of the wave function in the ε-Si cap (i.e. high V_{GT}), y decreases and the valence band splitting increases. The peak enhancement of 2.5 [Figure 5.6(b)] at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2} in the 60% Ge device represents the optimum combination of ε-Si and $Si_{0.4}Ge_{0.6}$ valence band characteristics. A slight loss in mobility enhancement takes place at $N_{inv} > 1 \times 10^{13}$ cm^{-2} for this device, echoing the trend seen in prior single-channel *p*-MOSFETs.

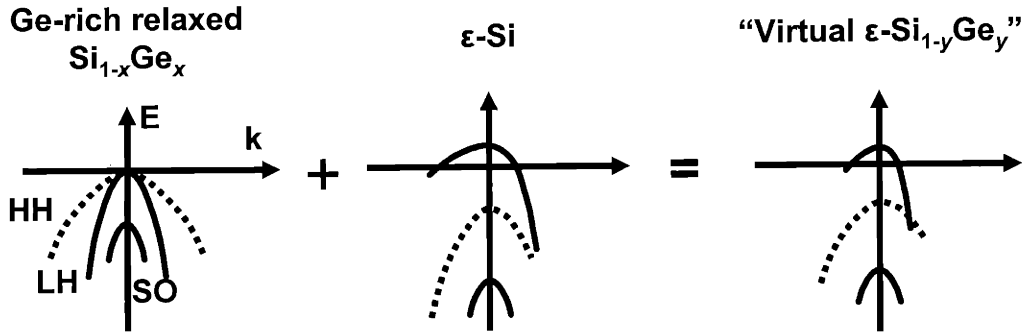


Figure 5.7 Schematic band diagram showing formation of the proposed hybrid valence band. Large gate overdrives cause the hole wave function to sample both the relaxed buffer and the ε-Si cap, resulting in a valence band that resembles compressively strained $Si_{1-y}Ge_y$.

One possible reason for the loss of mobility enhancement seen in ε-Si *p*-MOSFETs at large gate overdrives may be the penetration of the hole wave function into the gate SiO_2. While I have commented extensively on the low out-of-plane effective mass[90] of holes in ε-Si and the effect of the hole wave function sampling layers below the channel, I have, up to this point, ignored the detrimental effect of the hole wave function spreading into layers *above* the channel. The traditional boundary conditions for the solution of Schrödinger's equation in MOS inversion layers are that the wave function goes to zero both at the Si/SiO_2 interface and deep in the bulk of the semiconductor (i.e. $\Psi(0) = 0$ and $\Psi(\infty) = 0$, where $z = 0$ corresponds to the Si/SiO_2 interface).[91] However, the effects of the wave function penetrating the gate dielectric in deep-submicron bulk Si devices are coming under increased scrutiny due to the need for accurate calculations of direct tunneling currents through extremely thin gate oxides.[92,93] The slight shift (on the order of several angstroms) of the wave function towards the gate electrode also alters the inversion capacitance, leading to errors in the *C-V* extraction of device characteristics such as t_{ox} and N_{inv}.[94-96] The out-of-plane effective mass of carriers

is one of the primary variables in determining the extent of wave function penetration as was illustrated in a computational study of *n*-MOSFETs fabricated on (111) and (100) Si substrates.[97] In the (111) case, all of the cigar-shaped valleys are oriented in the same direction with respect to the vertical field, and the out-of-plane effective mass of all inversion electrons is $0.258m_o$.[91] However, in the (100) case, the majority of electrons populate the two transverse valleys in strong inversion,[98] and thus the out-of-plane effective mass is closer to $m_l = 0.98m_o$.[97] Since $0.258m_o$ is considerably less than m_l, simulations of Yunus *et al.*[97] showed that the upward shift of the centroid of charge towards the Si/SiO_2 interface was significantly larger for the (111) case.

Of course, due to the thick gate oxide used in the short-flow devices analyzed in this thesis, direct tunneling current through the gate is virtually impossible, and any error in the *C-V* estimation of the gate oxide thickness would be negligible. However, the rapid decrease in hole mobility (compared to bulk Si) at high vertical fields in ε-Si *p*-MOSFETs (Figure 5.2) could be caused by greater penetration of the wave function into the oxide. Since the out-of-plane effective mass of holes in ε-Si is lower than that in bulk Si, the degree of tunneling must be greater for ε-Si. According to the universal model, carrier mobility at high vertical field in bulk Si *p*- and *n*-MOSFETs is dominated by surface-roughness scattering. However, a paper by Polishchuk and Hu suggests that the mobility at high field may be better corroborated by the wave function penetrating into the dielectric and electrons experiencing the extremely low carrier mobility within the oxide; as vertical field increases, the number of carriers within the oxide increases, and the overall mobility drops (they found $\mu_{e,\,ox} = 0.12\ cm^2 / Vs$).[99] In another paper, researchers attempted to study the effect of intentional surface roughening on transport by

comparing *n*-MOSFETs fabricated on prime wafers with wafers that were subjected to 30 hours of rinsing in de-ionized water prior to gate oxidation. Despite an increase in surface-roughness from 0.5 nm to 2 nm caused by the rinsing process, no difference in high-field electron mobility was observed.[100]

Due to the fact that the barrier height for holes at the Si/SiO_2 interface is much larger for holes than for electrons (4.8 eV for holes, 3.1 eV for electrons), the penetration of the wave function is usually not considered in *p*-MOSFETs.[95] However, theoretical studies suggest that the centroid of charge in *p*-MOSFETs may shift into the oxide *more* than in *n*-MOSFETs. This unexpected result was explained by the fact that the out-of-plane effective mass of the dominant type of holes ($m_{HH} = 0.29m_o$) is lower than the effective mass of the dominant electrons ($m_l = 0.98m_o$);[95] again, the out-of-plane effective mass was shown to be the primary variable in determining the degree of wave function penetration. Since tensile strain considerably lowers the out-of-plane effective mass of holes, wave function penetration in ε-Si *p*-MOSFETs may be even more severe. In some cases, ε-Si *p*-MOSFETs have been observed to show lower μ_{eff} than Cz-Si devices at large gate overdrives.[51,52] ε-Si *n*-MOSFETs are immune from such an effect, because the out-of-plane effective mass of electrons in ε-Si is so large ($m_l = 0.98m_o$), and the resulting wave function is small in the vertical direction.

Figure 5.3 showed that the exponent of mobility degradation for $N_{inv} \geq 7 \times 10^{12}$ cm^{-2} exhibits a saddle-shaped dependence on x, with 35% Ge exhibiting the sharpest mobility degradation. The initial increase in c could be caused by the strain-induced decrease in the out-of-plane effective mass;[90] as the out-of-plane effective mass decreases, the potential for the wave function to penetrate into the gate oxide increases.

The drop in c at high Ge content might then be caused by the increasing ΔE_v between the ε-Si cap and the relaxed $Si_{1-x}Ge_x$; increasing the barrier to occupation in the ε-Si cap decreases the rate at which the hole wave function can penetrate into the oxide at high vertical field. This type of consideration could explain the mobility characteristics of the device grown on $Si_{0.3}Ge_{0.7}$, where the mobility enhancement never saturates in the testable range of the device, reaching a value of nearly 2.9 at $N_{inv}=1.35 \times 10^{13}$ cm^{-2} [Figure 5.6(b)]. The large ΔE_v and thin Si cap keep the centroid of charge below the surface, preventing the wave function from tunneling into the SiO_2. If larger gate overdrives could be achieved with our measurement setup, it can be anticipated that the mobility enhancement would eventually peak and decline just as in other cases where the hole wave function began to sample the gate dielectric. Alternately, in the limit of a much thicker Si cap, the wave function in the $Si_{0.3}Ge_{0.7}$ would go to zero, and the valence band offset would not be able to prevent the wave function from tunneling into the SiO_2.

A general strategy for maintaining high hole mobility in ε-Si p-MOSFETs might then be to prevent the hole wave function from ever penetrating into the amorphous SiO_2. One way of accomplishing this would be to avoid high vertical fields in the channel entirely through the implementation of a dual-gate device with an ultra-thin body. However, this work shows that an equivalent method of screening inversion charge from the oxide is to use a Ge-rich virtual substrate and a thin Si cap. The high barrier between the ε-Si and Ge-rich $Si_{1-x}Ge_x$ creates an additional barrier that prevents the hole wave function from entering the oxide. Additionally, the combination of band splitting in the ε-Si and low effective mass in the Ge-rich $Si_{1-x}Ge_x$ creates a band structure that simulates a compressively strained $Si_{1-y}Ge_y$ alloy without ever growing one; from the single-carrier

perspective, the uncertainty principle ensures that over time, the hole will experience both the valence band of the ε-Si and the Ge-rich $Si_{1-x}Ge_x$. A corollary of the theory presented here is that growing ε-Si on virtual substrates with even higher Ge content should allow even greater hole mobility enhancement at high vertical fields. However, I have found the mobility of single-channel heterostructures grown on $Si_{0.2}Ge_{0.8}$ as well as on pure Ge ($h_{Si\ cap}$ = 4 nm for both) to be depressed compared to $Si_{0.3}Ge_{0.7}$ (although significant enhancement was evident). The rampant dislocation nucleation that takes place for such high film/substrate mismatches is likely to be responsible for the lower hole mobility enhancements. Decreasing $h_{Si\ cap}$ to 2 nm might reduce the dislocation density sufficiently to allow higher mobility to be attained. Figure 5.8 plots hole mobility enhancement (computed at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2}) versus Ge content in the virtual substrate, and effectively summarizes all of the research done on single-channel ε-Si *p*-MOSFETs in the Fitzgerald group.

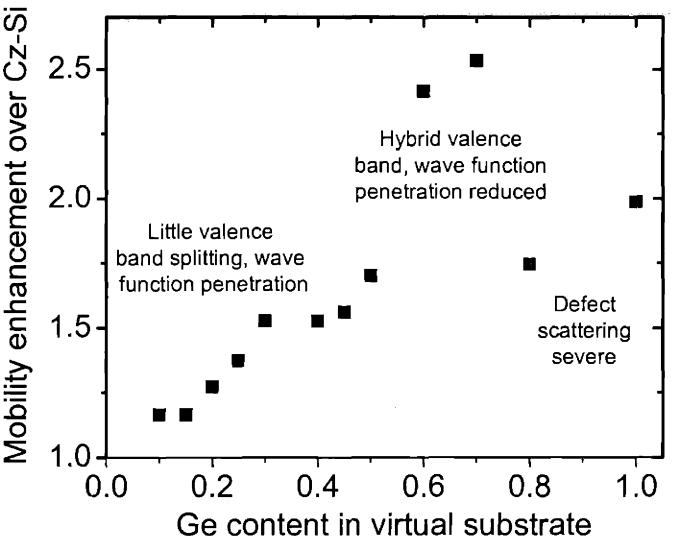


Figure 5.8 Hole mobility enhancement at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2} vs virtual substrate Ge content for single-channel *p*-MOS. For Si-rich buffers, little to no enhancement is observed due to the small valence band splitting and penetration of the hole wave function into the gate oxide. At $0.5 < x < 0.7$ enhancements are maximum, and at $x \geq 0.8$, defects limit the hole mobility.

5.4. Comparison of single-channel and dual-channel heterostructure *p*-MOSFETs

A direct comparison between the single-channel ε-Si device grown on relaxed $Si_{0.3}Ge_{0.7}$ from this study with a previously published dual-channel device[54] highlights the differences and similarities in the two configurations while summarizing the above discussion. The dual-channel heterostructure in consideration has a 7 nm ε-Si cap with an 8 nm buried $Si_{0.3}Ge_{0.7}$ layer grown on relaxed $Si_{0.6}Ge_{0.4}$. Figure 5.9(a) shows that for low N_{inv}, the dual-channel device displays much larger μ_{eff} than the single-channel device. Hole transport in both devices is dominated by the $Si_{0.3}Ge_{0.7}$, but the dual-channel device possesses the additional benefits of compressive strain. As N_{inv} is increased, the wave function is pulled into the ε-Si cap in both devices. Since the hole

wave function in the dual-channel device starts to lose contact with the buried layer, mobility enhancement decreases rapidly in response [Figure 5.9(b)]. In contrast, the mobility enhancement in the single-channel device increases because of the gradual addition of valence band splitting from the ε-Si cap to the degenerate $Si_{0.3}Ge_{0.7}$ valence band (i.e. formation of the "virtual ε-$Si_{1-y}Ge_y$ channel"). At $N_{inv} = 9 \times 10^{12}$ cm^{-2}, the mobility enhancement of the single-channel device actually exceeds that of the dual-channel device. This is caused by the fact that the ε-Si cap in the dual-channel device is ~7 nm thick, more than twice as thick as the single-channel device. Therefore, the dual channel device begins to resemble a single-channel device at large N_{inv} as the ε-$Si_{0.3}Ge_{0.7}$ layer gets screened out by the inversion charge in the ε-Si cap. Returning to the single-channel device on relaxed $Si_{0.3}Ge_{0.7}$, the ~2.5 nm ε-Si cap causes the hole wave function to stay in intimate contact with the relaxed $Si_{0.3}Ge_{0.7}$ buffer, allowing the hole to experience a Ge-rich valence band even for very large N_{inv}. A dual-channel device with a thinner ε-Si cap than the one shown here would be expected to maintain larger mobility enhancements at high N_{inv}.

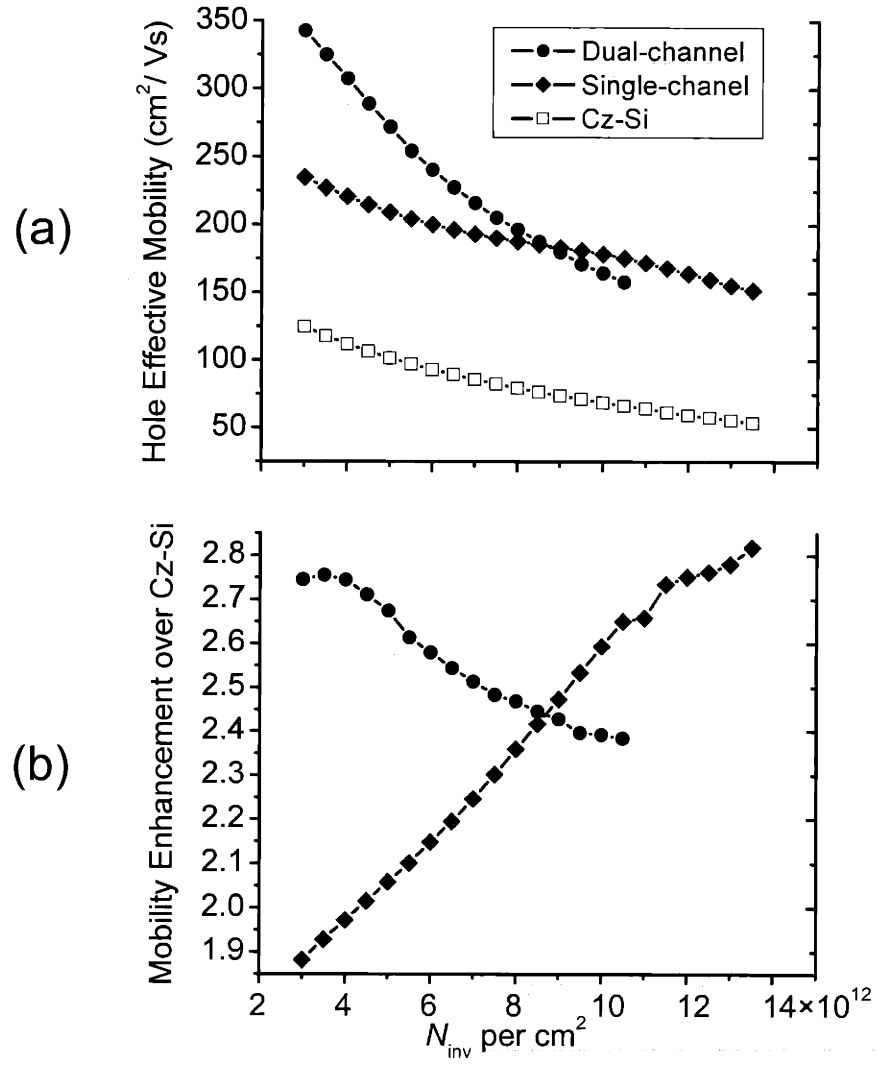


Fig. 5.9 (a) Hole effective mobility vs N_{inv} of a single-channel device grown upon relaxed $Si_{0.3}Ge_{0.7}$ and a dual-channel device with a ε-$Si_{0.3}Ge_{0.7}$ buried layer grown upon $Si_{0.6}Ge_{0.4}$. (b) Mobility enhancement vs N_{inv} of the same two devices. For $N_{inv} > 9 \times 10^{12}$ cm^{-2}, μ_{eff} of the single-channel device exceeds that of the dual-channel device.

5.5. Controlling the variation of hole mobility enhancement using ε-Si / relaxed $Si_{1-x}Ge_x$ digital-alloy channels

A final structure was grown on relaxed $Si_{0.3}Ge_{0.7}$, except that instead of simply capping the buffer with ε-Si, a digital alloy consisting of seven periods of alternating ε-Si

and relaxed $Si_{0.3}Ge_{0.7}$ layers was deposited. Each layer is approximately 8 Å thick, and the entire structure is capped with 2 nm of ε-Si in order to allow the use of an SiO_2 gate. Cross-sectional TEM done before and after device processing shows that the low thermal budget of our short-flow MOSFET process allowed the interfaces between layers to remain sharp, as seen in Figure 5.10(a) (a single-channel heterostructure is presented in Figure 5.10(b) for comparison).

