Heteroepitaxial Ge on Si via High-Bandgap III-V Buffers for Low-Power Electronic Applications

Peter D. Nguyen

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Electrical Engineering

Mantu K. Hudait. Chair
Marius K. Orlowski
Jean J. Heremans

May 04, 2016
Blacksburg, Virginia

Keywords: Germanium, III-V, Metal-Oxide-Semiconductor Devices, Metal-Oxide-Semiconductor Field-Effect Transistors, Heterogeneous Integration, Silicon

Heteroepitaxial Ge on Si via High-Bandgap III-V Buffers for Low-Power Electronic Applications

Peter D. Nguyen

ABSTRACT

Over the past four decades, aggressive scaling of silicon (Si) based complementary metal-oxide-semiconductor (CMOS) transistors has resulted in an exponential increase in device density, and thus an exponential increase in computing power. Increasing transistor density also results in increasing total power consumption and thus, necessitates supply voltage scaling in order to maintain low-power device operation. However, with increased supply voltage scaling, transistor drive current is significantly degraded due to the low carrier mobility of Si. To overcome the key challenges of device and voltage scaling required for low-power electronic operation without the degradation of transistor drive current requires the adoption of narrow bandgap channel materials with superior transport properties. However, the use of such materials as bulk substrates remains cost-prohibitive. Thus, another key challenge lies in the heterogeneous integration of high-mobility channel materials on affordable, established Si platform. Germanium (Ge) is an attractive candidate for next-generation low-power devices owing to its high electron and high hole mobility. Recently, AlAs/GaAs epilayers were demonstrated as a potential buffer platform for next-generation Ge-based electronics integrated on Si substrate. This research systematically investigates the structural characteristics of the Ge epitaxial layer heterogeneously integrated on Si using a composite III-V AlAs/GaAs buffer and the electrical characteristics of MOS capacitors (MOS-C's) fabricated on the aforementioned stack. Further passivation techniques and interface engineering is then pursued on MOS-C's fabricated from (100) and (110) crystallographically oriented epitaxial Ge integrated on AlAs/GaAs material stacks, balancing out effective oxide thickness (*EOT*) and reduction of oxide and interfacial traps in order to ensure a pristine interfacial quality for high-performance electronic applications. Further, work function tuning is demonstrated for the first time on the different crystallographically oriented epitaxial Ge integrated on AlAs/GaAs material stacks using two different gate metals, demonstrating the tunability of threshold voltage, V_{TH}, required for transistor applications. The research demonstrates the feasibility of future high-mobility channel material integration on Si via large bandgap buffer architectures for high-speed, low-power, high-performance CMOS logic applications.

Heteroepitaxial Ge on Si via High-Bandgap III-V Buffers for Low-Power Electronic Applications

Peter D. Nguyen

GENERAL AUDIENCE ABSTRACT

Over the past four decades, the shrinking of silicon (Si) based complementary metal-oxide-semiconductor (CMOS) transistors has allowed for a boost in electronic device performance, but also aggravates power consumption. To mitigate these power concerns, the industry is looking to switch over from silicon-based transistors to those that are alternative material based. A possible alternative material solution is Germanium (Ge), which has inherent material characteristics that are attractive for low-power, high-speed electronic applications. However, the use of Ge as a bulk material for transistor fabrication is cost-prohibitive. Thus, to address cost concerns, the integration of Ge on to Si via a buffer layer "bridge" is necessary. Recently, AlAs/GaAs buffer layers were demonstrated as a potential buffer solution for next-generation Ge-based electronic devices integrated on to Si. This research systematically investigates the material characteristics of Ge layers integrated on Si using the AlAs/GaAs buffer layer approach, as well as MOS capacitors (MOS-C's) fabricated on the aforementioned material stack. Furthermore, interface engineering is pursued on different crystallographically oriented Ge integrated on Si via the AlAs/GaAs buffer approach in order to refine the MOS capacitor devices in order to demonstrate the feasibility of the use of such material stacks for state-of-the-art FinFETs (Fin Field-Effect Transistors). Finally, the tunability of a key device characteristic known as threshold voltage, V_{TH}, is demonstrated via the use of two different gate metals for the Ge-based MOS capacitor devices integrated on Si using AlAs/GaAs buffer layers. Thus, this research demonstrates the feasibility of the use of Ge integrated on Si via AlAs/GaAs buffer layers for high-speed, low-power electronic devices.

Acknowledgements

First, I would like to thank my advisor Professor Mantu Hudait for his continued support, guidance, encouragement, and patience throughout the course of my research. I have continually been inspired by his positive attitude towards learning ever since I was an undergraduate student in his class. He has continued to be a mentor and a source of inspiration as I continued my post-graduate research studies under him. The skills and knowledge I have obtained under his supervision will serve me well throughout my life.

Second, I would like to thank Professors Marius Orlowski and Jean Heremans for serving as my M.S. Thesis committee members. Additionally, I am indebted to the manager of the MICRON cleanroom Don Leber. Without his patience, advice, and expertise, much of the experimental research work would not have been possible.

Third, I would like to express my sincere gratitude to all of my Advanced Devices & Sustainable Energy Lab (ADSEL) team members, namely Yan Zhu, Nikhil Jain, Michael Clavel, Patrick Goley, Jheng-Sin Liu, Aheli Ghosh, and Sarat Saluru, all of whom have become some of my closest friends. I'd also like to extend my thanks to all of the friends, professors, teachers, and colleagues who have all continually encouraged, guided, and supported me through all these years. Special thanks goes to my powerlifting friends who became like a second family outside of my graduate studies.

Finally, I am most grateful for the unconditional love and support from my parents Tho and Nathalie Nguyen. They taught me the importance of hard work and perseverance and laid the foundation to help me grow into the person I am now.

Table of Contents

List of Figures

List of Tables

Chapter 1 – Introduction

1.1   Silicon-based MOSFET Scaling History and Limitations

Over the past five decades, silicon (Si) metal-oxide-semiconductor field-effect transistors (MOSFETs) have dominated the semiconductor industry ever since its first experimental demonstration in 1959 [1]. Following the demonstration of planar processing in 1960 [2] and then its implementation to create the first commercial integrated circuits (ICs) based on Si MOSFETs in 1964 [3], Moore's Law [4] was conceptualized in 1965 to capitalize on the miniaturization of Si-based MOSFETs. Since then, transistor feature sizes have been drastically scaled down, resulting in an exponential increase in device density, and thus computing power. However, conventional scaling becomes increasingly difficult mainly due to increased power consumption in addition to fundamental limitations. Fig. 1.1 [5] shows a data conglomeration of subthreshold (red) and active (blue) power densities from commercial Si MOSFETs. The trend demonstrates that subthreshold power density increases much quicker than active power density with decreasing gate length. Additionally, at the cross-over point (black circle) below 20-nm gate length, subthreshold power density is suggested to surpass active power density. Thus, it is also equally as important to scale supply voltage (V_{dd}) in cadence with gate length scaling. The relationship between V_{dd} and total power dissipation is given by [6]:

$$P = n(CV_{dd}{}^2 f + V_{dd}I_{off})$$
(1.1)

where P is total power consumption, n is the number of transistors, C is the load capacitance, V_{dd} is supply voltage, f is the operation frequency, and I_{off} is the off-state current. It is easy to see that the reduction of V_{dd} impacts total power consumption greatly.
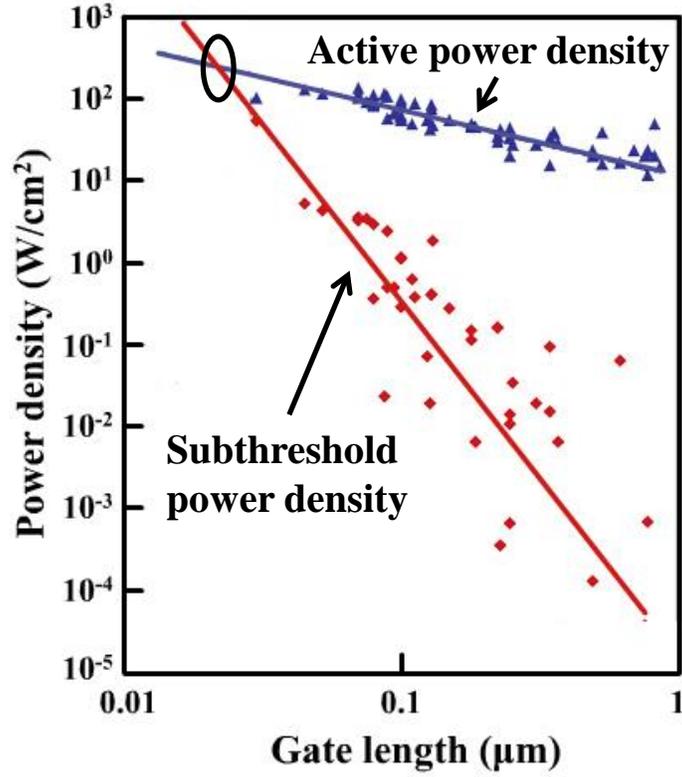
Figure 1.1 Active (blue) and subthreshold (red) power densities from commercial Si MOSFETs with gate lengths between 0.01 µm to 1 µm. The cross-over point (black circle) at sub 20-nm gate length indicates where subthreshold power density surpasses active power density [5].

However, V_{dd} scaling also greatly reduces the transistor on-state current, and thus, directly impacts device performance. The dependence between V_{dd} and I_{on} can be expressed by [6]:

$$I_{on} = \frac{W}{2L}\mu C_{ox}(V_{dd} - V_{th})^2 \tag{1.2}$$

where I_{on} is the on current, W is the transistor width, L is the transistor length, μ is carrier mobility, C_{ox} is the oxide capacitance, V_{dd} is the supply voltage, and V_{th} is the threshold

voltage. With the reduction of V_{dd} and subsequently V_{th}, other parameters must be increased in order to improve I_{on}.

Beyond the sub-100 nm node, conventional scaling was no longer possible due to the impact of V_{dd} on device performance and thus, novel architectural changes were introduced to the MOSFET as a way to continue Moore's Law [7]. At the 90 nm node, carrier mobility, μ, was increased through the implementation of uniaxial strain on the Si channel using Silicon-Germanium (SiGe) source/drains and silicon nitride (Si_3N_4) etch stops, thereby enhancing PMOS and NMOS transistor drive current, respectively [8]. At the 45 nm node, a metal gate and a scalable high-κ gate dielectric was used to further enhance transistor drive current by increasing oxide capacitance, C_{ox}, while reducing the off-state leakage current, an issue that had been plaguing the industry with silicon dioxide dielectric scaling [9]. At the 22 nm node, a novel FinFET design was implemented, yielding more gate control and thus better device performance at lower gate voltages [10]. Currently in production are 14 nm transistors, which implement a taller fin design to aid with increased drive current [11]. Fig. 1.2 shows the evolution of the Si-based MOSFET as the industry scaled from the 130 nm node down to the 22 nm node. Beginning with future smaller nodes, fundamental constraints will restrict the performance of Si-based MOSFETs. As a result, novel materials are being considered to continue Moore's Law. According to the 2014 ITRS, III-V, Germanium (Ge), carbon nanotube, and 2D materials are examples of such channel materials under consideration for future nodes [12].
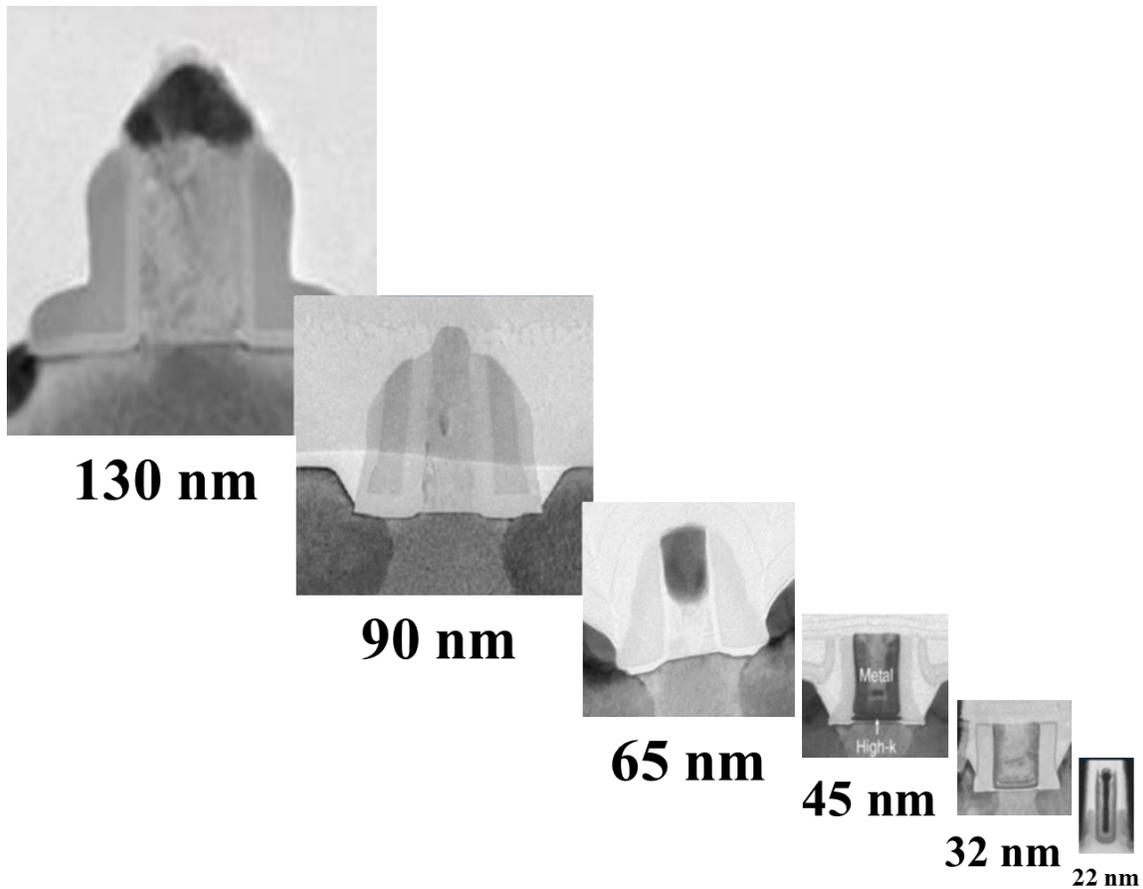
Figure 1.2 Fundamental changes in the Si-based MOSFET as starting at 130 nm node down to the 22 nm node [13].

1.2 Germanium as an Alternative Channel Material

High-mobility materials are under consideration for smaller future nodes. From eq. 1.2, because carrier mobility μ and I_{on} are directly proportional, replacing Si-based channels with high-mobility materials will have a large positive impact on I_{on}, and thus, overall device performance. The bulk mobilities and bandgaps of several semiconductors under consideration are compared in Table 1.1. For high-performance complementary MOS (CMOS) logic, it seems logical to have III-V materials and Ge serve as the NMOS and PMOS channel material, respectively. However, due to the challenges of integrating two

4

different types of materials in large 12" wafer form, this idea does not seem economically feasible for large-scale production. From a single-material standpoint, Ge is an excellent candidate for CMOS logic due to its 2.8× and 4.2× increase in electron and hole mobility, respectively, as compared to those of Si. Furthermore, its low bandgap of 0.66 eV makes Ge a suitable candidate for low-power devices operating at low supply voltage (≤ 0.5 V) while still mitigating leakage from band-to-band tunneling and thermionic emission that would otherwise plague very low-bandgap materials [14]. However, two major challenges must be dealt with before Ge can be considered a viable replacement as a future channel material: the heterogeneous integration of Ge on Si substrate and the development of a suitable gate stack for Ge.

1.2.1 Heterogeneous Integration of Ge on Si Substrate Using Buffers

To make Ge a feasible alternative channel material, it should be compatible with the modern manufacturing process. The most economically feasible solution is to heterogeneously integrate Ge on to Si substrate rather than developing processes using bulk large-diameter Ge wafers. The most difficult materials growth issue is that Ge's lattice constant is about 4.2% larger than that of Si. As a result of the lattice mismatch, direct Ge growth on to Si substrate results in the relaxation of the Ge film via the formation of misfit dislocations. These dislocations will deleteriously impact device performance by reducing both carrier mobility and carrier lifetime while increasing junction leakage [14]. In addition to the lattice mismatch problem, there is also a thermal mismatch issue between Ge and Si due to the large difference in the thermal expansion coefficients of Ge (6.1×10^{-6}/K) and Si (3×10^{-6}/K) [16]. The thermal mismatch will introduce a large density of thermal

Table 1.1 Carrier mobility and bandgap of Si, Ge, and common III-V materials [15].

Parameters	InSb	InAs	Ge	GaSb	In$_{0.53}$Ga$_{0.47}$As	Si	InP	GaAs
Electron mobility (cm^{2}Vs)	7.7×10^4	4×10^4	3900	3000	1.2×10^4	1400	5400	8500
Hole mobility (cm^{2}Vs)	850	500	1900	1000	300	450	200	400
Bandgap (eV)	0.17	0.354	0.661	0.726	0.74	1.12	1.344	1.424

mismatch induced defects which will also negatively impact overall device performance. The design of a proper buffer architecture to bridge the epitaxial Ge layer with the Si substrate is therefore imperative for heterogeneous integration of device-quality Ge on to Si substrate. A proper buffer utilizes wide bandgap materials and is low in both threading dislocation density (TDD) as well as root-mean-square (RMS) roughness [17]. Large bandgap materials provide for both device isolation and the suppression of parallel conduction whereas low TDD and low RMS allow for a smooth template from which to grow defect-minimal, device-quality epitaxial Ge. This introduction will focus on some common buffer architectures that have been pursued in order to heterogeneously integrate Ge on to Si substrate including the Si$_x$Ge$_{1-x}$ buffer and the germanium on insulator (GeOI) approach, as well as their shortcomings.

1.2.1.1 The Si$_x$Ge$_{1-x}$ Buffer

The Si$_x$Ge$_{1-x}$ buffer approach seeks to introduce the 4.2% lattice mismatch between Ge and Si gradually rather than abruptly with the hope of producing high-quality Ge films on

the top of the buffer. Mismatch strain is alleviated by the strain-relieving glide of threading dislocations, allowing for the threading dislocations present in the initial low-Ge composition layers to suppress the nucleation of additional dislocations in the higher-Ge composition layers. Currie *et al.* [18] were able to demonstrate epitaxial Ge films with TDD as low as 2.1×10^6 cm^{-2}, which is on the order of bulk Ge substrates. Their graded Si$_x$Ge$_{1-x}$ buffer approach introduced a chemical-mechanical-planarization (CMP) step at 50% Ge grading, which helped prevent dislocation pile-up due to a blocking of the threading dislocation glide, a complication observed in thicker grades to high Ge concentration. However, the Si$_x$Ge$_{1-x}$ buffer was 12 μm in thickness, which can add to overall process cost; has thermal mismatch induced cracks due to thicker buffer layer; and has RMS roughness of 24.2 nm, which is too rough for device definition. Since then, the Si$_x$Ge$_{1-x}$ buffer approach has been refined to yield improvements in buffer thickness, RMS roughness, and TDD. Loh *et al.* [19] demonstrated a Ge epilayer with TDD of 5×10^5 cm^{-2} and RMS roughness of 1.4 nm using only a 160 nm thick buffer layer. The buffer consists of a 30 nm Si$_{0.8}$Ge$_{0.2}$ buffer, followed by a 30 nm Ge seed layer grown at low-temperature, and then capped with a 100 nm Ge epitaxial layer grown at high-temperature. However, despite the recent advances in developing thinner and more optimal Si$_x$Ge$_{1-x}$ buffers, the Si$_x$Ge$_{1-x}$ buffer approach fails in that it does not provide suppression of parallel conduction nor effective device isolation. It has been demonstrated that high-bandgap buffer materials like InAlAs [20, 21] are necessary to eliminate parallel conduction to the active channel and also to electrically isolate multiple devices. The Si$_x$Ge$_{1-x}$ buffer's bandgap, which is between that of Ge and of Si (0.66 eV to 1.12 eV), is low enough to both 1) provide

additional higher conductivity paths not ideal for suppression of parallel conduction and 2) not ensure ideal electrical isolation of devices.

1.2.1.2   Germanium on Insulator (GeOI)

Through the incorporation of an oxide buffer, which is high bandgap in nature, the Germanium-on-insulator (GeOI) approach offers a way of heterogeneously integrating Ge onto Si substrate with the benefits of both suppression of parallel conduction as well as effective device isolation. There are various ways to form GeOI wafers, the first of which will be discussed is the Ge condensation method, proposed in 2000 by Tezuka *et al.* [22]. In this approach, a Si_xGe_{1-x} buffer of low Ge concentration is grown on a silicon-on-insulator (SOI) wafer. Through various cycles of dry, high-temperature oxidation, Ge atoms become trapped between an oxidizing interface and the SiO_2 underneath. Inert gas annealing is then introduced to homogenize the Ge layer. Thus, the final structure is a Ge-rich Si_xGe_{1-x} buffer between two layers of SiO_2. The top SiO_2 layer is then chemically etched away, yielding the Ge-rich Si_xGe_{1-x} buffer, where the Ge content can be as high as 100%. Hutin *et al.* [23] recently demonstrated pMOSFETs fabricated on GeOI substrates via the Ge condensation method, where, following an Ar anneal, they were able to achieve RMS roughness of 0.193 nm for their $Si_{0.05}Ge_{0.95}$ layer. The left image in Fig. 1.3 shows a cross-sectional transmission electron microscopy (TEM) image of their 30 nm gate-length pMOSFETs fabricated on GeOI substrates and the right image is an enlarged view to show their device gate stack. The Ge layer in both images is 25 nm thick. Their MOSFETs demonstrated an I_{on}/I_{off} ratio of ~ 5×10^5 and I_{off} < 1 nA/μm, showing that devices on GeOI substrates formed via the Ge condensation method exhibit excellent control of parallel

8

conduction as well as excellent device isolation. TDD is not reported, but can be assumed to be on the order of bulk substrates due to the high I_{on}/I_{off} ratio.
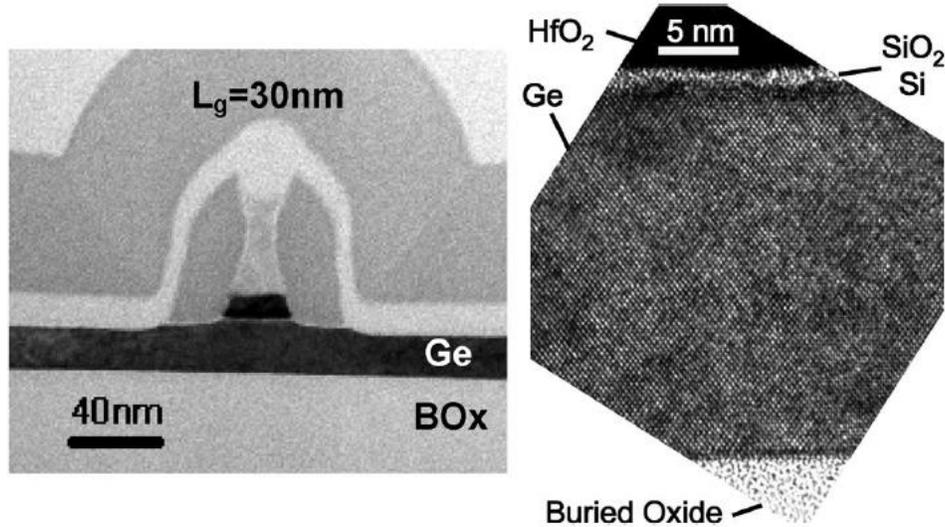
Figure 1.3 On the left is a cross-sectional transmission electron microscopy (TEM) image of a GeOI pMOSFET with gate length of 30 nm. On the right is an enlarged view to show the device gate stack. The Ge thickness is 25 nm [23]. © 2010 IEEE

Another popular method to form GeOI wafer is through wafer bonding. For on-insulator type wafers, the SOI wafer is typically Si with a layer of SiO_2. Ge wafers are then bonded onto the SOI wafer, thereby forming a GeOI wafer. Letertre *et al.* [24] demonstrated thin-film GeOI structures using Smart Cut™ technology. Plasma-enhanced chemical vapor deposition (PECVD) tetraethyl orthosilicate (TEOS) is first deposited on the Ge donor wafer. The structure is exposed to H^+ ions and then hydrophilically bonded to Si substrate at room-temperature. A layer of Ge is then detached from the donor wafer and polishing is done on the GeOI substrate to obtain a smooth Ge surface. They were able to achieve RMS roughness of less than 2Å and estimated TDD below 10^6 cm^{-2}. From a

material characteristic standpoint, the GeOI approach via bonding offers excellent Ge surfaces for device definition.

Despite the fact that the GeOI approach helps achieve high-quality Ge surfaces and has benefits that the Si_xGe_{1-x} buffer does not provide, namely effective parallel conduction as well as device isolation, the major shortcoming of the GeOI approach lies in the thermal mismatch between the Ge layer and the oxide layer. This is due to the fact that CMOS applications see usage in microprocessor logic, where high thermal budget are typically induced as a result of regular computing usage. Ge bonded to an SiO_2/Si wafer faces thermal issues due to the thermal coefficient of SiO_2 is 5.6×10^{-7}/K [25] whereas that of Ge is 6.1×10^{-6}/K [16]. As a result of the large mismatch, it is possible for thermal defects to propagate through repeated exposure to high thermal budgets. Such defects deleteriously affect device performance.

1.3   Gate Stack Engineering

The discovery of thermally-grown silicon dioxide (SiO_2) as a passivating mechanism for Si in 1959 [26] laid the foundation for scaling of Si-based MOSFETs. Dielectric scaling became an issue around the 90-nm node when it was discovered that SiO_2 thicknesses less than 2 nm would substantially introduce a large amount of gate leakage current due to direct tunneling of electrons, thereby significantly increasing circuit power dissipation [27]. To solve the dielectric scaling issue, the industry had to replace silicon dioxide. Gate capacitance, which plays a huge role in overall MOSFET performance, is explained by [27]:

$$C = \frac{\kappa \varepsilon_0 A}{t_{ox}} \tag{1.3}$$

where C is gate capacitance, κ is relative permittivity, ε_0 is permittivity of free space, A is area, and t_{ox} is oxide thickness. The tunneling issue can therefore be solved by replacing SiO$_2$ with high-κ dielectrics, which have intrinsically higher dielectric permittivity κ. Thus, it is possible to use thicker high-κ dielectrics in MOSFET gate stacks, which helps mitigate direct tunneling, while maintaining the same gate capacitance. Since the adoption of high-κ dielectrics, a convenient metric to define electrical thickness of high-κ films in terms of its equivalent SiO$_2$ thickness is the equivalent oxide thickness (*EOT*), defined by [27]:

$$EOT = \frac{3.9}{\kappa} t_{Hi-\kappa} \qquad (1.4)$$

where 3.9 is the dielectric constant of SiO$_2$, κ is the dielectric constant of the high-κ dielectric, and $t_{Hi-\kappa}$ is the thickness of the high-κ dielectric.

Since state-of-the-art Si-based MOSFETs use high-κ dielectrics, it will be especially important for Ge-based MOSFETs to also adopt high-κ dielectrics in order to be considered as a worthy replacement for Si-based MOSFETs. While high-κ dielectrics have successfully been demonstrated on Si-based MOSFETs and can now be found in consumer products [9], finding a gate-dielectric interface for Ge that is similar in quality to that for Si remains a challenge. The gate-dielectric interface is doubly important for device performance due to the fact that carriers induced in the gate are then induced within Angstroms of the semiconductor-oxide interface [27]. Therefore, an ideal interface must be atomically smooth and absent of interface defects. The mitigation of defects is especially important since defects can deleteriously affect device performance in the following four ways: 1) trapped charges in defects cause a shift in the threshold voltage of the transistor, which thereby affects the ability of the device to turn on; 2) trapped charge are transient and leads to unpredictability in the operating characteristics of devices; 3) trapped charges

can scatter carriers in the channel, thereby reducing carrier mobility, which in turn degrades device performance; and 4) defects can cause electrical failure and oxide breakdown [27]. Unfortunately, high-κ dielectrics are intrinsically high in defect concentration due to their high coordination number [27]. In other words, its bonds cannot relax and rebond broken bonds at defect sites, unlike dielectrics with low coordination number such as SiO_2 [27].

Beyond a defect-minimal oxide/Ge interface, a suitable high-κ dielectric will also require high conduction and valance band offsets (CBO and VBO, respectively). Sufficient CBO and VBO of at least 1 eV helps mitigate leakage current via suppression of Schottky emission of carriers into the oxide bands [27]. Fig. 1.4 shows calculated Ge CBO and VBO of several standard dielectrics.