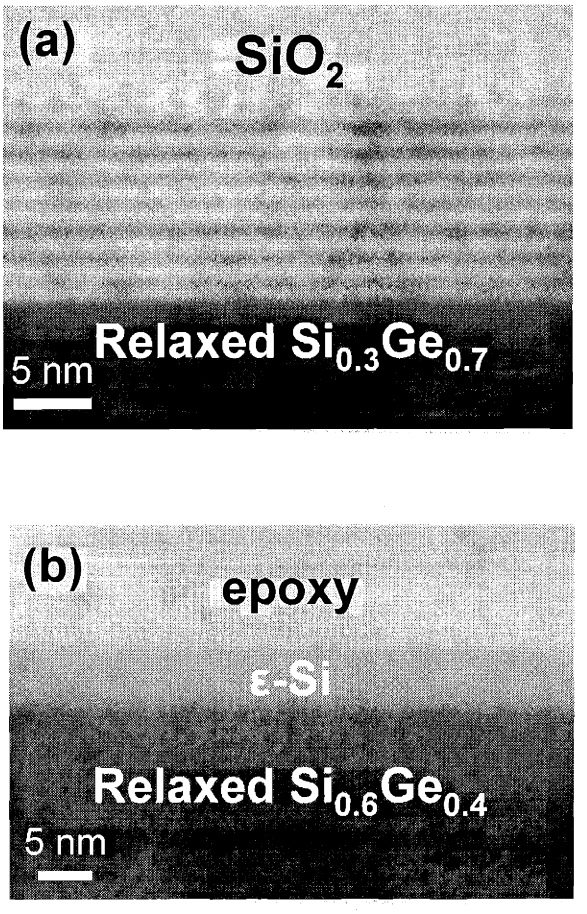


Figure 5.10 (a) XTEM of the digital-alloy channel *p*-MOSFET after device processing. Dark layers are relaxed $Si_{0.3}Ge_{0.7}$ and bright layers are ε-Si. (b) XTEM of a single-channel ε-Si heterostructure provided for comparison.

As discussed, mobility enhancements at low fields are dominated by transport in the relaxed $Si_{1-x}Ge_x$ for conventional single-channel devices, provided that the ε-Si cap is not so thick that the buffer is screened from the gate electrode entirely. In contrast, when

the inversion layer forms in the digital-alloy channel, holes experience valence band splitting even at low N_{inv} due to the numerous ε-Si layers below the surface. As the hole wave function is pulled towards the surface by the gate bias, the effective valence band that it experiences changes very little; each period of the digital alloy is a microcosm of the hybrid valence band that the single-channel device can only experience at high field. Since the valence band characteristics do not vary with vertical position, the mobility enhancement in the digital-alloy channel demonstrates little dependence on N_{inv} (Figure 5.11). Thus, unlike the single-channel ε-Si *p*-type MOSFET described above, the digital-alloy device does not need to be biased into strong inversion in order for the hole wave function to combine reduced intervalley scattering with the Ge-like effective mass of $Si_{0.3}Ge_{0.7}$. The digital-alloy device does not attain mobility enhancements as high as the single-channel device grown on 70% Ge, because the buried ε-Si layers dilute the effective Ge content of the virtual channel. Growing a digital alloy with thicker layers of $Si_{0.3}Ge_{0.7}$ would increase the effective Ge content, allowing the mobility enhancement to be fixed at a higher value. In nanometer-scale ε-Si heterostructures, N_{inv} and layer structure are interchangeable variables that govern the characteristics of the hybrid valence band and resulting hole mobility enhancements. For example, at a given value of N_{inv}, layer thickness and strain state determine the shape of the hybrid valence band that the hole wave function experiences. Analogously, for a given layer structure, N_{inv} determines which layers are being sampled by the hole wave function.

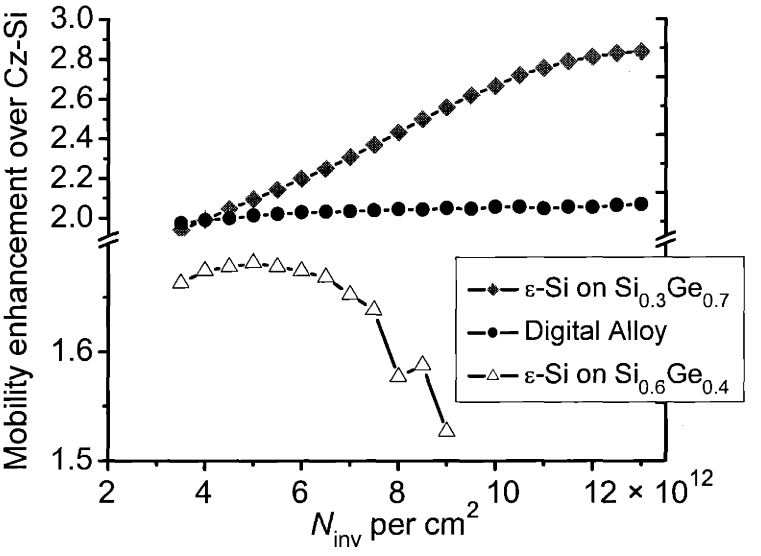


Figure 5.11 Hole mobility enhancement over Cz-Si for the digital-alloy channel device compared with other single-channel devices. The enhancement of the digital-alloy channel device changes only slightly over a wide range of N_{inv}.

Despite the fact that the individual layers of the digital alloy are extremely thin, holes do not experience the valence band of a random alloy. If the valence band structure were equivalent to the average composition of the layers taken as a whole, then the digital alloy could simply be replaced by a tensile $Si_{0.65}Ge_{0.35}$ layer on a relaxed $Si_{0.3}Ge_{0.7}$ buffer. However an experimental study of alloy scattering showed that tensile-strained random $Si_{1-x}Ge_x$ alloys with $x = 0.3$ to 0.4 have low hole mobilities resembling those in Cz-Si (Figure 5.12).[6] These random alloy channels are not equivalent to the digital-alloy channel discussed above; each material in the digital alloy must exist independently so that the hole can mix the valence band properties of each. The high mobility of the digital-alloy channel also indicates that removing randomness in one dimension (i.e. the digital alloy is ordered perpendicular to the growth surface) greatly reduces the $Si_{1-x}Ge_x$ alloy scattering potential for transport parallel to the surface.

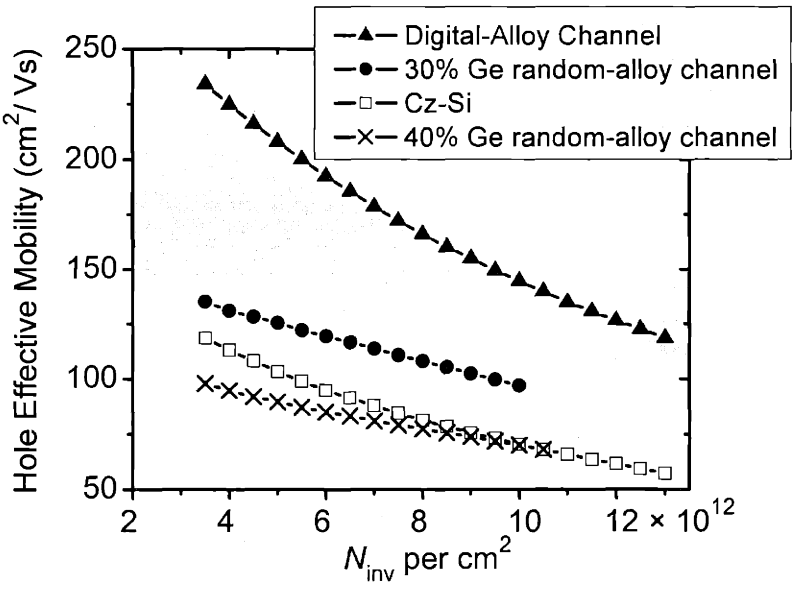


Figure 5.12 Hole effective mobility vs N_{inv} of the digital-alloy channel device compared with random-alloy channels of various compositions. The random-alloys are under 1% tensile strain, and all channels are capped with 2 nm Si layers in order to allow the use of an SiO_2 gate dielectric. Random-alloy data is adapted from Leitz *et al.*[6]

5.6. Summary

The high out-of-plane mobility of holes in strained heterostructures grown on relaxed $Si_{1-x}Ge_x$ causes the hole wave function to spread significantly below the surface; direct evidence of a hole wave function extending > 8 nm below the surface even at very large N_{inv} can be seen in dual-channel *p*-MOSFETs. The size and position of the hole wave function in the *out-of-plane* direction are therefore key factors in determining the *in-plane* mobility of holes. The results presented here indicate that the use of a Ge-rich virtual substrate and a thin ε-Si cap allow the formation of a hybrid valence band structure at large N_{inv} where the ε-Si surface layer contributes valence band splitting and the relaxed buffer contributes a low, Ge-like effective mass. The steady increase in enhancement seen over a wide range of N_{inv} in the novel single-channel devices presented

here represents the gradual formation of a "virtual $Si_{1-y}Ge_y$" channel. In this way, the large hole mobility enhancements seen in dual-channel heterostructures can be attained in a single-channel heterostructure; devices with a thin ε-Si single-channel grown on relaxed $Si_{0.3}Ge_{0.7}$ demonstrate hole mobility enhancements that increase with gate overdrive, peaking at a value of nearly 3 times. Re-analysis of previously published mobility data suggests that the penetration of the hole wave function into the gate oxide may be responsible for the rapid drop in mobility that takes place at large N_{inv} in ε-Si *p*-MOSFETs grown on Si-rich virtual substrates. The use of a thin ε-Si cap on a Ge-rich virtual substrate prevents this phenomenon. A digital-alloy channel consisting of alternating layers of relaxed $Si_{0.3}Ge_{0.7}$ and ε-Si upon a $Si_{0.3}Ge_{0.7}$ buffer was grown to study how the hole wave function responds to nanometer-scale composition/band structure modulation. The periodic layer structure allows the hole wave function to experience a hybrid valence band structure for a wide range of N_{inv}. Digital-alloy channel *p*-MOSFETs therefore behave more like ε-Si *n*-MOSFETs in that their mobility enhancement shows little dependence on gate overdrive; a nearly constant mobility enhancement of 2.0 was observed over inversion densities ranging from 3 to 14×10^{12} cm^{-2}. Even though the layers comprising the digital-alloy are on the order of several atomic layers thick, the hole does not suffer from the depressed mobility seen in random-alloy channels and instead combines the unique benefits intrinsic to each component of the layer structure. For strained-layer *p*-MOSFETs, any layer within 10-15 nm from the SiO_2/Si interface (including, possibly, the SiO_2 itself) is part of a hybrid virtual channel for inversion carriers due to the large hole wave function. In this chapter, I have demonstrated that by engineering the virtual channel, *p*-MOSFET mobility enhancements

that increase, decrease, or even remain constant can be realized. Although the present results are interpreted by solely considering the band structure and layer effects on a single carrier, it must be stated that transport in inversion layers, and particularly mobility, depends on many more factors such as carrier-carrier screening.[101] Still, single particle considerations appear to provide an excellent empirical guideline for designing layer structures and attaining desirable mobility characteristics.

Chapter 6. Strained Si / Strained Ge dual-channel heterostructure *p*-MOSFETs

Single-channel and dual-channel *p*-MOSFETs grown on $Si_{1-x}Ge_x$ provide large hole mobility enhancements that scale upwards with increasing Ge content. The trend seen in all of the prior work indicates that compressively strained Ge will lead to the highest *p*-channel mobilities. In contrast to previous work with ε-Ge FETs reviewed in Chapter 3, all of the heterostructures investigated in this thesis are capped with epitaxial ε-Si in order to allow the use of an SiO_2 gate dielectric.[102,103] The ε-Ge channel devices presented here are enhancement-mode *p*-MOSFETs that can be operated at very large inversion charge densities, and thus should not be confused with conventional buried-channel FETs that are typically limited to small gate voltages. Note also that the hole mobilities reported in this chapter are not deduced from Hall measurements, but are instead true MOSFET effective mobilities extracted from the linear-regime transfer characteristics. In this sense, the mobility enhancements shown here are actually long-channel drain-current enhancements. In this chapter, the effects of all major growth variables on hole mobility in ε-Si / ε-Ge dual-channel heterostructures are presented, including Si and Ge layer thickness ($h_{Si\ cap}$ and h_{Ge}, respectively) and strain level. Just as was the case in single-channel devices, the tendency of the hole wave function to spread out in the vertical direction has significant consequences for in-plane hole transport.

6.1. Effect of compressive strain on hole mobility

Figure 6.1 plots hole effective mobility vs N_{inv} for ε-Si / ε-Ge dual-channel heterostructures grown on $Si_{1-x}Ge_x$ virtual substrates with $x = 0.5$ to 1.0, or equivalently, compressive strains in the Ge channel ranging from 2% down to 0%. The strain is not the only parameter that differs in these devices, and Table 6.1 shows that the layer thicknesses vary significantly from sample to sample. The mobility results are therefore

"contaminated" by the variation of other factors, and it is impossible to deduce a precise relationship between hole mobility and strain. In general, it is still evident that high compressive strain levels favor high hole mobility due to the increased splitting of the LH, HH and SO bands, as well as the strain-induced reduction of the effective mass.

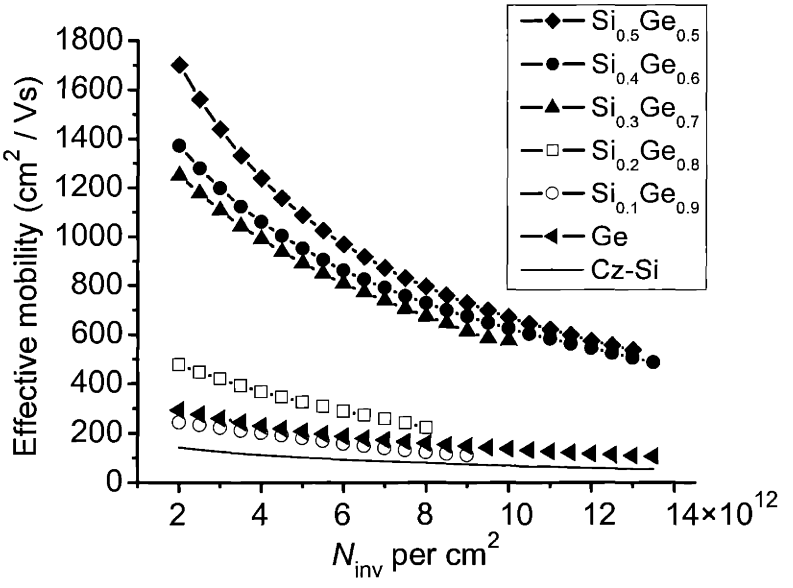


Figure 6.1 Hole effective mobility vs N_{inv} of ε-Si / ε-Ge dual-channel heterostructures grown on $Si_{1-x}Ge_x$ virtual substrates with x = 0.5 to 1.0.

Virtual Substrate Ge content	% Compressive strain in Ge layer	$h_{Si\ cap}$	h_{Ge}
50	2	4 nm	12 nm
60	1.6	3 nm	8 nm
70	1.2	4.5 nm	12 nm
80	0.8	6 nm	14 nm
90	0.4	8 nm	15 nm
100	0	4 nm	N/A

Table 6.1 As-grown layer thicknesses for the heterostructures plotted in Figure 6.1. Layer thicknesses are ±0.5 nm

Figure 6.2 shows that the experimentally extracted hole mobility enhancements (computed at $N_{inv} = 8 \times 10^{12}$ cm^{-2}) are in relatively close agreement with those predicted by Fischetti and Laux for compressive strains ranging from 1.2 to 2%.[58] For the highest compressive strains, both the experimental and theoretical results show that the rate of change of the in-plane hole mobility with strain drops. Fischetti and Laux attribute the reduced slope to a saturation in the splitting of the valence bands.[58] Of course, the experimental mobility data represents an overall inversion layer hole mobility that, in addition to the holes in the ε-Ge, also includes carriers sampling the ε-Si cap and even the relaxed $Si_{1-x}Ge_x$ buffer. Since the theoretical calculations only consider low-field phonon-limited mobility and do not take extrinsic factors into account, the experimental enhancements are expected to be lower. For compressive strains less than 1%, Figure 6.2 shows that the experimentally measured hole mobility enhancement degrades substantially. In Chapter 4, it was hypothesized that the drop in hole mobility enhancement for single-channel devices grown on $Si_{1-x}Ge_x$ virtual substrates with $x \geq 0.8$ was caused by defect scattering in the Si cap. Likewise, the hole mobility in the ε-Si / ε-Ge dual-channel devices fabricated on relaxed $Si_{1-x}Ge_x$ with $x \geq 0.8$ is also limited by defect scattering in the Si cap. Note that the slight improvement in experimental hole mobility enhancement as strain is decreased from 0.4% to 0% does not necessarily corroborate the theoretical calculations, since the Si cap thicknesses vary by a factor of two in these samples. The effect of the Si cap is described in greater detail in Section 6.3.