Another requirement for high-κ dielectrics is a need for good thermodynamic and kinetic stability. Essentially, the high-κ dielectric must not react with Ge when the two are in intimate contact, even at high thermal budgets that are commonplace during processing. Although Ge requires much lower processing temperatures as compared to Si (400 °C [28, 29] vs. 1000 °C [27]), it's been demonstrated that HfO_2 and ZrO_2, two popular high-κ dielectrics, can recrystallize when exposed to typical dopant activation temperatures in Ge (400-500 °C) and thereby promote indiffusion of Ge into the dielectric, resulting in undesirable electronic states as well as the formation of undesirable interfacial layers (IL) [29].
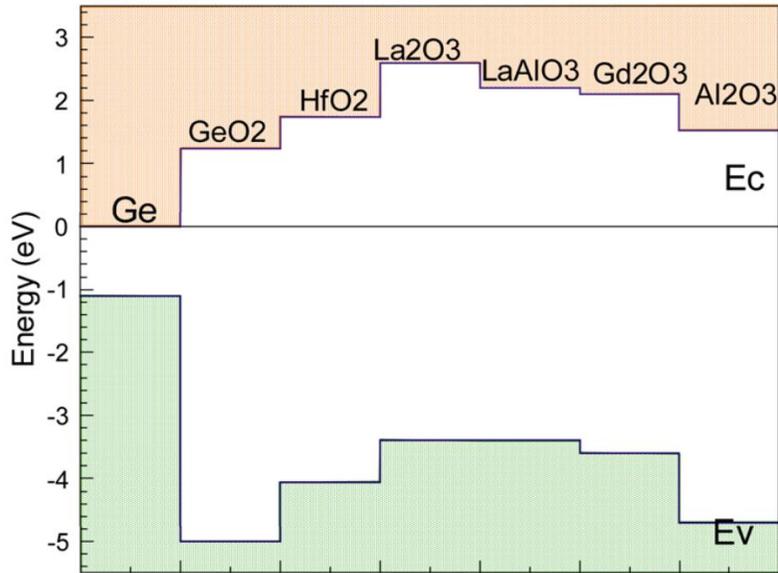
Figure 1.4 Calculated conduction and valance band offsets of several oxides on Ge [30]. Reproduced from "Band offsets, Schottky barrier heights, and their effects on electronic devices," with the permission of the American Vacuum Society.

A list of requirements when considering a suitable high-κ dielectric are listed below:

1. Low EOT (i.e., higher κ-value)

2. Low leakage current density ($< 1.5 \times 10^{-2}$ A/cm^2) [27]

3. CBO/VBO of at least 1 eV

4. Thermodynamic and kinetic stability

5. Good electrical interface with Ge

Table 1.2 is a list of some candidate dielectrics for Ge, their dielectric constant κ, and their bandgaps. HfO$_2$, ZrO$_2$, and Al$_2$O$_3$ are common dielectrics that have been implemented onto Ge and are chosen because they have decent CBO/VBO's, good thermodynamic/kinetic stability, and reasonable κ values. Al$_2$O$_3$ is on the lower end of what is acceptable for κ values (25-30 is preferable [27]), but to aid with EOT scaling, it

has also successfully been implemented in bilayer stacks with higher-κ dielectrics such as HfO$_2$ for MOSFETs [31].

1.3.1 Ge Surface Passivation via Interfacial Layers

An interfacial layer (IL), which is purposefully inserted between the high-κ dielectric and the Ge channel, can be a solution to help improve the oxide/Ge interface. The IL helps passivates the surface by reducing defect concentration, thereby forming a higher quality interface with Ge [17, 27]. Two well-researched IL methods that have been implemented for Ge devices include the use of Si IL [32-35] and GeO$_2$ IL [36-39], the latter of which will be discussed in more detail later. However, the IL negatively affects scaling due to its lower dielectric constant κ, which in turn adds to the EOT, as shown below [17]:

$$EOT_{total} = \frac{3.9}{\kappa_{Hi-\kappa}} t_{Hi-\kappa} + \frac{3.9}{\kappa_{IL}} t_{IL} \tag{1.5}$$

Table 1.2 Dielectric constant κ and bandgap for common gate dielectrics [27].

	κ	Gap (eV)
SiO_2	3.9	9
Si_3N_4	7	5.3
Al_2O_3	9	8.8
Ta_2O_5	22	4.4
TiO_2	80	3.5
$SrTiO_3$	2000	3.2
ZrO_2	25	5.8
HfO_2	25	5.8
$HfSiO_4$	11	6.5
La_2O_3	30	6
Y_2O_3	15	6
α-$LaAlO_3$	30	5.6

where $\kappa_{Hi\text{-}\kappa}$ and κ_{IL} are the dielectric constants of the high-κ and IL layers, respectively, and t_{IL} is the thickness of the interfacial layer. Therefore, there must be a fine balance between using thick enough IL layers to passivate the surface and not sacrificing EOT.

The use of GeO_2 as a passivating IL is possible but care must be taken. This is because Ge's native oxide (commonly written as GeO_x due to the fact that GeO_2 is not dominant oxide) is both thermodynamically unstable [40] and also soluble in water, thereby placing

a bare and exposed GeO$_x$/Ge surface very susceptible to the environment. As a passivating IL, it is common to have the high-κ act as a capping layer for the GeO$_x$/Ge surface, but proper selection of the high-κ is important, as it has been demonstrated that Ge can updiffuse into certain high-κ material, thereby forming interfacial defects [29]. Fortunately, certain high-κ dielectric such as Al$_2$O$_3$ can act as a barrier for Ge updiffusion [31]. The formation of the GeO$_x$ IL has been well-studied and various methods exist, including thermal oxidation [35-38, 40-43], high-pressure oxidation [39], ozone oxidation [44], radical oxidation [45], and plasma oxidation [31, 46]. A wide variety of methods exist due to the ongoing research efforts in developing a thin GeO$_x$ layer for minimum EOT while maintaining an excellent, defect-minimal oxide/Ge interface.

Thermal oxidation has been a relatively popular method for the formation of high-quality GeO$_x$/Ge interfaces. By careful selection and control of the oxidation temperature and oxidation time, Nakakita *et al.* [42] were able to demonstrate MOS capacitors (MOS-C) and MOSFET gate stacks consisting of Al$_2$O$_3$/GeO$_x$/Ge with low interface trap densities (D_{it}) of approximately $2 \times 10^{11} - 4 \times 10^{11}$ eV^{-1} cm^{-2}, where the GeO$_x$ was thermally formed at 550 °C. Typically, D_{it} values lesser than 10^{11} cm^{-2}eV^{-1} are considered good [17], so for these devices, these D_{it} values are within the range of what is acceptable. They were able to correlate their good oxide/Ge interfaces with a high peak hole mobility of 575 cm^2 V^{-1} s^{-1}, as well as a high I_{on}/I_{off} ratio of $\sim 10^5$, as measured from their MOSFETs implementing the same gate stack, demonstrating that thermally-oxidized GeO$_x$ ILs are promising for the development of high-κ gate stacks for future Ge p-MOSFET implementation.

GeO$_x$ ILs formed by plasma postoxidation is a novel but very promising path for the formation of highly-scaled Ge gate stacks. Zhang *et al.* [31] demonstrated MOSFETs

16

implementing gate stacks with EOT of 0.76 nm, where 0.2 nm Al_2O_3 followed by 2.2 nm HfO_2 are first deposited on to Ge via ALD. Afterwards, a 0.35 nm GeO_x layer is formed under the Al_2O_3 via post plasma oxidation, where the Al_2O_3 serves as a suppression barrier between HfO_2 and GeO_x. Comparable D_{it} values were measured as with Nakakita *et al.*'s work [42], but with extremely scaled oxide thicknesses. I_{on}/I_{off} ratios of ~10^4 and ~10^3 were reached for PMOS and NMOS devices, respectively, using the highly-scaled gate stack. Furthermore, peak hole and electron mobilities of 546 cm^2 V^{-1} s^{-1} and 689 cm^2 V^{-1} s^{-1}, respectively, were obtained.

1.4   Thesis Objective and Organization

The objective of this research is to systematically and comprehensively investigate the material and electrical properties of epitaxial Ge films heterogeneously integrated onto large-bandgap III-V buffers, including two material stacks: (i) Ge heterogeneously integrated onto Si substrate via large-bandgap AlAs/GaAs buffers and (ii) different-oriented Ge films heterogeneously integrated onto GaAs substrate via large-bandgap AlAs buffers. The thesis will detail the structural design and characterization in addition to the electrical properties of the first material stack to understand its impacts on low-power electronic applications. Building upon this, the thesis will then comprehensively detail the interface engineering and passivation of the second material stack via the use of various annealing schemes. Finally, the thesis will detail work metal engineering and its extent on flat-band voltage shift on the second material stack.

This thesis is organized in to seven chapters. ***Chapter 2*** presents the experimental techniques related to material characterization used throughout this research, including the

basic principles of molecular beam epitaxy (MBE) growth and various material characterization methods.

Chapter 3 presents the experimental techniques related to electrical characterization of MOS capacitors used throughout this research and includes a brief synopsis on the operating principles of MOS capacitors.

Chapter 4 presents a comprehensive investigation of the structural and electrical properties of Ge heterogeneously integrated on to Si substrate via high-bandgap AlAs/GaAs buffers. The epitaxial Ge heterointerface of these structures was first examined using various material characterization methods, including x-ray rocking curves, secondary ion mass spectrometry, and energy-dispersive spectroscopy to determine the interface quality, coherence, uniformity, and abruptness of the epitaxial Ge material layer. MOS capacitors are then fabricated from these material stacks, where transmission electron microscopy is performed to study the structural quality of the oxide/Ge heterointerface. Electrical characterization is then performed to quantitatively analyze the interface quality of the MOS capacitors fabricated from the aforementioned material stack. Based upon these results, the suitability of Ge heterogeneously integrated onto Si substrate via large-bandgap AlAs/GaAs buffers for low-power electronic applications is discussed.

Chapter 5 builds upon the work presented in *Chapter 4,* and presents the detailed analysis of the passivation methods using various annealing schemes for (100) and (110) epitaxial Ge heterogeneously integrated on GaAs substrates via the large-bandgap AlAs buffer. X-ray photoelectron spectroscopy is first used to analyze the composition of GeO_x thermally grown on the aforementioned material stacks. MOS capacitors are then fabricated from the aforementioned material stacks and then electrically characterized

following various annealing schemes in order to quantitatively analyze the interface quality and provide insight into the optimal annealing scheme in order to demonstrate suitability for the use of (100) and (110) epitaxial Ge with a large-bandgap AlAs buffer for novel low-power electronic applications.

Chapter 6 describes work metal function engineering for the material stack discussed in *Chapter 5* to demonstrate flat-band voltage tunability. MOS capacitors fabricated from the material stack described in *Chapter 5* using two different gate metals are electrical analyzed to demonstrate the extent of flat-band voltage shift. *Chapter 6* provides the last critical MOS engineering step required in order to demonstrate the feasibility of epitaxial Ge heterogeneously integrated on large-bandgap AlAs buffers for low-power electronic applications.

Lastly, *Chapter 7* summarizes the conclusions of this work and presents the prospects for future avenues of research and investigation based upon this thesis' current findings.

References

[1]   D. Kahng and M. M. Atalla, "Silicon-silicon dioxide field induced surface devices," in *IRE-AIEE Solid-state Device Res. Conf.*, (Carnegie Inst. of Technol., Pittsburgh, PA), 1960.

[2]   J. A. Hoerni, "Planar silicon diodes and transistors," *IRE Trans. Electron Devices*, vol. ED-8, p. 178, Mar. 1961.

[3]   R. H. Norman and H. E. Stephenson, Shift Register Employing Insulated Gate Field Effect Transistors, U.S. Patent 3 454 785, Jul. 8, 1969.

[4]   G. E. Moore, "Cramming more components onto integrated circuits," *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, Jan. 1998.

[5]   E. J. Nowak, "Maintaining the benefits of CMOS scaling when scaling bogs down," in *IBM J. Res. Develop.*, vol. 46, Mar./May 2000, pp. 169–180.

[6]   C. Hu, *Modern Semiconductor Devices for Integrated Circuits*, 1st Ed., Prentice Hall: New Jersey (2010).

[7]   G. Moore. "No exponential is forever: but "forever" can be delayed! [semiconductor industry]," in *Solid-State Circuits Conference*, 2003.

[8]   S. Thompson *et al.*, "A 90-nm logic technology featuring 50-nm strained silicon channel transistors, 7 layers of Cu interconnects, low k ILD, and 1 μm^2 SRAM cell," in *IEDM Tech. Dig.*, 2002, pp. 61–64.

[9]   K. Mistry *et al.*, "A 45 nm logic technology with high-k+ metal gate transistors, strained silicon, 9 cu interconnect layers, 193 nm dry patterning, and 100% Pb-free packaging," in *IEDM Tech. Dig.*, 2007, pp. 247–250.

[10]   C. Auth *et al.*, "A 22 nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," in *Proc. VLSI Symp. Technol.*, Jun. 2012, pp. 131–132.

[11]   S. Natarajan *et al.*, "A 14 nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm^2 SRAM cell size," in *IEDM Tech. Dig.*, Dec. 2014, pp. 3.7.1–3.7.3.

[12]   ITRS 2014 edition [Online]. Available: www.itrs2.net.

[13]   K. Mistry. (2014). *Transistor Scaling: The Age of Innovation*. [Online]. Available: http://www.nanohub.org/resources/20935/download/2014.04.08-Mistry-NEEDS.pdf

[14] R. Pillarisetty, "Academic and industry research progress in germanium nanodevices," *Nature*, vol. 479, no. 7373, pp. 324–328, Nov. 2011.

[15] *Semiconductors on NSM* [Online]. Available: www.ioffe.ru/SVA/NSM/Semicond

[16] *Coefficients of Linear Thermal Expansion* [Online]. Available: http://www.engineeringtoolbox.com/linear-expansion-coefficients-d_95.html.

[17] P. S. Goley and M. K. Hudait, "Germanium based field-effect transistors: Challenges and opportunities," *Materials*, vol. 7, no. 3, pp. 2301–2339, Mar. 2014.

[18] M. T. Currie *et al.*, "Controlling threading dislocation densities in Ge on Si using graded SiGe layers and chemical-mechanical polishing," *Appl. Phys. Lett.*, vol. 72, pp. 1718–1720, 1998.

[19] T. H. Loh *et al.*, "Ultrathin low temperature SiGe buffer for the growth of high quality Ge epilayer on Si(100) by ultrahigh vacuum chemical vapor deposition," *Appl. Phys. Lett.*, vol. 90, no. 9, p. 092 108, Feb. 2007.

[20] M. K. Hudait, "Heterogeneously integrated III-V on silicon for future nanoelectronics," *ECS Trans.*, vol. 45, no. 3, pp. 581–594, May 2012.

[21] M. K. Hudait *et al.*, "Heterogeneous integration of enhancement mode In0.7Ga0.3As quantum well transistor on silicon substrate using thin ($\leq$ 2 μm) composite buffer architecture for high-speed and low-voltage (0.5 V) logic applications," in *IEDM Tech. Dig.*, 2007, pp. 625–628.

[22] T. Tezuka *et al.*, "A novel fabrication technique of ultra-thin and relaxed SiGe buffer layers with high Ge fraction for sub-100 nm strained silicon-on-insulator MOSFET's," *Jpn. J. Appl. Phys.*, vol. 40, p. 2866, 2001.

[23]   L. Hutin *et al.*, "GeOI pMOSFETs scaled down to 30-nm gate length with record off-state current," *IEEE Electron Device Lett.*, vol. 31, no. 3, pp. 234–236, Mar. 2010.

[24]   F. Letertre *et al.*, "Germanium-On-Insulator (GeOI) structure realized by the Smart Cut™ technology," in *Mat. Res. Soc. Symp. Proc.*, vol. 809, pp. B4.4.1–B4.4.6, 2004.

[25]   *2.1   The   Material   Silicon   Dioxide*   [Online].   Available: http://www.iue.tuwien.ac.at/phd/hollauer/node11.html.

[26]   D. Kahng and M. M. Atalla, "Silicon-silicon dioxide field induced surface devices," in *Proc. IRE-AIEE Solid-State Device Res. Conf.*, Pittsburgh, PA, 1960.

[27]   J. Robertson, "High dielectric constant oxides," *Eur. Phys. J. Appl. Phys.*, vol. 28, no. 3, pp. 265–291, Dec. 2004.

[28]   C. Chui, S. Ramanathan, B. Triplett, P. McIntyre, and K. Saraswat, "A sub-400 C germanium MOSFET technology with high-κ dielectric and metal gate," in *IEDM Tech. Dig.*, 2002, pp. 437–439.

[29]   Y. Kamata, "High-k/Ge MOSFETs for future nanoelectronics," *Mater. Today*, vol. 11, no. 1/2, pp. 30–38, Jan./Feb. 2008.

[30]   J. Robertson, "Band offsets, Schottky barrier heights, and their effects on electronic devices," *J. Vacuum Sci. Technol. A*, vol. 31, no. 5, pp. 1-19, Aug. 2013.

[31]   R. Zhang *et al.*, "High-mobility Ge p- and n-MOSFET with 0.7-nm EOT using $HfO_2/Al_2O_3/GeOx$ /Ge gate stacks fabricated by plasma postoxidation," *IEEE Trans. Electron Devices*, vol. 60, no. 3, pp. 927–933, Mar. 2013.

[32]   N. Wu *et al.*, "Alternative surface passivation on Germanium for metal-oxide-semiconductor applications with high-κ gate dielectric," *Appl. Phys. Lett.*, vol. 85, pp. 4127–4129, 2004.

[33]   W. P. Bai, N. Lu, and D.-L. Kwong, "Si interlayer passivation on Germanium MOS capacitors with high-κ dielectric and metal gate," *IEEE Electron Device Lett.*, vol. 26, no. 6, pp. 378–380, Jun. 2005.

[34]   B. De Jaeger *et al.*, "Optimisation of a thin epitaxial Si layer as Ge passivation layer to demonstrate deep sub-micron n- and p-FETs on Ge-on-insulator substrates," in *Proc. 14th Conf. Insul. Films Semicond.*, 2005, pp. 26–29.

[35]   F. E. Leys *et al.*, "Thin epitaxial Si films as a passivation method for Ge (100): Influence of deposition temperature on Ge surface segregation and the high-k/Ge interface quality," *Mater. Sci. Semicond. Process.*, vol. 9, no. 4/5, pp. 679–684, Aug.–Oct. 2006.

[36]   A. Delabie *et al.*, "Effective electrical passivation of Ge(100) for high-K gate dielectric layers using germanium oxide," *App. Phys. Lett.*, vol. 91, no. 8, p. 082 904, Aug. 2007.

[37]   H. Matsubara, T. Sasada, M. Takenaka, and S. Takagi, "Evidence of low interface trap density in GeO2/Ge metal–oxide–semiconductor structures fabricated by thermal oxidation," *Appl. Phys. Lett.*, vol. 93, no. 3, p. 032104, Jul. 2008.

[38]   F. Bellenger *et al.*, "Passivation of Ge (100)/GeO2/high-k gate stacks using thermal oxide treatments," *J. Electrochem. Soc.*, vol. 155, no. 2, pp. G33–G38, Feb. 2008.

[39]   A. Toriumi *et al.*, "Opportunities and challenges for Ge CMOS—Control of interfacing field on Ge is a key," *Microelectron. Eng.*, vol. 86, no. 7–9, pp. 1571–1576, Jul.–Sep. 2009.

[40]   K. Prabhakaran *et al.*, "Distinctly different thermal decomposition pathways of ultrathin oxide layer on Ge and Si surfaces," *Appl. Phys. Lett.*, vol. 76, no. 16, pp. 2244–2247, Apr. 2000.

[41]   T. Sasada *et al.*, "Surface orientation dependence of interface properties of GeO2/Ge metaloxidesemiconductor structures fabricated by thermal oxidation," *J. Appl. Phys.*, vol. 106, no. 7, pp. 073716-1–073716-7, Oct. 2009.

[42]   Y. Nakakita *et al.*, "Interface-controlled self-align source/drain Ge p-channel metal–oxide– semiconductor field-effect transistors fabricated using thermally oxidized GeO2 interfacial layers," *Jap. J. Appl. Phys.*, vol. 50, no. 1, pp. 010109- 1–010109-7, Jan. 2011.

[43]   P.D. Nguyen *et al.*, "Heteroepitaxial Ge MOS Devices on Si Using Composite AlAs/GaAs Buffer," *IEEE J. Electron Devices Soc.*, vol. 3, no. 4, pp. 341-348, July 2015.

[44]   D. Kuzum *et al.*, "Ge-interface engineering with ozone oxidation for low interface-state density," *IEEE Electron Device Lett.*, vol. 29, no. 4, pp. 328–330, Apr. 2008.

[45]   M. Kobayashi *et al.*, "Radical oxidation of germanium for interface gate dielectric GeO2 formation in metal–insulator– semiconductor gate stack," *J. Appl. Phys.*, vol. 106, no. 10, pp. 104 117- 1–104 117-7, Nov. 2009.

[46]   R. Zhang *et al.*, "Al2O3/GeOx/Ge gate stacks with low interface trap density fabricated by electron cyclotron resonance plasma postoxidation," *Appl. Phys. Lett.*, vol. 98, no. 11, pp. 112 902-1–112 902-3, Mar. 2011.

Chapter 2 – Experimental Methodology:

Materials Synthesis and Materials Characterization

2.1   Molecular Beam Epitaxy

Solid-source molecular beam epitaxy (MBE) is a physical-vapor deposition technique used in heteroepitaxial research and commercial production that allows for the growth of thin epitaxial films as well as complex heterostructures with precise control and excellent uniformity [1, 2]. Molecular beams consisting of neutral atoms or molecules are impinged on to a heated substrate where epitaxial growth occurs [1, 2]. These molecular beams are formed via thermal evaporation of highly pure, elemental liquid or crystalline sources [1, 2]. A key characteristic of the MBE growth technique is the utilization of an ultra-high vacuum (UHV) environment (where residual-gas pressure is below 10^{-9} torr), which allows for maximal flux of material impinging on to the heated substrate and a consistent growth environment [1, 2]. Furthermore, the UHV environment also allows for the growth of high-purity, defect-minimal epitaxial films due to the increase in free-mean path of atomic or molecular species, thereby minimizing the collisions between the particles and background vapor [1]. Fig. 2.1 depicts what a typical MBE growth chamber may look like [1].

Several measures are taken to maintain the UHV environment required for growth of high-purity, defect-minimal epitaxial films. Firstly, UHV is achieved by a combination of oil-free roughing, turbo, ion, and cryogenic pumps [1]. Prior to growth, the MBE chamber is baked at 200 °C for at least two weeks after having exposed the system to atmosphere.
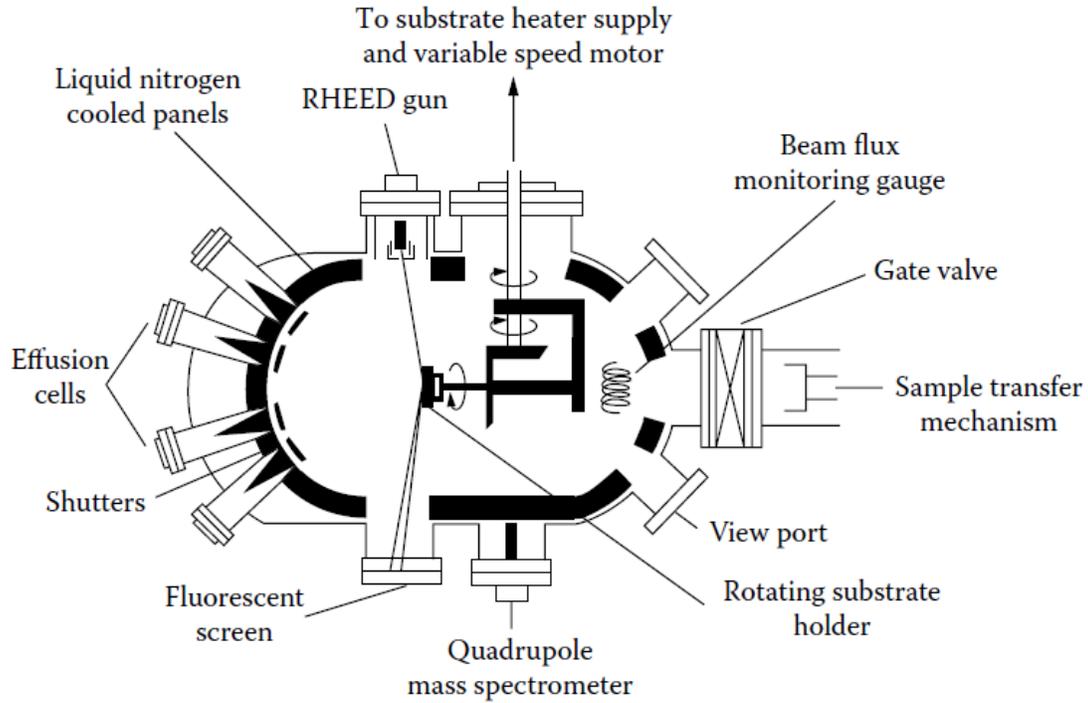
Figure 2.1 Schematic representation of a typical MBE growth chamber [1].

During this baking period, each effusion or cracking cell is also baked to degas the growth materials and minimize their contamination. To avoid contamination during growth, MBE components are typically machined from highly stable, non-reactive, low-vapor pressure materials such as tantalum or molybdenum. The growth chamber walls are additionally cooled to cryogenic temperature using a liquid N_2 shroud, which helps reduce the walls' evaporation, thereby minimizing flux generation of atomic or molecular species coating the walls [1]. To ensure a pristine growth surface prior to epitaxial growth, growth substrates undergo several degassing steps before being loaded into the growth chamber: (i) A three-hour long bake at 180 °C is done in a loading chamber to remove moisture from the substrate surface, substrate holder, and intro-chamber; (ii) a two-hour long bake with temperatures varying from 300-400 °C is done in a buffer chamber to degas higher-

temperature surface contamination; and (iii), a high temperature oxide desorption is done in the growth chamber to remove native oxide.

Following the degassing of the substrate, thin-film epitaxy growth is done by heating of the effusion or cracking cells, thereby generating an elemental flux. Growth thicknesses are controlled using individual motorized shutters, which controls the exposure of the elemental flux on to the substrate. Using an ion gauge, beam-equivalent pressure (BEP) for each source is measured to calibrate the flux levels during growth [2]. Growth uniformity is obtained by rotation of the substrate holder [2] at a speed of six to eight revolutions per minute (RPM).

In-situ surface monitoring is performed using reflection high-energy electron diffraction (RHEED) to monitor oxide-desorption and epitaxial growth. Electron beams nearly parallel to the growth surface are reflected on a phosphorous-coated screen, forming a surface reconstruction in reciprocal space known as a RHEED pattern [1]. Streaky RHEED patterns are obtained when ideal 2D, smooth, layer-by-layer growth is obtained (known as Frank-van der Merwe growth) [1]. In the case of non-smooth, island-like growth (known as Volmer-Weber growth), spotty RHEED patterns can be observed. In addition to surface monitoring, RHEED can also be used to precisely determine growth rate [1]. This is because RHEED intensity oscillation frequency is proportional to the epitaxial monolayer thickness [1]. With this relationship, a linear relation between the growth rate and BEP can be generated.

2.2   X-ray Diffraction: Thin Film Strain and Relaxation Analysis

High-resolution x-ray diffraction (HR-XRD) is an important tool in the structural characterization and the extraction of strain relaxation properties in heteroepitaxial layers. A typical HR-XRD set up is shown in Fig. 2.2 [1]. The x-ray tube (usually made of copper) serves as the x-ray source and produces a broad-spectrum, divergent x-ray beam [1]. The beam is restricted in angular divergence and to a specific wavelength (a process known as conditioning) by four diffracting surfaces arranged in a pattern called Bartels monochromator. The conditioned beam is then diffracted by the specimen crystal and detected via a scintillation detector. In a rocking curve measurement, the sample is rotated about the ω-axis, resulting in a spectrum of diffracted intensities which are measured as a function of ω, the specimen angle. The spectrum (known as the diffraction profile or rocking curve) consists of intensity peaks that vary in position, intensity, and width. This rocking curve can then be used to analyze the strain relaxation properties and structural quality of the specimen.
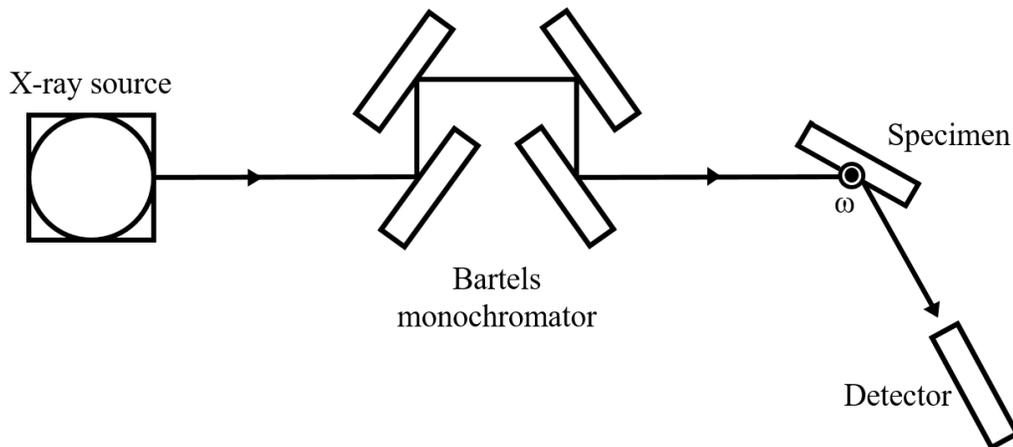
Figure 2.2 Typical HR-XRD setup, including X-ray source, Bartels monochromator, multi-axis specimen stage, and scintillation detector.

XRD analysis is built on the principle of real and reciprocal geometric equivalences for angular diffraction. Real space geometric equations are expressed using Bragg equations, whereas reciprocal space geometric equations are expressed using Laue equations [1]. Fig. 2.3 is a depiction of the Bragg condition for diffraction, where d is the lattice spacing. In the special condition where incident and reflection angles are equivalent to specular reflection θ, the path difference Δ between beams a and b then is equal to *2dsinθ* [1]. The constructive interference condition is $\Delta = n\lambda$, where *n* is an integer and λ is the x-ray wavelength [1]. Thus, the Bragg diffraction condition is given by [1]:

$$2dsin\theta_B = n\lambda \tag{2.1}$$

where *n* is the order of reflection, θ_B is the Bragg angle, *d* is the inter-atomic plane spacing, and λ is the x-ray wavelength. The equivalent Laue equation for reciprocal space is given by [1]:

$$d(hkl) = \frac{a}{\sqrt{h^2+k^2+l^2}} \tag{2.2}$$

where *a* is the crystal's lattice spacing and *h*, *k*, and *l* are Miller indices of the Bragg plane. Thus, the *hkl* Bragg angle, θ_B (*hkl*) is simply [1]:

$$\theta_B(hkl) = sin^{-1}\left(\frac{\sqrt{h^2+k^2+l^2}}{2a}\right) \tag{2.3}$$
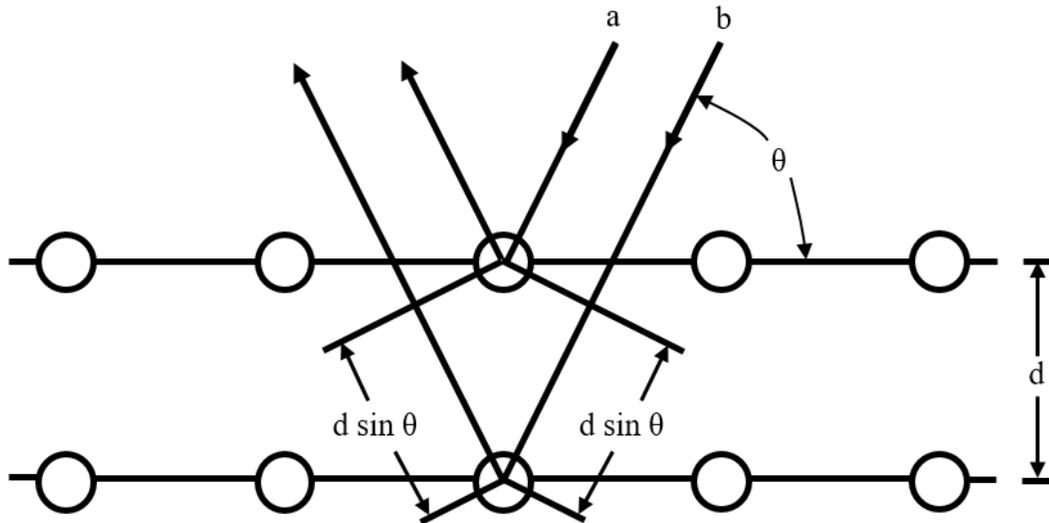
Figure 2.3 Bragg diffraction condition.

2.3 Transmission Electron Microscopy: Defect and Structural Analysis

Cross-sectional high-resolution transmission electron microscopy (HR-TEM) is a useful technique to characterize short- and long-range structural quality of samples. More specifically, for short-range analysis, HR-TEM is valuable to study the abruptness of heterointerfaces. For semiconductor applications, HR-TEM requires thinned cross-sectional samples known as foils. These foils must be thinned to 100 nm or less in order to become electron transparent, or in other words, able to transmit electrons [1]. Furthermore, the foils must also be able to survive exposure to high-energy electron beams (which are typically anywhere from ~100 keV to ~1 MeV) in ultra-high vacuum [1].

A schematic diagram for a conventional TEM instrument is shown in Fig. 2.4 [1]. Using magnetic condenser lenses, collimated high-energy electrons impinge and transmit through the TEM specimen [1]. The electrons are scattered into particular directions, and especially

for crystalline samples, the scattering is ruled by Bragg's law for diffraction [1]. The diffracted beams are then brought into focus at the focal plane for the objective lens [1].

The TEM operates in two different modes: diffraction and imaging mode. In diffraction mode, the first intermediate lens is focused on the back focal plane of the objective lens, thereby creating a diffraction pattern, which is then magnified and projected through a combination of intermediate and projection lenses [1]. The pattern displayed on the phosphorous screen is made up of an array of spots which each correspond to a particular diffraction vector $\mathbf{g}$ [1]. The spots in the diffraction pattern can be used as an index for the diffraction beams and can also assist in the production of the final TEM micrograph [1].

In imaging mode, the intermediate lens is now focused on a magnified (up to 10^6) inverted image of the sample formed by the objective lens and the image is projected onto the screen [1]. The back focal plane aperture of the objective lens is used to select only one of the diffracted beams to form the image [1]. A bright-field image is only displayed if $\mathbf{g}$ = [000] is chosen; otherwise, if any other diffracted beam is chosen, a dark-field image is displayed instead [1].

Information can also be collected from inelastically-scattered electrons. Energy from the incident electron beam is transferred to the host atoms and when these host atoms radiatively relax, they emit x-rays. These x-rays correspond to separation between orbitals, which is different for every atom. Thus, the emitted x-rays can be used to determine the elemental composition of the sample. This technique of analysis is called energy-dispersive spectroscopy (EDS).
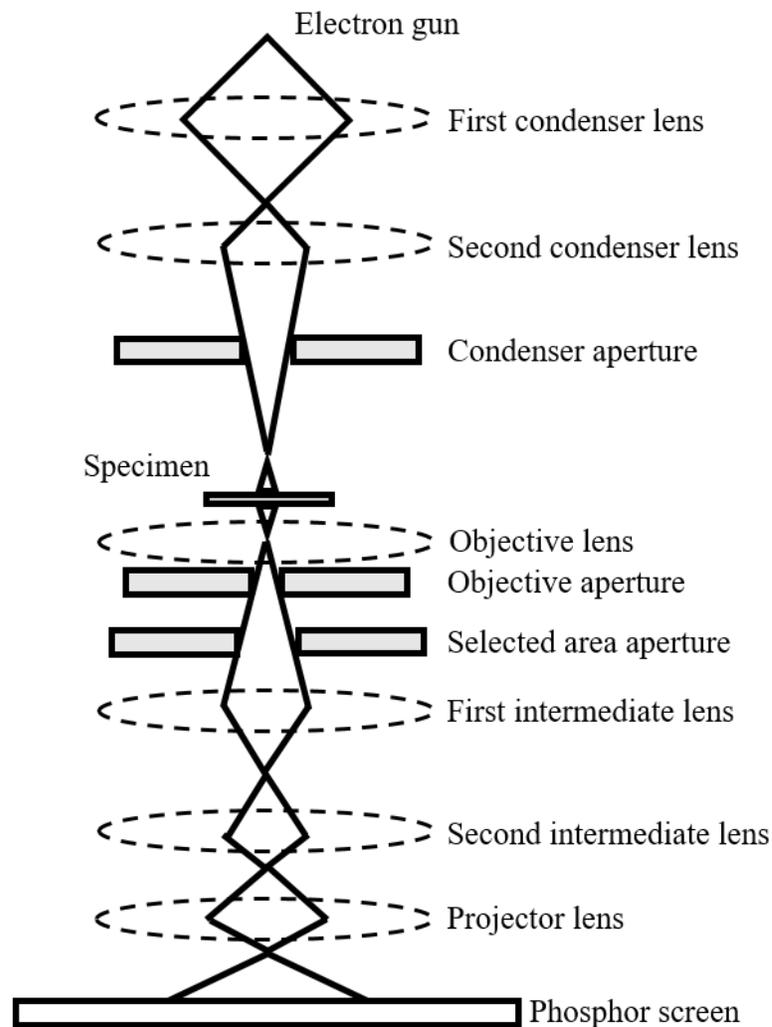
Figure 2.4 Schematic diagram of a conventional transmission electron microscope.

2.4 X-ray Photoelectron Spectroscopy: Oxide Chemical Structure Analysis

The elemental nature of samples (and more specifically in this thesis, the oxide chemical structure) can be analyzed using x-ray photoelectron spectroscopy (XPS). Fig. 2.5a depicts a conventional XPS set-up, which consists of an electron gun, an aluminum (Al) anode, a crystal x-ray monochromator, an electron lens, a hemispherical energy analyzer, and a photoelectron detector [3]. The electron gun generates Al-Kα x-rays from

the Al anode target, which are then monochromated and focused by the crystal on to the sample surface [3]. The x-rays transfer enough energy to eject photoelectrons from the core level of the atoms in the sample, as depicted in Fig. 2.5b [3]. The photoelectrons emitted and passed through the electron lens then enter the hemispherical energy analyzer, where they are then filtered by their kinetic energies [3]. The photoelectron's kinetic energy, KE, can be calculated by $KE = h\upsilon - BE$ where $h\upsilon$ is the X-ray energy and BE is the binding energy of a given atomic orbital [3]. Because each atom has unique energy separations between their electron orbitals (thereby unique BE), the measured photoelectron's KE can be used to determine the elemental composition of a sample [3]. However, XPS is a surface-sensitive characterization technique, in that the escape depth of photoelectrons is only about 5 nm [3]. Photoelectrons generated by the x-ray beam beyond 5 nm lose a too much energy to escape the sample surface due to in-elastic collision [3]. These photoelectrons thus appear in the XPS spectrum in the background at lower KE's, whereas the photoelectrons that are able to escape the sample surface appear as high-intensity peaks.
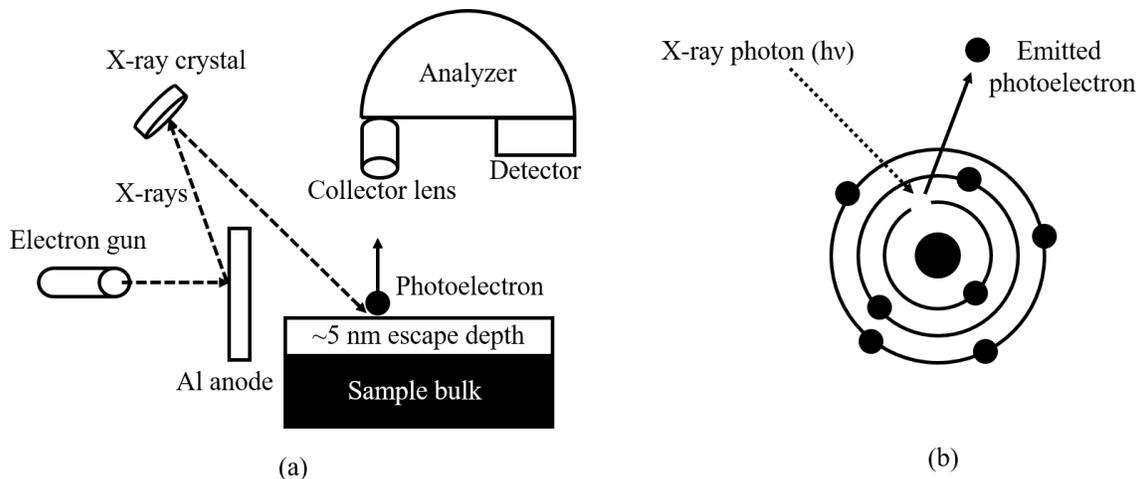
Figure 2.5 (a) Conventional x-ray photoelectron spectroscopy set-up. (b) Schematic diagram showing emission of photoelectron.

2.5   Secondary Ion Mass Spectrometry: Interface and Doping Profile Abruptness

Dynamic secondary ion mass spectrometry (SIMS) provides insight into doping profile abruptness and heterointerface sharpness. In SIMS, a 10 – 40 kEV high energy beam of primary particles (i.e. electrons, ions, neutrals, or photons) bombards a surface [3]. By way of a nearly elastic scattering process, the primary particle's kinetic energy is then transferred into the secondary particles (i.e. electrons, neutral species, atoms, molecules, atomic and cluster ions), which are then emitted from the sample surface [3]. Out of all the emitted species, which are mostly neutral, only the secondary ions are detected and subsequently separated by mass/charge ratio by a mass spectrometer [3]. By continuous sputtering of the sample surface, an elemental profile can be generated. However, the accuracy of SIMS is limited by a phenomenon known as the *matrix effect*, which is due to the participation of the matrix when ionization occurs at or close to the emission of particles from the surface [3]. Fortunately, the use of SIMS standards helps alleviate analysis complications caused by the *matrix effect*.

References

[1]   J. E. Ayers, *Heteroepitaxy of Semiconductors: Theory, Growth, and Characterization*, 1st Ed., CRC Press: Boca Rolton, Florida (2007).

[2]   U. W. Pohl, *Epitaxy of Semiconductors: Introduction to Physical Principles*, 1st Ed., Springer: Berlin, Germany (2013).

[3]   J. C. Vickerman and I. S. Gilmore, *Surface Analysis – The Principal Techniques*, 2nd Ed., John Wiley & Sons: Chichester, United Kingdom (2009).

Chapter 3 –Experimental Methodology 2:

Interface Characterization Techniques

3.1 Ideal Metal-Oxide-Semiconductor Capacitors

The study of semiconductor surfaces is crucial because most practical problems that deal with the reliability and stability of all semiconductor devices ultimately comes down to the semiconductor's surface conditions, especially in regards to the oxide/semiconductor heterointerface. By conducting capacitance-voltage (C-V) and conductance-voltage (G-V) measurements on metal-oxide-semiconductor capacitors (MOS-C), standard metrics including flatband voltage, doping level, fixed charge, effective oxide thickness, work function, and interface trap density can be extracted. The interface trap density (D_{it}) is one of the most important metrics to quantitatively measure the quality of the oxide/semiconductor heterointerface. A typical MOS-C is shown below in Fig. 3.1, where V_G is the gate bias and t_{OX} is the oxide thickness.
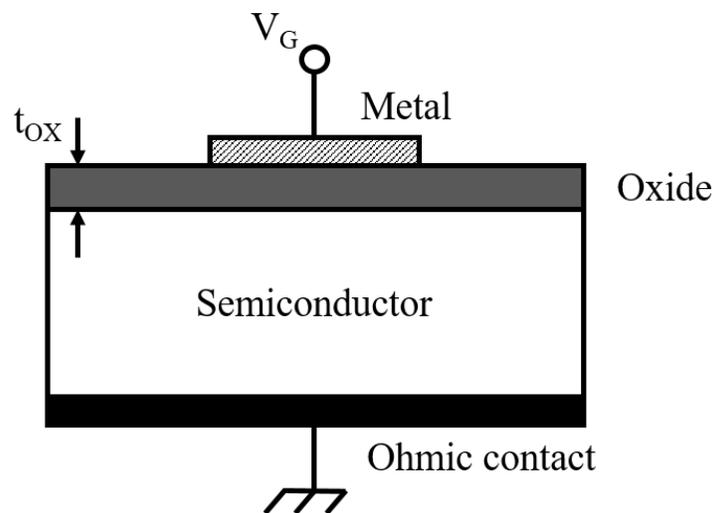
Figure 3.1 Schematic of the most simplistic, conventional MOS-C.

3.1.1 Ideal Energy-Band Diagrams and Modes of Operation

Understanding ideal MOS-C theory is important to exploring semiconductor surface physics and also critical before going into a full discussion of metric extraction. An ideal MOS-C is defined as such: (1) There are no interface traps nor trapped oxide charges and therefore, the only charges that exist are those that are intrinsic to the semiconductor and those that exist in the metal, but are of opposite polarity to the semiconductor's charges; (2) the oxide is a perfect insulator with infinite resistivity; (3) the difference between the metal's work function ϕ_M and the semiconductor's work function ϕ_S is 0, or in other words, $\phi_{MS} = 0$ [1]. The energy-band diagram of ideal MOS-C's at equilibrium (i.e. $V_G = 0$) is shown below for n-type and p-type semiconductors in Fig. 3.2a and Fig. 3.2b, respectively.
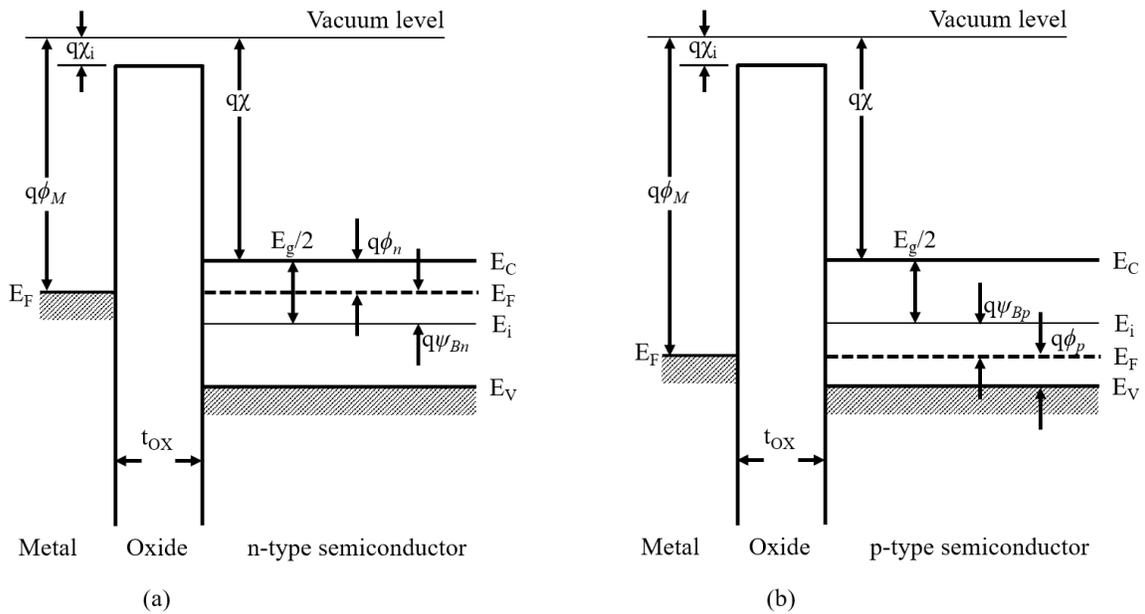
Figure 3.2 Energy-band diagram of ideal MOS-C at equilibrium ($V_G = 0$) for (a) n-type semiconductor and for (b) p-type semiconductor.

The condition for equilibrium is demonstrated by the following equations [1]:

$$\phi_{MS} = \phi_M - \left(\chi + \frac{Eg}{2q} - \psi_{Bn}\right) = \phi_M - (\chi + \phi_n) = 0 \qquad \text{for n-type} \qquad (3.1a)$$

$$\phi_{MS} = \phi_M - \left(\chi + \frac{Eg}{2q} + \psi_{Bp}\right) = \phi_M - \left(\chi + \frac{Eg}{q} - \phi_p\right) = 0 \qquad \text{for p-type} \qquad (3.1b)$$

where χ and χ_i are electron affinities for the semiconductor and the insulator respectively and ϕ_n and ψ_{Bn} are the Fermi potentials with respect to the band edges and midgap for n-type semiconductors and ϕ_p and ψ_{Bp} are the Fermi potentials with respect to the band edges and midgap for p-type semiconductors. The energy levels E_C, E_i, E_F, and E_V denote the conduction band energy level, the intrinsic energy level, the Fermi energy level, and the valance band energy level, respectively. In an ideal MOS-C, the condition of equilibrium where no gate voltage is applied is also called flatband because all the bands in the energy-band diagram are flat.

Three modes exist for the ideal MOS-C when it is biased with either positive or negative voltage: accumulation, depletion, and inversion. The energy-band diagrams for n-type MOS-Cs for the accumulation, depletion, and inversion modes are shown below in Fig. 3.3a, 3.3b, and 3.3c, respectively. The mode's corresponding schematics are also shown below in Fig. 3.3d, 3.3e, and 3.3f, respectively. For n-type MOS-C's under the accumulation mode, a positive bias ($V_G > 0$) is applied to the gate. E_C bends downward toward E_F, as shown in Fig. 3.3a. This downward bending results in an accumulation of majority carriers (for n-type, majority carriers are electrons) at the semiconductor surface, as clearly reflected in Fig. 3.3d. When a negative bias is applied ($V_G < 0$) to the gate, the bands begin to bend upward and the majority carriers are depleted, as shown in Fig. 3.3b. As seen in Fig. 3.3e, a depletion layer forms (the lighter grey layer underneath the oxide layer denoted with W), the semiconductor surface begins to look more positive as ionized

37

donor ions start to appear. The MOS-C has now entered the depletion mode. When larger

negative voltage is applied, the bands bend even more upward and E_i now crosses E_F, as

shown in Fig 3.3c. Minority carriers (holes in this case) now exceed majority carriers,

resulting in an inverted surface and thus, the MOS-C is now in the inversion mode. In

inversion mode, as shown in Fig. 3.3f, the depletion layer from Fig. 3.3e grows larger as

more ionized donor ions and holes (white circles) begin to appear, resulting in a highly

positive semiconductor surface.

The same modes are observed for p-type semiconductors as well, but the polarity is

switched. For p-type MOS-C's under the accumulation mode, a negative bias ($V_G < 0$) is

applied to the gate. E_V bends upward toward E_F, as shown in Fig. 3.4a, resulting in an
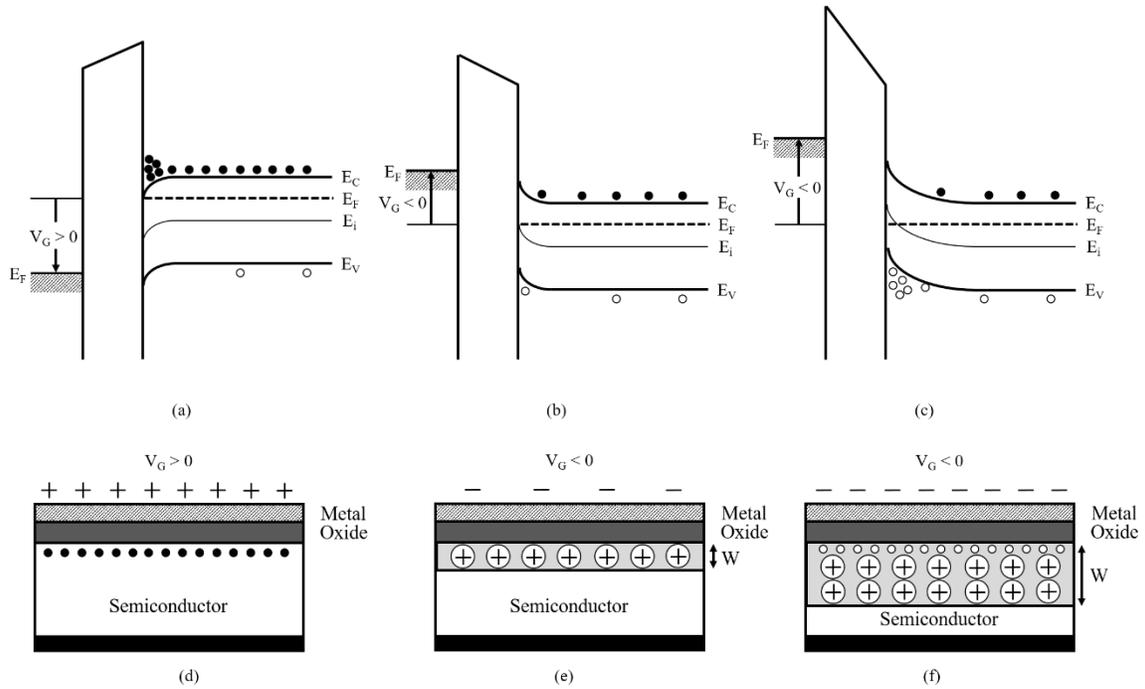
Figure 3.3 Energy-band diagrams of ideal n-type MOS-C for the modes of (a)

accumulation, (b) depletion, and (c) inversion. Below are the schematics for n-type MOS-

C for the modes of (d) accumulation, (e) depletion, and (f) inversion.

accumulation of holes at the semiconductor surface, which are the majority carriers in p-type MOS-C's, as shown in Fig. 3.4d. When a positive bias is applied ($V_G > 0$) to the gate, the bands begin to bend downward, resulting in a depletion of the holes, as shown in Fig. 3.4b. As with the n-type MOS-C case in depletion mode, a depletion layer also forms and the semiconductor surface begins to look more negative as ionized acceptor ions start to appear. With more positive bias applied, the bands downward further until E_i crosses E_F, as shown in Fig 3.4c. The MOS-C now enters inversion mode, where the electrons, which are the minority carriers in this case, now exceed holes, which are the majority carriers, resulting in an inverted surface. Similarly to the n-type MOS-C, the depletion layer grows larger as more ionized acceptor ions and electrons (black circles) begin to appear, resulting in a highly negative semiconductor surface, as shown in Fig. 3.4f.
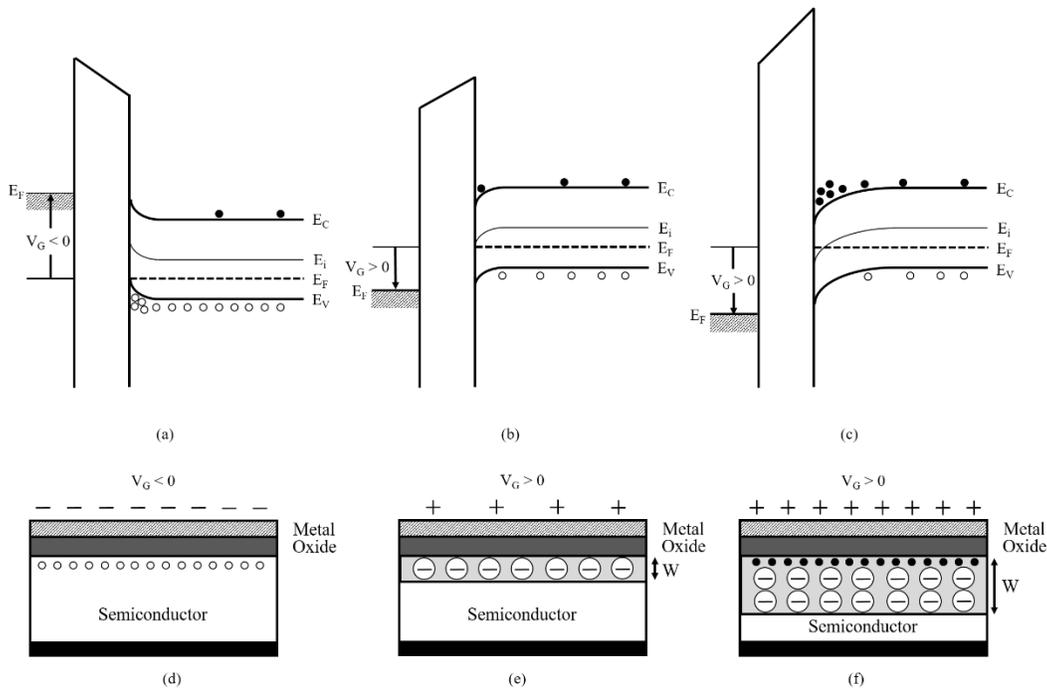
Figure 3.4 Energy-band diagrams of ideal p-type MOS-C for the modes of (a) accumulation, (b) depletion, and (c) inversion. Below are the schematics for p-type MOS-C for the modes of (d) accumulation, (e) depletion, and (f) inversion.

3.1.2 Ideal Capacitance-Voltage Curves

Capacitance-voltage (C-V) curves help visualize the effect of direct-current (DC) voltage sweep with a supersetted alternating-current (AC) voltage of various frequencies on the capacitance of the semiconductor surface. In an ideal p-type MOS-C, as mentioned in the previous section, negative voltage bias results in an accumulation of holes at the semiconductor surface. This large accumulation of holes results in a very large capacitance in series with the oxide capacitance, C_{ox}, which results in a net capacitance of simply C_{ox}, as shown in the C-V curve of an ideal p-type MOS-C in Fig. 3.5. The capacitance remains independent of further applied gate voltage until the gate bias on the MOS-C approaches $V_G = 0$, which in an ideal MOS-C is assumed to the flatband condition, where the surface

potential, $\psi_s = 0$, as pointed out in Fig. 3.5. At flatband condition, there is a distinct capacitance called the flatband capacitance, C_{FB}, which is defined below [1]:

$$C_{FB} = \frac{\varepsilon_s \varepsilon_{ox}}{\varepsilon_s t_{ox} + \varepsilon_{ox}\sqrt{kT\varepsilon_s/N_A q^2}}$$ (3.2)

where ε_{ox} and ε_s are the permittivities of the oxide and the semiconductor, respectively, k is the Boltzmann constant, T is the temperature, N_A is the acceptor concentration, q is the electron charge, and t_{ox} is the oxide thickness. From flatband to the condition where $\psi_s = \psi_{Bp}$, which was defined previously to be the Fermi potential with respect to the midgap, is the mode of depletion. During this mode, the depletion capacitance C_D comes into play and is defined below [1]:

$$C_D = \frac{\varepsilon_s}{W}$$ (3.3)

where W is defined to be the depletion width previously seen in Fig. 3.3 and 3.4. W is defined below [1]:

$$W = \sqrt{\frac{\varepsilon_s^2}{C_{ox}^2} + \frac{2\epsilon_s V}{qN_D}} - \frac{\varepsilon_s}{C_{ox}}$$ (3.4)

where N_D is the donor concentration and V is the applied voltage. As equation 3.4 points out, W is directly proportional to V, meaning that with further applied gate voltage, the depletion width grows. Since C_D is inversely proportional to W, C_D therefore shrinks in response to further applied gate voltage. Within the mode of depletion, where C_{ox} is in series with C_D, the diminishing C_D begins to take more of an effect on the overall series capacitance, which is why there is a drop in overall capacitance. Beyond the point where $\psi_s = \psi_{Bp}$, the MOS-C enters the state of weak inversion and the MOS-C begins to display two different phenomenon, which is affected by the frequency of the supersetted AC voltage. ψ_{Bp} is defined below [1]:

$$\psi_{Bp} = \frac{kT}{q}\ln(\frac{N_A}{n_i}) \tag{3.5}$$

The first phenomenon discussed will deal with the MOS-C behavior when an AC voltage of low frequency is applied. With low frequencies, the recombination-generation rates of minority carriers are able to keep up with the small-signal variation, which helps with the charge exchange in the inversion layer and therefore, the appearance of minority carriers is possible at the semiconductor surface [1]. From weak inversion to strong inversion (defined as $\psi_s = 2\psi_{Bp}$) the semiconductor surface begins to see more accumulation of minority carriers. In addition, at $\psi_s = 2\psi_{Bp}$, $V_G = V_{TH}$, which is defined as the threshold voltage. V_{TH} is defined below [1]:

$$V_{TH} = \frac{\sqrt{2\varepsilon_s q N_A(2\psi_{Bp})}}{C_{ox}} + 2\psi_{Bp} \tag{3.6}$$

From weak inversion to strong inversion, the situation then becomes akin to the situation observed in accumulation mode, where with a larger increase in the capacitance associated with the carriers (minority carriers in this situation), C_{ox} starts to take more of an effect in the overall series capacitance. This is clearly seen in Fig. 3.5, where after the initial dip in overall capacitance, the capacitance (marked in black for the case of low frequency), begins to increase as more minority carriers begins to populate at the semiconductor surface.