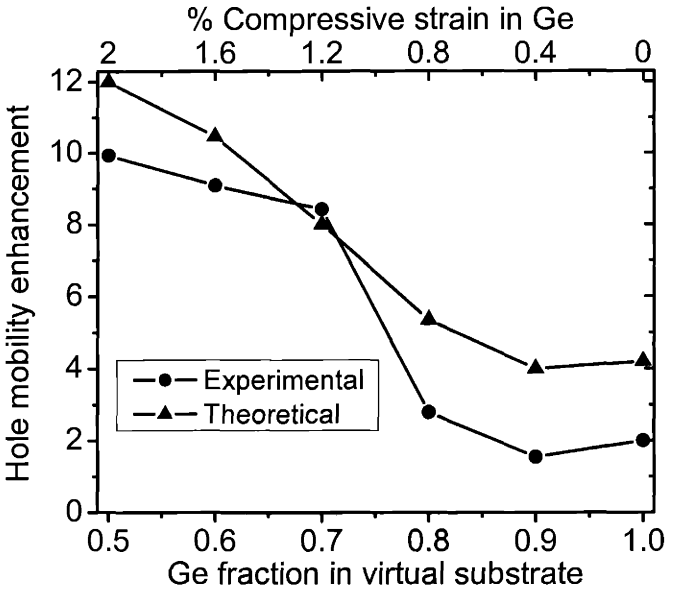


Figure 6.2 Experimental and theoretical[58] hole mobility enhancements vs compressive strain in the ε-Ge layer. Experimental enhancements computed at N_{inv} = 8 × 10^{12} cm^{-2}. For high strains, the experimental results track the theoretical results quite closely. For low strains, defect scattering in the Si cap significantly degrades the experimentally observed enhancements.

6.2. Effect of ε-Ge thickness

To study the effect of ε-Ge thickness on hole mobility, h_{Ge} was varied while holding the strain level and $h_{Si\ cap}$ constant. Due to the uncertainties in growth rate outlined in Chapter 3, the Si cap thicknesses vary by around 1 nm from sample to sample (Table 6.2). While these small differences cannot be neglected entirely, it is important to recall that the practical resolution of XTEM measurements done in this study was on the order of 1 nm. This experiment was repeated on both $Si_{0.5}Ge_{0.5}$ and $Si_{0.3}Ge_{0.7}$ virtual substrates to form a 2×2 experimental matrix (Table 6.2).

Ge content in virtual substrate	$h_{Si\ cap}$	h_{Ge}
50%	6 nm	6 nm
50%	4 nm	12 nm
70%	5.5 nm	6 nm
70%	4.5 nm	12 nm

Table 6.2 As-grown layer thicknesses for 4 samples used to examine the effects of h_{Ge} and virtual substrate composition (strain) on hole mobility. Layer thicknesses are ± 0.5 nm.

Theory would indicate that a ε-Ge layer grown on $Si_{0.5}Ge_{0.5}$ would exhibit higher mobility than one grown on $Si_{0.3}Ge_{0.7}$, both because of the lowered in-plane effective mass[58] and the larger valence band offset (type-I band alignment) between the ε-Ge and relaxed $Si_{1-x}Ge_x$. However, for the two devices with thin (h_{Ge} = 6 nm) ε-Ge layers, the more highly strained sample demonstrates *lower* μ_{eff} across the entire range of N_{inv} (Figure 6.3). In Ge, compressive strain greatly reduces the vertical effective mass of holes, and the out-of-plane mobility is predicted to exceed the already large in-plane mobility in a manner that is qualitatively similar to the case of tensile strained Si.[58] However, the out-of-plane hole mobilities in ε-Ge are predicted to be much larger than those in ε-Si, as shown in Figure 6.4. Thus, the hole wave function can spread into the layers above and below, despite the deep potential well for holes that forms in the ε-Ge. For the sample with thin ε-Ge on $Si_{0.5}Ge_{0.5}$, the 2% compressive strain and resultant high out-of-plane mobility cause the hole wave function to penetrate considerably into the ε-Si cap and the lower mobility buffer below. While the same spreading must take place in the thin ε-Ge sample on $Si_{0.3}Ge_{0.7}$, the vertical extent of the wave function is lessened due to the lower strain. Increasing compressive strain in the Ge can therefore serve to

decrease the overall μ_{eff} of the inversion layer if the ε-Ge is not thick enough to confine the hole wave function vertically.

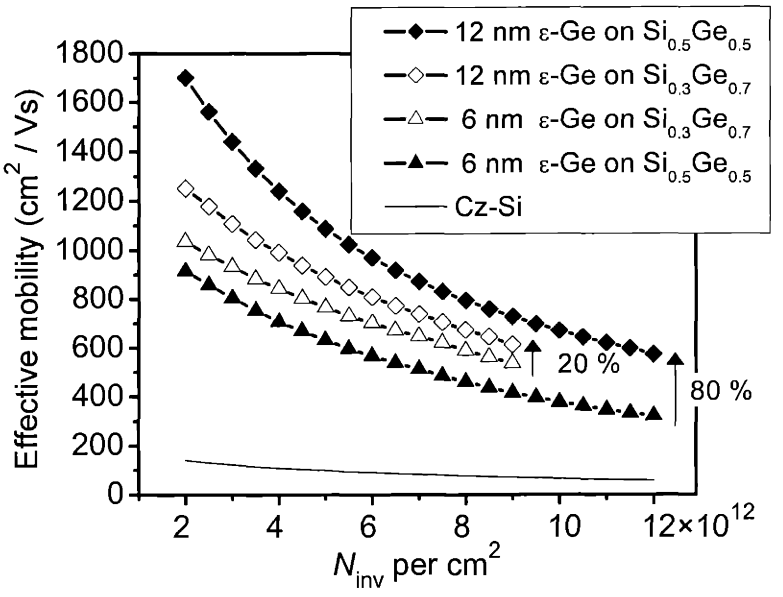


Figure 6.3 Hole effective mobility vs N_{inv} for the samples described in Table 6.2. The use of a thicker ε-Ge layer gave rise to a 20% increase in μ_{eff} for the samples grown on $Si_{0.3}Ge_{0.7}$ and an 80% increase for the samples grown on $Si_{0.5}Ge_{0.5}$.

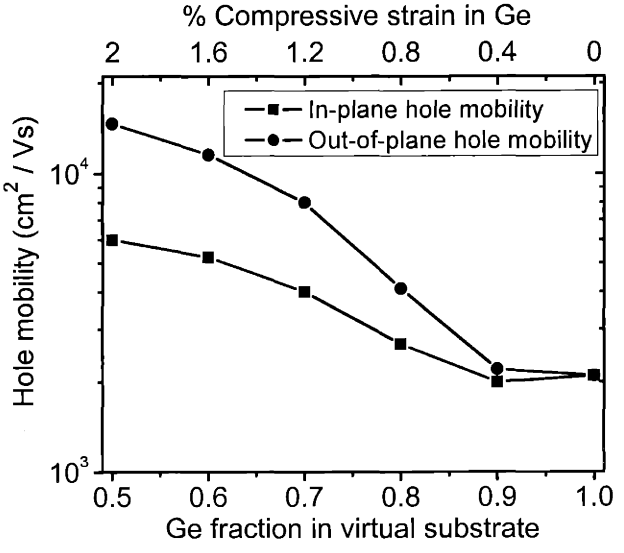


Figure 6.4 Calculated in-plane and out-of-plane hole mobilities in ε-Ge. The extremely large out-of-plane hole mobilities lead to a hole wave function that can significantly penetrate into both the Si cap above and the relaxed $Si_{1-x}Ge_x$ below.

The samples with thick (h_{Ge} = 12 nm) ε-Ge demonstrate markedly different behavior from the samples with thin ε-Ge. As Figure 6.3 shows, increasing h_{Ge} gave rise to an 80% increase in μ_{eff} for the devices on $Si_{0.5}Ge_{0.5}$ and only a 20% increase for the devices grown on $Si_{0.3}Ge_{0.7}$. μ_{eff} increases considerably for the thick ε-Ge layer on $Si_{0.5}Ge_{0.5}$ because the 12 nm ε-Ge layer allows the large hole wave function to be better confined in the ε-Ge. While 12 nm significantly exceeds the expected critical thickness of Ge on $Si_{0.5}Ge_{0.5}$,[13] the results presented here show that confinement in the ε-Ge is the primary factor in determining hole mobility, overwhelming the potentially deleterious effects of partial strain relaxation and misfit dislocation scattering. Conversely, it may be deduced that 6 nm is nearly sufficient to confine the hole wave function when ε-Ge is grown on $Si_{0.3}Ge_{0.7}$ based upon the relatively small gain in μ_{eff} caused by doubling h_{Ge}. Table 6.3 summarizes the results presented in Section 6.2 while emphasizing that high strain *and* a relatively thick ε-Ge layer are necessary for attaining the highest hole mobilities.

Strain	**h_{Ge} = 6 nm**	**h_{Ge} = 12 nm**
ε_{Ge} = 1.2 %	7.2 ×	8.2 ×
ε_{Ge} = 2 %	5.6 ×	9.8 ×

Table 6.3 Experimental matrix summarizing hole mobility enhancements with varied h_{Ge} and ε_{Ge} (enhancement computed at $N_{inv} = 9 \times 10^{12}$ cm^{-2}). The combination of high ε_{Ge} *and* high h_{Ge} are required to obtain the highest μ_{eff}.

The unexpected result of hole mobility decreasing with higher compressive strain was also seen by Leitz *et al.* in dual-channel heterostructures where the buried layer was compressively strained $Si_{0.2}Ge_{0.8}$.[14] The heterostructures were grown on $Si_{0.5}Ge_{0.5}$ and $Si_{0.7}Ge_{0.3}$ virtual substrates with ε-$Si_{0.2}Ge_{0.8}$ thicknesses of 6.5 nm and 12 nm respectively

($h_{Si\ cap}$ for these two samples varied by ~1 nm). As Figure 6.5 shows, the hole mobility of the more highly strained sample was depressed across the entire range of N_{inv}. Increasing the thickness of the ε-$Si_{0.2}Ge_{0.8}$ layer to 12 nm in the more highly strained case would be expected to provide a considerably higher hole mobility.

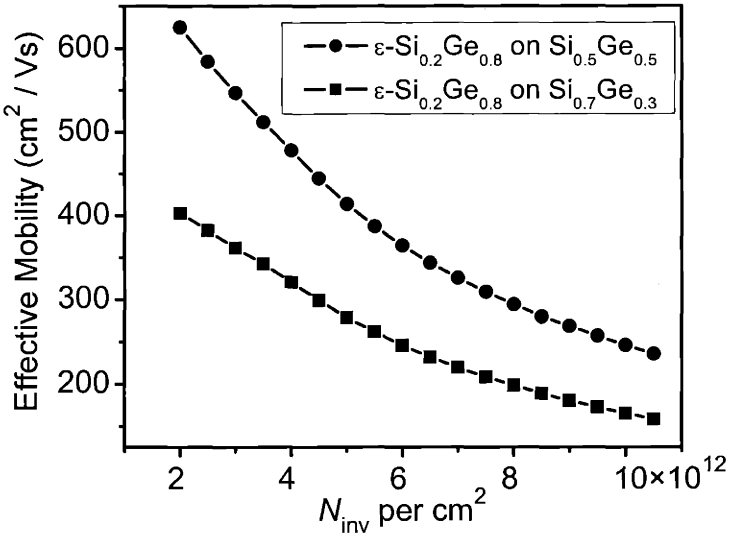


Figure 6.5 Hole effective mobility vs N_{inv} for ε-Si / ε-$Si_{0.2}Ge_{0.8}$ dual-channel heterostructures grown on $Si_{0.5}Ge_{0.5}$ and $Si_{0.7}Ge_{0.3}$ virtual substrates.[14]

6.3. Effect of ε-Si cap thickness

The previous sections underscore the effect that ε-Ge channel thickness and strain have on hole mobility, as well as the interplay between the two variables; for example, changing h_{Ge} had vastly different effects on 50% and 70% Ge virtual substrates. Therefore, to isolate the effect of varying $h_{Si\ cap}$, it is important to compare only heterostructures where h_{Ge} and x in the virtual substrate are held constant.

A simple model of the mobility characteristics in dual-channel devices can lead to fundamental insight on the inversion mobility in these heterostructures. For now, I return to a semi-classical description of the inversion carriers, foregoing the previous hypothesis

on the hybrid nature of the valence band when a large hole wave function experiences spatially frequent changes in composition (Chapter 5). Note however that to arrive at quantitatively accurate simulations of these devices, it is likely that far more detailed information on band offsets and carrier mobilities would be required, as well as the collection of empirical data to corroborate the model. Still, a simple model of inversion layer mobility is a useful tool in the design of ε-Si / ε-Ge dual-channel heterostructures. Due to the dual-channel band alignment, the total inversion charge is comprised of carriers in both the surface and buried layers:

$$N_{inv} = N_{Si} + N_{Ge} + N_{SiGe} \tag{6.1}$$

where N_{Si} is the inversion density in the Si cap, N_{Ge} is the inversion density in the buried Ge layer, N_{SiGe} is the inversion density in the relaxed buffer, and N_{inv} is the total inversion density. The effective mobility of the inversion layer should therefore be expressed as the number-weighted average mobility of each layer:

$$\mu_{eff} = (\mu_{Si} N_{Si} + \mu_{Ge} N_{Ge} + \mu_{SiGe} N_{SiGe}) / N_{inv} \tag{6.2}$$

The inversion density in the Si cap will depend on many factors, including the valence band offset between the buried layer and the cap, the channel doping, the gate overdrive, and the cap thickness. In the limit of an extremely thick Si cap, the vertical electrical field in the ε-Ge layer will be negligible and N_{Ge} and N_{SiGe} will approach zero. Consequently, N_{inv} will approach N_{Si}, and μ_{eff} will approach μ_{Si}. In the converse limit of $h_{Si\ cap} = 0$, all of the inversion charge would be found in the ε-Ge and $Si_{1-x}Ge_x$, and μ_{Si} would make no contribution to μ_{eff}. In weak inversion, only a small portion of the inversion charge is likely to reside in the Si cap due to the deep hole well in the ε-Ge layer. The above intuitive reasoning leads to a simple formulation for N_{Si} :

(6.3) $$N_{Si} \propto V_{GT} \times h_{Si\ cap}$$

6.3.1. $Si_{0.3}Ge_{0.7}$ virtual substrate, h_{Ge} = 6 nm

Figure 6.6 plots the hole mobility of two different heterostructures grown on $Si_{0.3}Ge_{0.7}$ with h_{Ge} = 6 nm and slightly different $h_{Si\ cap}$ (5.5 nm and 6.5 nm as-grown). At low N_{inv}, N_{Si} is expected to be very low, and thus μ_{eff} is determined almost solely by holes in the ε-Ge and $Si_{0.3}Ge_{0.7}$. Therefore, μ_{eff} for $N_{inv} \leq 4 \times 10^{12}$ $cm^{-2}$of the two devices is nearly identical. As gate overdrive is increased, N_{Si} in the $h_{Si\ cap}$ = 6.5 nm case exceeds N_{Si} in the $h_{Si\ cap}$ = 5.5 nm case, because the cap is thicker, or equivalently, the electrical field in the ε-Ge layer is lower. μ_{eff} of the $h_{Si\ cap}$ = 5.5 nm sample is therefore higher, because proportionally more of the inversion charge is in the high mobility ε-Ge. Clearly μ_{eff} in these heterostructures can be significantly altered by varying $h_{Si\ cap}$, and this difference becomes most pronounced for large gate overdrives. For $N_{inv} = 3 \times 10^{12}$ cm^{-2}, the extracted mobility of the two samples are within 1% of each other, while at N_{inv} = 9×10^{12} cm^{-2}, the difference has jumped to more than 10%.

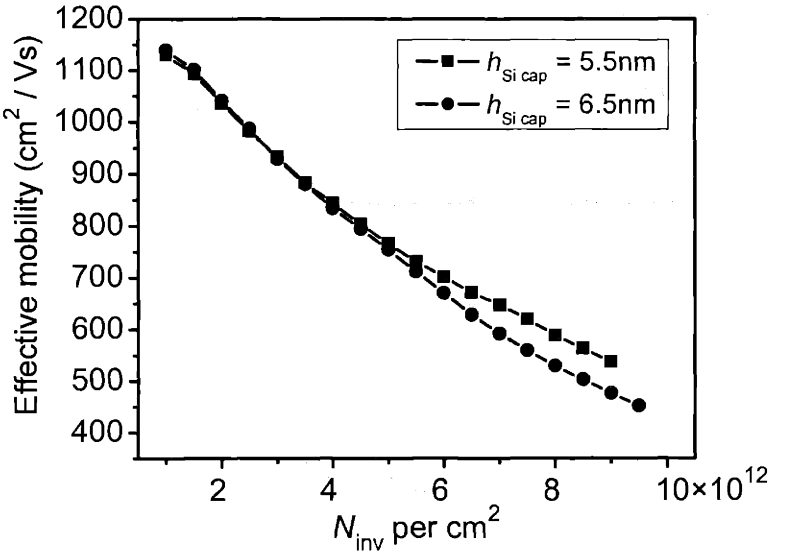


Figure 6.6 Hole effective mobility vs N_{inv} for samples with 6 nm ε-Ge grown on $Si_{0.3}Ge_{0.7}$. A slight variation in $h_{Si\ cap}$ causes μ_{eff} to deviate significantly at large N_{inv}.