The second phenomenon discussed deals with the MOS-C behavior when an AC voltage of high frequency is applied. In this situation, the recombination-generation rates of the minority carriers are not able to keep up with the small-signal variation, which therefore results in no accumulation of minority carriers at the semiconductor surface. Instead, the depletion capacitance begins to continually shrink with increasing gate voltage as the depletion width continually grows. Eventually, there is a saturation point beyond strong inversion where the depletion width reaches its max (W_{max}) and the capacitance

reaches its minimum at C_{min}, as shown by the red line in Fig. 3.5. W_{max} is defined below [1]:

$$W_{max} \approx \sqrt{\frac{4\varepsilon_s kT ln(\frac{N_A}{n_i})}{q^2 N_A}}$$

(3.7)

and C_{min} is defined below [1]:

$$C_{min} = \frac{\varepsilon_i \varepsilon_s}{\varepsilon_s t_{ox} + \varepsilon_i W_{max}}$$

(3.8)

It should be worth mentioning that all the equations derived in this section are only relevant for ideal p-type MOS-C's. However, simply replacing N_D for N_A and vice versa will yield the parameters of interest for n-type MOS-C's.
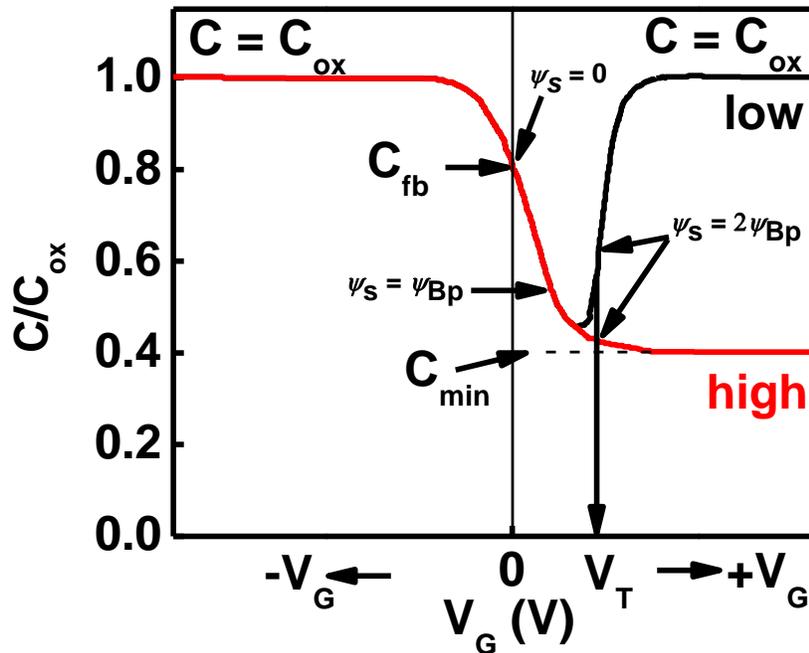
Figure 3.5 Capacitance-voltage curves of ideal p-type MOS-C where the black line resembles MOS-C behavior with low AC voltage frequency and the red line resembles the MOS-C behavior with high AC voltage frequency.

43

3.2 Non-idealities and Their Effect on Capacitance-Voltage Behavior

The previous section dealt entirely with ideal MOS-C's. However, non-idealities that are introduced by the growth of an oxide layer on a terminated single-crystal semiconductor surface are present for the case of practical MOS-C's. These non-idealities come in the form of traps and charges and will affect ideal MOS-C characteristics. The four types of traps and charges are:

(1) Interface trapped charges (Q_{it}, N_{it}, D_{it}): Positive or negative charges that exist at the oxide/semiconductor heterointerface. They exist due to structural defects, oxidation-induced defects, metal impurities, or other defects caused by radiation damage or other bond breaking processes [1, 2]. They are states within the semiconductor's forbidden bandgap and are in electrical communication with the underlying semiconductor [1, 2].

(2) Fixed oxide charges (Q_f, N_f): Positive charges near the oxide/semiconductor heterointerface, whose density of charges depends on oxidation ambient and temperature, cooling conditions, and the semiconductor's orientation [2]. Unlike the interface trap charges mentioned above, fixed oxide charges are not in electrical communication with the underlying semiconductor [1, 2].

(3) Oxide trapped charges (Q_{ot}, N_{ot}): Positive or negative charges due to carriers (i.e. holes or electrons, respectively) being trapped in the oxide [1, 2]. These carriers are trapped by mechanisms such as radiation, avalanche injection, hot-carrier injection, and other mechanisms [1, 2].

(4) Mobile oxide charge (Q_m, N_m): Charges that exist due to ionic impurities such as

Na$^+$, Li$^+$, K$^+$, and H$^+$ [2]. They are mobile within the oxide under bias-temperature

stress conditions [1].

A pictorial schematic of the above-mentioned traps and charges and their approximate

location within the oxide is shown below:
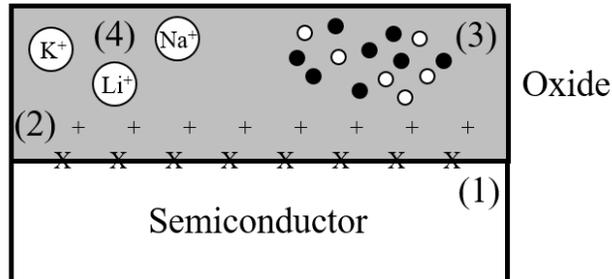
Figure 3.6 Pictorial schematic of traps and charges and their approximate locations within

the oxide where the numbers are associated with the charges mentioned above.

Having discussed interfacial traps in the previous section, it is now important to include

them in the different C-V modes discussed in section 4.1.1. Fig. 3.7 summarizes the

different modes and what capacitances are in effect at each mode for p-type MOS-C:
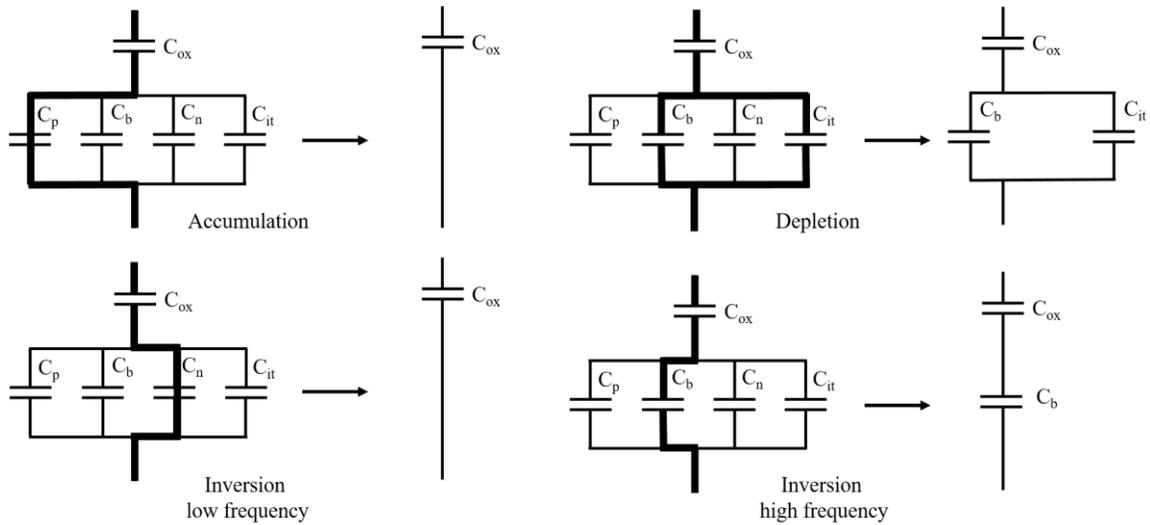
Figure 3.7 Effective capacitances at several modes in p-type MOS-C capacitance-voltage measurements.

where C_p, C_b, C_n, and C_{it} are the capacitances associated with holes, the bulk substrate, electrons, and interfacial traps. Essentially, all modes are exactly as described in section 4.1.1 with the exception of the depletion mode. Interfacial traps contribute a capacitance C_{it} which is in parallel with C_b. The effect of interfacial traps on C-V curves will be explained in a later section.

3.2.1 Maserjian Extraction Method of Oxide Capacitance C_{ox}

As can be seen in Fig. 3.7, C_{ox} plays a role in all the C-V modes. In addition, as will be later seen in the following sections, C_{ox} also plays an indubitable role in the extraction of various important parameters used to quantitatively and qualitatively gauge the quality of the oxide/semiconductor heterointerface. The method used in experimentally extracting C_{ox} from measured C-V curves in this thesis is called the Maserjian *et al.* method [3]. The governing equation is [3]:

$$\frac{1}{C_{hf,acc}} = \frac{1}{C_{ox}} + s \left| \frac{d(\frac{1}{C_{hf,acc}^2})}{dV} \right|^{1/4} \quad (3.9)$$

where $C_{hf,acc}$ is the high-frequency accumulation capacitance and s is a constant. Plotting $1/C_{hf,acc}$ versus $(d(1/C_{hf,acc}^2)/dV)^{1/4}$ and then linearly fitting the data point yields a linear line with y-intercept on the $1/C_{hf,acc}$ line equivalent to $1/C_{ox}$. Shown below in Fig. 3.8 is a generic example of the Maserjian Method being applied:
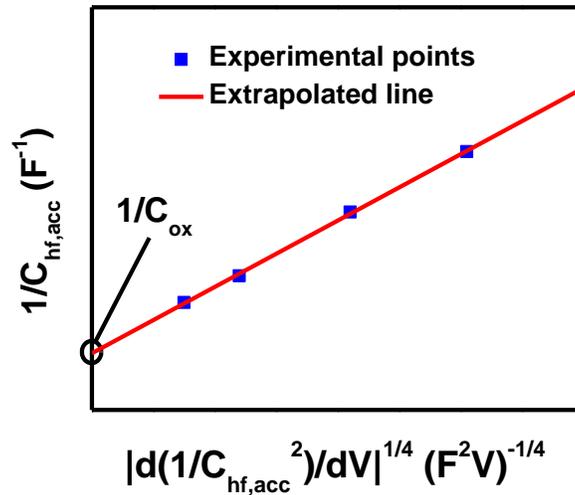
Figure 3.8   A generic example of the Maserjian *et al.* method [3] being applied, where blue is experimental points that would be extracted from C-V measurements and red is the extrapolated linear line. The encircled y-intercept is equivalent to $1/C_{ox}$.

Experimental points, shown in blue, are extracted from high-frequency C-V measurements and then fitted with a linear line, shown in red, which is then extrapolated to determine the y-intercept. The y-intercept, encircled in black, is equivalent to $1/C_{ox}$.

3.2.2 Extraction of Flatband Voltage V_{FB}

Another parameter of interest that will help extraction of important parameters is the flatband voltage V_{FB}. The method as discussed by Hillard *et al.* [4] requires plotting $d^2(1/C_{hf}^2)/dV^2$ against V_G, where C_{hf} is the high-frequency C-V, which in this thesis was the 1 MHz C-V curve. The double differentiated curve has a peak that coincides with $V_G = V_{FB}$. Fig. 3.9 shows a generic example of the application of this method to determine V_{FB}:
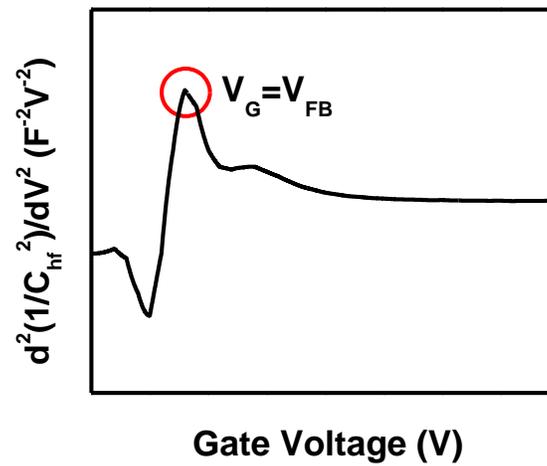
Figure 3.9   Hillard *et al.* [4] method being used to extract V_{FB} from generic 1 MHz C-V data.

Application of the Hillard *et al.* [4] method yields a sharp peak (encircled in red) whose location correlates to $V_G = V_{FB}$.

3.2.3 Extraction of Oxide Trapped Charge Density N_{ot}

With knowledge of both C_{ox} and V_{FB}, oxide trapped charge density N_{ot} can be easily obtained. By doing a forward and reverse sweep of a high-frequency curve, there is a shift in V_{FB} (ΔV_{FB}) between the two sweeps known as hysteresis. With knowledge of ΔV_{FB}, N_{ot} can be calculated using [2]:

$$N_{ot} = \frac{\Delta V_{FB} C_{ox}}{q} \qquad (3.10)$$

Fig. 3.10 shows a generic demonstration of hysteresis between a forward and reverse sweep of a high-frequency curve:
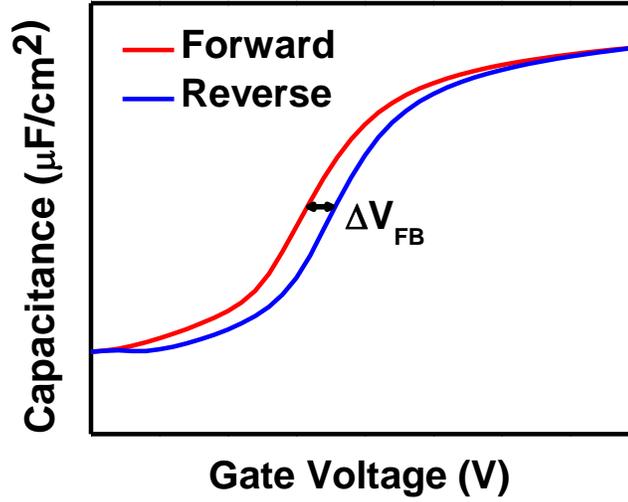
Figure 3.10   Hysteresis (ΔV_{FB}) between forward sweep (red curve) and reverse sweep (blue curve) of a generic high-frequency C-V curve.

3.2.4 Extraction of Doping Density

The extraction of doping density is based on an iterative approach called the maximum-minimum capacitance technique. In this technique, the doping density N can be derived using the following formula [2]:

$$N = \frac{4\varphi_F}{q K_s \varepsilon_0 A^2} \frac{C_{inv}^2}{(1-\frac{C_{inv}}{C_{ox}})^2} \qquad (3.11)$$

where C_{inv} is the minimum capacitance from the high-frequency C-V curve, K_s is the permittivity of the semiconductor, ε_0 is the vacuum permittivity, A is the gate electrode area, C_{ox} is the oxide capacitance, and ϕ_F is the Fermi surface potential. ϕ_F is also dependent on N and is given by the following formula [2]:

$$\varphi_F = \frac{kT}{q} \ln \frac{N}{n_i} \tag{3.12}$$

where k is the Boltzmann constant, T is temperature, q is the electron charge, and n_i is the intrinsic carrier concentration of the semiconductor. Since both 3.10 and 3.11 both depend on N, N can be iteratively found by continuously plugging in different values into eq. 3.11, which is then plugged into eq. 3.10. N is then found by comparing the difference between the N being plugged in and the N derived using eq. 3.10.

3.2.5 Extraction of Interfacial Trap Density D_{it}

A key metric used to quantitatively measure the quality of an oxide/semiconductor heterointerface is the interfacial trap density, D_{it}. Because these traps are in electrical communication with the underlying semiconductor, their influence on C-V curves are readily noticeable. The figure shown below highlight the effect of D_{it} on high- and low-frequency C-V curves:
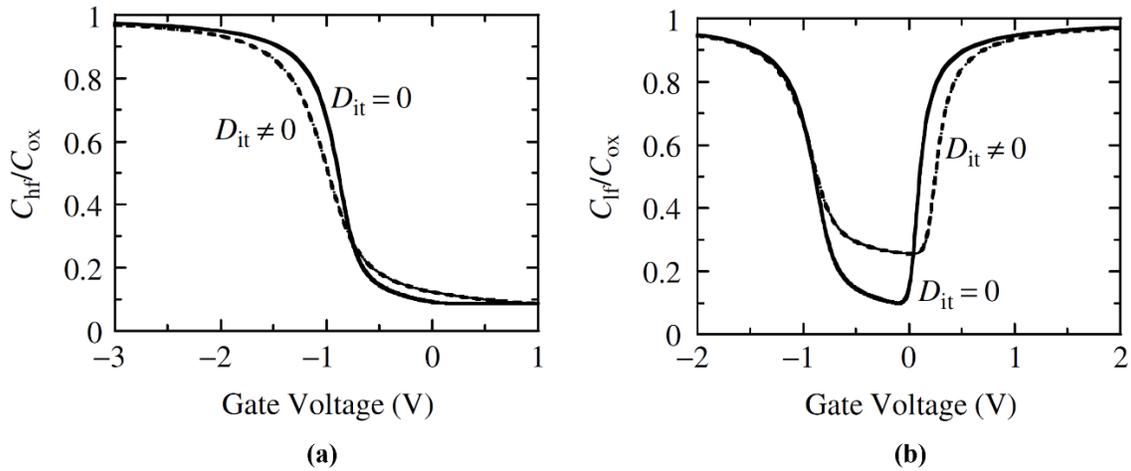
Figure 3.11 Theoretical normalized capacitance with and without the contribution of D_{it} for (a) high frequency and (b) low frequency [2].

For the high-frequency case, interface traps cannot follow the AC probe frequency and thus do not contribute any capacitance. However, they do follow the slowly varying DC bias.

Thus, as gate voltage is swept from accumulation to inversion, there is now a contribution of charge Q_{it} to the gate charge. As a result, the gate charge is now $Q_G = -(Q_S + Q_{it})$, meaning that now both semiconductor and interface traps must be charged [2]. As a result, the high-frequency C-V curve now stretches along the gate voltage axis. On the other hand, at lower frequencies, interface traps do respond to the AC probe frequency and contribute a capacitance C_{it}, resulting in some stretch-out along the gate voltage axis as well as higher minimum capacitance C_{min}.

3.2.5.1   Conductance Method Overview

Various methods exist to extract D_{it} including the quasi-static method, Terman method, Gray-Brown and Jenq method, charge pumping, and many more. The method that will be discussed in this thesis is the conductance method because it is the most sensitive method to measure D_{it} and it is also the most complete method, yielding D_{it} in both depletion and weak inversion portions of the band gap [2, 5]. Before delving into the extraction method itself, it's important to understand the MOS-C equivalent circuit models that are used in the conductance method, which are shown below:
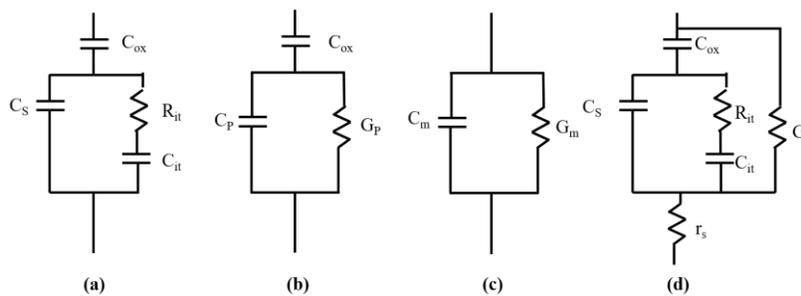
Figure 3.12   Equivalent circuits for the conductance method; (a) MOS-C with interface trap time constant $\tau_{it} = R_{it}C_{it}$, (b) simplified circuit of (a), (c) measured circuit, (d) (a) with included series resistance r_s and tunnel conductance G_t.

The equivalent circuit of a MOS-C for the conductance method is shown in Fig. 3.12a, where C_S is the semiconductor capacitance, C_{ox} is the oxide capacitance, C_{it} is the interface trap capacitance, and R_{it} is the resistance representing the lossy process of carrier capture-emission. The circuit in Fig. 3.12a can be simplified to the circuit shown in Fig. 3.12b to yield C_p and G_p, which are the equivalent parallel capacitance and conductance. C_p and G_p are given by the following equations [2]:

$$C_p = C_S + \frac{C_{it}}{1+(\omega\tau_{it})^2} \tag{3.13}$$

$$\frac{G_p}{\omega} = \frac{qD_{it}}{2\omega\tau_{it}}\ln[1 + (\omega\tau_{it})^2] \tag{3.14}$$

where $C_{it} = q^2 D_{it}$, $\omega = 2\pi f$ (f is measurement frequency), and $\tau_{it} = R_{it}C_{it}$. However, it's worth noting that capacitance meters generally assume the MOS-C to consist of the parallel combination of measured capacitance, C_m, and measured conductance, G_m, as shown in Fig. 3.12c. Thus, a more practical form for G_p/ω with the inclusion of C_m and G_m is given by the following equation [2]:

$$\frac{G_p}{\omega} = \frac{\omega G_m C_{ox}^2}{G_m^2 + \omega^2(C_{ox} - C_m)^2} \tag{3.15}$$

This form of G_p/ω assumes negligible series resistance, r_s. Accounting for r_s and tunnel conductance, G_t, gives the most complete circuit of the MOS-C, as shown in Fig. 3.12d. Including r_s and G_t parameters into the derivation of G_p/ω yields [2]:

$$\frac{G_p}{\omega} = \frac{\omega(G_c - G_t)C_{ox}^2}{G_c^2 + \omega^2(C_{ox} - C_c)^2} \tag{3.16}$$

where C_c and G_c are the corrected capacitance and conductance, respectively, where r_s is not negligible. C_c and G_c can be derived using the following equations [2]:

$$C_c = \frac{C_m}{(1 - r_s G_m)^2 + (\omega r_s C_m)^2} \tag{3.17}$$

$$G_c = \frac{\omega^2 r_s C_m C_c - G_m}{r_s G_m - 1} \tag{3.18}$$

G_t can be determined from eq. 3.17 by having $\omega \rightarrow 0$. The derivation of r_s is done using the following equation [2]:

$$r_s = \frac{G_{ma}}{G_{ma}^2 + \omega^2 C_{ma}^2} \tag{3.19}$$

where G_{ma} and C_{ma} are measured conductance and capacitance, respectively, in accumulation.

3.2.5.2   Accounting for Surface Potential Fluctuation

The conductance is measured as a function of frequency ω and plotted as G_p/ω. G_p/ω has a maximum at $\omega = 1/\tau_{it}$ and at that maximum, $D_{it} \approx 2.5 G_p/q\omega$ [2]. In addition to obtaining D_{it}, τ_{it} can also be obtained by the ω at the peak G_p/ω location on the ω-axis. However, experimental G_p/ω curves are typically much broader than the theoretical curves as predicted by eq. 3.14 due to surface potential fluctuations that are a result of non-uniform oxide and interface trap charges as well as doping density [2]. Thus, a more accurate extraction of D_{it} takes into account this peak broadening. Brews showed that with a single MOS conductance curve, some fitting parameters, and very little calculation, it is possible to accurately extract D_{it} while accounting for surface potential fluctuations [6]. The figure below shows a generic G_p/ω curve where by accounting for the width of the curve, interfacial broadening σ_s is considered in the derivation of D_{it} [6]:
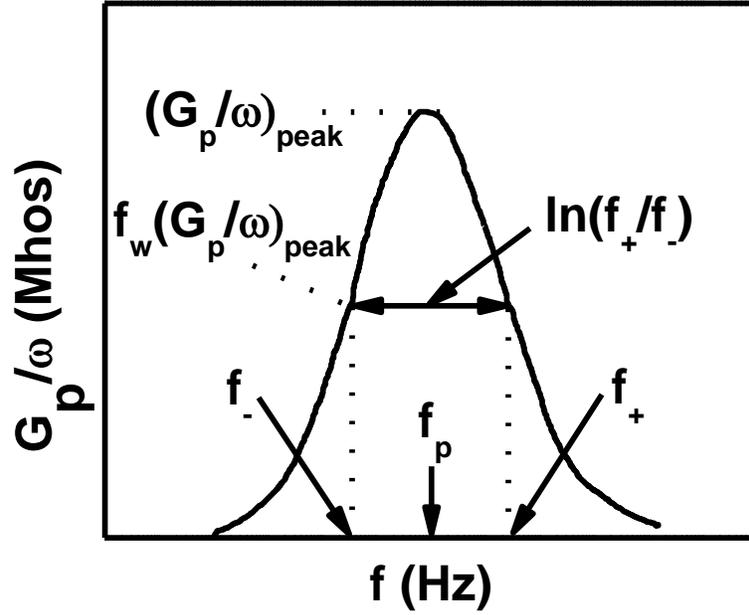
Figure 3.13   Conductance peak used in the analysis of D_{it} while accounting for surface potential fluctuation.

To find σ_s, an estimate of the width of the peak of the conductance curve is taken into account with $f_w(G_p/\omega)_{peak}$, or some fraction of $(G_p/\omega)_{peak}$. Here, $(G_p/\omega)_{peak}$ is the max value of the conductance curve and f_w is some arbitrarily chosen fraction. G_p/ω is a function of the parameter $\xi = \omega\tau$, where ω is radian frequency and τ is the capture time for a majority carrier by interface traps [6]. Because the influence of surface potential on peak broadening is regarded, width of the curve is an important parameter. Thus, there must be two points for this parameter (i.e. ξ_+ and ξ_-) and the width of the curve is simply $ln(\xi_+/\xi_-)$. To correlate the parameter ξ with a G_p/ω vs. f curve, $ln(\xi_+/\xi_-)$ is equivalent to $ln(f_+/f_-)$, where f_+ and f_- are the upper and lower bound of frequencies that correspond to $f_w(G_p/\omega)_{peak}$. Once $ln(\xi_+/\xi_-)$ is found, it can be fitted against various f_w to yield σ_s, as shown by Fig. 3.14:
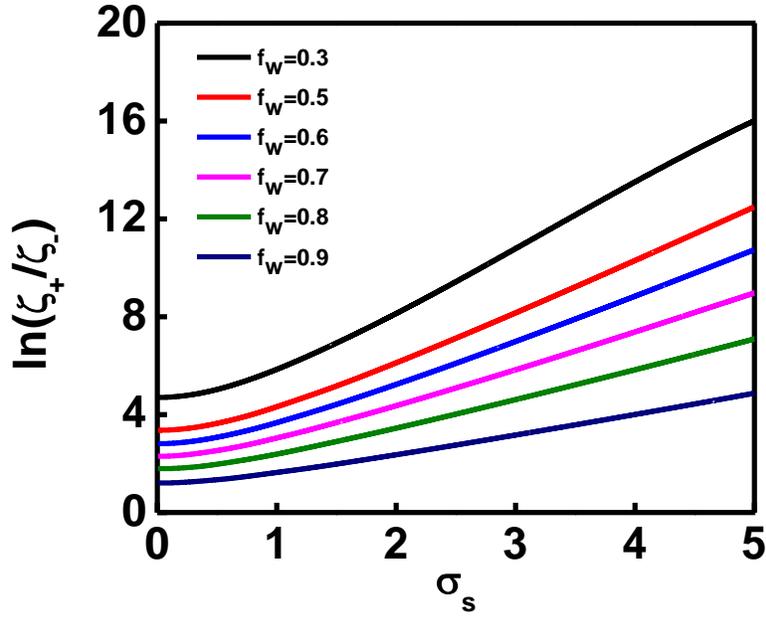
Figure 3.14 Plotted experimental width parameter $ln(\xi_+/\xi_-)$ against interfacial broadening parameter σ_s for various arbitrarily chosen fraction of conductance curve width f_w [6].

Once σ_s is determined, the ratio f_d can be determined, which is stated as:

$$f_D = \frac{(\frac{G_p}{\omega})_{peak}}{qD_{it}A}$$ (3.20)

where q is the electron charge and A is the device area. Thus, with the determination of f_D, D_{ikt} can be derived:

$$D_{it} = \frac{(\frac{G_p}{\omega})_{peak}}{qf_DA}$$ (3.21)

The correlation between f_D and σ_s is shown in the below figure:
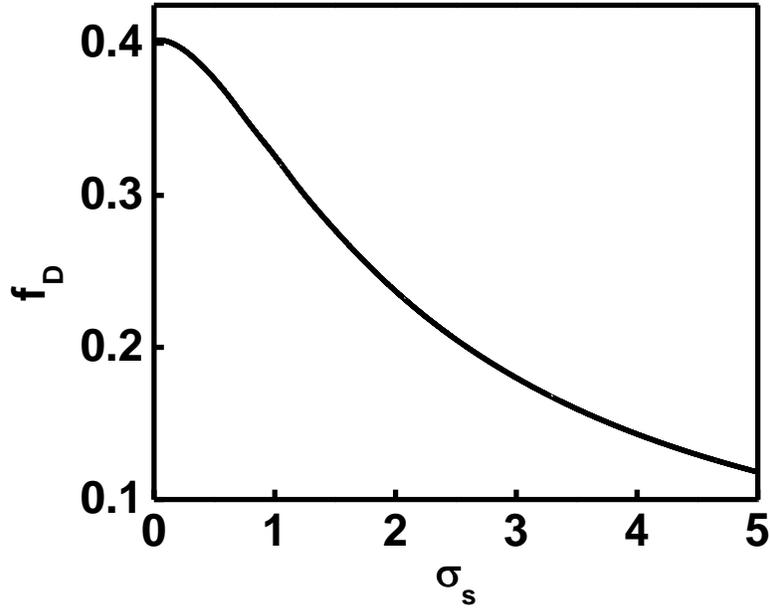
Figure 3.15 Plotted ratio f_D against interfacial broadening parameter σ_s. f_D can be used to determine D_{it} [6].