6.3.2. $Si_{0.5}Ge_{0.5}$ virtual substrate, h_{Ge} = 6 nm

Next, the above Si cap variation experiment was repeated but with increased strain in the Ge channel. For this series, I grew 6 nm of ε-Ge on $Si_{0.5}Ge_{0.5}$ instead of $Si_{0.3}Ge_{0.7}$ and varied $h_{Si\ cap}$ from 6 nm in one sample to 7.5 nm in the other. Figure 6.7 shows that in contrast to the samples grown on $Si_{0.3}Ge_{0.7}$, μ_{eff} of the two samples grown on $Si_{0.5}Ge_{0.5}$ are extremely close. Furthermore, the shapes of the two curves track each other very closely across the entire range of N_{inv}. Returning to Eq. 6.2, note that if $\mu_{Ge} >> \mu_{Si}$, then small changes in N_{Si} will cause little change in μ_{eff}. In Section 6.2, it was confirmed that increasing the compressive strain in the Ge layer by 0.8% caused μ_{Ge} to increase considerably. Therefore, $\mu_{Ge}N_{Ge}$ becomes the dominant term in Eq. 6.2 for large compressive strains, and the dependence of μ_{eff} on $h_{Si\ cap}$ is proportionately reduced. In this sense, one strategy for making μ_{eff} less sensitive to growth variations is to increase the difference in carrier mobilities.

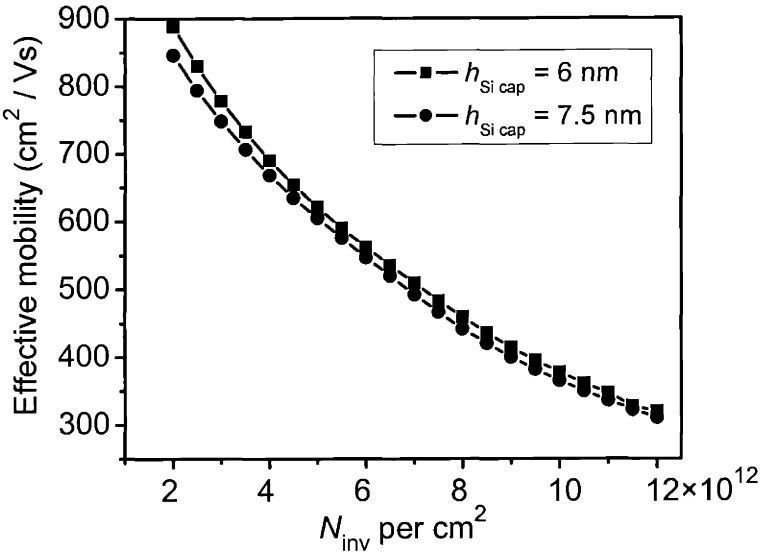


Figure 6.7 Hole effective mobility vs N_{inv} for samples with 6 nm ε-Ge grown on $Si_{0.5}Ge_{0.5}$. μ_{eff} in samples grown on $Si_{0.5}Ge_{0.5}$ with slightly different $h_{Si\ cap}$ match one another closely across a wide range of N_{inv}.

6.3.3. $Si_{0.3}Ge_{0.7}$ virtual substrate, h_{Ge} = 12 nm - Investigating hole transport with an extremely thin Si cap

In the third series of experiments investigating the effect of $h_{Si\ cap}$, two heterostructures with h_{Ge} = 12 nm on $Si_{0.3}Ge_{0.7}$ were grown ($h_{Si\ cap}$ was 4.5 nm in one sample and 1.5 nm in the other). While the previous experiments indicated that a thinner cap would lead to higher μ_{eff} due to the reduced N_{Si}, Figure 6.8 shows the exact opposite trend. In fact, μ_{eff} of the sample with $h_{Si\ cap}$ = 1.5 nm is roughly half that of the sample with $h_{Si\ cap}$ = 4.5 nm across the entire range of N_{inv}. Recall that cleaning steps prior to device processing remove roughly 1 nm of the Si cap, and therefore in the sample with $h_{Si\ cap}$ =1.5 nm (as-grown), the ε-Ge layer is practically a surface channel. Leitz *et al.* observed a degradation in Si/SiO_2 interface quality for extremely thin Si caps caused by diffusion of Ge atoms,[6] and hence the positive slope of μ_{eff} in weak-to-moderate

inversion for the $h_{\text{Si cap}}$ = 1.5 nm case likely indicates Coulomb scattering-limited mobility. The low vertical effective mass of holes in ε-Ge also increases the possibility of the hole wave function interacting with the low conductivity SiO_2 gate, as I have argued may be the case in ε-Si single-channel *p*-MOSFETs. While the true mobility-limiting mechanism is unclear, for the thinnest Si caps, μ_{eff} becomes sharply depressed, and the simple relation given by Eq. 6.3 breaks down. This experiment also underscores the fact that the Si cap plays a role in shielding holes from the Si/SiO_2 interface and cannot simply be treated as a lower mobility parasitic channel above the ε-Ge. Viewing all of the results together shows that the highest hole mobilities are only attained when $h_{\text{Si cap}}$ is at least 3 nm.

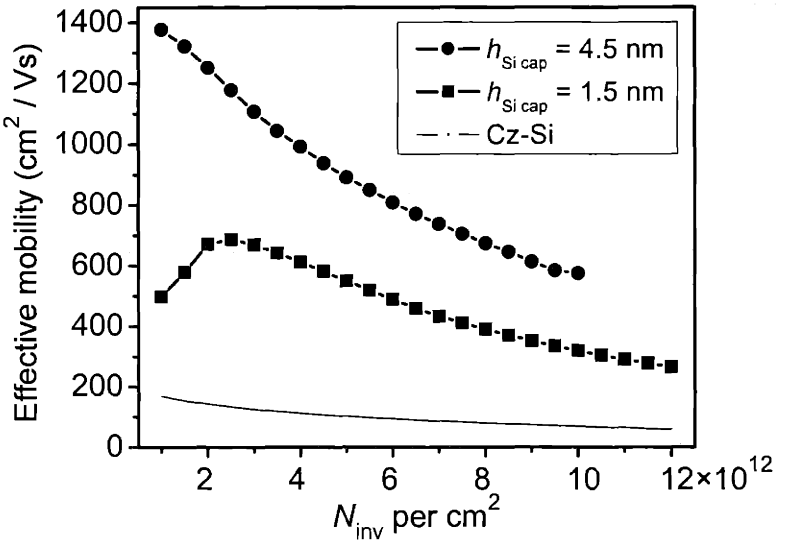


Figure 6.8 Hole effective mobility vs N_{inv} for samples with 12 nm ε-Ge grown on $Si_{0.3}Ge_{0.7}$. When the Si cap is made too thin, the hole mobility becomes depressed across the entire range of N_{inv}.

While an extremely thin Si cap may not lead to the highest hole mobilities, the sample with $h_{\text{Si cap}} = 1.5$ nm still maintains mobility enhancements of 4.5 to 5.5 over Cz-Si across most of the range of N_{inv}. An extremely thin Si cap also allows a steeper subthreshold slope, because the gate electrode is closer to the inversion carriers and the deleterious effects associated with buried-channel operation are reduced. In application, some of the loss in performance due to the lower μ_{eff} could be compensated by superior subthreshold characteristics and an attendant reduction in I_{off}. These issues, while somewhat beyond the scope of this thesis, should be examined in future work.

Though the effect of varying Si cap thickness on hole mobility can be somewhat unpredictable, in general it appears that μ_{eff} has a saddle-shaped dependence on $h_{\text{Si cap}}$. Thus, the intended application may dictate the optimal cap thickness (Figure 6.9). For example, a relatively large $h_{\text{Si cap}}$ will impart lower noise and higher μ_{eff} for analog applications, at the cost of degraded gate control and higher static power dissipation. At the opposite extreme, a thin Si cap may be favorable for digital applications where improved subthreshold slope and $I_{\text{on}} / I_{\text{off}}$ are primary performance metrics.

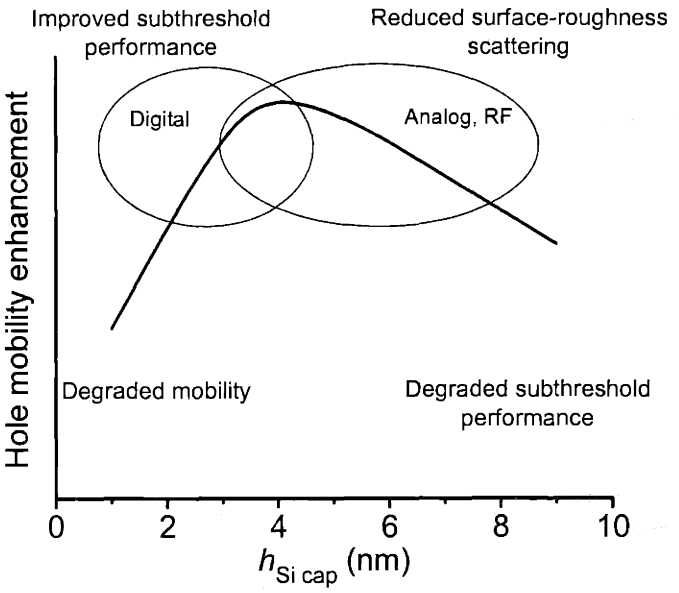


Figure 6.9 Schematic chart of hole mobility enhancements vs $h_{Si\ cap}$ along with possible application space.

6.4. Summary

A large number of different ε-Si / ε-Ge dual-channel heterostructures was studied in order to understand hole transport in this class of *p*-MOSFETs. However, the division of this chapter into sections describing different growth variables is somewhat artificial, since all of the variables were shown to be interrelated. For example, the effects of varying h_{Ge} and $h_{Si\ cap}$ were very different for heterostructures grown on different $Si_{1-x}Ge_x$ virtual substrates. Still the key findings of this chapter can be summarized by stating that the highest hole mobilities were favored by high strain (> 1%) in a relatively thick Ge layer (≥ 8 nm) and $h_{Si\ cap}$ = 3-5 nm. As a case-in-point, out of all heterostructures studied, the one with the largest measured hole mobility had h_{Ge} = 12 nm and $h_{Si\ cap}$ = 4 nm grown upon $Si_{0.5}Ge_{0.5}$. Figure 6.10 shows that this heterostructure had a mobility enhancement of 10 times or more across a wide range of N_{inv}. The room-temperature peak μ_{eff}

extracted from these devices was 2170 cm^2/Vs (at low V_{GT}), indicating potential for low-noise analog applications. The challenges in successfully implementing ε-Si / ε-Ge dual-channel heterostructures on a commercial scale would undoubtedly be formidable due to the potential difficulty of engineering devices with acceptable V_T behavior and subthreshold characteristics. However, the prospect of attaining *p*-MOSFETs with such high hole mobility may eventually justify the added cost and complexity of using ε-Ge.

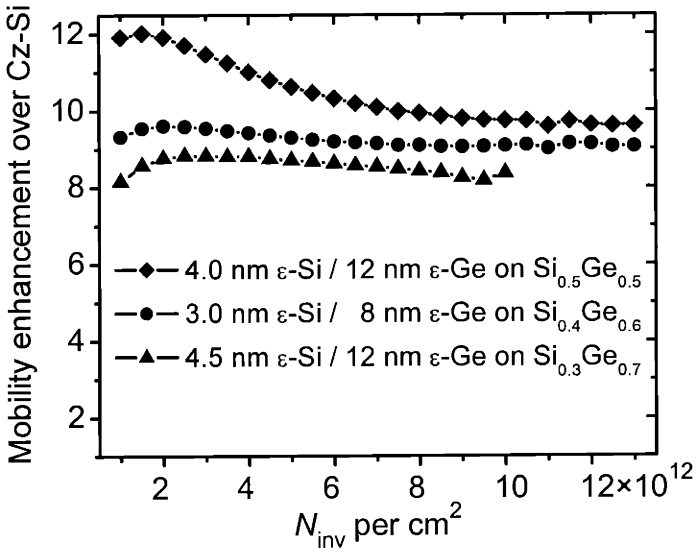


Figure 6.10 Hole mobility enhancement vs N_{inv} for the highest mobility ε-Si / ε-Ge dual-channel heterostructures studied in this thesis.

Chapter 7. *n*-type MOSFETs fabricated on Ge-rich $Si_{1-x}Ge_x$ heterostructures

Chapters 5 and 6 show that very large hole mobility enhancements can be attained in Ge-rich heterostructure *p*-MOSFETs in both single- and dual-channel configurations. However for CMOS applications, it is equally important to integrate high mobility *n*-MOSFETs on the same wafers. In this chapter, I present electron mobility characteristics in single-, dual-, and digital-alloy channel heterostructures grown on Ge-rich virtual substrates.

7.1. Single-Channel Heterostructures

While the enhanced device characteristics of single-channel ε-Si *n*-MOSFETs grown on Si-rich $Si_{1-x}Ge_x$ virtual substrates have been well established, little work with Ge-rich virtual substrates ($x > 0.5$) has been shown. As described in Chapter 3, work by Currie *et al.*[7] outlined two basic criteria for attaining maximum electron mobility enhancement in single-channel ε-Si heterostructures: first, $h_{Si\ cap} > 5$-6 nm to confine the electron wave function away from the relaxed buffer, and second, at least 0.8% biaxial tensile strain is necessary to completely suppress intervalley scattering between the Δ_2 (out-of-plane) and Δ_4 (in-plane) valleys. However, the requirement of 5-6 nm of ε-Si is problematic for growth on Ge-rich $Si_{1-x}Ge_x$ because of the low h_c of Si at such large mismatches (for $Si_{0.4}Ge_{0.6}$ and $Si_{0.3}Ge_{0.7}$, $h_{c,\ Si}$ is ~4 nm and ~3 nm, respectively). Therefore, growing ε-Si layers at or near h_c will force a portion of the electron wave function into the relaxed $Si_{1-x}Ge_x$. On the other hand, exceeding the critical thickness for such a high film/substrate mismatch may result in dislocation densities in the ε-Si cap that are high enough to strongly scatter electrons; recall that in early attempts to study ε-Si, the electron mobility was sharply limited by defects inherited from the highly defective uniform $Si_{1-x}Ge_x$ buffer layers that were used.[43]

Figure 7.1(a) plots electron effective mobility vs N_{inv} of single-channel ε-Si *n*-MOSFETs on $Si_{0.75}Ge_{0.25}$ (hereafter referred to as the ε-Si control), $Si_{0.4}Ge_{0.6}$, and $Si_{0.3}Ge_{0.7}$. The ε-Si control sample, with a 10 nm ε-Si cap in 1% biaxial tension, meets the two simple design criterion described above and exhibits the expected mobility enhancement of 1.7 to 1.9 times [Figure 7.1(b)]. For the sample on $Si_{0.3}Ge_{0.7}$, a 6 nm ε-Si cap was grown in order to ensure confinement of the electron wave function in the Si cap. Due to the 2.8% lattice mismatch of Si with $Si_{0.3}Ge_{0.7}$, an extremely high density of defects (~10^9 cm^{-2}) nucleates in the cap once the critical thickness is exceeded. However, the high mismatch and low thickness of the Si layer make complete relaxation nearly impossible,[13] and a large residual tensile strain (at least 1.4%) can be assumed. The 6 nm Si layer should be sufficient to confine electrons, and furthermore, the 1.4% strain is more than enough to completely suppress intervalley scattering.[7] While significant enhancement is evident [Figure 7.1(b)], the typical 1.8 times electron mobility enhancement is not attained. The degraded μ_{eff} can be explained by the fact that the type-II band alignment between ε-Si and $Si_{1-x}Ge_x$ forces electrons into the cap where threading dislocations are most concentrated. Defect scattering is thereby maximized because both the carriers and dislocations are confined to such a small physical volume. Since growing the Si cap thicker than 3 nm will always produce a high density of defects (at typical growth temperatures), and at least 5 nm is required for electron confinement, an optimal ε-Si *n*-MOSFET cannot be realized on a $Si_{0.3}Ge_{0.7}$ buffer.

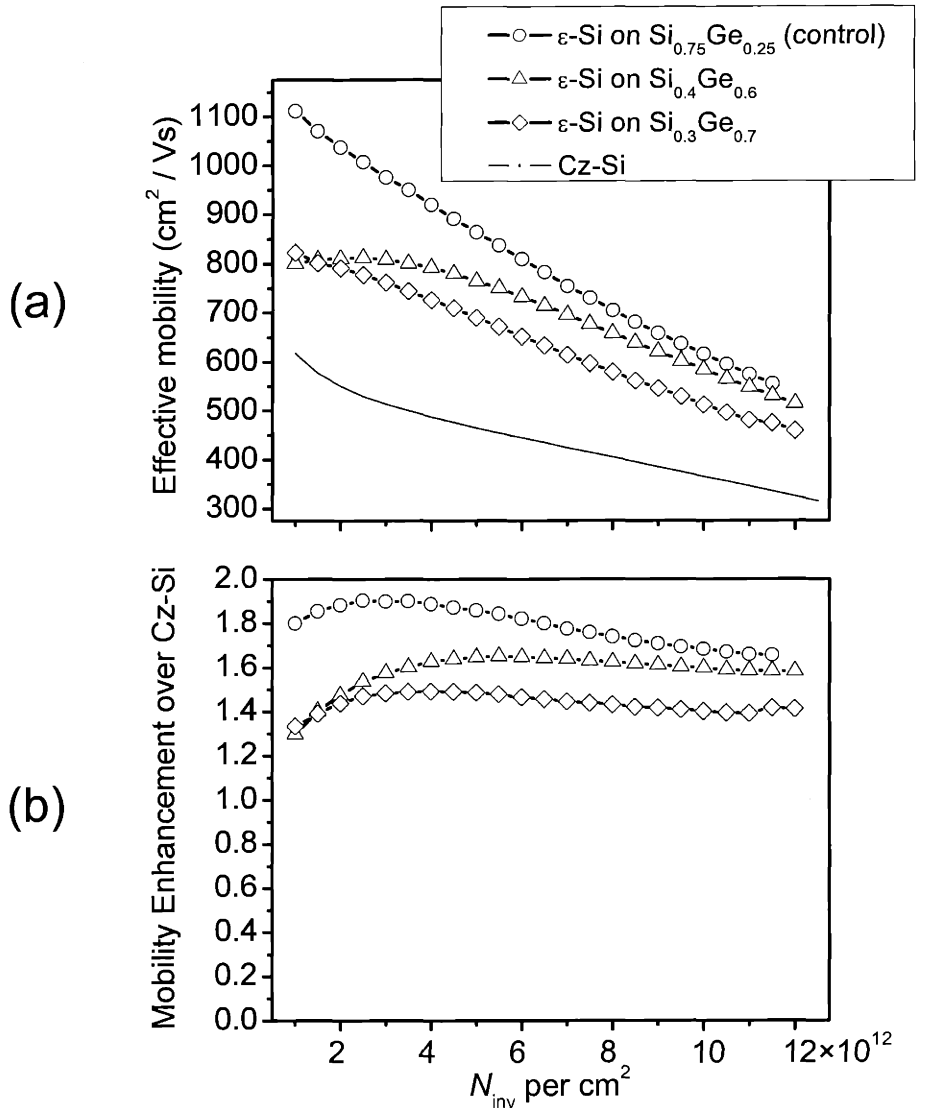


Figure 7.1 (a) Electron effective mobility vs N_{inv} of single-channel *n*-MOSFETs grown on $Si_{0.75}Ge_{0.25}$ ($h_{Si\ cap}$ = 10 nm), $Si_{0.4}Ge_{0.6}$ ($h_{Si\ cap}$ = 3 nm), and $Si_{0.3}Ge_{0.7}$ ($h_{Si\ cap}$ = 6 nm). (b) Electron mobility enhancement of the same devices.

For the sample on $Si_{0.4}Ge_{0.6}$, a 3 nm ε-Si cap was grown, and thus the critical thickness was not greatly exceeded at the cost of incomplete confinement of the electron wave function. As Figure 7.1(b) shows, the electron mobility for this device is enhanced by around 60% across a wide range of N_{inv}. This result is in stark contrast with Currie's work where devices grown on $Si_{0.8}Ge_{0.2}$ with $h_{Si\ cap} < 5$ nm displayed enhancements of

only 10%; with $h_{Si\ cap}$ = 2 nm, the electron mobility in his devices was actually degraded below that of a Cz-Si control device.[7] However, Currie's process flow utilized a 1 second RTA at 1000-1100°C to activate the source-drain-gate implant, while a 30 minute furnace anneal at 600°C was used for activation in the present work. For the thinnest ε-Si channels, even a 1 second RTA at high temperature could have a highly detrimental effect on electron mobility due to interdiffusion of Ge atoms at the ε-Si / $Si_{0.8}Ge_{0.2}$ interface. While a small amount of interdiffusion would have a negligible impact on a 10 nm ε-Si channel, the same amount of interdiffusion would be highly detrimental for a 3 nm ε-Si channel.

7.2. Dual-channel heterostructures

Since the highest hole mobilities have been fabricated using dual-channel heterostructures, it is particularly important to investigate the effect of a Ge-rich buried layer on electron mobility. As stated previously, the type-II band offset allows electrons to be confined in the ε-Si cap. Therefore, it has been predicted that for a sufficiently thick Si cap, the electron mobility will only experience the ε-Si band structure and not be affected by the presence of the buried channel.

7.2.1. High electron mobility in dual-channel heterostructures

Figure 7.2 shows electron effective mobility vs N_{inv} for several ε-Si / ε-Ge dual-channel heterostructures grown on $Si_{0.5}Ge_{0.5}$ with various Si cap thicknesses. For the sample with a 7.5 nm Si cap, the extracted mobility is nearly identical to the ε-Si control. In the samples with thinner Si caps, some of the mobility enhancement is lost, as some portion of the electron wave function is forced to populate the ε-Ge below. A mixture of

conduction through the buried Ge and surface Si might be expected to benefit electron transport for a very thin Si cap due to Ge's high electron mobility ($\mu_{e,\,Ge}$ = 3900 cm^2 / Vs [104]). However, Ge's conduction band minima lie in <111> directions while Si's lie in <100> directions. Therefore, it is likely that electrons near the ε-Ge / ε-Si interface must undergo strong *X*-valley to *L*-valley scattering during transport, in a manner analogous to the *Γ* to *L*-valley scattering used in GaAs negative differential conductivity devices.[25,29] Furthermore, theoretical studies by Fischetti suggest that electron mobility in Ge is degraded by compressive strain.[58] Regardless of the mechanism of mobility degradation, Figure 7.2 shows that the only way to achieve high electron mobility in a dual-channel heterostructure is to completely isolate the electron wave function from the buried compressive layer.

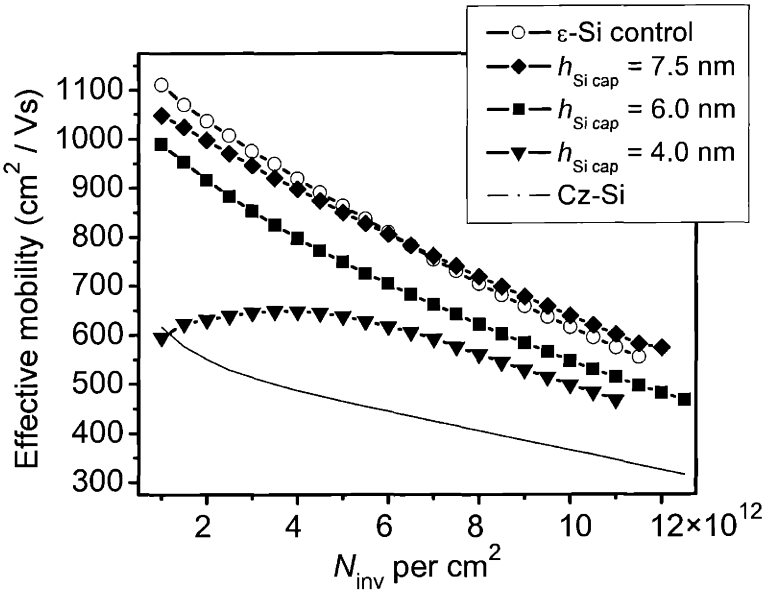


Figure 7.2 Electron effective mobility vs N_{inv} of ε-Si / ε-Ge dual-channel *n*-MOSFETs grown on $Si_{0.5}Ge_{0.5}$ with $h_{Si\ cap}$ varied. Only the sample with $h_{Si\ cap}$ = 7.5 nm matches the mobility of the ε-Si control.

7.2.2. Investigating the effect of *X*-valley to *L*-valley scattering in inversion-layer transport

Direct comparison of NMOS mobility in samples with and without a buried Ge layer

In order to test the possibility of *X*-valley to *L*-valley scattering in *n*-channel inversion layers, the mobility characteristics of two heterostructures grown on $Si_{0.4}Ge_{0.6}$ virtual substrates were compared: one with a buried ε-Ge layer and one without (hereafter referred to as the dual-channel and single-channel devices, respectively). Since $h_{Si\ cap}$ = 3 nm in both heterostructures, the same proportion of the electron wave function conducts below the cap layer. Figure 7.3(a) shows that μ_{eff} of the single-channel device greatly exceeds that of the dual-channel, proving that the presence of a buried ε-Ge layer degrades *n*-channel mobility when the Si cap is too thin. Note also that at low N_{inv} the electron mobility in the dual-channel device drops much more rapidly than either the single-channel or the Cz-Si device as N_{inv} is increased. Replotting Figure 7.3(a) as a log-log chart reveals that $\mu_{eff} \propto N_{inv}^{-0.26}$ for the dual-channel device at relatively low N_{inv}, where Coulomb scattering should dominate. According to the universal mobility model, phonon-limited electron mobility follows a dependence on $E_{eff}^{-0.3}$, while Coulomb-scattering limited mobility has a positive exponential dependence on E_{eff}. Recall that for low substrate doping levels, E_{eff} and N_{inv} are essentially proportional to each other, because nearly all of the charge that goes into creating the vertical electrical field originates from the inversion layer itself. Therefore the observation of a phonon scattering-like slope at low N_{inv} supports the theory that *X*-valley to *L*-valley scattering dominates in the dual-channel device; in order for an electron to be transferred from Ge's conduction band to Si's conduction band, a phonon-scattering event must take place by

definition. At $N_{inv} > 1 \times 10^{13}$ cm^{-2}, the slope of all of the devices in Figure 7.3 becomes uniform, indicating that a more "universal" phenomenon limits the electron mobility at large vertical fields regardless of channel composition.

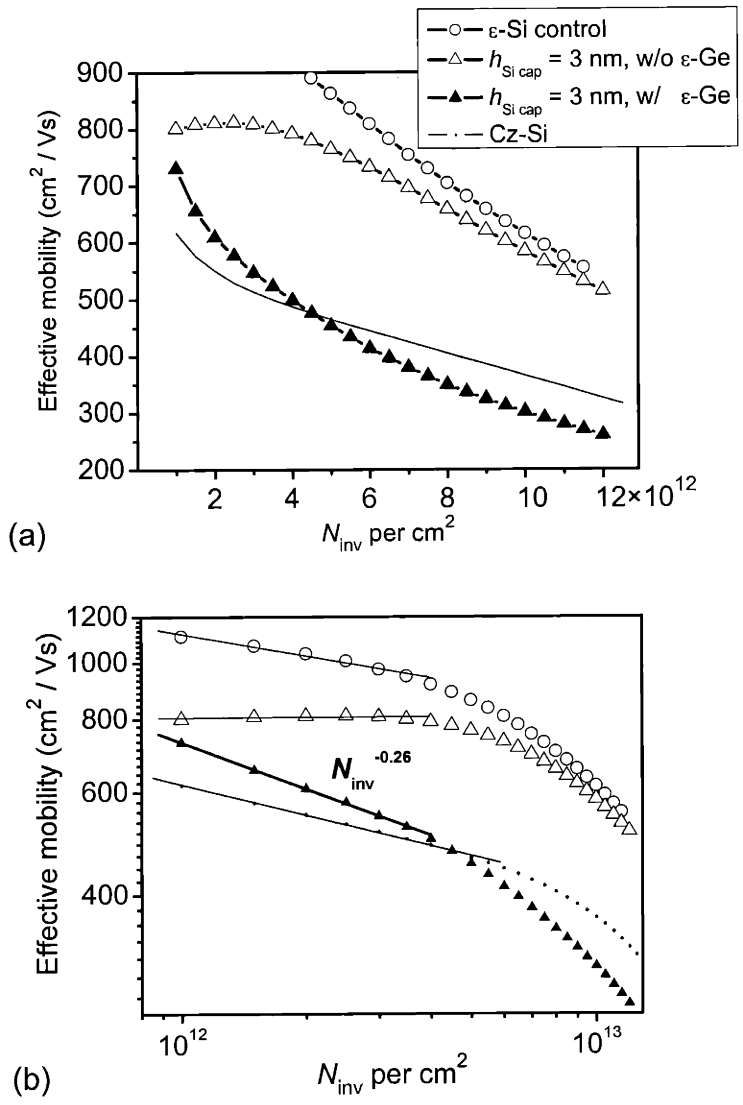


Figure 7.3 (a) Linear and (b) log-log plots of electron effective mobility vs N_{inv} for single-channel and ε-Si / ε-Ge dual-channel *n*-MOSFETs grown on $Si_{0.4}Ge_{0.6}$. $h_{Si\ cap}$ = 3 nm for both devices.

Dependence on X-to-L-valley scattering potential on buried-layer Ge content

In order to investigate the effect of Ge content on *X*-to-*L*-valley scattering in *n*-type inversion layers, a dual-channel heterostructure with an 80% Ge buried layer was grown. The Si cap was held at 3 nm, just as in the previous samples, and the Ge content of the virtual substrate was reduced to 0.4 in order to place the same amount of compressive strain in the buried layer. Figure 7.4 shows that the effective electron mobility of the ε-Si / ε-$Si_{0.2}Ge_{0.8}$ dual-channel device is considerably higher than that of the ε-Si / ε-Ge dual-channel device. The relatively high mobility in the alloy case compared to pure Ge indicates that lowering the Ge-content of the buried layer greatly reduces the *X*-to-*L*-valley scattering potential. As described previously, the band structure of $Si_{1-x}Ge_x$ alloys does not become Ge-like until $x \sim 0.7$-0.8,[32] and therefore the mismatch in conduction band minima between the buried $Si_{0.2}Ge_{0.8}$ and ε-Si cap is greatly reduced compared to the case of a pure Ge buried layer.

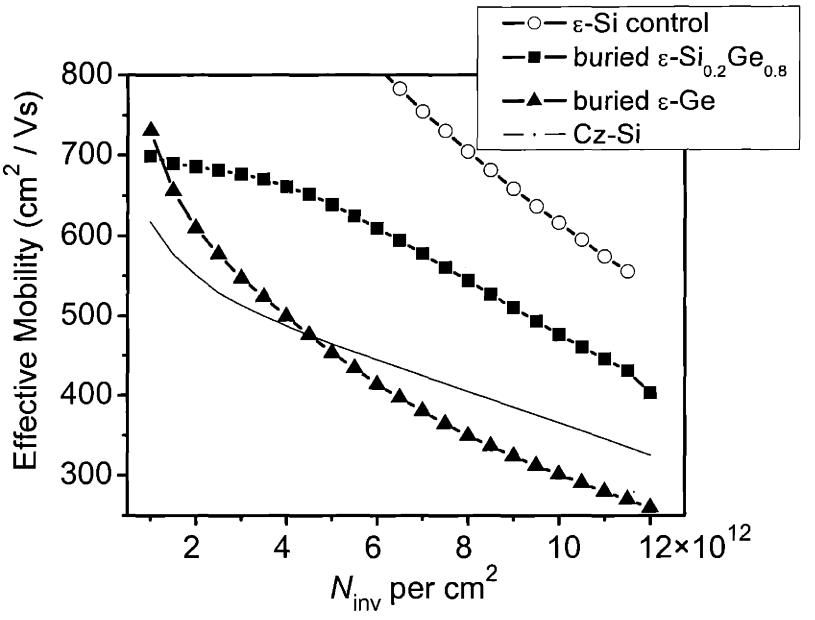


Figure 7.4 Electron effective mobility vs N_{inv} for dual-channel *n*-MOSFETs with buried layers of Ge and $Si_{0.2}Ge_{0.8}$. The amount of compressive strain in the buried layer is 1.2% in both cases.

Investigating the effect of compressive strain on transport in ε-Si / ε-Ge dual-channel heterostructures

The results presented here demonstrate that mixed electron transport through both ε-Si and ε-Ge introduces a strong phonon scattering phenomenon that greatly reduces inversion layer mobility. Another possible reason for the low electron mobility of ε-Si / ε-Ge heterostructures is the changes in Ge's band structure that are predicted to take place as a result of applied stress. According to Fischetti and Laux, compressive strain decreases the bandgap of the Δ(100) valleys (the in-plane lobes), which in turn increases the rate of intervalley scattering between the *L*-valleys and the Δ(100) valleys. As a result, the electron mobility is predicted to be decreased by compressive strain (Figure 7.5).[58] For compressive strain greater than 1.6%, Fischetti and Laux predict that

electrons *only* populate the Δ(100) valleys, where they experience a large effective mass ($m_l \approx 1.0m_o$), and an electron mobility 5 times lower than bulk Ge.[58]

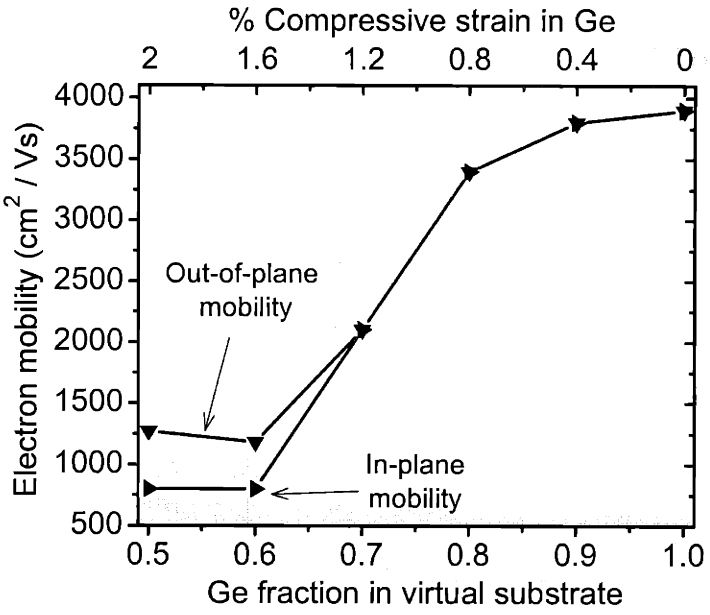


Figure 7.5 Calculated low-field electron mobility in Ge as a function of compressive strain.[58]

Comparing ε-Si / ε -Ge heterostructures grown on relaxed $Si_{0.3}Ge_{0.7}$ and $Si_{0.5}Ge_{0.5}$ virtual substrates may allow this theory to be tested. Both samples have a 4 nm ε-Si cap and a 12 nm ε-Ge layer. As Figure 7.6 shows, neither sample matches the mobility of the ε-Si control device presented earlier, but the sample grown on $Si_{0.5}Ge_{0.5}$ shows markedly higher mobility than the sample grown on $Si_{0.3}Ge_{0.7}$. While this seems to contradict the theoretical calculations, the Si cap in both samples is 4 nm thick (as-grown), so only a relatively small fraction of the overall inversion mobility can be attributed to the ε-Ge. Recall also that the single-channel electron mobility on $Si_{0.3}Ge_{0.7}$ was found to be limited by defect scattering, so it is likely that defect scattering also plays a strong role in depressing the mobility of the ε-Si / ε-Ge dual-channel heterostructure grown on

$Si_{0.3}Ge_{0.7}$. Another reason that the electron mobility of the sample grown on $Si_{0.5}Ge_{0.5}$ is higher may be that electrons are predicted to solely populate the Δ(100) valleys in Ge with such a large compressive strain. Recall that in ε-Si, only the Δ(001) valleys (out-of-plane lobes) are populated. Therefore, electrons in the ε-Si / ε-Ge dual-channel heterostructure on $Si_{0.5}Ge_{0.5}$ should "see" a six-fold degenerate conduction band: two from the ε-Si and four from the ε-Ge. For lower strain, the electrons in the ε-Ge may be able to scatter to both the eight-fold degenerate *L*-valleys *and* the four-fold degenerate Δ(100) valleys. Adding on the two-fold degenerate ε-Si valleys, electrons in the ε-Si / ε-Ge dual-channel heterostructure on $Si_{0.3}Ge_{0.7}$ may "see" a fourteen-fold (twelve from the ε-Ge and two from the ε-Si) degeneracy! However, the present results cannot be used to determine whether defect-scattering or rapid intervalley scattering is the dominant mobility-limiting mechanism.