Once D_{iti} is extracted, each point, which is a function of frequency, can be plotted as a distribution across the bandgap using the characteristic trapping time relation [7]:

$$\tau = \frac{1}{2\pi f} = \frac{1}{\sigma v_{th} n_i} \exp(-\frac{E_t - E_i}{kT})$$ (3.22)

where τ is the characteristic trapping time, f is the measurement frequency, σ is the trap capture cross section, v_{th} is the carrier thermal velocity, n_I is the semiconductor intrinsic carrier concentration, E_t is the trap energy level, E_i is the midgap energy level, k is the Boltzmann's constant, and T is temperature.

3.2.6 Extraction of Fermi-level Efficiency by Application of Conductance Method

Asides from the extraction of D_{it}, another application of the conductance method is the application of the Fermi-level efficiency (*FLE*) method [8]. The *FLE* method utilizes

contours of the conductance peaks extracted from the conductance method and allows for the visualization of conductance peaks as both a function of gate voltage and frequency. Moreover, it allows for the tracing of the Fermi-level displacement as a function of gate voltage, where the change in this Fermi-level position or band bending is a reflectance of the trap density levels at the oxide-semiconductor heterointerface [8]. *FLE* is defined as follows [8]:

$$FLE = \frac{d\phi_s(V_G)}{dV_G} = \ln(\frac{f_2}{f_1})(\frac{kT/q}{V_1 - V_2}) \quad (\%) \tag{3.23}$$

where $\phi_s(V_G)$ is the semiconductor band bending with respect to gate voltage bias V_G and f_1 and f_2 are frequencies at which G_p/ω conductance peaks occur under voltage bias V_1 and V_2, respectively. A generic contour is shown below, where the dotted black line is the tracing of the Fermi-level:
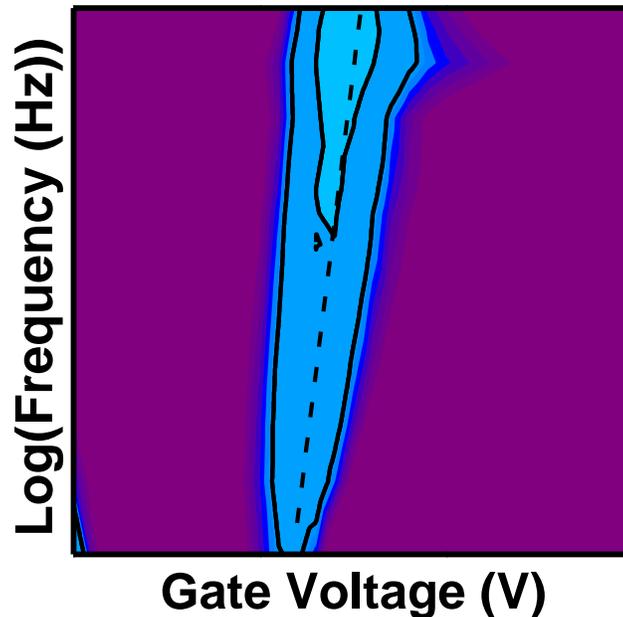
Figure 3.16   Generic conductance contour to demonstrate Fermi-level efficiency of a MOS-C. The dotted line traces the Fermi-level.

As can be seen in Fig. 3.16, the dotted line tracing denotes the Fermi-level tracing. The steeper the dotted line tracing, the more efficient the MOS-C can modulate the Fermi level with respect to gate voltage, which is an indicator of low D_{it}.

3.2.7 Low-temperature Extension of the Conductance Method for Full D_{it} Mapping

The conductance method is one of the most sensitive and complete techniques to measure D_{it} in MOS-C's, but when performed at room-temperature, the technique becomes limited in its scope for low-bandgap materials such as Ge. Unlike Si, low-bandgap materials have short time constants for the capture/emission processes of carriers through interface traps. As a result, C-V and G-V measurements done at room-temperature do not reflect the typical interface trap behavior of Si MOS-C's due to weak-inversion responses and thermal generation of minority carriers in Ge [9]. Because the conductance method was originally developed for Si MOS-C devices and thus, does not take into account minority carrier interaction, application of the conventional conductance method at room temperature can often lead to both under- and overestimation of D_{it} by more than an order of magnitude [9]. From eq. 3.22, it can be seen that trap energy level is dependent on variables such as temperature, σ, v_{th}, and n_i. The later three variables are typically not more than a magnitude in difference from semiconductor to semiconductor, which implies that the trap time constants is relatively similar from semiconductor to semiconductor. For all semiconductors at 300 K, it is impossible to see D_{it} at the band edges. Therefore, in order to obtain the full distribution of D_{it}, it is necessary to probe the MOS-C using a wide range of temperatures, including low temperatures down as low as 78 K for select semiconductors such as Ge.

References

[1] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, Hoboken, NJ, USA: Wiley, 2007.

[2] D. K. Schroder, *Semiconductor Material and Device Characterization*, 3rd ed. Hoboken, NJ, USA: Wiley, 2006.

[3] J. Maserjian, G. Peterson, and C. Svensson, "Saturation capacitance of thin oxide MOS structures and the effective surface state density of silicon," *Solid-State Electron.*, vol. 17, pp. 335-339, Apr. 1974.

[4] R. Hillard *et al*., "Direct and rapid method for determining flatband voltage from non-equilibrium capacitance voltage data," in *Diagnostic Techniques for Semiconductor Materials and Devices*. Pennington, NJ: ECS, 1992, pp. 261-274.

[5] E. H. Nicollian and A. Goetzberger, "The Si-SiO$_2$ Interface – Electrical Properties as Determined by the Metal-Insulator-Silicon Conductance Technique," *Bell Syst. Tech. J.*, vol. 46, pp. 1055-1133, Jul./Aug. 1967.

[6] J. R. Brews, "Rapid interface parametrization using a single MOS conductance curve," *Solid State Electron.*, vol. 26, no. 8, pp. 711–716, Aug. 1983.

[7] W. Shockley and W. T. Read, "Statistics of the recombination of holes and electrons," *Phys. Rev.*, vol. 87, no. 5, pp. 835–842, Sep. 1952.

[8] H. C. Lin *et al.*, "The Fermi-level efficiency method and its applications on high interface trap density oxide-semiconductor interfaces," *Applied Physics Letters*, vol. 94, no. 15, Apr. 2009.

[9] K. Martens *et al.*, "On the correct extraction of interface trap density of MOS devices with high-mobility semiconductor substrates," *IEEE Trans. Electron Devices*, vol. 55, no. 2, pp. 547–556, Feb. 2008.

Chapter 4 – Heteroepitaxial Ge MOS Devices on Si Using Composite AlAs/GaAs Buffer

Chapter 1 has demonstrated that Ge is an attractive candidate for next-generation low-power devices. However, there remains a need to integrate Ge on to the standard and established Si platform, as the integration of large-diameter, bulk Ge wafers with the current CMOS process is cost-prohibitive. Two methods to integrate Ge were mentioned in *chapter 1* including the use of Si_xGe_{1-x} buffer as well as GeOI techniques. Further methods not discussed in this thesis also exist to integrate Ge on to Si substrate [1]. However, none of these methods have addressed the utilization of high conduction and valance band offsets to confine carriers nor the elimination of parallel conduction within the buffer layers. A novel method was discussed in [2] where a large-bandgap III-V composite AlAs/GaAs buffer was used to integrate device-quality epitaxial Ge on to Si substrate. This AlAs/GaAs buffer has several advantages, namely i) the AlAs buffer acts as an insulating "oxide" layer due to its large bandgap (2.15 eV), thereby preventing parallel conduction to the active Ge layer; ii) high conduction and valance band offsets between the Ge/AlAs heterojunction of ~1 eV and ~0.5 eV helps confine electrons and holes, respectively, to the active Ge layer, and (iii) the GaAs buffer helps to reduce dislocations within the active Ge layer by acting as a "virtual" substrate for active layer integration onto the Si platform [2]. This chapter presents structural characteristics of the Ge epitaxial layer heterogeneously integrated on to Si substrate using a composite III-V AlAs/GaAs buffer as well as electrical characteristics of MOS-Cs fabricated from the aforementioned material stack.

4.1 MBE Growth of Ge/AlAs/GaAs Heterostructures on Si

In this chapter, the Ge on AlAs/GaAs buffer heterostructure was grown *in-situ* by solid source MBE utilizing separate III-V and Ge growth chambers connected *via* an ultra-high vacuum transfer chamber. The undoped epitaxial 240 nm thick Ge was heterogeneously integrated onto Si substrate by way of a large bandgap III-V composite buffer layer stack consisting of i) a 170 nm AlAs layer and ii) a 2.2 µm GaAs layer. The similar lattice constants of Ge, AlAs, and GaAs (5.658 Å, 5.660 Å, and 5.653 Å, respectively) help to minimize defect and dislocation propagation through the layers of interest, with the majority of the defects confined between the heterointerface of GaAs/Si, thereby demonstrating the virtual substrate characteristic of the GaAs buffer layer.

The GaAs nucleation and buffer layer were grown on (100) Si substrates that were 6° offcut towards the [110] direction using a two-step growth process involving i) a low-temperature growth stage of < 450 °C with low growth rate of 0.25 Å/s and ii) a high-temperature growth stage of ≥ 600 °C with high growth rate of 2.5 Å/s. The low-temperature stage creates two-dimensional growth of GaAs whereas the high-temperature stage reduces dislocations within the GaAs layer that are a result of the lattice mismatch between the GaAs layer and the Si substrate. Following the 2.2 µm GaAs virtual substrate layer growth, a large bandgap 170 nm AlAs layer was grown at a growth temperature of ≥ 600 °C. After the growth of the GaAs and AlAs buffer layers within the III-V compound semiconductor MBE chamber, the sample was cooled down from the growth temperature of ≥ 600 °C to below 200 °C before being transferred to the separate Ge MBE growth chamber using an ultra-high vacuum transfer chamber of ~5 × 10^{-10} torr. The 240 nm Ge layer growth was done at 400 °C at a low growth rate of 0.1 Å/s. The base pressure of the

Ge growth chamber is maintained at $4\text{-}5 \times 10^{-11}$ torr and was $\sim 2.4 \times 10^{-9}$ torr during the Ge growth. Concluding the Ge layer growth, the sample was slowly cooled down at a ramp rate of 5 °C/min to prevent thermal cracking of the structure.

The structural quality and relaxation state of the epitaxial Ge and underlying AlAs/GaAs buffer were evaluated by high-resolution triple-axis x-ray rocking curve from HR-XRD using a Panalytical X-pert Pro system equipped with both PIXel and proportional detectors. Panalytical Epitaxy simulation software was used to simulate the triple-axis x-ray rocking curve. Dynamic SIMS was used to determine the depth profile of Al, As, Ga, Ge, and Si atoms at the interfaces of the Ge/AlAs/GaAs heterostructure grown on Si substrate. The SIMS analysis was performed using a Cameca IMS-7f GEO with 5 kV Cs+ bombardment and MCs+ detection to reduce matrix effects. To analyze the elemental composition of the Ge/AlAs/GaAs heterostructure grown on Si substrate, EDS was performed using a JEOL 2100 TEM operating in scanning TEM mode. The electron transparent foils required for EDS were prepared by mechanical polishing and subsequent low-temperature Ar+ ion beam milling.

4.2   Fabrication of MOS-Cs on Epitaxial Ge Material Stack

N-type MOS-Cs were fabricated on the epitaxial Ge material stack. Fabrication of the devices began with a degrease using acetone, isopropanol, and deionized (DI) water, followed by a three minute native oxide removal in dilute (1:10) hydrofluoric acid. A high-quality, native GeO_x interfacial passivating layer was then formed by thermal oxidation at 450 °C for 10 minutes in an O_2 ambient. Immediately afterwards, a 4 nm Al_2O_3 gate oxide was deposited at 250 °C using a Cambridge NanoTech (ALD) system with trimethylaluminum and DI water as precursors for Al and oxygen, respectively. The 100

nm Al gate electrodes and ohmic contacts were subsequently deposited using a Kurt J. Lesker PVD250 electron beam deposition chamber. The devices were then annealed at 250 °C in forming gas (N$_2$:H$_2$ 95%:5%) for two minutes. A detailed process flow is shown below in Fig. 4.1:
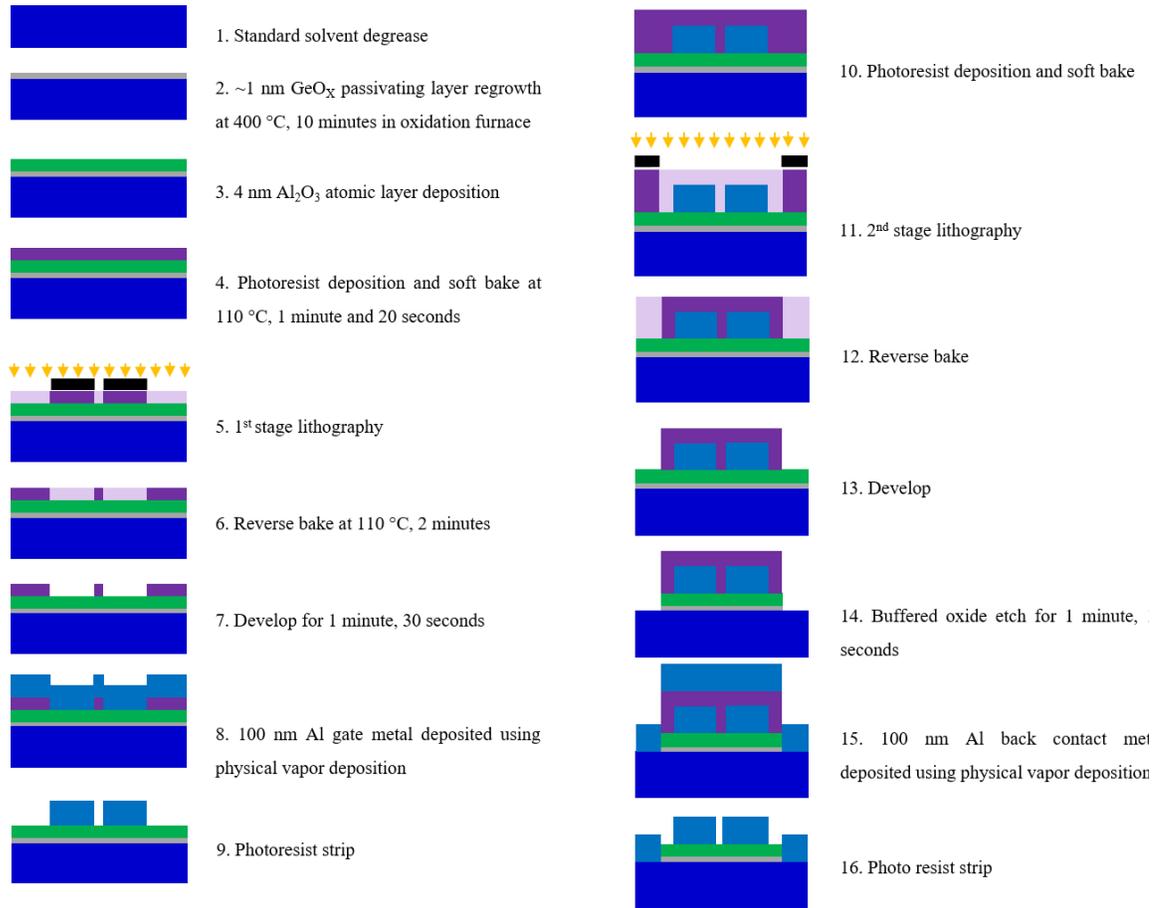
Figure 4.1 N-type MOS-C process flow on Ge/AlAs/GaAs heterostructures grown on Si substrate. Shown in blue is the Ge layer.

Fig. 4.2a shows a cross-sectional schematic of the fabricated Ge MOS-C device structure and Fig. 4.2b is a top-down optical image of the fabricated Ge MOS-C devices.
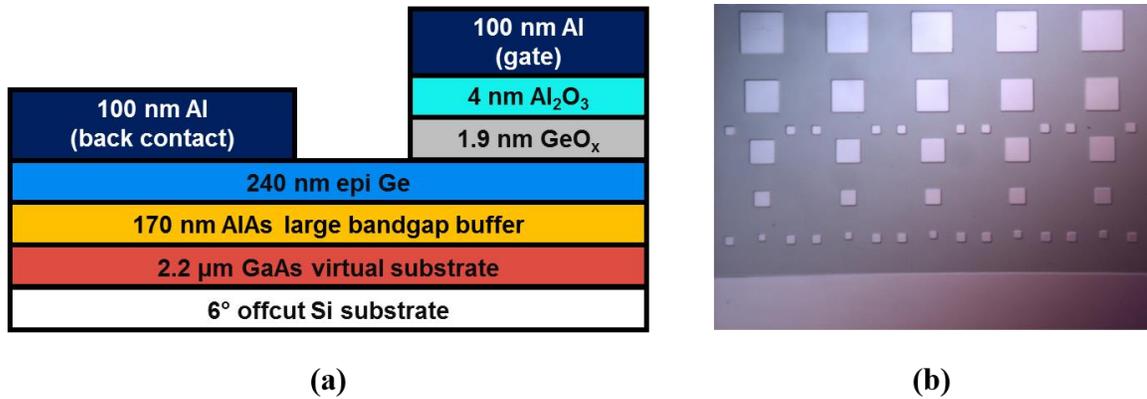
Figure 4.2 (a) cross-sectional schematic of the Ge MOS-C device structure and (b) top-down optical image of the Ge MOS-C devices.

In Fig. 4.2b, the device areas ranging from largest to smallest square are 1.5625×10^4 μm^2, 1.0×10^4 μm^2, 5.625×10^3 μm^2, and 2.5×10^3 μm^2. The electrical characteristics presented in this chapter are from measured devices with area 1.0×10^4 μm^2. The smallest squares seen in the optical image are alignment marks and are not measurable devices.

To evaluate the electrical characteristics of the fabricated MOS-C devices, low- and room-temperature multi-frequency C-V and conductance-voltage (G-V) were performed using an HP4284A precision LCR meter with frequencies ranging from 1 kHz to 1 MHz. Accurate measurements were obtained with the removal of series resistance, as discussed in [3]. TCAD Sentaurus software was used to simulate high frequency C-V curves to compare against the experimental data. To study the structural quality of the MOS-C interfaces, cross-sectional high-resolution TEM was performed using a JEOL 2100 TEM. The foils prepared for EDS were also used for cross-sectional high-resolution TEM.

4.3 Strain Relaxation Properties of the Ge/AlAs/GaAs Heterostructures on Si

To determine the structural quality and relaxation state of the epitaxial Ge active layer on AlAs/GaAs buffer layers grown on Si, high-resolution triple-axis x-ray (004) rocking curves were measured. Fig. 4.3 shows both experimental and simulated symmetric (004) rocking curve for the Ge/AlAs/GaAs heterostructure grown on Si [2]. From the experimental rocking curve, shown in blue in Fig. 4.3, distinct peaks of the Ge, AlAs, and GaAs layers, with angular separations resulting from differences in their respective lattice plane spacing, are clearly observed. The GaAs is almost fully relaxed with respect to Si and thus, serves as a virtual substrate for the epitaxial Ge. Additionally, the quasi-lattice-matched nature of the Ge, AlAs, and GaAs layers is confirmed due to the Ge and AlAs peak positions with respect to GaAs. A simulated rocking curve, shown in red in Fig. 4.3, shows distinct Ge, AlAs, and GaAs peaks fitting well with the experimental peaks. The small peak on the right side of the experimentally measured Si peak is due to a measurement artifact. To ensure accuracy of the peak positions, the simulation takes into account the quasi-lattice matched nature ($\sim$0.07% lattice mismatch) of the Ge/GaAs heterostructure. However, the simulation also assumes that the AlAs/GaAs buffer layer is fully relaxed with respect to Si. Thus, the simulated GaAs peak has a much narrower full width at half maximum as compared to that in the experimental measurement. Furthermore, the close proximity of the measured and simulated full width at half maximums (69.7 arcsec and 79.5 arcsec, respectively), where the simulation assumes a perfect Ge epitaxy, demonstrates that high-quality epitaxial Ge film growth was achieved.
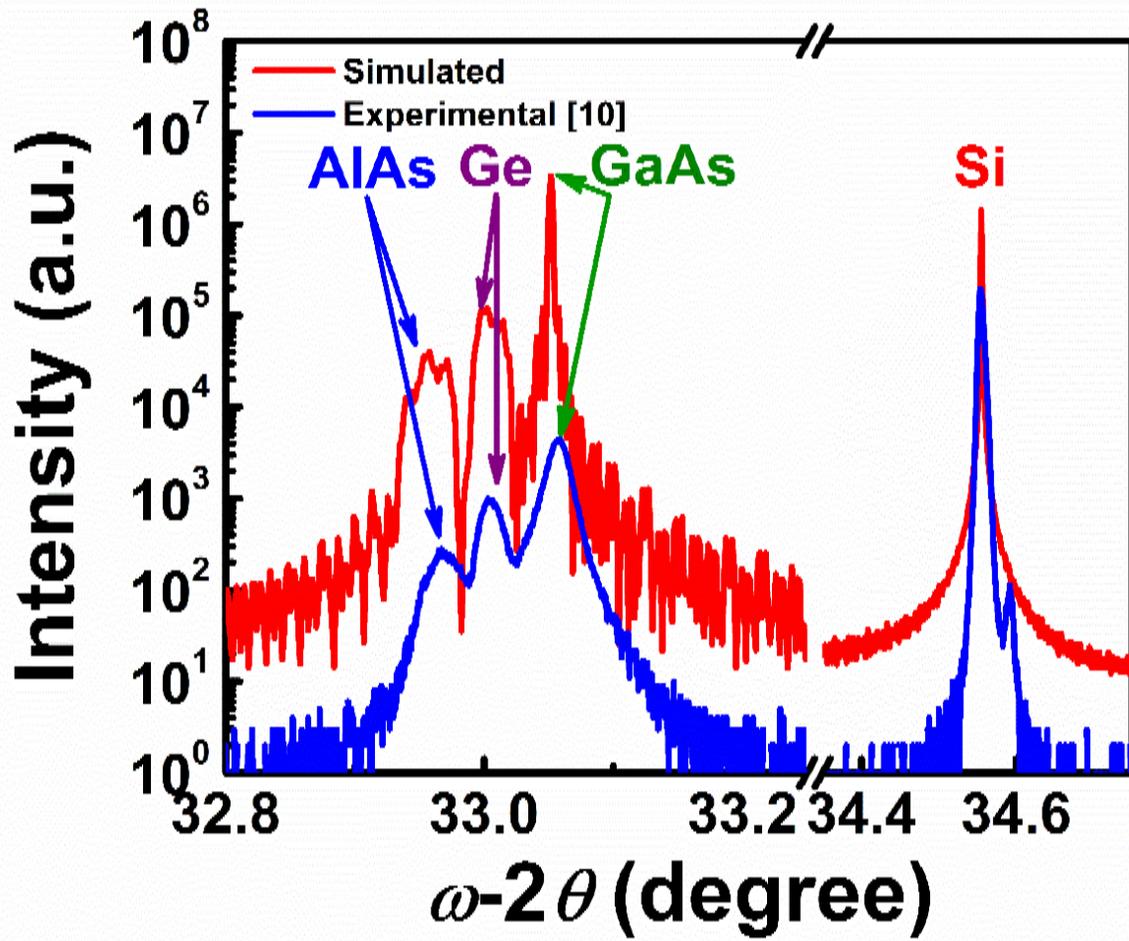
Figure 4.3   Simulated (red) and experimental [2] (blue) symmetric (004) X-ray rocking curve of the Ge/AlAs/GaAs heterostructure grown on Si substrate.

4.4 SIMS Depth Profile and EDS Analysis of the Ge/AlAs/GaAs Heterostructures on Si

Dynamic SIMS depth profiling provides insight into the extent of elemental diffusion between the epitaxial Ge active layer and the underlying AlAs/GaAs buffer layers grown on Si. Higher growth temperature promotes diffusion of adatoms across surfaces and heterointerfaces. Thus, a low Ge growth temperature (400 °C) was selected to allow epitaxial Ge growth with negligible interdiffusion of species at the Ge/AlAs heterointerface. Low Ge growth temperatures (≤ 450 °C) have been demonstrated to yield high-quality Ge epitaxial films with limited interdiffusion between the Ge and corresponding buffer layers [4]. Such optimization of Ge growth temperatures minimizes the alteration of the doping characteristics of the undoped Ge layer as a result of interdiffusion of species in the underlying buffer layers. Fig. 4.4a shows the Ge, Al, As, Ga, and Si compositional profiles in the Ge/AlAs/GaAs heterostructure grown on Si substrate. Constant Ge, Al, As, and Ga intensities within each layer demonstrates good growth uniformity of the epitaxial Ge layer as well as the underlying AlAs and GaAs buffer layers. Moreover, very low levels of Al, As, and Ga are detected within the Ge and a sharp, abrupt Ge/AlAs heterointerface is observed. The apparent interdiffusion between each interface is due to cascade mixing and the matrix effect during SIMS analysis. Thus, both Ge outdiffusion into the underlying AlAs buffer layer as well as indiffusion of Al, As, and Ga into the Ge layer are minimal, further reinforcing the successful growth of a high-quality Ge epitaxial layer on AlAs/GaAs buffer layers grown on Si. To obtain further insight into the elemental distribution and growth behavior of the Ge/AlAs/GaAs heterostructure grown on Si, EDS elemental mapping was performed. Fig. 4.4b shows a

three-element overlay of Ge, Al, and Ga, where the red corresponds to the Ge content, the blue corresponds to the Al content, and the green corresponds to the Ga content. Four distinct, uniform color regions corresponding to the Ge (red), AlAs (blue), GaAs (green), and Si (black) layers are clearly visible with no evidence of interdiffusion between the heterointerfaces, which would have been signified by color mixing. Fig. 4.4c shows the EDS line profile of the same Ge/AlAs/GaAs heterostructure, which depicts sharp transitions between the heterostructures as well as low levels of Al, As, and Ga within the epitaxial Ge layer. The apparent rise in the Ge and Si signal is an artifact of the EDS line profile scan. The EDS line profile serves as further reinforcement to the SIMS data to demonstrate minimal intermixing and interdiffusion of species.

4.5   MOS Capacitor TEM

The structural quality of the oxide and Ge surface was investigated by HR-TEM. The TEM micrograph shown in Fig. 4.5a shows an atomically abrupt and uniform oxide/Ge heterointerface, suggesting minimal interdiffusion across the interface, thereby minimizing the creation of interfacial defect states that would deleteriously affect device performance. Due to the low contrast in the TEM micrograph, there is no clear distinction between the ALD deposited Al_2O_3 and the thermally grown GeO_x layer shown in Fig. 4.5a. The total thickness of the bilayer oxide is about 5.9 nm, yielding an approximate GeO_x thickness of 1.9 nm. Fig. 4.5b shows the full material stack incorporating epitaxial Ge integrated on Si substrate via high-bandgap AlAs/GaAs buffers, where the TEM was performed after Al metallization.
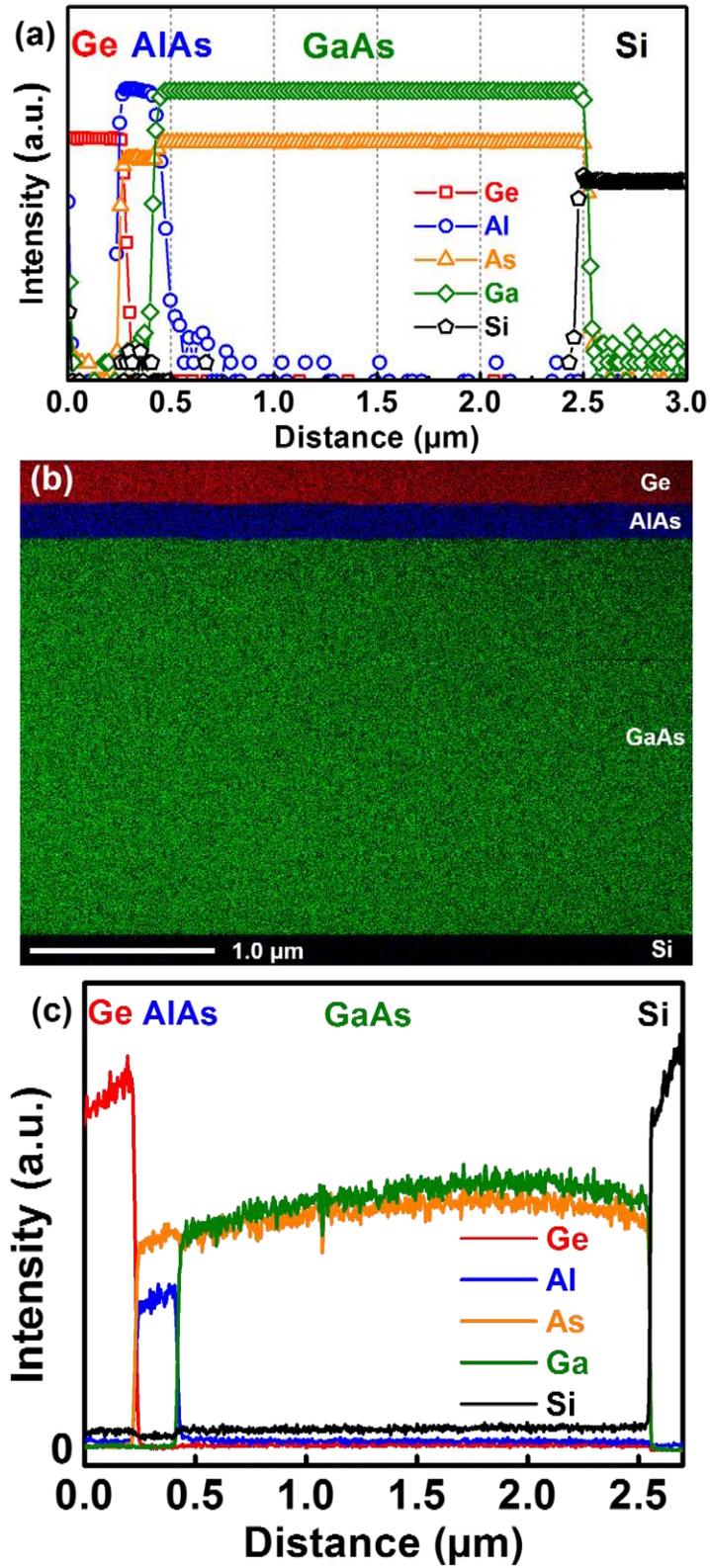
Figure 4.4   (a) SIMS profile, (b) EDS elemental mapping, and (c) EDS line profile of the

Ge/AlAs/GaAs heterostructure grown on Si.