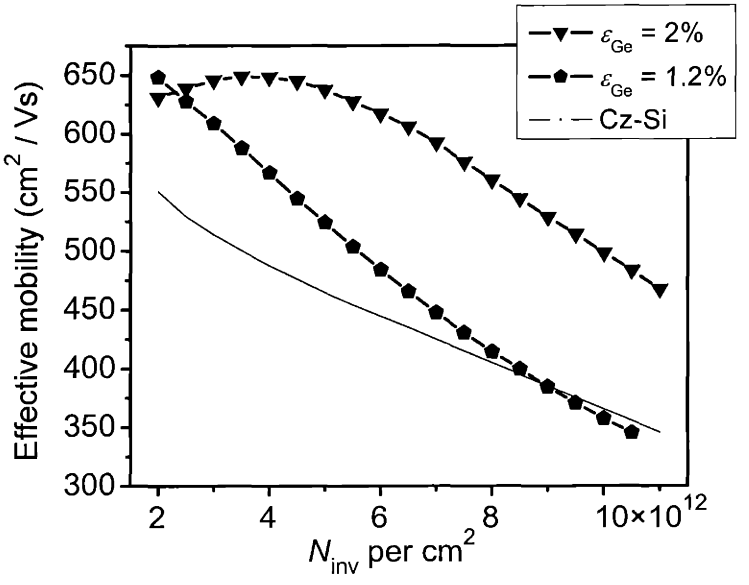


Figure 7.6 Electron effective mobility vs N_{inv} for ε-Si / ε-Ge dual-channel *n*-MOSFETs where the amount of strain in the Ge layer is 1.2% and 2%.

7.3. Digital-alloy channel *n*-MOSFETs

In Chapter 4, it was found that the hole mobility of digital-alloy channel *p*-MOSFETs was much larger than what could be expected from a stoichiometrically equivalent random-alloy. *n*-MOSFETs were also fabricated from the same digital-alloy heterostructure (alternating 8 Å layers of ε-Si and relaxed $Si_{0.3}Ge_{0.7}$, capped with 2 nm ε-Si) in order to study alloy scattering in *n*-channels. Leitz *et al.* found that random $Si_{1-x}Ge_x$ channels with x = 0.3 to 0.4 demonstrate electron mobilities considerably lower than Cz-Si.[6] As Figure 7.7 shows, the digital-alloy channel *n*-MOSFET demonstrates considerably higher electron mobility than either the $Si_{0.7}Ge_{0.3}$ or $Si_{0.6}Ge_{0.4}$ random-alloy channels. While no mobility enhancement over Cz-Si is seen for the digital-alloy channel *n*-MOSFET, the digital-alloy structure considerably reduces the random alloy scattering potential for electrons, just as was the case for the *p*-channel devices.

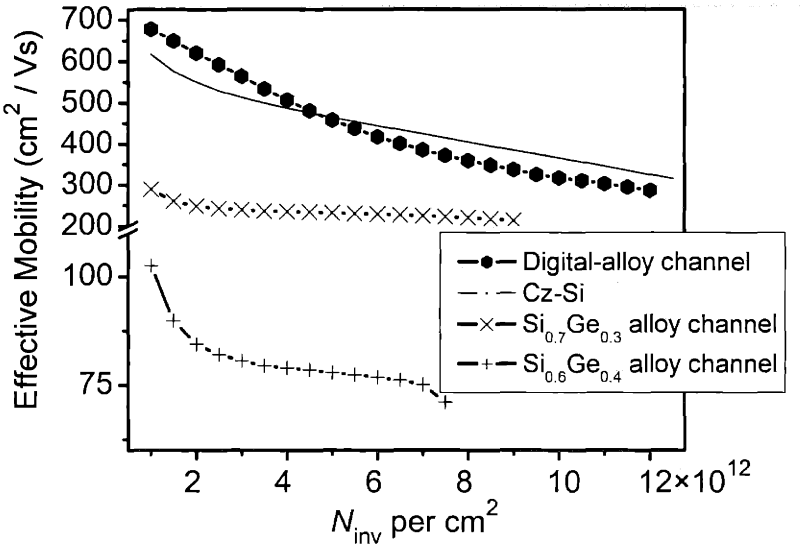


Figure 7.7 Electron effective mobility of $Si_{1-x}Ge_x$ random-alloy channel *n*-MOSFETs[6] compared with the digital-alloy channel *n*-MOSFET.

7.4. Summary

An investigation of electron mobility characteristics in single-, dual- and digital-alloy channel *n*-MOSFETs grown on Ge-rich virtual substrates has been presented. In the single-channel case, it was found that considerable mobility enhancements could be attained, but that defect scattering limits the electron mobility for devices fabricated on $Si_{1-x}Ge_x$ with $x \geq 0.7$. In the dual-channel case, it was confirmed that for a ε-Si cap thicker than 6 nm, the presence of the buried Ge-rich layer does not affect electron mobility; the type-II conduction band offset and large out-of-plane effective mass prevent electrons from ever "seeing" the buried layer. This fact may have practical significance, since it means that high mobility *n*-MOSFETs can be implemented on the same wafers as high hole mobility dual-channel heterostructures. As was shown in Figure 7.2, *n*-MOSFETs fabricated on ε-Si / ε-Ge heterostructures ($Si_{0.5}Ge_{0.5}$ virtual substrate) with $h_{Si\ cap}$ = 7.5 nm display practically identical μ_{eff} to single-channel ε-Si *n*-MOSFETs on $Si_{0.75}Ge_{0.25}$. Taking this result together with the highest mobility ε-Si / ε-Ge *p*-MOSFETs presented in Chapter 6, it is shown that nearly symmetric mobility *n*- and *p*-MOSFETs can be fabricated on a single epitaxial structure. Since the *n*-MOSFET demands $h_{Si\ cap}$ ~ 7 nm and the highest mobility *p*-MOSFETs had $h_{Si\ cap}$ = 3-5 nm, an extremely precise etching step might be needed to achieve optimal results; the cap could be slightly thinned in areas of the die intended for *p*MOS. Figure 7.8 emphasizes that the hole mobility always lags behind the electron mobility in Cz-Si CMOS and surface ε-Si CMOS. However, the *p*-channel μ_{eff} in the ε-Si / ε-Ge case actually matches or exceeds the *n*-channel μ_{eff} across the entire tested range of N_{inv}.

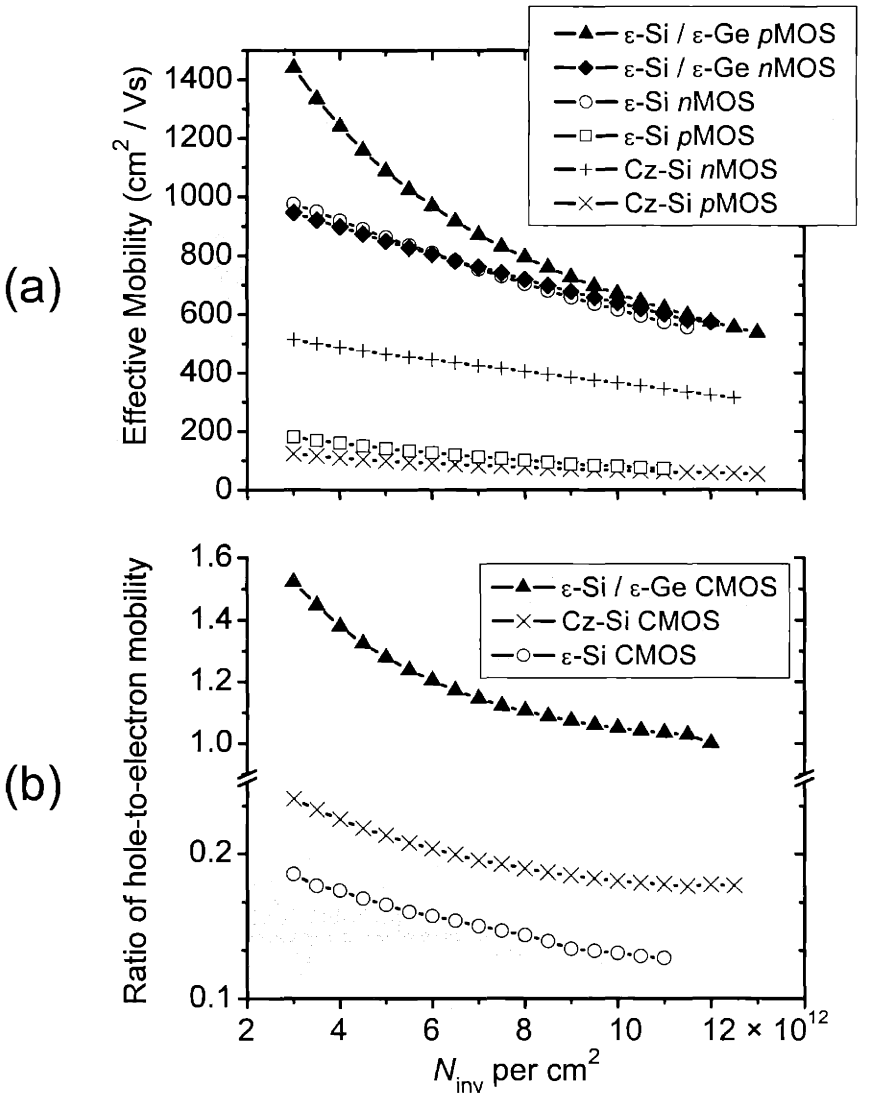


Figure 7.8 (a) *n*- and *p*-MOSFET effective mobility and (b) hole-to-electron mobility ratio vs N_{inv} for MOSFETs fabricated on Cz-Si, ε-Si[7], and ε-Si / ε-Ge on $Si_{0.5}Ge_{0.5}$.

Dual-channel heterostructures where the Si cap is *not* thick enough to confine the electron wave function were investigated in order to understand the effect of electron conduction through both ε-Ge and ε-Si. A direct comparison of heterostructures with and without a ε-Ge layer proved that the presence of ε-Ge can significantly degrade the electron mobility. The mismatch in conduction band minima between Ge and Si has been suggested as an electron mobility-limiting mechanism due to *X*-to-*L*-valley

scattering of inversion carriers. The results presented here show that the X-to-L-valley scattering potential can be reduced by lowering the Ge-content in the buried layer. Finally, it was found that SiGe digital-alloy channels possess much higher electron mobility than random-alloy channels with comparable net Ge content, indicating that the alloy scattering potential is greatly reduced by ordering of the layer structure in one dimension (i.e. the growth direction), just as was the case for holes.

Chapter 8. Conclusions and Suggestions for Future Work

8.1. Summary of Experimental Results

This thesis illustrates that after the initial deployment of strained Si in the CMOS industry, significant performance benefits can continue to be extracted from the relaxed-$Si_{1-x}Ge_x$ platform both by increasing the Ge content in the buffer of single-channel devices *and* by employing dual-channel heterostructures. Developing the growth technique needed for depositing planar, highly strained compressive and tensile films on virtual substrates was the critical first step underlying the device results presented in Chapters 5 through 7. Extremely low temperatures for growth were explored, and planar ε-Ge films with as much as 2% compressive strain were attained. Several methods for capping compressively strained layers with Si were developed, and the relative strengths and weaknesses of each were explored. The extremely clean background of the UHVCVD reactor was crucial to enabling growth at very low temperatures with no increase in impurities.

In Chapter 5, previously published results on ε-Si *p*-MOSFETs were reviewed, and it was found that at high N_{inv}, the exponent of the mobility curve, c, demonstrates an unexpected dependence on Ge content in the $Si_{1-x}Ge_x$ virtual substrate. For $0 \leq x \leq 0.35$, c increases with x, while for $0.35 \leq x \leq 0.5$, c decreases back down to the value for bulk Si. In the low-Ge samples, the strain-induced reduction of the out-of-plane effective mass may allow the hole wave function to penetrate into the gate oxide more readily, resulting in a more rapid drop in hole mobility at high fields. When the Ge content is made sufficiently high, the barrier between the relaxed $Si_{1-x}Ge_x$ and ε-Si helps to prevent the hole wave function from sampling the low-mobility SiO_2. Samples with Ge-rich virtual substrates (60% Ge and 70% Ge) exhibited much higher mobility enhancements

than any previously published single-channel ε-Si *p*-MOSFETs. The large enhancement seen at high field in these devices was attributed to the formation of a hybrid valence band, where the ε-Si layer provides valence band splitting and the Ge-rich buffer provides a low, Ge-like effective mass. A digital-alloy channel *p*-MOSFET that exhibited constant mobility enhancement as a function of N_{inv} was also demonstrated, proving that the slope of the mobility curve can be greatly engineered in nanometer-scale ε-Si heterostructures.

Strained Si / Strained Ge dual-channel heterostructures were explored extensively in Chapter 6. Strain, Si cap thickness, and ε-Ge thickness were seen to be intertwined variables that could each dominate the mobility characteristics of a device. However, it was found that the combination of a highly strained Ge layer (> 1%), that was also relatively thick (8-12 nm), with a Si cap 3-5 nm in thickness, led to the highest hole mobilities.

In Chapter 7, ε-Si single-channel *n*-MOSFETs were demonstrated on virtual substrates with Ge contents as high as 70%. While the mobility was somewhat depressed compared to lower-Ge content samples, enhancements of 40-60% were still observed. For dual-channel devices, the highest electron mobility was attained when the electron wave function could be completely isolated from the buried layer (i.e. by growing the Si cap thicker than 6 nm). This led to the demonstration that nearly symmetric mobility *n*- and *p*-MOSFETs could be fabricated on the same ε-Si / ε-Ge / relaxed-$Si_{0.5}Ge_{0.5}$ heterostructure. In devices where the Si cap is extremely thin (~ 3 nm), electrons undergo strong intervalley phonon scattering due to the difference in conduction band minima between Ge and Si, and the resulting effective mobility was observed to be lower

than that of Cz-Si. Therefore despite the high electron mobility of bulk Ge, *n*-MOSFETs where the electron wave function can sample both pure Ge and pure Si will always have low mobility.

8.2. Future Work

8.2.1. Further study of highly strained heterostructures

While Figure 8.1 presents an extremely broad overview of possible mobility enhancements in single-channel and dual-channel heterostructures, it is likely that a wide range of new insight and phenomena could be discovered if the parameter space were extended even further. For example, all of the experimental evidence in Chapter 5 suggests that single-channel heterostructures grown on $Si_{1-x}Ge_x$ with $x \geq 0.8$ should lead to even larger mobility enhancements. To date, however, this has not been achieved, most likely due to large defect densities in the Si cap. Growing extremely thin Si caps (i.e. < 3 nm) on $Si_{1-x}Ge_x$ with $x \geq 0.8$ should somewhat mitigate the defect nucleation problem, and such heterostructures should be tested for hole mobility enhancement.

The concept of digital-alloy MOSFET channels should also be investigated in greater detail. For example, symmetrically strained (net zero strain) heterostructures could be grown using only Ge and Si layers (e.g. 1 nm ε-Ge alternated with 1 nm ε-Si to simulate a $Si_{0.5}Ge_{0.5}$ alloy). By varying the periodicity, great insight about the fundamental nature of random-alloy scattering of both holes and electrons could be gained.

The limits of ε-Si / ε-Ge dual-channel heterostructures should be tested by utilizing even lower Ge-content virtual substrates to attain higher strain in the Ge layer.

Again, all of the evidence seen in this thesis indicates that the hole mobility should continue to rise with strain. However, at a certain point, defect scattering is expected to dominate the mobility characteristics. Another potentially interesting modification to the dual-channel heterostructure would be to insert a ε-Si layer *below* the ε-Ge layer (in addition to the one on top), forming a triple heterostructure. Growing a lower ε-Si "cladding layer" may encourage confinement of the hole wave function in the ε-Ge, obviating the need for growing thick (i.e. ~10 nm) ε-Ge channels.

8.2.2. Advanced devices

While the MOSFETs fabricated in this work demonstrate extremely high effective mobilities, the barriers to incorporating what has been learned here in any commercial setting are formidable. For example, the physical juxtaposition of ε-Si and ε-Ge will lead to massive potential for interdiffusion during high temperature processing steps. Furthermore, the highest hole mobility structures shown here contain layers of super-critical thickness (e.g. 12 nm ε-Ge on $Si_{0.5}Ge_{0.5}$). Understanding the impact of misfit dislocation nucleation on device reliability and overall yields will be critical to determining the commercial prospects of devices with ε-Ge channels. The subthreshold problems that are anticipated for dual-channel heterostructures must also be analyzed in detail. While it is well-known that buried-channel operation leads to increased subthreshold swing, the design of MOSFETs in two dimensions (i.e. through doping and junction profiles, etc.) often controls the subthreshold performance of deeply scaled devices. Therefore, it must be determined which aspect will dominate the subthreshold characteristics before a definitive judgment can be made. As always, detailed device models are a necessary complement to ongoing experimental work.

One area of great scientific and technological interest would be to implement MOSFETs that combine the high mobility of a ε-Ge channel with the scaling benefits of a high-κ gate dielectric such as HfO_2 or ZrO_2. While the difficulties of deploying such a technology would be immense, the growing knowledge-base in the processing of ε-Si devices combined with a concurrently developing knowledge-base on the processing of high-κ films may ultimately lead to a roadmap that includes both ε-Ge and high-κ gate dielectrics. With suitable high-κ processing technology, the need for the Si cap may be eliminated, allowing the creation of "single-channel" ε-Ge devices. If this is the case, the strategy for attaining the best mobility enhancements for both *n*- and *p*-channel devices changes slightly. For dual-channel devices, since the ε-Si is generally meant to serve as the electron channel, the strain in the Ge layer should be maximized to attain the best hole mobility. However, if the theory of Fischetti is correct, then the electron mobility will be greatly degraded in highly strained Ge.[58] Thus, in Ge single-channel devices, any increase in hole mobility may carry with it a loss in electron mobility.

8.3. Conclusions

The work in this thesis has continued to expose the useful phase space of ε-Si, ε-SiGe, and ε-Ge heterostructures grown on relaxed $Si_{1-x}Ge_x$ through the use of the short-flow MOSFET process. Figure 8.1 illustrates five years of progress made in the Fitzgerald group on high mobility single- and dual-channel *n*- and *p*-MOSFETs.

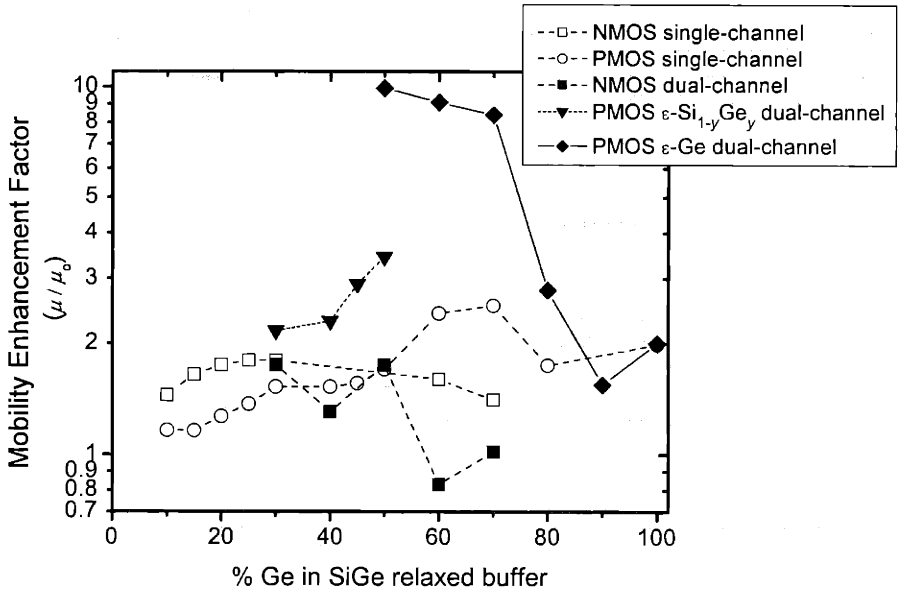


Figure 8.1 Mobility enhancement of single- and dual-channel heterostructures explored in the Fitzgerald group since 1998. All enhancements were computed at $N_{inv} = 9.5 \times 10^{12}$ cm^{-2}.

Appendix A. Short Flow MOSFET Fabrication Sequence

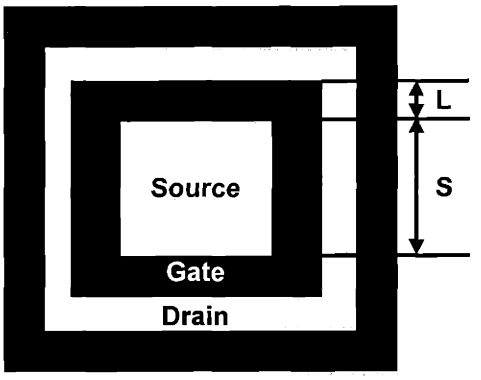
The following section outlines the MTL machines and recipes used in short flow MOSFET fabrication. Both MIT's Integrated Circuits Laboratory (ICL) and Technology Research Laboratory (TRL) were used in this process. In cases where more than one machine could be utilized for a given step, the preferred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

Process Module	Machine	Description
Gate Stack Deposition	ICL RCA	Modified RCA Clean (5-10 min. piranha clean, 15 s 50:1 H_2O:HF dip, 2.5-15 min. SC-2 clean)
Gate Stack Deposition	ICL tubeA7	Deposition of 3000 Å LTO (recipe 462, 400°C; average deposition rate ~ 100 Å/min.)
Gate Stack Deposition	ICL tubeA6	Deposition of 500 Å poly-Si (recipe 705, 560°C; average deposition rate ~16 Å/min.)
Backside Clear	ICL HMDS	Coat wafers with HMDS; standard recipe
Backside Clear	ICL coater6	Coat wafer fronts with 1 μm positive resist and post-bake; standard recipe
Backside Clear*	ICL oxide	Remove backside native oxide with 5 s BOE dip
Backside Clear	ICL AME5000	Remove backside poly-Si with poly_std etch
Backside Clear	ICL oxide	Remove backside LTO with 30 s BOE dip
Backside Clear	ICL pre-metal or *ICL asher*	Remove photoresist with either standard asher recipe or 10 min. piranha clean
Frontside Patterning	ICL HMDS	Coat wafers with HMDS; standard recipe
Frontside Patterning	ICL coater6	Coat wafer fronts with 1 μm positive resist; standard recipe
Frontside Patterning	ICL stepper2 or *TRL EV1*	Expose photoresist through MOBIL mask
Frontside Patterning	ICL developer or *TRL photowet-1*	Develop photoresist
Frontside Patterning	ICL coater6 or *TRL postbake*	Post-bake photoresist
Frontside Patterning*	ICL oxide	Remove frontside native oxide in exposed regions with 5 s BOE dip
Frontside Patterning	ICL AME5000	Etch frontside poly-Si in exposed regions with poly_std etch
Frontside Patterning	ICL AME5000	Etch frontside LTO in exposed regions with Matt_LTO etch, leaving ~250 Å LTO remaining

Frontside Patterning	ICL UV1280	Use ellipsometry to verify pre- and post-etch LTO thicknesses and calibrate LTO etch rates
Frontside Patterning	*ICL asher* or TRL asher	Remove photoresist with standard asher recipe
Frontside Patterning	*ICL oxide* or TRL acid-hood	Remove remaining LTO and undercut gates with 30 s BOE dip
Implant	Outside vendor	BF_2 or As, 35 keV, 1×10^{15} ions cm^{-2}; 4 identical implants at 90° rotation
Implant	TRL acid hood	Post-implant clean; 2 × 60 s piranha clean
Implant	TRL tubeA2	30 min. anneal at 600°C in N_2 ambient
Metallization	ICL pre-metal	Pre-metal clean; 60s piranha clean, 20 s 50:1 H_2O:HF dip
Metallization	ICL e-beam	Deposit 500 Å Ti + 1000 Å Al on wafer fronts
Metallization*	ICL e-beam	Deposit 5000 Å Al on wafer backsides
Sinter	TRL tubeA3	30 min. anneal at 400°C in N_2/H_2 ambient

* Optional

Appendix B Transistor layout and mobility extraction



Planar view of MOSFETs measured in this thesis

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

Relevant dimensions of MOSFETs characterized in this work.

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

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

where W/L is replaced by the inverse of the geometry factor, G. $|V_{DS}|$ was set to 0.1 V for all measurements. A more detailed description of the derivation of G was given by Armstrong.[85]

Bibliography

1. M. Bulsara, "Strained silicon joins the drive to keep CMOS chips on course." *Compound Semiconductor*, 39-40 (2002).
2. D. Lammers. "Industry reacts to Intel's strained silicon move", **2002,** http://www.eetimes.com/story/OEG20020816S0048 (2002).
3. K. Rim, S. Koester, M. Hargrove, J. O. Chu, P. M. Mooney, J. Ott, T. Kanarsky, P. Ronsheim, M. Ieong, A. Grill, and H.-S. P. Wong, "Strained Si NMOSFETs for High Performance CMOS Technology." Symposium on VLSI Technology, Kyoto, Japan(2001).
4. D. K. Nayak, J. C. S. Woo, J. S. Park, K. L. Wang, and K. P. MacWilliams, "High-mobility *p*-channel metal-oxide-semiconductor field-effect transistor on strained Si." *Applied Physics Letters* **62**, 2853-2855 (1993).
5. K. Rim, J. Welser, J. L. Hoyt, and J. F. Gibbons, "Enhanced Hole Mobilites in Surface-channel Strained-Si p-MOSFETs." IEEE IEDM, 517-519 (1995).
6. C. W. Leitz, M. T. Currie, M. L. Lee, Z. Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, "Hole Mobility Enhancements and Alloy Scattering-Limited Mobility in Tensile Strained Si/SiGe Surface Channel Metal-Oxide-Semiconductor Field-Effect Transistors." *Journal of Applied Physics* **92**, 3745-3751 (2002).
7. M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, D. A. Antoniadis, and E. A. Fitzgerald, "Carrier mobilities and process stability of strained Si *n*- and *p*-MOSFETs on SiGe virtual substrates." *Journal of Vacuum Science and Technology B* **19**, 2268-2279 (2001).

8. K. Rim, J. Chu, H. Chen, K. A. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Ieong, and H.-S. Wong, "Characteristics and Device Design of Sub-100 nm Strained Si N- and PMOSFETs." Symposium on VLSI Technology, San Francisco, CA(2002).

9. E. A. Fitzgerald, "GeSi/Si Nanostructures." *Annual Review of Materials Science* **25**, 417-54 (1995).

10. S. K. Gandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*. New York, John Wiley & Sons (1994).

11. E. A. Fitzgerald, Y.-H. Xie, M. L. Green, D. Brasen, A. R. Kortan, J. Michel, Y.-J. Mii, and B. E. Weir, "Totally relaxed Ge_xSi_{1-x} layers with low threading dislocation densities grown on Si substrates." *Applied Physics Letters* **59**, 811-813 (1991).

12. J. W. Matthews, in *Epitaxial Growth Part B*, edited by J. W. Matthews (Academic Press, New York, 1975), p. 559-607.

13. E. A. Fitzgerald, "Dislocations in strained-layer epitaxy:theory, experiment, and applications." *Materials Science Reports* **7**, 87-142 (1991).

14. C. W. Leitz, "High mobility strained Si/SiGe heterostructure MOSFETs: Channel engineering and virtual substrate optimization," Ph.D., MIT (2002).

15. E. A. Fitzgerald, Y.-H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, A. R. Kortan, F. A. Thiel, and B. E. Weir, "Relaxed Ge_xSi_{1-x} structures for III-V integration with Si and high mobility two-dimenstional electron gases in Si." *Journal of Vacuum Science and Technology B* **10**, 1807-1819 (1992).

16. E. A. Fitzgerald, A. Y. Kim, M. T. Currie, T. A. Langdo, G. Taraschi, and M. T. Bulsara, "Dislocation Dynamics in Relaxed Graded Composition Semiconductors." *Materials Science and Engineering* **B67**, 53-61 (1999).

17. S. B. Samavedam and E. A. Fitzgerald, "Novel dislocation structure and surface morphology effects in relaxed Ge/Si-Ge(graded)/Si structures." *Journal of Applied Physics* **81**, 3108-3116 (1997).

18. M. T. Currie, S. B. Samavedam, T. A. Langdo, C. W. Leitz, and E. A. Fitzgerald, "Controlling threading dislocation densities in Ge on Si using graded SiGe layers and chemical-mechanical polishing." *Applied Physics Letters* **72**, 1718-1720 (1998).

19. M. E. Groenert, C. W. Leitz, A. J. Pitera, V. Yang, H. Lee, R. J. Ram, and E. A. Fitzgerald, "Monolithic integration of room-temperature cw GaAs/AlGaAs lasers on Si substrates via relaxed graded GeSi buffer layers." *Journal of Applied Physics* **93**, 362-367 (2003).

20. J. A. Carlin, S. A. Ringel, E. A. Fitzgerald, M. Bulsara, and B. M. Keyes, "Impact of GaAs buffer thickness on electronic quality of GaAs grown on graded Ge/GeSi/Si substrates." *Applied Physics Letters* **76**, 1884-1886 (2000).

21. V. K. Yang, M. E. Groenert, G. Taraschi, C. W. Leitz, A. J. Pitera, M. T. Currie, Z.-Y. Cheng, and E. A. Fitzgerald, "Monolithic integration of III-V optical interconnects on Si using SiGe virtual substrates." *Journal of Materials Science: Materials in Electronics* **13**, 377-380 (2002).

22. Y. H. Xie, E. A. Fitzgerald, D. Monroe, P. J. Silverman, and G. P. Watson, "Fabrication of high mobility two-dimensional electron and hole gases in GeSi/Si." *Journal of Applied Physics* **73**, 8364-8370 (1993).

23. B. S. Meyerson, "Low-temperature silicon epitaxy by ultrahigh vacuum/chemical vapor deposition." *Applied Physics Letters* **48**, 797-799 (1986).

24. S. M. Sze, *Physics of Semiconductor Devices*. New York, John Wiley and Sons (1981).

25. B. G. Streetman, *Solid State Electronic Devices*. Englewood Cliffs, NJ, Prentice Hall (1995).

26. H. Veendrick, *Deep-Submicron CMOS ICs*. Boston, USA, Kluwer academic publishers (2000).

27. S.-M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits-Analysis and Design*. Boston, McGraw Hill (2003).

28. S.-i. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the Universality of Inversion Layer Mobility in Si MOSFET's:Part I-Effects of Substrate Impurity Concentration." *IEEE Transactions on Electron Devices* **41**, 2357-2362 (1994).

29. M. Lundstrom, *Fundamentals of Carrier Transport*. Reading, MA, Addison-Wesley Publishing Company (1990).

30. J. A. Del Alamo, *Integrated Microelectronic Devices: Physics & Modeling*. New York, Prentice Hall (2000).

31. M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, "Effective electron mobility in Si inversion layers in metal-oxide-semiconductor systems with a high-K

insulator: The role of remote phonon scattering." *Journal of Applied Physics* **90**, 4587-4608 (2001).

32. T. Fromherz and G. Bauer, in *Properties of strained and relaxed Silicon Germanium; Vol. 12*, edited by E. Kasper (INSPEC, London, U.K., 1994), p. 87-93.

33. Y. H. Xie, "SiGe field effect transistors." *Materials Science and Engineering Reports* **R25**, 89-119 (1999).

34. W. Lu, A. Kuliev, S. J. Koester, X.-W. Wang, J. O. Chu, T. P. Ma, and I. Adesida, "High Performance 0.1 um Gate-Length P-Type SiGe MODFET's and MOS-MODFET's." *IEEE Transactions on Electron Devices* **47**, 1645-1652 (2000).

35. P. M. Garone, V. Venkataraman, and J. C. Sturm, "Hole Mobility Enhancement in MOS-Gated Ge_xSi_{1-x}/Si Heterostructure Inversion Layers." *IEEE Electron Device Letters* **13**, 56-58 (1990).

36. S. K. Chun and K. L. Wang, "Effective Mass and Mobility of Holes in Strained $Si_{1-x}Ge_x$ Layers on (001) $Si_{1-y}Ge_y$ Substrate." *IEEE Transactions on Electron Devices* **39**, 2153-2164 (1992).

37. R. J. P. Lander, Y. V. Ponomarev, v. B. J.G.M., and W. B. de Boer, "High Hole Mobilities in Fully-Strained $Si_{1-x}Ge_x$ Layers ($0.3<x<0.4$) and their significance for SiGe pMOSFET performance." *IEEE Transactions on Electron Devices* **48**, 1826-1832 (2001).