70
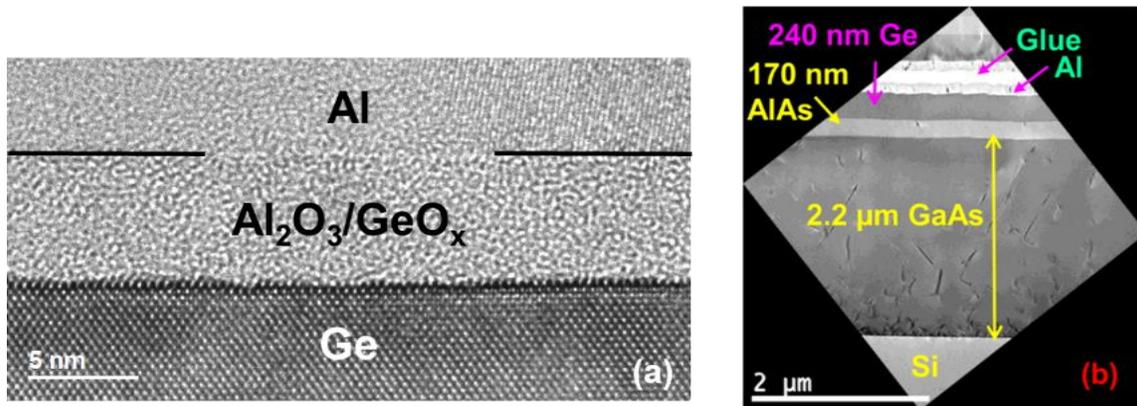
Figure 4.5 (a) Cross-sectional TEM micrograph of the oxide-Ge heterointerface of the Al/Al$_2$O$_3$/GeO$_x$/Ge MOS-C on Si via a composite AlAs/GaAs buffer architecture after post-metal annealing. (b) Low magnification TEM micrograph of the Ge/AlAs/GaAs/Si structure following Al metallization.

4.6 MOS Capacitor C-V and Conductance Characteristics

Fig. 4.6a-4.6c shows the C-V characteristics of the Ge MOS-C at (a) 89 K, (b) 150 K, and (c) 300 K, respectively. The low-frequency weak inversion response due to minority carriers, a feature of narrow bandgap materials, is suppressed at 89 K, thus resembling room temperature Si MOS C-V characteristics [5]. A moderate V_{FB} shift in the overall C-V curves was observed across all temperatures, which was attributed to the high unintentional doped (UID) of the epitaxial Ge, as previously reported [2]. The devices demonstrated hysteresis in the 100 kHz C-V measurements of 0.485 V at 300 K, which reduces to 0.165 V at 89 K. The large hysteresis is a result of bulk oxide traps and is not related to the passivation of the oxide/Ge heterointerface [6-7]. Furthermore, the bulk oxide traps may be due to different deposition techniques of high-k oxides, where the hysteresis in thermal ALD Al$_2$O$_3$ on GeO$_x$/Ge MOSCAPs [6-7] is much higher as compared to that in plasma-enhanced ALD (PE-ALD) Al$_2$O$_3$ on GeO$_x$/Ge MOSCAPs [8]. The reason for the

71

discrepancy may be due to the fact that, unlike PE-ALD, thermal ALD requires water as a pre-cursor for oxygen. Since GeO_x is hygroscopic and thus, very susceptible to moisture, this might be a reason for the larger oxide charge trapping responsible for the hysteresis as a result of water exposure during thermal ALD. From the hysteresis, N_{ot} can be extracted using eq. 3.10. C_{ox}, which is a necessary parameter to determine N_{ot}, was calculated to be 1.4 µF/cm^2 using the Maserjian et al. method [9] accounting for quantization effects, as seen in eq. 3.9. The calculated C_{ox} corresponds to an *EOT* of 2.5 nm. From eq. 3.10, N_{ot} of 4.25×10^{12} cm^{-2} and 1.21×10^{12} cm^{-2} was extracted from 300 K and 89 K, respectively. Additionally, the MOS-Cs exhibited limited C-V stretch-out, as shown in the inset of Fig. 4.6c, where a simulated 300 K 1 MHz curve with zero D_{it} is shown compared to the experimental 300 K 1 MHz curve. The parameters used for the simulated 1 MHz curve include a doping density of 1.24×10^{18} cm^{-3}, as estimated from eq. 3.11; N_{ot} of 4.25×10^{12} cm^{-2}, as extracted earlier from the 300 K C-V data from eq. 3.10; and D_{it} of 0 cm^{-2} eV^{-1} uniformly distributed throughout the bandgap. Limited C-V stretch-out of the experimental C-V curve as compared to the simulated C-V curve is indicative of low interface states throughout the upper half of the Ge bandgap [10]. The MOS-C's also demonstrate low frequency dispersion in the accumulation regime of 1.80% per decade over three decades at 300 K, which reduces down to 0.87% per decade over three decades at 89 K. The low frequency dispersion of the MOS-C's at 300 K suggests low interface states near midgap [10]. We speculate that this reduction in frequency dispersion is a reflection of the weak temperature dependence of the supply of carriers to the accumulation layer. Suggestive of low interface states near the conduction band edge is a low frequency-dependent V_{FB} shift [10-11], which was calculated to be 59 mV, 77 mV, and 153 mV for 89 K, 150 K, and 300

K C-V measurements, respectively. Finally, there is an apparent small frequency dispersion kink in the depletion regime due to interfacial traps, as observed clearly in Fig. 4.6a and 4.6b. Fig. 4.7a–4.7c shows the conductance contours corresponding to G-V sweeps measured at (a) 89 K, (b) 150 K, and (c) 300 K, respectively, which demonstrates the *FLE* of the MOS-Cs. The Fermi level trace (dotted black line) at each measurement temperature follows the conductance peaks under different frequency and bias conditions. Efficient biasing of the Fermi level from near midgap to close to the conduction band edge is observed in the MOS-C, as demonstrated by the steepness of the Fermi level traces with respect to gate bias change. Quantitative calculation of the *FLE* [12] was performed using eq. 3.23. Shockley–Read–Hall statistics can be applied to demonstrate FLE as a function of energy within the Ge bandgap using [13] eq. 3.22. σ_n was assumed to be a constant value [5] of 10^{-16} cm^{-2} [14]. Fig. 4.8 shows the *FLE* of the MOS-C as a function of E_t away from E_i, where E_C is the conduction band edge of Ge. A peak *FLE* of 27.5% was observed 0.19 eV away from E_i, which demonstrates good modulation of the Fermi level with respect to gate voltage, suggesting low D_{it} throughout the bandgap. The poor *FLE* closer to E_i, which is partially attributable to the additional conductance contribution by the minority carrier response [15], is clearly demonstrated by the horizontal "arcing" of the Fermi level trace, which can be clearly seen in Fig. 4.7c. Conductance contours extracted from G-V measurements performed at various temperatures from 89 K - 300 K were utilized in the extraction of *FLE*, as each temperature allows the sampling of a limited region of the Ge bandgap, as shown in the inset of Fig. 4.8.
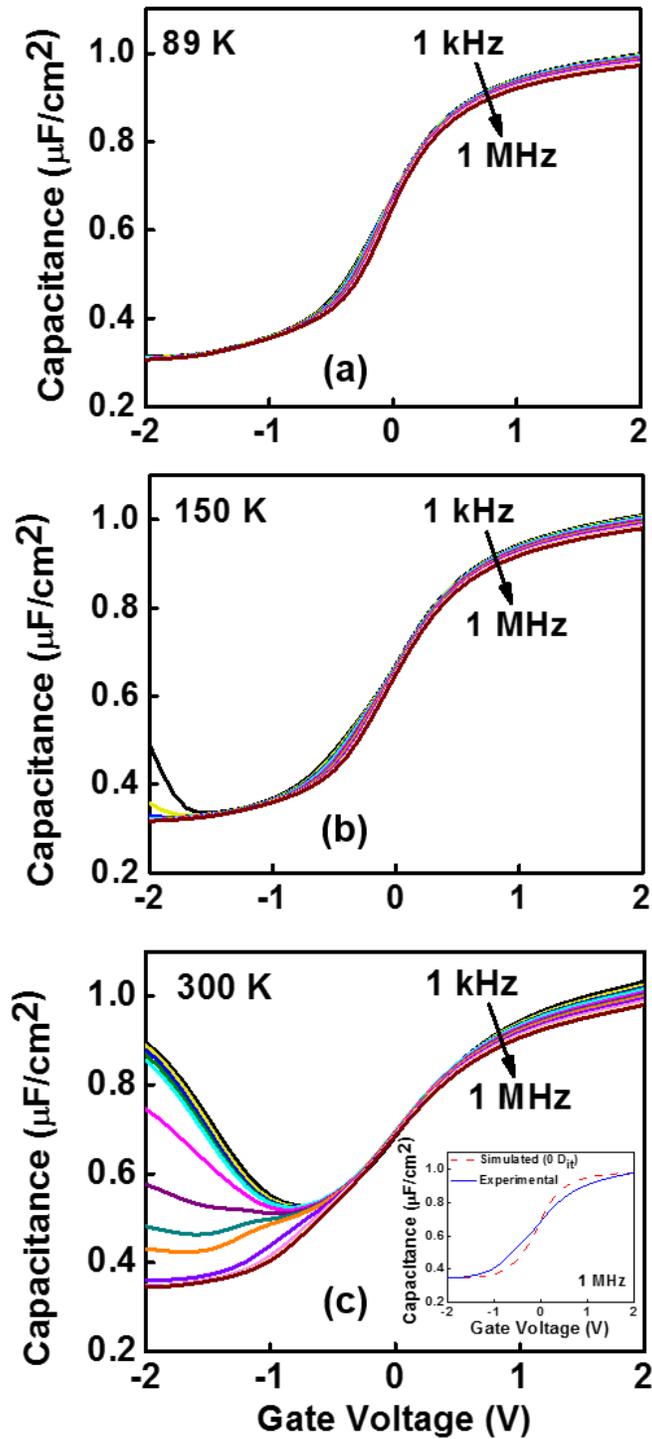
Figure 4.6 C-V characteristics of the Al/Al$_2$O$_3$/GeO$_x$/Ge MOS-C on Si via a composite AlAs/GaAs buffer architecture at (a) 89 K, (b) 150 K, and (c) 300 K. (Inset) Comparison of 1 MHz simulated (with zero D_{it}) and experimental curve showing limited C-V stretch-out.
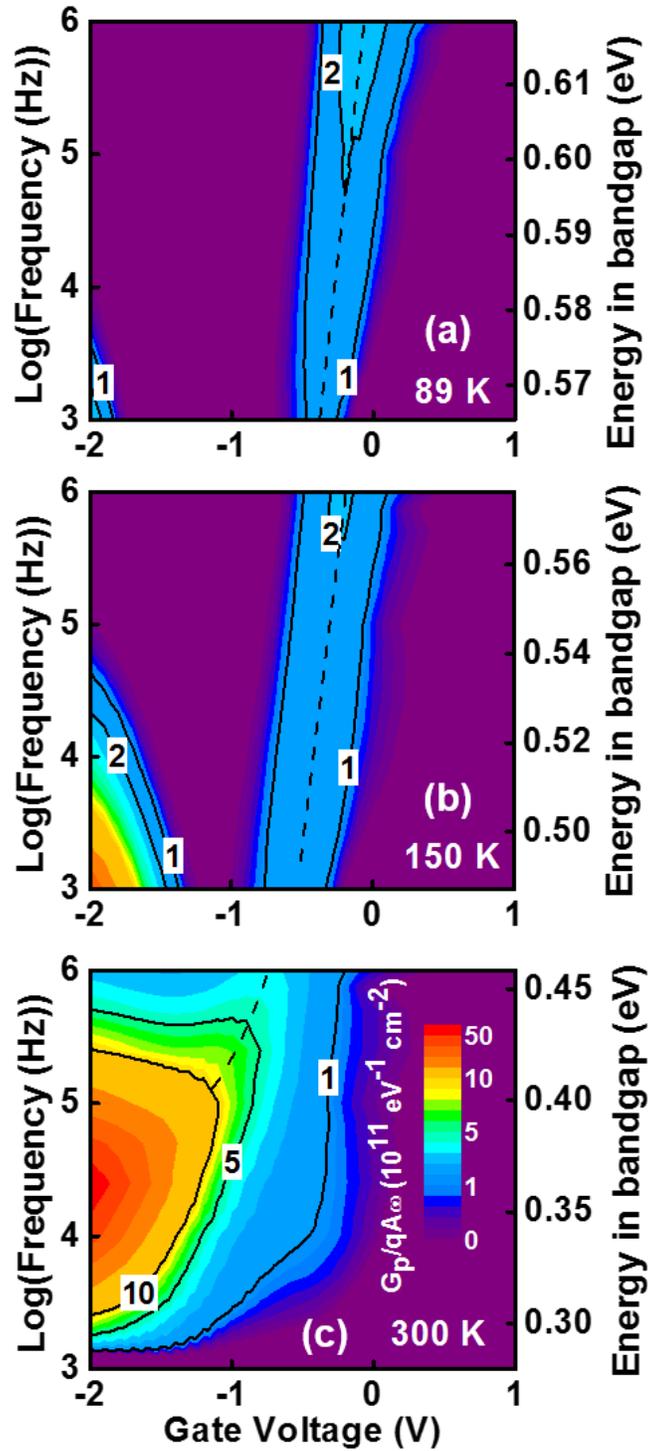
Figure 4.7 Conductance (G_p/ω) contours of the Al/Al$_2$O$_3$/GeO$_x$/Ge MOS-C on Si via a composite AlAs/GaAs buffer architecture at (a) 89 K, (b) 150 K, and (c) 300 K. All figures follow the color scale shown in (c).
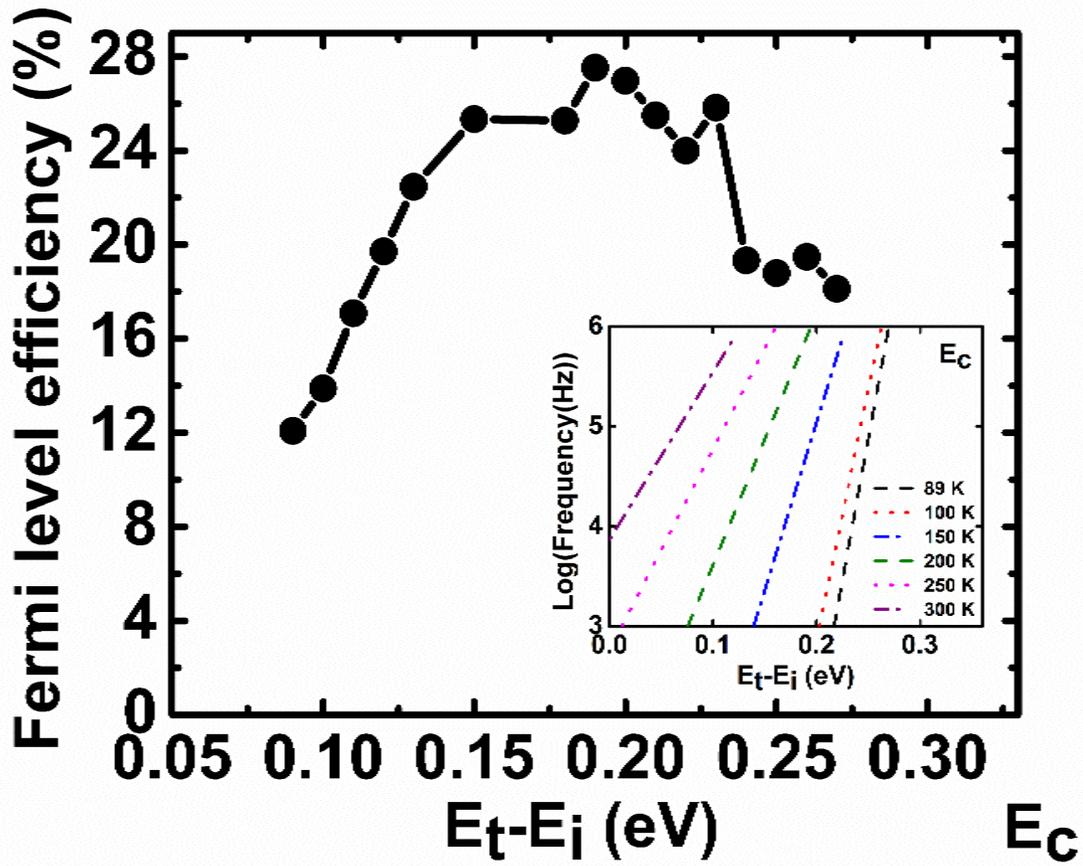
Figure 4.8 *FLE* as a function of energy of the Al/Al₂O₃/GeO*ₓ*/Ge MOS capacitor on Si via a composite AlAs/GaAs buffer architecture. (Inset) Ge bandgap energy ranges accessible at various measurement temperatures (89–300 K).

4.7   MOS Capacitor D_{it} Distribution

D_{it} was extracted from the conductance method with surface potential fluctuation correction [16-17] using eq. 3.21. The conductance method was performed from 89 K - 300 K to allow sampling of the D_{it} distribution at various ranges of the bandgap, where eq. 3.22 can be applied to show the distribution of D_{it} as a function of energy within the Ge bandgap. $f_D(\sigma_s)$ from eq. 3.21 is determined by fitting the approximate width of the conductance (G_p/ω) plot to established metrics, as discussed in [16]. Fig. 4.9 shows the extracted D_{it} as a function of $E_t - E_i$. A peak D_{it} of 1.09×10^{12} cm^{-2} ev^{-1} was observed close to the conduction band edge, E_C, while a minimum D_{it} of 8.55×10^{11} cm^{-2} eV^{-1} was observed approximately 0.15 eV away from E_C. Moreover, the D_{it} numbers are consistent with the *FLE* calculations, with higher D_{it} being responsible for Fermi level pinning, and therefore a lower *FLE*. Furthermore, the correlation between the extracted D_{it} and *FLE* are consistent with correlations reported elsewhere [12, 18]. The extracted D_{it} numbers are comparable to some of the values reported for the state-of-the-art bulk Ge MOS devices [19], which suggests both excellent passivation of the Ge surface as well as a defect-minimal Ge epitaxy on Si. The higher D_{it} extracted from the 200 K - 300 K measurements reflect an upper bound due to the temperature-dependent supply of minority carriers to the inversion layer which contribute to higher conductance, and thus lead to an overestimation of D_{it} [5, 14]. The insets of Fig. 4.9 show the inversion response of the device at both 200 and 250 K, which is due to the minority carrier response. The inversion response is also apparent from C-V measurements at 300 K, as shown in Fig. 4.6c.
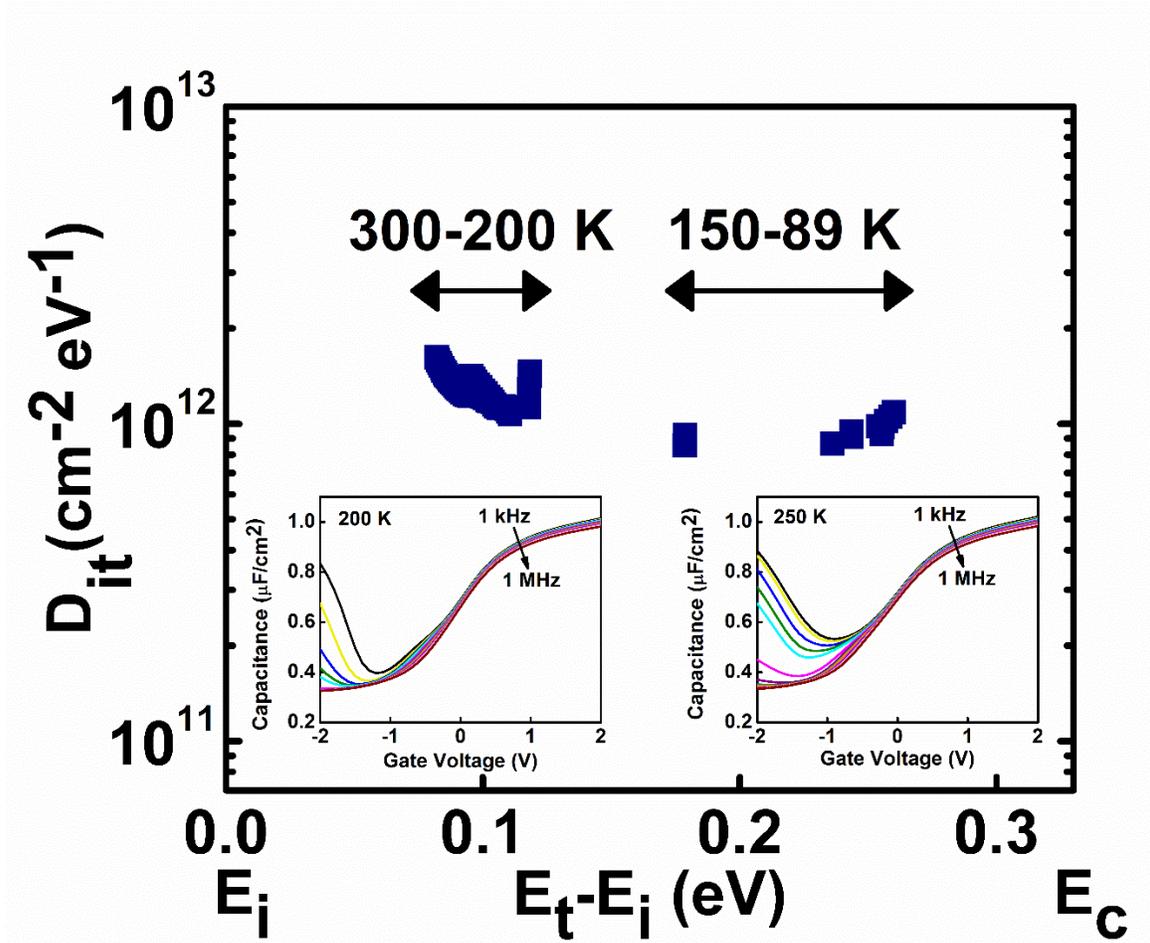
Figure 4.9 D_{it} as a function of energy of the Al/Al$_2$O$_3$/GeO$_x$/Ge MOS capacitor on Si via a composite AlAs/GaAs buffer architecture. (Insets) C-V measurement of the device at 200 and 250 K, showing inversion response.

References

[1]   P. S. Goley and M. K. Hudait, "Germanium based field-effect transistors: Challenges and opportunities," *Materials*, vol. 7, no. 3, pp. 2301–2339, Mar. 2014.

[2]   M. K. Hudait, M. Clavel, P. Goley, N. Jain, and Y. Zhu, "Heterogeneous integration of epitaxial Ge on Si using AlAs/GaAs buffer architecture: Suitability for low-power fin field-effect transistors," *Sci. Rep.*, vol. 4, pp. 6964–6970, Nov. 2014.

[3]   D. K. Schroder, Semiconductor Material and Device Characterization, 3rd ed. Hoboken, NJ, USA: Wiley, 2006.

[4]   M. K. Hudait, Y. Zhu, N. Jain, and J. L. Hunter, Jr., "Structural, morphological, and band alignment properties of GaAs/Ge/GaAs heterostructures on (100), (110), and (111)A GaAs substrates," *J. Vac. Sci. Technol. B*, vol. 31, no. 1, Jan./Feb. 2013.

[5]   K. Martens *et al.*, "On the correct extraction of interface trap density of MOS devices with high-mobility semiconductor substrates," IEEE Trans. Electron Devices, vol. 55, no. 2, pp. 547–556, Feb. 2008.

[6]   A. Delabie *et al.*, "Effective electrical passivation of Ge(100) for high-k gate dielectrics layers using germanium oxide," *Appl. Phys. Lett.*, vol. 91, no. 8, Aug. 2007.

[7]   F. Bellenger *et al.*, "Passivation of Ge(100)/GeO$_2$/high-k gate stacks using thermal oxide treatments," *J. Electrochem. Soc.*, vol. 155, no. 2, pp. G33–G38, 2008.

[8]   R. Zhang, T. Iwasaki, N. Taoka, M. Takenaka, and S. Takagi, "High-mobility Ge pMOSFET with 1-nm EOT Al$_2$O$_3$/GeO$_x$/Ge gate stack fabricated by plasma post oxidation," *IEEE Trans. Electron Devices*, vol. 59, no. 2, pp. 335–341, Feb. 2012.

[9]   J. Maserjian, G. Peterson, and C. Svensson, "Saturation capacitance of thin oxide MOS structures and the effective surface state density of silicon," *Solid-State Electron.*, vol. 17, pp. 335–339, Apr. 1974.

[10]   C.-W. Cheng, G. Apostolopoulos, and E. A. Fitzgerald, "The effect of interface processing on the distribution of interfacial defect states and the C-V characteristics of III-V metal-oxide-semiconductor field effect transistors," *J. Appl. Phys.*, vol. 102, no. 2, Jan. 2011.

[11]   C. Lin *et al.*, "Achieving a low interfacial density of states with a flat distribution in high-k Ga_2O_3 (Gd_2O_3) directly deposited on Ge," *Appl. Phys. Exp.*, vol. 4, no. 11, Nov. 2011.

[12]   H. C. Lin *et al.*, "The Fermi level efficiency method and its applications on high interface trap density oxide-semiconductor interfaces," *Appl. Phys. Lett.*, vol. 94, no. 15, Apr. 2009.

[13]   W. Shockley and W. T. Read, "Statistics of the recombination of holes and electrons," Phys. Rev., vol. 87, no. 5, pp. 835–842, Sep. 1952.

[14]   D. Kuzum, "Interface-engineered Ge MOSFETs for future high performance CMOS applications," Ph.D. dissertation, Dept. Electr. Eng., Stanford Univ., Stanford, CA, USA, Dec. 2009.

[15]   S. H. Tang *et al.*, "High quality Ge thin film growth by ultrahigh vacuum chemical vapor deposition on GaAs substrate," *Appl. Phys. Lett.*, vol. 98, no. 16, Apr. 2011.

[16]   J. R. Brews, "Rapid interface parametrization using a single MOS conductance curve," *Solid State Electron.*, vol. 26, no. 8, pp. 711–716, Aug. 1983.

[17]   E. H. Nicollian and J. R. Brews, *MOS Physics and Technology*. Hoboken, NJ, USA: Wiley, 2003.

[18]   N. Miyata, A. Ohtake, M. Ichikawa, T. Mori, and T. Yasuda, "Electrical characteristics and thermal stability of HfO2 metal-oxide-semiconductor capacitors fabricated on clean reconstructed GaSb surfaces," *Appl. Phys. Lett.*, vol. 104, no. 23, Jun. 2014.

[19]   Q. Xie *et al.*, "Germanium surface passivation and atomic layer deposition of high-k dielectrics—A tutorial review on Ge-based MOS capacitors," *Semicond. Sci. Technol.*, vol. 27, no. 7, Jun. 2012.

Chapter 5 – Study of Interface Passivation on Different Crystallographically Oriented Epitaxial Ge and Its Effects on MOS Capacitor Behavior

The study of surface orientation and its effects on device behavior is paramount for FinFET devices, where the sidewalls are of (110) orientation while the top of the fin is of (100) orientation [1]. Accordingly, it is also important to study the interfacial passivation of both surface orientations. Prior work has been done on the interfacial passivation of different crystallographically oriented bulk Ge MOS capacitor samples [2]. However, there is no prior work on the study of interfacial passivation on different crystallographically oriented epitaxial Ge MOS capacitors. This chapter continues upon the work of the previous chapter, where the structural and electrical properties of epitaxial Ge integrated on large-bandgap AlAs/GaAs buffer heterogeneously integrated on Si substrate was discussed. An in-depth analysis of different annealing schemes and their passivation role on the interface of epitaxial Ge heterogeneously integrated on to AlAs buffer on GaAs substrate is discussed in this chapter.

5.1 MBE Growth of (100) Ge/AlAs and (110) Ge/AlAs Heterostructures on GaAs Substrate

In this chapter, two heterostructures consisting of (100) and (110) Ge on AlAs buffer layer integrated on GaAs substrate were grown *in-situ* by solid source MBE utilizing separate III-V and Ge growth chambers connected *via* an ultra-high vacuum transfer chamber. Similar to the work from the previous chapter, the similar lattice constants of Ge,

AlAs, and GaAs (5.658 Å, 5.660 Å, and 5.653 Å, respectively) help to minimize defect and dislocation propagation through the layers of interest.

For the (100) oriented epitaxial Ge growth, a 250 nm GaAs nucleation layer was grown on (100) GaAs substrates that were 6° offcut towards the [110] direction at a growth temperature of $\geq$ 600 °C with a growth rate of 0.5 µm/hr. Following the GaAs epitaxial layer growth, a large bandgap 170 nm AlAs layer was grown at a growth temperature of $\geq$ 600 °C. After the growth of the GaAs and AlAs buffer layers within the III-V compound semiconductor MBE chamber, the sample was cooled down from the growth temperature of $\geq$ 600 °C to below 200 °C before being transferred to the separate Ge MBE growth chamber using an ultra-high vacuum transfer chamber of ~5 × 10^{-10} torr. The undoped 270 nm Ge layer growth was done at 400 °C at a low growth rate of 0.067 Å/s. The base pressure of the Ge growth chamber is maintained at 4-5 × 10^{-11} torr and was ~2.1 × 10^{-9} torr during the Ge growth. Concluding the Ge layer growth, the sample was slowly cooled down at a ramp rate of 5 °C/min to prevent thermal cracking of the structure.

For the (110) oriented epitaxial Ge growth, a 250 nm GaAs epitaxial layer was grown on (110) GaAs substrates at a growth temperature of $\geq$ 600 °C with a growth rate of 0.5 µm/hr. Similar to the (100) epitaxial Ge growth, a large bandgap 170 nm AlAs layer was grown at a growth temperature of $\geq$ 600 °C following the GaAs epitaxial layer growth. The sample followed a similar cool down procedure as the (100) epitaxial Ge growth, where following the cool-down, the sample was then transferred to the separate Ge MBE growth chamber using an ultra-high vacuum transfer chamber of ~5 × 10^{-10} torr. The undoped 270 nm Ge layer growth was done at 400 °C at a low growth rate of 0.067 Å/s. The base pressure of the Ge growth chamber is maintained at 4-5 × 10^{-11} torr and was ~2.45

$\times 10^{-9}$ torr during the Ge growth. Concluding the Ge layer growth, the sample was slowly cooled down at a ramp rate of 5 °C/min to prevent thermal cracking of the structure.

5.2   Fabrication of MOS-Cs on Different Crystallographically Oriented Epitaxial Ge Material Stacks

N-type MOS-Cs were fabricated on both the (100) and (110) crystallographically oriented epitaxial Ge material stacks, where they followed identical fabrication processes. Fabrication of the devices began with a degrease using acetone, isopropanol, and deionized (DI) water, followed by a one minute native oxide removal in dilute (1:10) hydrofluoric acid. A high-quality, native GeO_x interfacial passivating layer was then formed by thermal oxidation at 450 °C for 40 minutes in an O_2 ambient. Immediately afterwards, a 4 nm Al_2O_3 gate oxide was deposited at 250 °C using a Cambridge NanoTech ALD system with trimethylaluminum and DI water as precursors for Al and oxygen, respectively. The 0.8 nm TiN/100 nm Al gate electrodes and 0.8 nm TiN/100 nm Al/10 nm Ti/30 nm Ni ohmic contacts were subsequently deposited using a Kurt J. Lesker PVD250 electron beam deposition chamber, Ni is known to form excellent ohmic contacts to Ge by forming nickel germanide [3, 4] and Ti helps prevent the unwanted formation of NiGeAl alloy, which is lower resistance compared to pure NiGe alloy [5]. A process flow similar to that shown in Fig. 4.1 was followed for the fabrication of the devices studied in this chapter. Below, Fig. 5.1a shows a cross-sectional schematic of the fabricated (100) Ge MOS-C device structure and Fig. 5.1b shows a cross-sectional schematic of the fabricated (110) Ge MOS-C device structure.
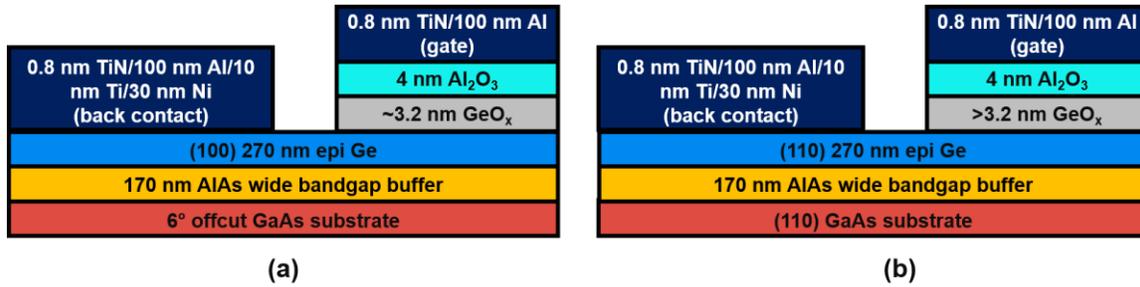
Figure 5.1 (a) cross-sectional schematic of the (100) Ge MOS-C device structure and (b) cross-sectional schematic of the (110) Ge MOS-C device structure.

As shown in Fig. 5.1a, ~3.2 nm was estimated for the thickness of GeO_x for the (100) sample [6]. However, >3.2 nm was estimated for the thickness of GeO_x for the (110) sample, as shown in Fig. 5.1b, due to the higher thermal oxidation rates for (110) Ge [2].

Three different annealing schemes and their role on the passivation of both (100) and (110) crystallographically oriented Ge were studied in this chapter and are summarized in table 5.1 below:

Table 5.1 Annealing schemes utilized for the fabricated devices.

Annealing scheme	Time (min)	Temperature (°C)	Annealing Ambient
Post deposition anneal (PDA)	30	405	O_2
Post metal anneal (PMA) 1	30	405	O_2
PMA 2	2	300	Forming gas (N_2:H_2 95%:5%)

The post deposition anneal (PDA) was performed following the deposition of 4 nm Al_2O_3 using the Cambridge ALD system, where the devices were placed back into the same furnace used for the GeO_x regrowth for 30 min at 405 °C under an O_2 ambient. The other annealing schemes, PMA and PMA 2, were performed following the 0.8 nm TiN/100 nm Al/10 nm Ti/30 nm Ni ohmic contact formation, where PMA 1 (which hereon shall be referred to as O_2 PMA) annealing scheme was performed under an O_2 ambient for 30 min at 405 °C using the same furnace for the GeO_x regrowth and PMA 2 annealing scheme (which hereon shall be referred to as FGA) was performed under a forming gas ambient for 2 min at 300 °C, which was found to be the optimal FGA condition for both (100) and (110) samples, as will be demonstrated in the section below.

The oxygen composition of GeO_x for both the (100) and (110) epitaxial Ge heterostructures was investigated using a PHI Quantera SXM XPS system with a monochromatic Al-Ka (energy of 1486.7 eV) x-ray source. The Ge 3d binding energy spectra was collected with a pass energy of 26 eV and an exit angle of 45u. The binding energy was corrected by adjusting the carbon 1 s CL peak position to 285.0 eV for each sample surface. Curve fitting was performed by CasaXPS 2.3.14 using a Lorentzian convolution with a Shirley-type background.

To evaluate the electrical characteristics of the fabricated MOS-C devices, low- and room-temperature multi-frequency C-V and conductance-voltage (G-V) were performed using an HP4284A precision LCR meter with frequencies ranging from 1 kHz to 1 MHz. Accurate measurements were obtained with the removal of series resistance, as discussed in [7].

5.3 XPS Analysis of GeO$_x$ Composition

To study the GeO$_x$ composition, XPS measurements were performed for the Ge oxides on the epitaxial Ge (100) and (110) surfaces oxidized at 450 °C prior to annealing. Fig. 5.2a-b shows the XPS spectra for the Ge $3d$ orbital for the (100) and (110) epitaxial Ge surface, respectively, where the curves corresponding to Ge^{4+}, Ge^{3+}, Ge^{2+}, Ge $3d_{3/2}$, and Ge $3d_{5/2}$ were determined by fitting with the measured result, where the measured result is denoted by black circles and the fit is shown as the red line. As shown in Fig. 5.2a for the (100) epitaxial Ge surface case, the Ge^{4+} peak dominates, while the suboxide peaks are much smaller, denoting that the GeO$_x$ thermally grown on (100) epitaxial Ge surface is composed dominantly of GeO$_2$. However, for the (110) epitaxial Ge surface case, as shown in Fig. 5.2b, the Ge^{3+} peak dominates, which is why there is less of a distinction between the GeO$_x$ and the Ge peaks as compared to the case for (100) epitaxial Ge surfaces. Therefore, the GeO$_x$ grown on (110) epitaxial Ge surfaces is not composed dominantly of GeO$_2$ unlike the case for GeO$_x$ grown on (100) epitaxial Ge surfaces, yielding an oxide predominantly made of Ge$_2$O$_3$ rather than the preferable GeO$_2$. This differs from the bulk Ge case, where, irrespective of surface orientation, the thermally grown GeO$_x$ is composed dominantly of GeO$_2$ [2]. Therefore, further optimization of thermal oxidation times and temperatures in the future will be necessary to produce better quality of thermally grown GeO$_x$ on (110) epitaxial Ge surfaces.
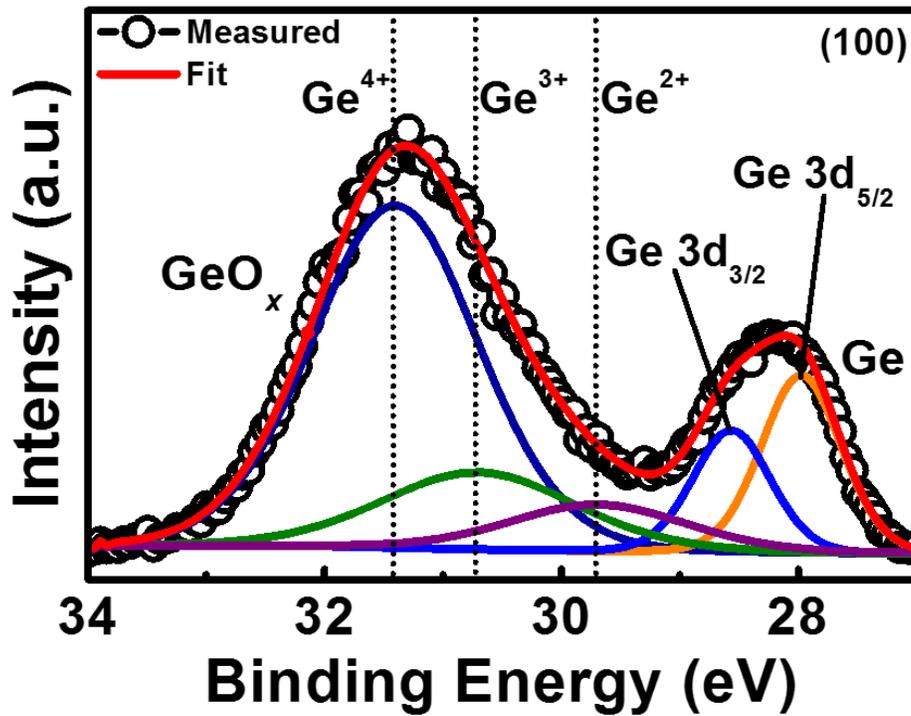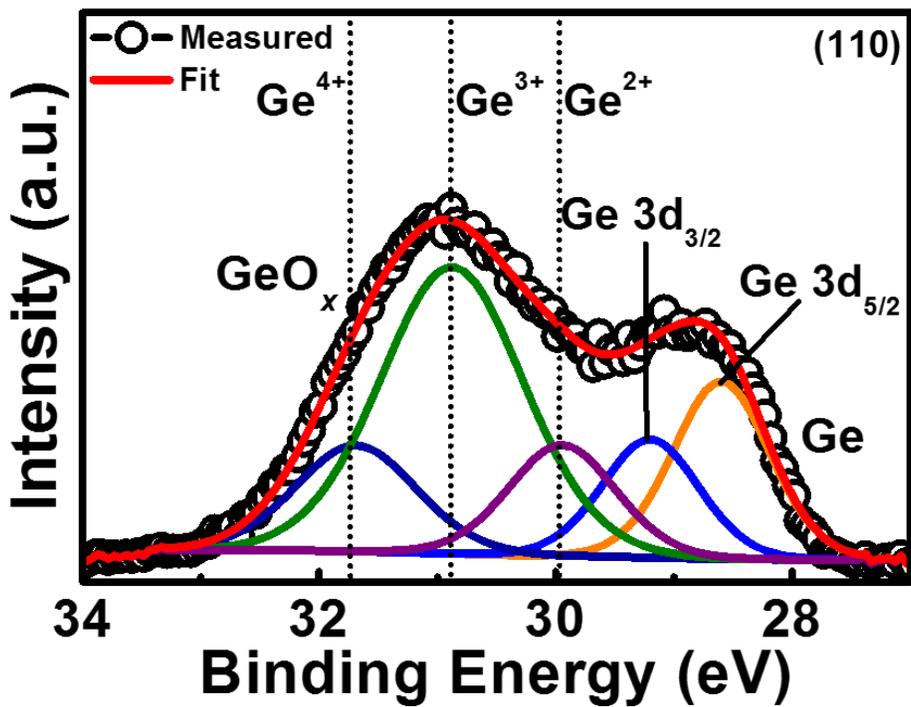
Figure 5.2 XPS analysis on the oxygen composition of GeO$_x$ thermally grown on (a) (100) epitaxial Ge surfaces and (b) (110) epitaxial Ge surfaces.

5.4   Optimization of FGA

Optimization of the FGA annealing scheme was done on both (100) and (110) epitaxial Ge devices. The four conditions compared include (1) pre-anneal control, (2) FGA for 250 °C for two minutes, (3) FGA for 300 °C for two minutes, and (4) FGA for 350 °C for two minutes.

5.4.1 FGA Optimization for (100) Ge MOS Devices

Fig. 5.3a-d below shows the room-temperature (300 K) C-V characteristics of the (100) Ge MOS-C under different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes. With increasing addition, the kinks in the higher frequency curves, which correspond to D_{it} response at weak inversion bias [8], flatten out with increasing thermal budget, demonstrating improved passivation of the GeO$_x$/Ge interface. However, with increasing thermal budget, there is also a decrease in the C_{max}, which as a result, will reduce C_{ox}, and therefore EOT, as will be demonstrated later in the section. In addition, the 750 kHz and 1 MHz curves, which are pointed out in the figures, are unusually higher than the lower frequency curves, which is believed to be a contact issue.
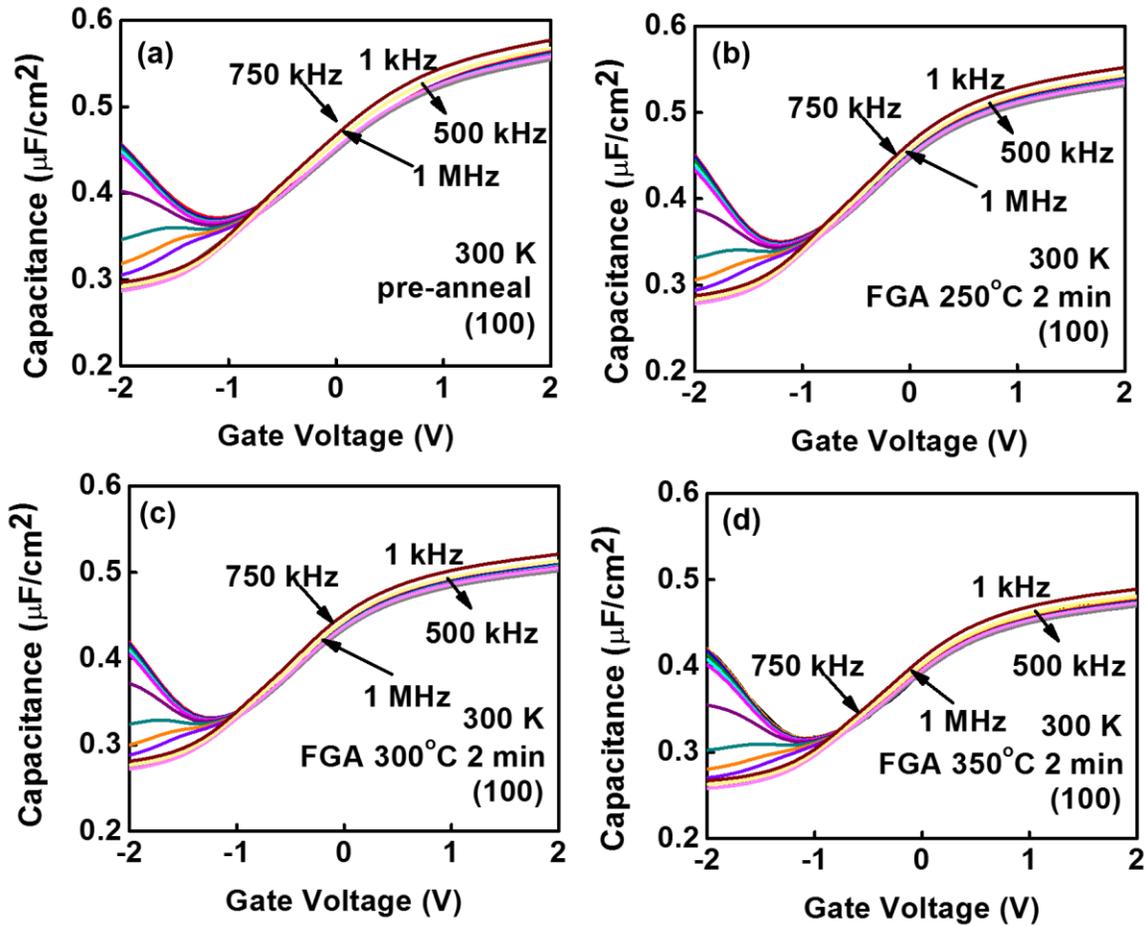
Figure 5.3 C-V characteristics of the (100) Ge MOS-C on GaAs via an AlAs buffer architecture for different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes.

Fig. 5.4a-d below shows the 300 K conductance contours of the (100) Ge MOS-C under different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes, where all four figures follow the same color scale. With increasing thermal budget, the magnitude of conductance within the depletion region begins to taper significantly, as seen by the larger blue region in Fig. 5.4a disappears gradually into the background purple color. This reduction of conductance magnitude signifies that the GeO_x/Ge heterointerface is being
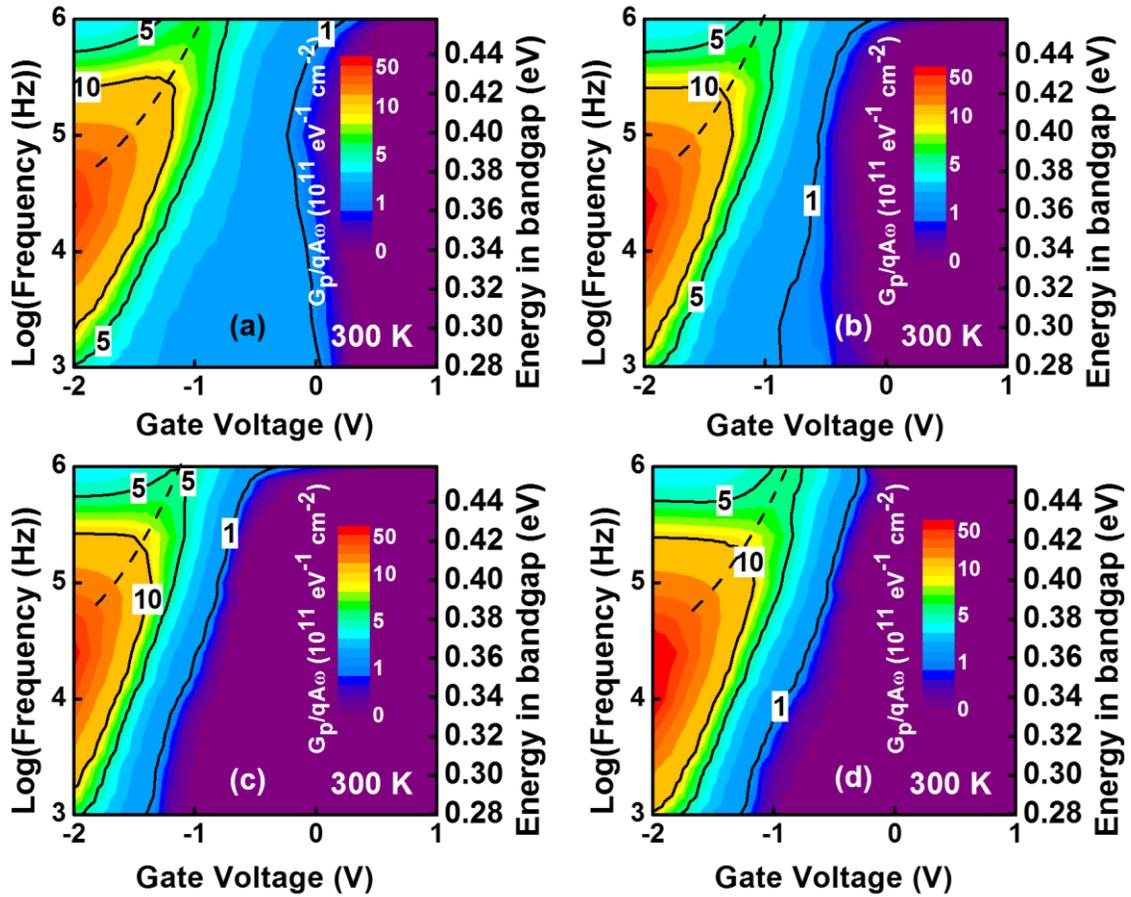
Figure 5.4 Conductance (G_p/ω) contours of the (100) Ge MOS-C on GaAs via an AlAs buffer architecture for different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes.

passivated by the reduction of D_{it} because D_{it} is directly proportional to the magnitude of the conductance peaks. However, from 5.4c to 5.4d, the colored conductance contours begins to broaden, signifying an increase in electron trapping due to a rise in D_{it} [9]. The black line, which demonstrates the modulation of the Fermi-level and therefore, the overall *FLE* of the device [10], does not appear to significantly reduce in arcing, which would be indicative of more efficient modulation of the Fermi-level with gate voltage bias, and therefore an increase in *FLE*. However, this is typically due to minority carrier response,

which is due to the additional conductance contribution by the minority carrier response [11].

Fig 5.5 shows *EOT* (marked in black) and N_{ot} (marked in blue) of the (100) Ge devices as a function of FGA temperature. *EOT* is calculated from extracted C_{ox} values obtained using the Maserjian *et al.* method [12], as summarized in eq. 3.9. N_{ot} is derived from the hysteresis of the 100 kHz C-V curves and can be calculated using eq. 3.10. N_{ot} shows a relatively linear decline with increasing annealing temperature, demonstrating proper passivation of trapped oxide charges, thereby resulting in a decrease in hysteresis. However, of concern is the increasing *EOT* with increasing annealing temperature. *EOT* shows a significant increase beyond an annealing temperature of 250 °C.
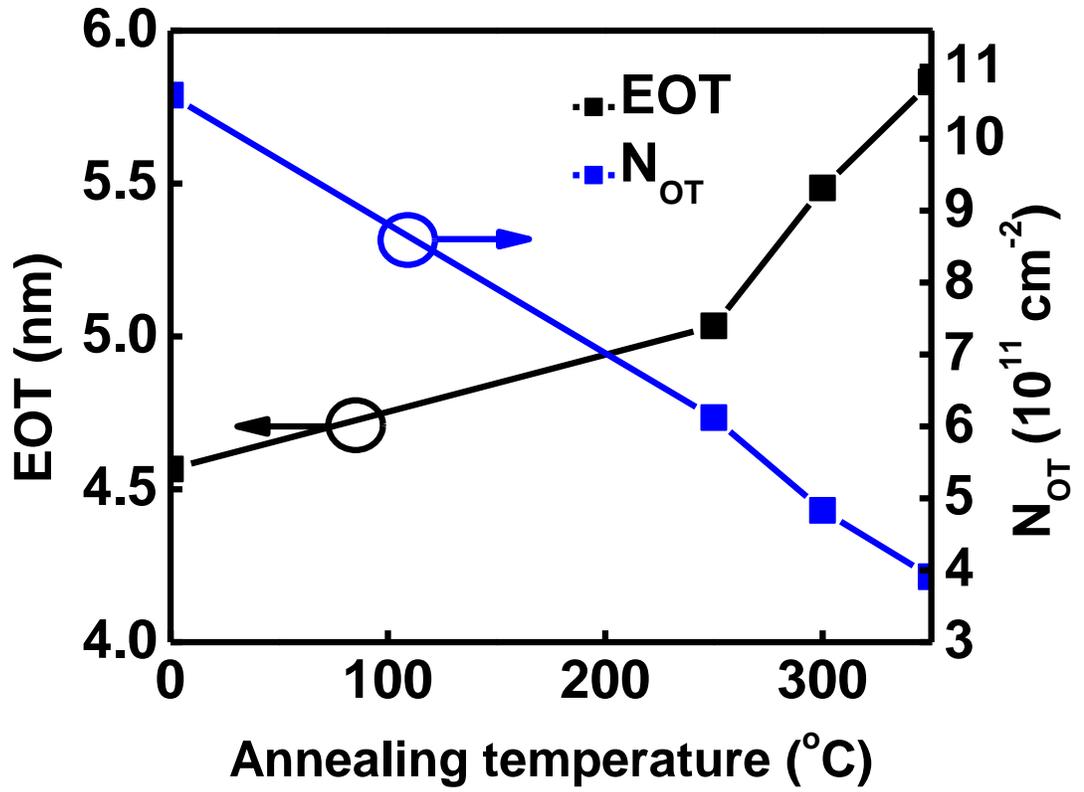
Figure 5.5 *EOT* (black) and *N$_{ot}$* (blue) as a function of annealing temperature of the (100) Ge MOS-C on GaAs via an AlAs buffer architecture.

Fig. 5.6 shows the extracted D_{it} as a function of $E_t - E_i$ for the (100) Ge devices. D_{it} is calculated using the conductance method accounting for surface potential [13-14], which is summarized through eq. 3.21. Typically D_{it} values extracted from 300 K measurements are overestimated due to the temperature-dependent supply of minority carriers to the inversion layer which contribute to higher conductance, and thus lead to an overestimation of D_{it} [8, 11, 15]. However, Fig 5.5 still offers a fair comparison standpoint between the different FGA schemes. Relatively, 300 °C and 350 °C offers lower D_{it} values compared to the pre-anneal and 250 °C case. However, as for the case of 350 °C annealing temperature, there is a slight increase in D_{it}. Thus, it was found that for (100) epi Ge, the FGA scheme of 300 °C, 2 min proved to be the most optimal in suppression of N_{ot} and D_{it}. Although *EOT* significantly increases above 250 °C temperature, the benefits of passivation offered by the 300 °C annealing scheme outweigh the *EOT* cost.
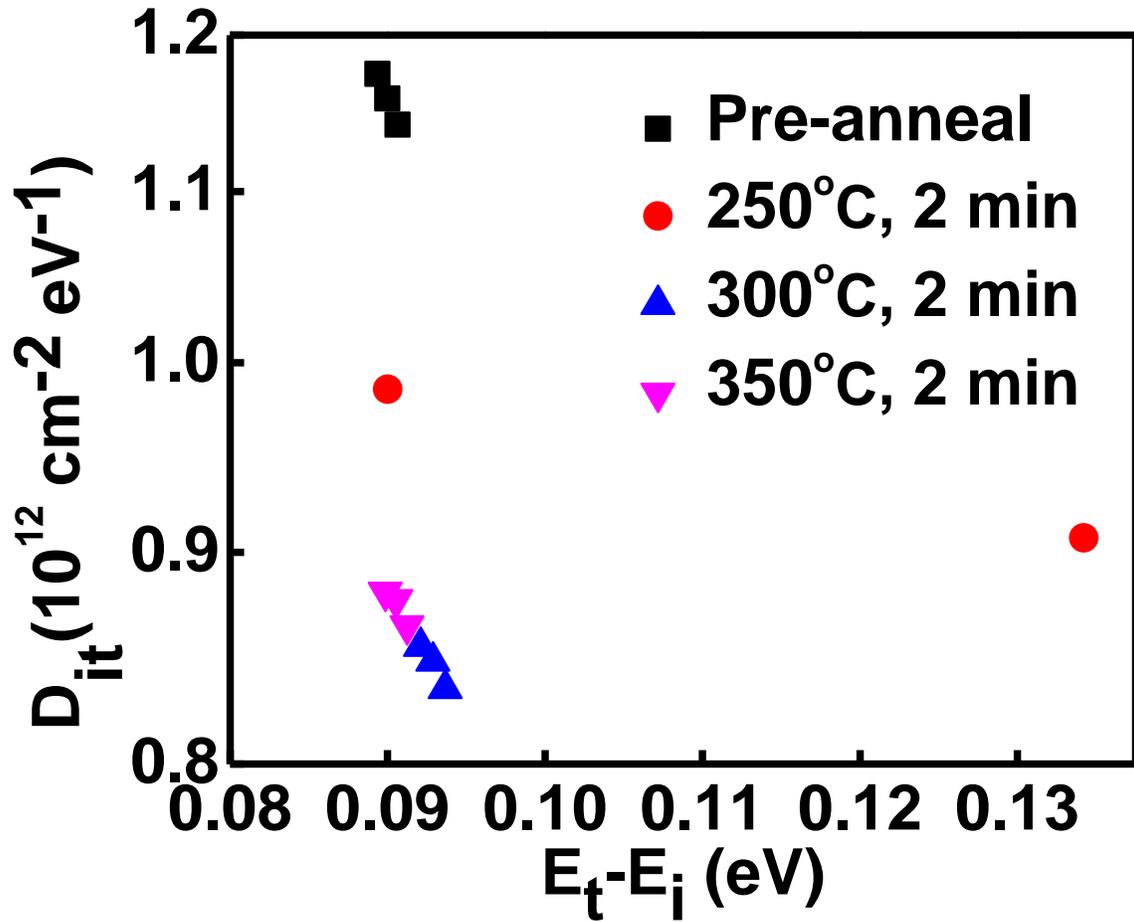
Figure 5.6 *D$_{it}$* as a function of energy of the (100) Ge MOS-C on GaAs via an AlAs buffer architecture for various FGA schemes.