38. R. People and J. C. Bean, "Band Alignments of Coherently Strained $Ge_x/Si_{1-x}/Si$ Heterostructures on <001> Ge_ySi_{1-y} Substrates." *Applied Physics Letters* **48**, 538-540 (1986).

39. D. Monroe, Y. H. Xie, E. A. Fitzgerald, P. J. Silverman, and G. P. Watson, "Comparison of mobility-limiting mechanisms in high-mobility $Si_{1-x}Ge_x$ heterostructures." *Journal of Vacuum Science and Technology B* **11**, 1731-1737 (1993).

40. G. Abstreiter, H. Brugger, T. Wolf, H. Jorke, and H. J. Herzog, "Strain-Induced Two-Dimensional Electron Gas in Selectively Doped Si/Si_xGe_{1-x} Superlattices." *Physical Review Letters* **54**, 2441-2444 (1985).

41. G. Schuberth, F. Schaffler, M. Besson, G. Abstreiter, and E. Gornik, "High electron mobility in modulation-doped Si/SiGe quantum well structures." *Applied Physics Letters* **59**, 3318-3320 (1991).

42. Y. J. Mii, Y. H. Xie, E. A. Fitzgerald, D. Monroe, F. A. Thiel, B. E. Weir, and L. C. Feldman, "Extremely high electron mobility in Si/Ge_xSi_{1-x} structures grown by molecular beam epitaxy." *Applied Physics Letters* **59**, 1611-1613 (1991).

43. F. Schaffler, "High-Mobility Si and Ge Structures." *Semiconductor Science and Technology* **12**, 1515-1549 (1997).

44. F. Schaffler, D. Tobben, H. J. Herzog, G. Abstreiter, and B. Hollander, "High-electron mobility Si/SiGe heterostructures: influence of the relaxed SiGe buffer layer." *Semiconductor Science and Technology* **7**, 260-266 (1992).

45. J. Welser, J. L. Hoyt, and J. F. Gibbons, "Electron Mobility Enhancement in Strained-Si N-Type Metal-Oxide-Semiconductor Field-Effect Transistors." *IEEE Electron Device Letters* **15**, 100-102 (1994).

46. A. G. O'Neill, P. Routley, P. K. Gurry, P. A. Clifton, H. Kemhadjian, J. Fernandez, A. G. Cullis, and A. Benedetti, "SiGe virtual substrate N-channel heterojunction MOSFETs." *Semiconductor Science and Technology* **14**, 784-789 (1999).

47. J. Welser, J. Hoyt, S.-i. Takagi, and J. F. Gibbons, "Strain Dependence on the Performance Enhancement in Strained-Si n-MOSFETs." IEEE IEDM, San Francisco, CA(1994).

48. H. M. Nayfeh, J. L. Hoyt, C. W. Leitz, A. J. Pitera, E. A. Fitzgerald, and D. A. Antoniadis, "Electron inversion layer mobility in strained-Si n-MOSFET's with high channel doping concentration achieved by ion implantation." Device Research Conference, Santa Barbara, CA., 43-44 (2002).

49. A. Lochtefeld and D. A. Antoniadis, "Investigating the Relationship Between Electron Mobility and Velocity in Deeply Scaled NMOS via Mechanical Stress." *IEEE Electron Device Letters* **22**, 591-593 (2001).

50. R. Oberhuber, G. Zandler, and P. Vogl, "Subband structure and mobility of two-dimensional holes in strained Si/SiGe MOSFET's." *Physical Review B* **58**, 9941-9948 (1998).

51. D. K. Nayak, K. Goto, A. Yutani, J. Murota, and Y. Shiraki, "High Mobility Strained-Si PMOSFET's." *IEEE Transactions on Electron Devices* **43**, 1709-1716 (1996).

52. C. K. Maiti, L. K. Bera, S. S. Dey, D. K. Nayak, and N. B. Chakrabarti, "Hole mobility enhancement in strained-Si p-MOSFETs under high vertical field." *Solid-State Electronics* **41**, 1863-1869 (1997).

53. T. Mizuno, N. Sugiyama, T. Tezuka, T. Numata, and S. I. Takagi, "High Performance CMOS Operation of Strained-SOI MOSFETs using Thin Film SiGe-on-Insulator Substrate." Symposium on VLSI Technology, San Francisco, CA(2002).

54. C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, "Hole mobility enhancements in strained $Si/Si_{1-y}Ge_y$ *p*-type metal-oxide-semiconductor field-effect transistors grown on relaxed $Si_{1-x}Ge_x$ ($x<y$) virtual substrates." *Applied Physics Letters* **79**, 4246-4248 (2001).

55. G. Hock, E. Kohn, C. Rosenblad, H. von Kanel, H.-J. Herzog, and U. Konig, "High hole mobility in $Si_{0.17}Ge_{0.83}$ channel metal-oxide-semiconductor field-effect transistors grown by plasma-enhanced chemical vapor deposition." *Applied Physics Letters* **76**, 3920-3922 (2000).

56. M. Arafa, K. Ismail, P. Fay, J. Chu, B. S. Meyerson, and I. Adesida, "High transconductance p-type SiGe modulation-doped field-effect transistor." *Electronics Letters* **31**, 680-681 (1995).

57. F. M. Bufler and B. Meinerzhagen, "Hole transport in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates." *Journal of Applied Physics* **84**, 5597-5602 (1998).

58. M. V. Fischetti and S. E. Laux, "Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys." *Journal of Applied Physics* **80**, 2234-2252 (1996).

59. M. Arafa, P. Fay, K. Ismail, J. O. Chu, B. S. Meyerson, and I. Adesida, "High Speed P-Type SiGe Modulation-Doped Field-Effect Transistors." *IEEE Electron Device Letters* **17**, 124-126 (1996).

60. M. Arafa, P. Fay, K. Ismail, J. O. Chu, B. S. Meyerson, and I. Adesida, "DC and RF Performance of 0.25 um p-Type SiGe MODFET." *IEEE Electron Device Letters* **17**, 449-451 (1996).

61. T. Hackbarth, G. Hock, H. J. Herzog, and M. Zeuner, "Strain relieved SiGe buffers for Si-based heterostructure field-effect transistors." *Journal of Crystal Growth* **201/202**, 734-738 (1999).

62. K. Ismail, J. O. Chu, and B. S. Meyerson, "High hole mobility in SiGe alloys for device applications." *Applied Physics Letters* **64**, 3124-3126 (1994).

63. C. W. Leitz, M. T. Currie, Z.-Y. Cheng, M. L. Lee, D. A. Antoniadis, and E. A. Fitzgerald, "Optimized Heterostructures for SiGe-Based CMOS Applications." MRS, Boston, MA, A3.10.1-A.3.10.6 (2001).

64. J.-W. Jung, M. L. Lee, S. F. Yu, E. A. Fitzgerald, and D. A. Antoniadis, "Implementation of both high hole and electron mobility in strained Si /Strained Si_{1-y} Ge_y on relaxed $Si_{1-x}Ge_x$ ($x<y$) virtual substrate." *IEEE Electron Device Letters* **To be published** (2003).

65. D. J. Hymes and J. J. Rosenberg, "Growth and Materials Characterization of Native Germanium Oxynitride Thin Films on Germanium." *Journal of the Electrochemical Society* **135**, 961-965 (1988).

66. D. C. Paine, J. J. Rosenberg, S. C. Martin, D. Luo, and M. Kawasaki, "Evaluation of device quality germanium-germanium oxynitride interfaces by high-resolution

transmission electron microscopy." *Applied Physics Letters* **57**, 1443-1445 (1990).

67. D. Reinking, M. Kammler, N. Hoffmann, M. Horn-von Hoegen, and K. R. Hofmann, "Fabrication of high-mobility Ge *p*-channel MOSFETs on Si substrates." *Electronics Letters* **35**, 503-504 (1999).

68. T. N. Jackson, C. M. Ransom, and J. F. DeGelormo, "Gate-Self-Aligned p-Channel Germanium MISFET's." *IEEE Electron Device Letters* **12**, 605-607 (1991).

69. S. C. Martin, L. M. Hitt, and J. J. Rosenberg, "p-Channel Germanium MOSFET's with High Channel Mobility." *IEEE Electron Device Letters* **10**, 325-326 (1989).

70. J. J. Rosenberg and S. C. Martin, "Self-Aligned Germanium MOSFET's Using a Nitrided Native Oxide Gate Insulator." *IEEE Electron Device Letters* **9**, 639-640 (1988).

71. H. Shang, H. Okorn-Schmidt, K. K. Chan, M. Copel, J. A. Ott, P. M. Kozlowski, S. E. Steen, S. A. Cordes, H.-S. P. Wong, E. C. Jones, and W. E. Haensch, "High-mobility p-channel Germanium MOSFETs with a Thin Ge Oxynitride Gate Dielectric." IEDM, San Francsico, CA.(2002).

72. C. O. Chui, S. Ramanathan, B. B. Triplett, P. C. McIntyre, and K. C. Saraswat, "Germanium MOS Capacitors Incorporating Ultrathin High-K Gate Dielectric." *IEEE Electron Device Letters* **23**, 473-475 (2002).

73. C. O. Chui, H. Kim, D. Chi, B. B. Triplett, P. C. McIntyre, and K. C. Saraswat, "A sub-400 degrees C germanium MOSFET technology with high-K dielectric and metal gate." IEDM, San Francisco, CA.(2002).

80. R. Hammond, S. J. Koester, and J. O. Chu, "High performance 0.1um gate-length Ge/$Si_{0.4}Ge_{0.6}$ *p*-channel MODFETs." *Electronics Letters* **35**, 1590-1591 (1999).

81. G. Hock, T. Hackbarth, U. Erben, E. Kohn, and U. Konig, "High performance 0.25um *p*-type Ge/SiGe MODFETs." *Electronics Letters* **34**, 1888-1889 (1998).

82. E. Murakami, K. Nakagawa, A. Nishida, and M. Miyao, "Fabrication of a Strain-Controlled SiGe/Ge MODFET with Ultrahigh Hole Mobility." *IEEE Transactions on Electron Devices* **41**, 857-861 (1994).

83. G. Hock, T. Hackbarth, N. Kab, H. J. Herzog, M. Enciso, F. Aniel, P. Crozat, R. Adde, E. Kohn, and U. Konig, "0.1 um gate length *p*-type Ge/$Si_{0.4}Ge_{0.6}$ MODFET with 135 GHz f_{max}." *Electronics Letters* **36**, 1428-1429 (2000).

84. S. J. Koester, R. Hammond, and J. O. Chu, "Extremely High Transconducatnce Ge/$Si_{0.4}Ge_{0.6}$ p-MODFET's Grown by UHV-CVD." *IEEE Electron Device Letters* **21**, 110-112 (2000).

85. M. Armstrong, "Technology for SiGe Heterostructure-Based CMOS Devices," PhD, Massachusetts Institute of Technology (1999).

86. D. J. Tweet, T. Tatsumi, H. Hirayama, K. Miyanaga, and K. Terashima, "Factors determining the composition of strained GeSi layers grown with disilane and germane." *Applied Physics Letters* **65**, 2579-2581 (1994).

87. D. J. Eaglesham and C. M., "Dislocation-Free Stranski-Krastanov Growth of Ge on Si(100)." *Physical Review Letters* **64**, 1943-1950 (1990).

88. D. K. Nayak and S. K. Chun, "Low-field hole mobility of strained Si on (100) $Si_{1-x}Ge_x$ substrate." *Applied Physics Letters* **64**, 2514-2516 (1994).

74. W. P. Bai, N. Lu, J. Liu, A. Ramirez, D. L. Kwong, D. Wristers, A. Ritenour, M. L. Lee, and D. A. Antoniadis, "Ge MOS Characteristics with CVD HfO_2 Gate Dielectrics and TaN Gate Electrode." 2003 Symposium on VLSI Technology, Kyoto, Japan(2003).

75. Y. H. Xie, D. Monroe, E. A. Fitzgerald, P. J. Silverman, F. A. Thiel, and G. P. Watson, "Very high mobility two-dimensional hole gas in $Si/Ge_xSi_{1-x}/Ge$ structures grown by molecular beam epitaxy." *Applied Physics Letters* **63**, 2263-2264 (1993).

76. U. Konig and F. Schaffler, "p-Type Ge-Channel MODFET's with High Transconductance Grown on Si Substrates." *IEEE Electron Device Letters* **14**, 205-207 (1993).

77. M. Myronov, T. Irisawa, O. A. Mironov, S. Koh, Y. Shiraki, T. E. Whall, and H. C. Parker, "Extremely high room-temperature two-dimensional hole gas mobility in $Ge/Si_{0.33}Ge_{0.67}/Si(001)$ *p*-type modulation-doped heterostructures." *Applied Physics Letters* **80**, 3117-3119 (2002).

78. T. Irisawa, S. Tokumitsu, T. Hattori, K. Nakagawa, S. Koh, and Y. Shiraki, "Ultrahigh room-temperature hole Hall and effective mobility in $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$ heterostructures." *Applied Physics Letters* **81**, 847-849 (2002).

79. T. Irisawa, H. Miura, T. Ueno, and Y. Shiraki, "Channel Width Dependence on Mobility in Ge Channel Modulation-Doped Structures." *Japan Journal of Applied Physics* **40**, 2694-2696 (2001).

89. Y. H. Xie, G. H. Gilmer, C. Roland, P. J. Silverman, S. K. Buratto, J. Y. Cheng, E. A. Fitzgerald, A. R. Kortan, S. Schuppler, M. A. Marcus, and P. H. Citrin, "Semiconductor Surface Roughness: Dependence on Sign and Magnitude of Bulk Strain." *Physical Review Letters* **73**, 3006-3009 (1994).

90. J. C. Hensel and G. Feher, "Cyclotron Resonance Experiments in Uniaxially Stressed Silicon: Valence Band Inverse Mass Parameters and Deformation Potentials." *Physical Review* **129**, 1041-1066 (1963).

91. F. Stern, "Self-Consistent Results for n-Type Si Inversion Layers." *Physical Review B* **5**, 4891-4899 (1972).

92. F. Rana, S. Tiwari, and D. A. Buchanan, "Self-consistent modeling of accumulation layers and tunneling currents through very thin oxides." *Applied Physics Letters* **69**, 1104-1106 (1996).

93. M. M. A. Hakim and A. Haque, "Effects of Neglecting Carrier Tunneling on Electrostatic Potential in Calculating Direct Tunneling Gate Current in Deep Submicron MOSFETs." *IEEE Transactions on Electron Devices* **49**, 1669-1671 (2002).

94. M. Z. Kauser, M. S. Hasan, and A. Haque, "Effects of Wave Function Penetration Into the Gate-Oxide on Self-Consistent Modeling of Scaled MOSFETs." *IEEE Transactions on Electron Devices* **49**, 693-695 (2002).

95. A. Haque and M. Z. Kauser, "A Comparison of Wave-Function Penetration Effects on Gate Capacitance in Deep Submicron n- and p-MOSFETs." *IEEE Transactions on Electron Devices* **49**, 1580-1587 (2002).

96. S. Mudanai, L. F. Register, A. F. Tasch, and S. K. Banerjee, "Understanding the Effects of Wave Function Penetration on the Inversion Layer Capacitance of NMOSFETs." *IEEE Electron Device Letters* **22**, 145-147 (2001).

97. M. Yunus and A. Haque, "Wave function penetration effects on current-voltage characteristics of ballistic metal-oxide-semiconductor transistors." *Journal of Applied Physics* **93**, 600-604 (2003).

98. M. V. Fischetti, F. Gamiz, and W. E. Haensch, "On the enhanced electron mobility in strained-silicon inversion layers." *Journal of Applied Physics* **92**, 7320-7324 (2002).

99. I. Polishchuk and C. Hu, "Electron Wavefunction Penetration into Gate Dielectric and Interface Scattering - An Alternative to Surface Roughness Scattering Model." Symposium on VLSI Technology(2001).

100. F. Assaderaghi, D. Sinitsky, J. Bokor, P. K. Ko, H. Gaw, and C. Hu, "High-Field Transport of Inversion-Layer Electrons and Holes Including Velocity Overshoot." *IEEE Transactions on Electron Devices* **44**, 664-671 (1997).

101. T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems." *Review of Modern Physics* **54**, 437-621 (1982).

102. M. L. Lee, C. W. Leitz, Z.-Y. Cheng, A. J. Pitera, G. Taraschi, D. A. Antoniadis, and E. A. Fitzgerald, " Strained Ge Channel p-type MOSFETs Fabricated on $Si_{1-x}Ge_x$/Si Virtual Substrates." Materials Research Society Symposium, Boston, MA, A.1.9.1-A.1.9.5 (2001).

103. M. L. Lee, Z.-Y. Cheng, C. W. Leitz, A. J. Pitera, T. A. Langdo, M. T. Currie, G. Taraschi, E. A. Fitzgerald, and D. A. Antoniadis, "Strained Ge channel *p*-type

metal-oxide-semiconductor field-effect transistors grown on $Si_{1-x}Ge_x$/Si virtual substrates." *Applied Physics Letters* **79**, 3344-3346 (2001).

104. F. Schaffler, in *Properties of strained and relaxed Silicon Germanium*; *Vol. 12*, edited by E. Kasper (INSPEC, London, U.K., 1994), p. 135-144.