5.4.2 FGA Optimization for (110) Ge MOS Devices

Fig. 5.7a-d below shows the 300 K C-V characteristics of the (110) Ge MOS-C under different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes. Regardless of FGA conditions, the (110) devices demonstrate an unusual inversion behavior at all frequencies, including the 1 MHz frequency, which is typically flat at 300 K measurements for Ge [7, 8]. This inversion behavior is also not typical of (110) devices, which usually resemble (100) C-V curves [2].
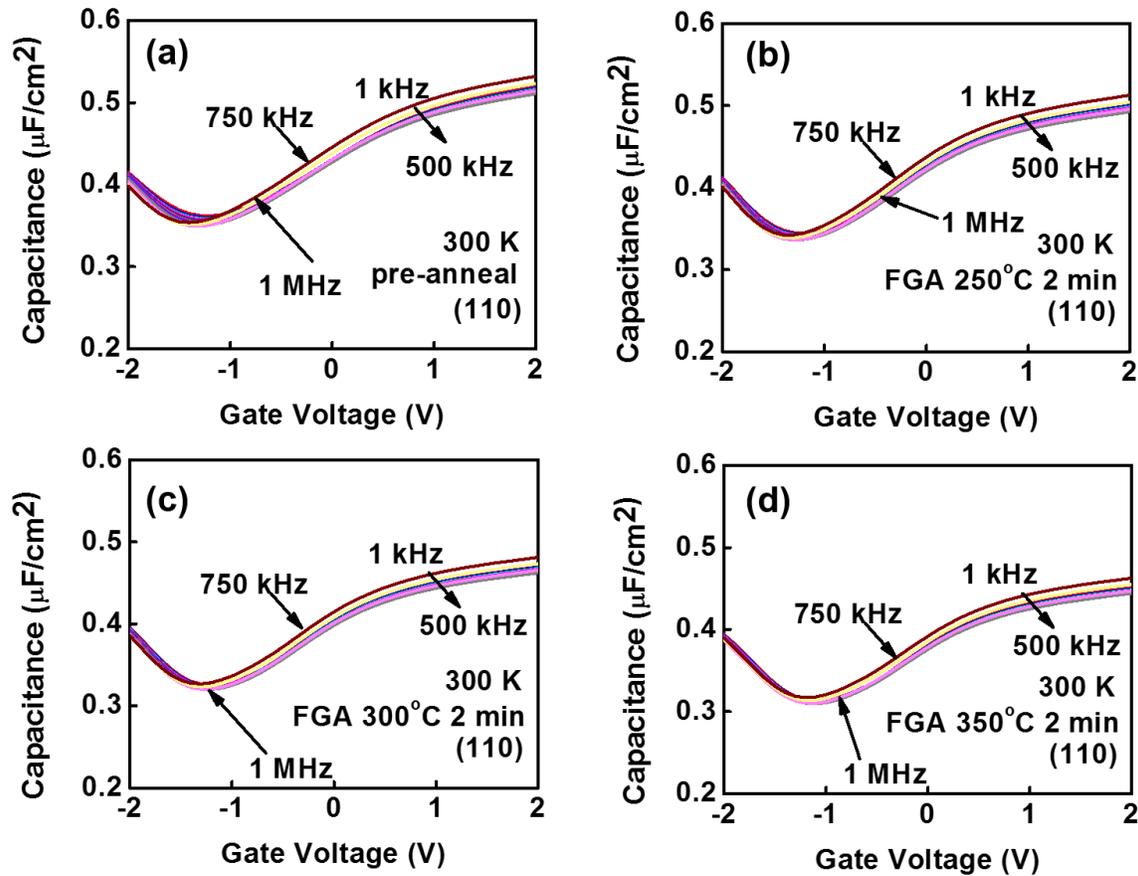

Figure 5.7   C-V characteristics of the (110) Ge MOS-C on GaAs via an AlAs buffer architecture for different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes.

The inversion behavior is suspected to be due to shallow traps distributed throughout the bandgap with fast response times, as they are still responsive even at higher frequencies. This inversion behavior may also be due to the fact that Ge_2O_3 is found to be the dominant species in the thermally grown GeO_x, as demonstrated in Fig. 5.2b. However, more in-depth studies would need to be done to conclude both of these facts. Asides from the unusual inversion behavior, C-V stretch-out is reduced in the inversion regime with increasing thermal budget, demonstrating easy modulation of charge with voltage bias. Reduction of C-V stretch-out with increasing thermal budget is indicative of passivation of interfacial defects throughout the upper half of the bandgap [16].

Fig. 5.8a-d below shows the 300 K conductance contours of the (110) Ge MOS-C under different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes, where all four figures follow the same color scale. The higher magnitude conductance contours (shown in red, orange, and yellow), which typically correlate to inversion behavior [9-11], are shifted up in frequencies, as compared to the higher magnitude conductance contours shown in Fig. 5.4a-d. This shift in position correlates well with the unusual inversion behavior in the (110) C-V graphs, as even at higher frequencies, inversion still occurs. Asides from the shift in higher magnitude conductance contours, the blue region within the weak inversion regime from log(frequency) = 3-4 decreases into gradually into the background purple color, which, similar to the Fig. 5.4a-d for the (100) devices, demonstrates the passivation of D_{it} at the GeO_x/Ge heterointerface.

Fig 5.9 shows *EOT* (marked in black) and N_{ot} (marked in blue) of the (110) Ge devices as a function of FGA temperature. As was the case for the (100) devices, N_{ot} shows a
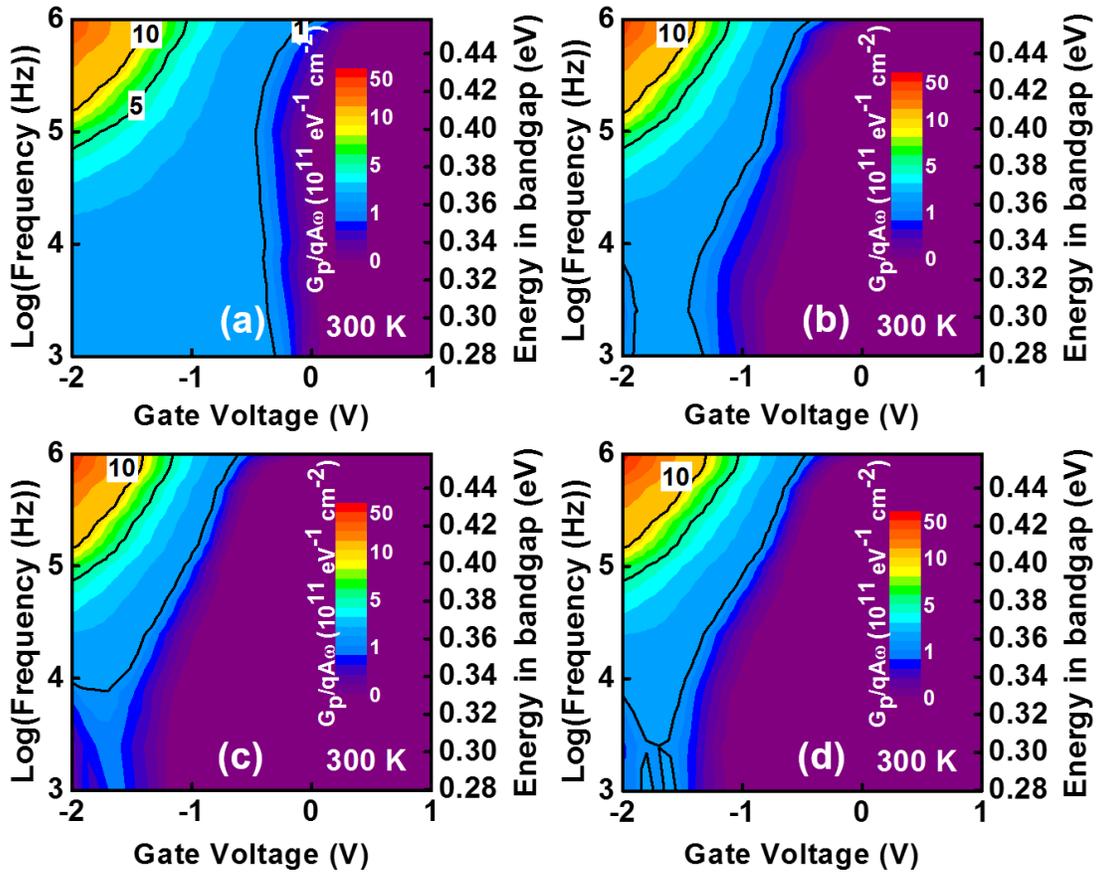
97

Figure 5.8 G_p/ω contours of the (110) Ge MOS-C on GaAs via an AlAs buffer architecture for different FGA conditions: (a) pre-anneal control; (b) 250 °C, two minutes; (c) 300 °C, two minutes; and (d) 350 °C, two minutes.

decline with increasing annealing temperature, although not as linear as for the case of the (100) devices. Again, this demonstrates the proper passivation of trapped oxide charges with increasing annealing temperature, thereby resulting in a decrease in hysteresis. Similar to the (100) case as well, increasing annealing temperature also increases *EOT*, with a significant increase beyond an annealing temperature of 250 °C.
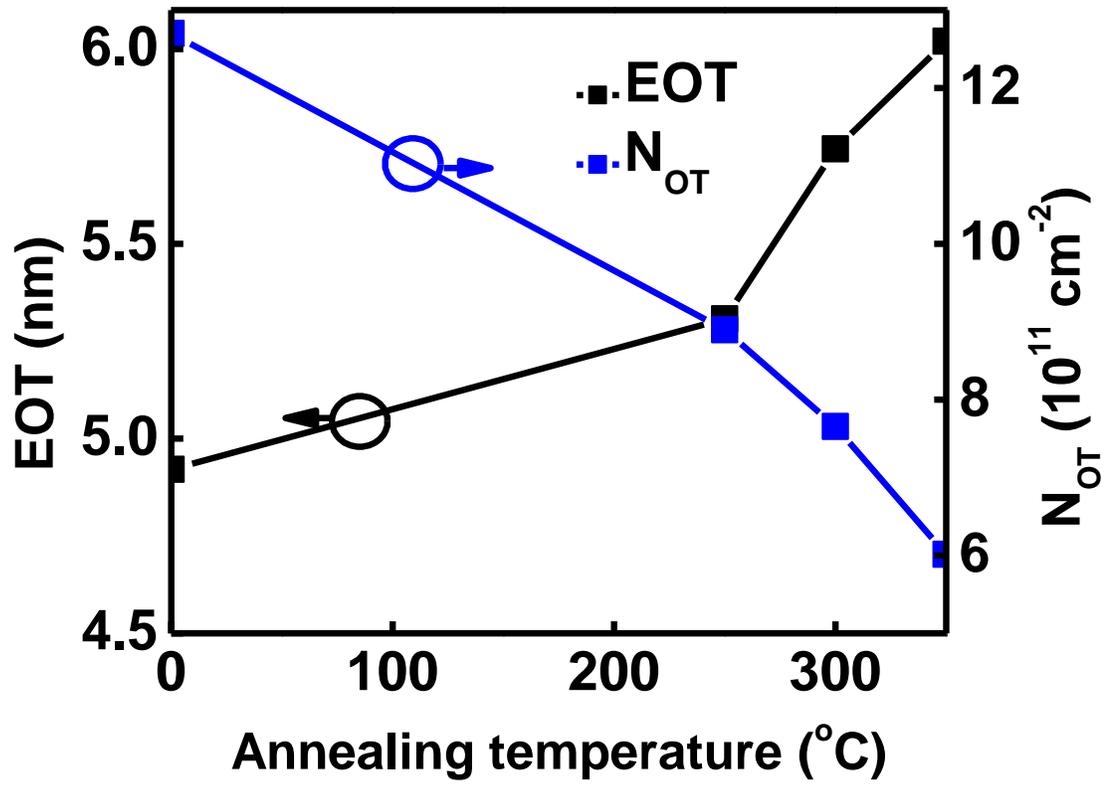
Figure 5.9 *EOT* (black) and *N_{ot}* (blue) as a function of annealing temperature of the (110)

Ge MOS-C on GaAs via an AlAs buffer architecture.

Fig. 5.10 shows the extracted D_{it} as a function of $E_t - E_i$ for the (110) Ge devices. There is a significant reduction of D_{it} with increasing FGA temperature and unlike the (100) Ge device case, D_{it} also slightly reduces with an increase in FGA temperature from 300 °C to 350 °C. However, despite the decrease in D_{it}, the further increase in *EOT* was not warranted. Thus, it was decided that for (110) epi Ge, the FGA scheme of 300 °C, 2 min proved to be the most optimal in suppression of N_{ot} and D_{it} while not significantly sacrificing *EOT*.
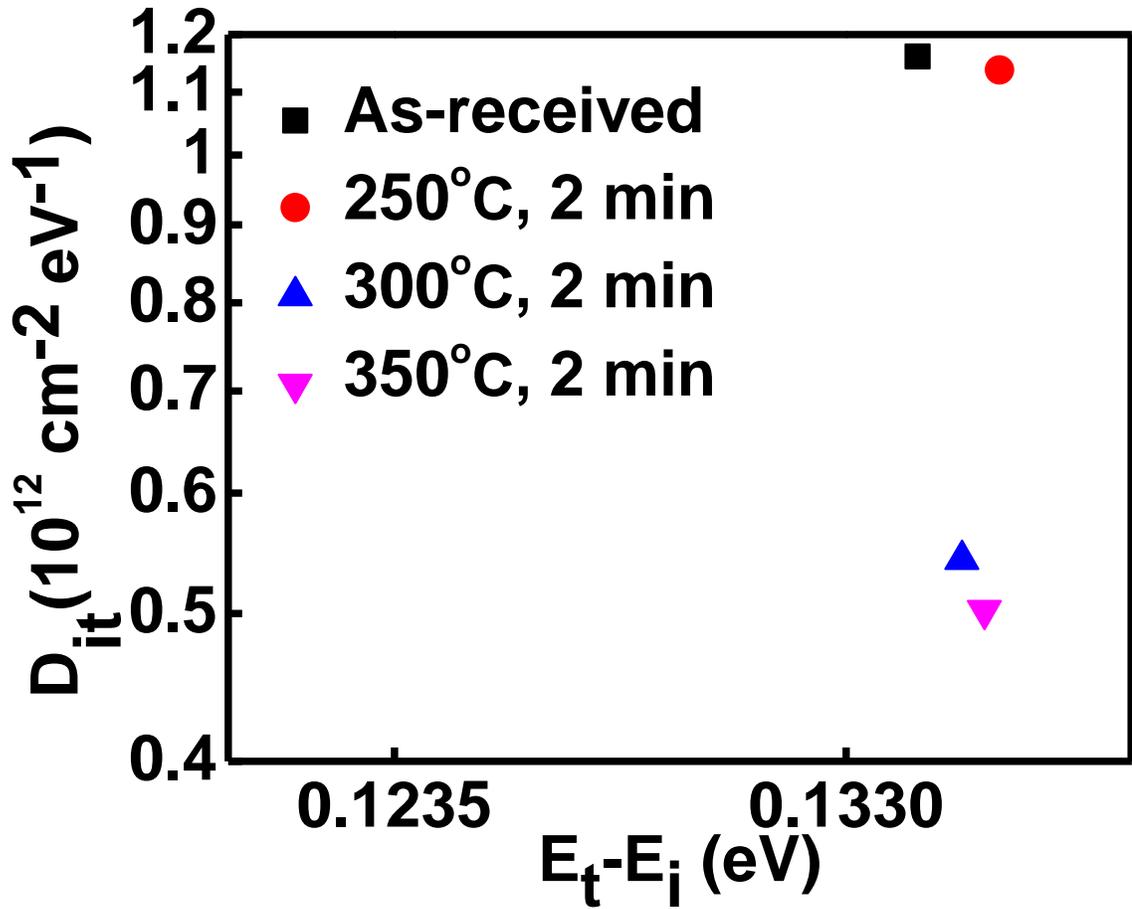
Figure 5.10 D_{it} as a function of energy of the (110) Ge MOS-C on GaAs via an AlAs buffer architecture for various FGA schemes.

5.5   Interface Passivation via Different Annealing Schemes

Following the optimization of the FGA scheme, all three annealing schemes shown in Table 5.1 were applied to both orientations of epitaxial Ge.

5.5.1 Interface Passivation via Different Annealing Schemes for (100) Ge MOS Devices

Fig. 5.11a-l below shows the C-V characteristics of the (100) Ge MOS-C, including pre-anneal C-V curves, as well as curves measured after annealing using the annealing schemes mentioned in Table 5.1. For the pre-anneal case and all three annealing schemes, the samples were measured at temperatures from $78 - 300$ K. The C-V curves shown are from 78, 150, and 300 K measurements in order to demonstrate the suppression of minority carrier response with decreasing temperature [8-9, 11, 15]. As with the case for all annealing schemes used, as shown in Fig. 5.11d-l, C_{max} decreases, which demonstrates an increase in EOT with annealing, although EOT is significantly increased with O_2 PMA. For the pre-anneal case, there is a small frequency dispersion kink in the depletion regime, which is apparent even at the low temperatures of 78 K and 150 K, demonstrating that this hump is not due to weak inversion response from minority carriers [8], but rather D_{it}. This hump is not suppressed with the PDA case, as shown in Fig. 5.11d-e. On the other hand, this small frequency dispersion kink is much better suppressed with FGA, as shown in Fig. 5.11g-h, but is still slightly apparent at 78 K, as shown in Fig. 5.11g. With the O_2 PMA, the kink is completely suppressed, demonstrating effective D_{it} passivation with O_2 PMA. Finally, unlike the case for FGA and O_2 PMA, minority carrier is not completely suppressed at 78 K or 150 K for the PDA case, which suggests that traps with quick response times are still not passivated with PDA.
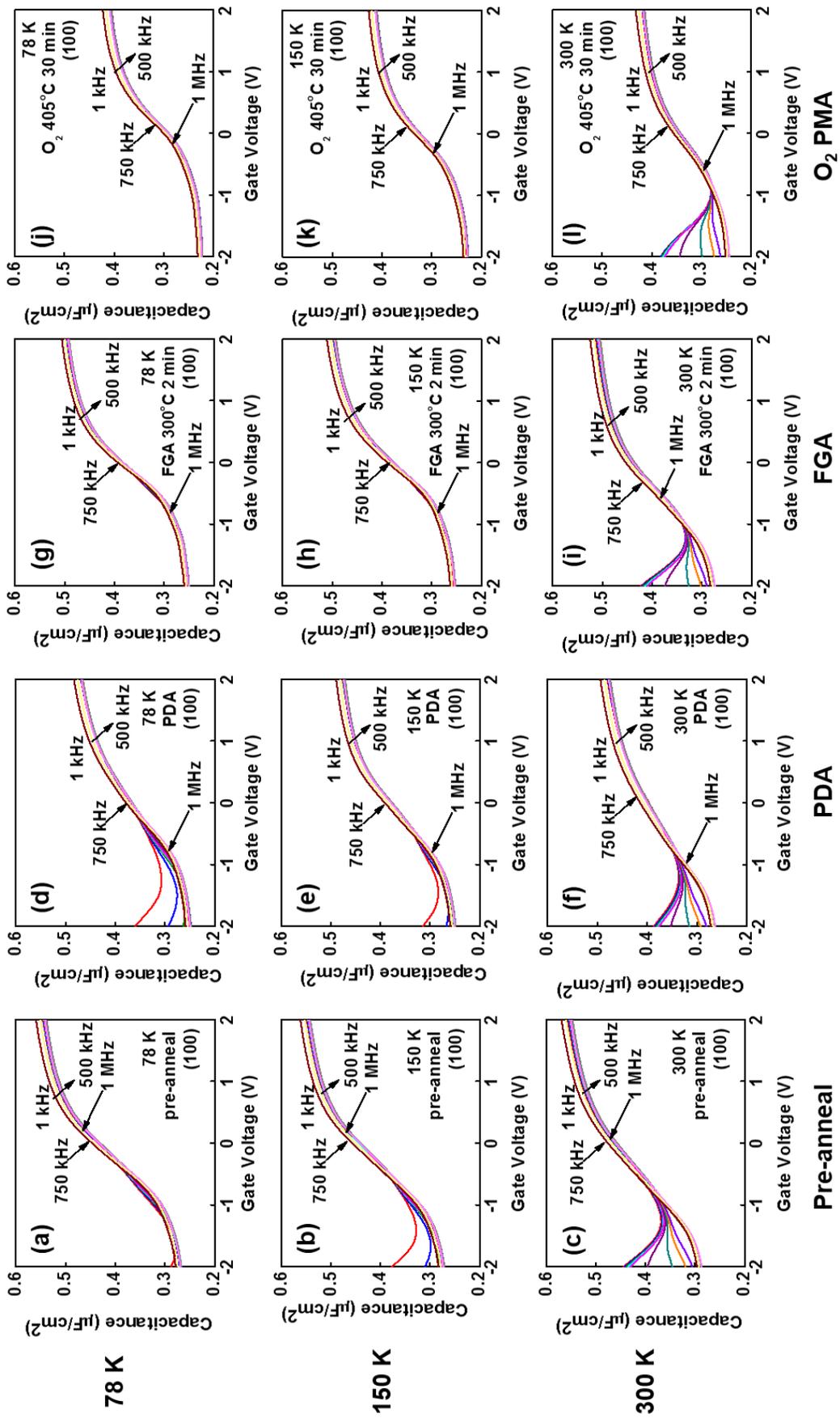
Figure 5.11 C-V characteristics of the (100) Ge MOS-C on GaAs via an AlAs buffer architecture, where (a-c) are C-V curves measured at 78, 150, and 300 K for the pre-anneal case, (d-f) are C-V curves measured at 78, 150, and 300 K for the PDA case, (g-i) are C-V curves measured at 78, 150, and 300 K for the FGA case, and (j-l) are C-V curves measured at 78, 150, and 300 K for the O₂ PMA case.

Fig. 5.12a-l below shows the G_p/ω contours of the (100) Ge MOS-C, including pre-anneal G_p/ω contours, as well as contours measured after annealing using the annealing schemes mentioned in Table 5.1, where the contours followed the same color scale as shown in Fig. 5.12c, 5.12f, 5.12i, and 5.12j. For the pre-anneal case and all three annealing schemes, the samples were measured at temperatures from 78 – 300 K. The G_p/ω contours shown are from measurements at 78, 150, and 300 K in order to separate minority carrier response [8-9, 11, 15] from the Fermi-level tracing to yield a more accurate representation of the devices' *FLE's*. There is significant broadening of the contours at 78, 150, and 300 K for the pre-anneal and PDA case, as shown in Fig. 5.12a-c and Fig. 5.12d-f. In fact, this broadening worsens for the PDA case, as demonstrated in Fig. 5.12d, where the blue region grows in width across voltage bias. This broadening demonstrates that there is an increase in electron trapping due to a large presence of D_{it} [9]. On the other-hand, the contour widths taper significantly for the FGA and O_2 PMA case, as shown in Fig. 5.12g-i and Fig. 5.12j-l. Furthermore, the Fermi-level traces are also steeper for both FGA and O_2 PMA compared to traces in the pre-anneal and PDA case, which demonstrates good modulation of the Fermi-level with respect to gate voltage bias due to the passivation of D_{it} [10]. The Fermi-level contours for both FGA and O_2 PMA become darker in blue color, and for the O_2 PMA case, especially in Fig. 5.12k, the contours begin to disappear into the purple background, which demonstrates very good passivation using these two annealing schemes because of the decrease in conductance magnitude, correlating with a decrease in D_{it} as a result.
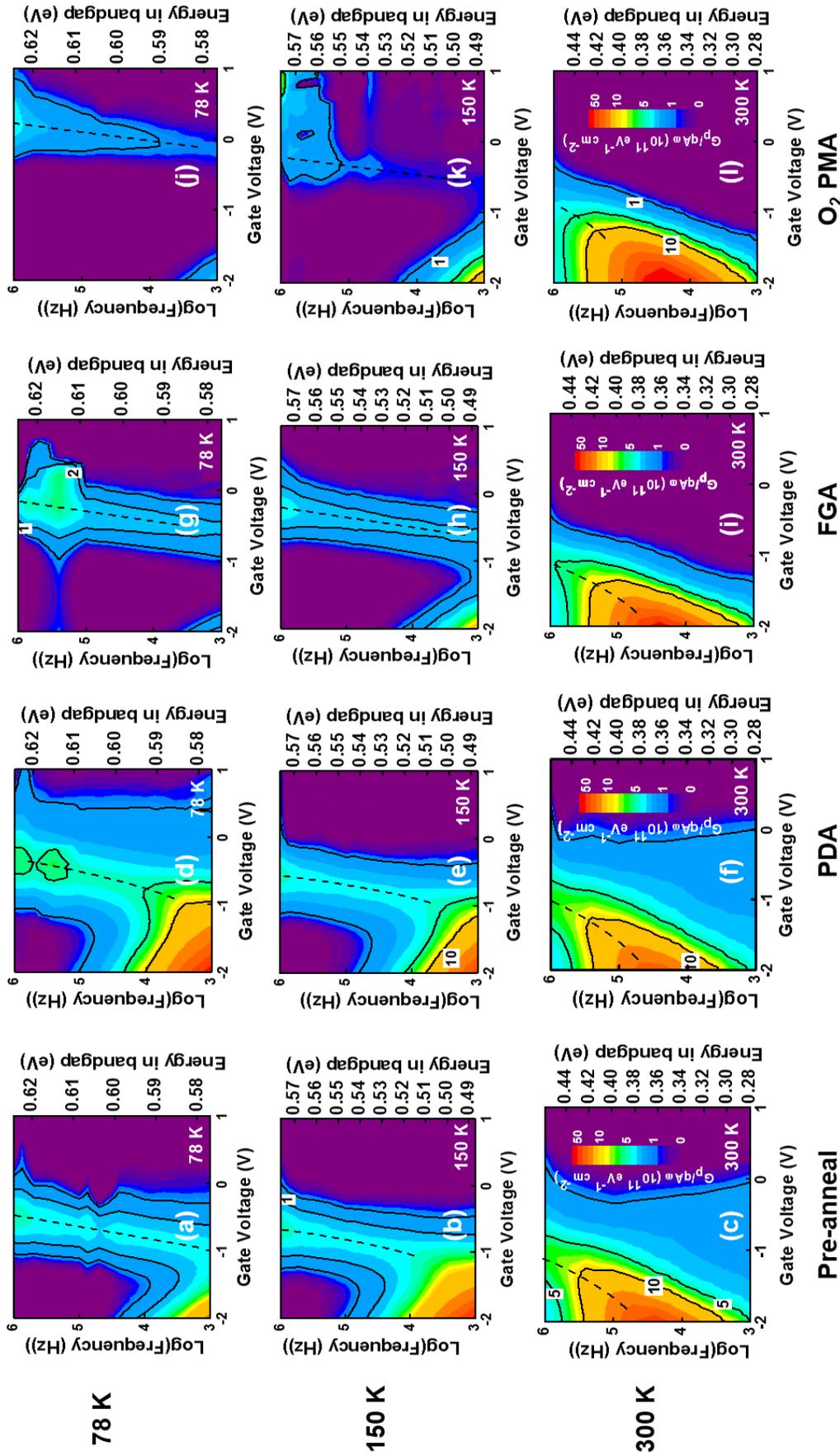
Figure 5.12 G_p/ω contours of the (100) Ge MOS-C on GaAs via an AlAs buffer architecture, where (a-c) are G_p/ω contours measured at 78, 150, and 300 K for the pre-anneal case, (d-f) are G_p/ω contours measured at 78, 150, and 300 K for the PDA case, (g-i) are G_p/ω contours measured at 78, 150, and 300 K for the FGA case, and (j-l) are G_p/ω contours measured at 78, 150, and 300 K for the O2 PMA case.

Table 5.2 summarizes the *EOT* and N_{ot} values extracted from the (100) Ge MOS C-V measurements for the pre-anneal case as well as all the annealing schemes mentioned in Table 5.1. As is expected, *EOT* is highest for the case of O_2 PMA, which correlates well with the significant decrease in C_{max}, as shown in Fig. 5.11(j-l). N_{ot} also decreases with annealing and is best for the case of O_2 PMA, with FGA yielding the second best.

Fig. 5.13 shows D_{it} extracted from the measured devices for the pre-anneal case and all annealing schemes mentioned in Table 5.1. D_{it} is highest across the bandgap for the pre-anneal sample, shown as a black square, and is improved using the different annealing schemes, with the best D_{it} distribution being achieved with the O_2 PMA scheme. The second best D_{it} distribution was achieved using FGA, with a minimum D_{it} value of 3.79×10^{11} cm^{-2} eV^{-1}. The increase in D_{it} closer to mid-gap observed across all samples is due to the temperature-dependent supply of minority carriers to the inversion layer which contribute to higher conductance, and thus lead to an overestimation of D_{it} [8, 11, 15].

Table 5.2 *EOT* and N_{ot} extracted for pre-anneal and the three annealing schemes of the (100) Ge MOS-C devices.

Annealing scheme	*EOT* (nm)	N_{OT} (10^{11} cm^{-2})
Pre-anneal	4.95	8.82
PDA	5.56	8.01
FGA	5.45	6.15
O_2 PMA	6.62	4.81
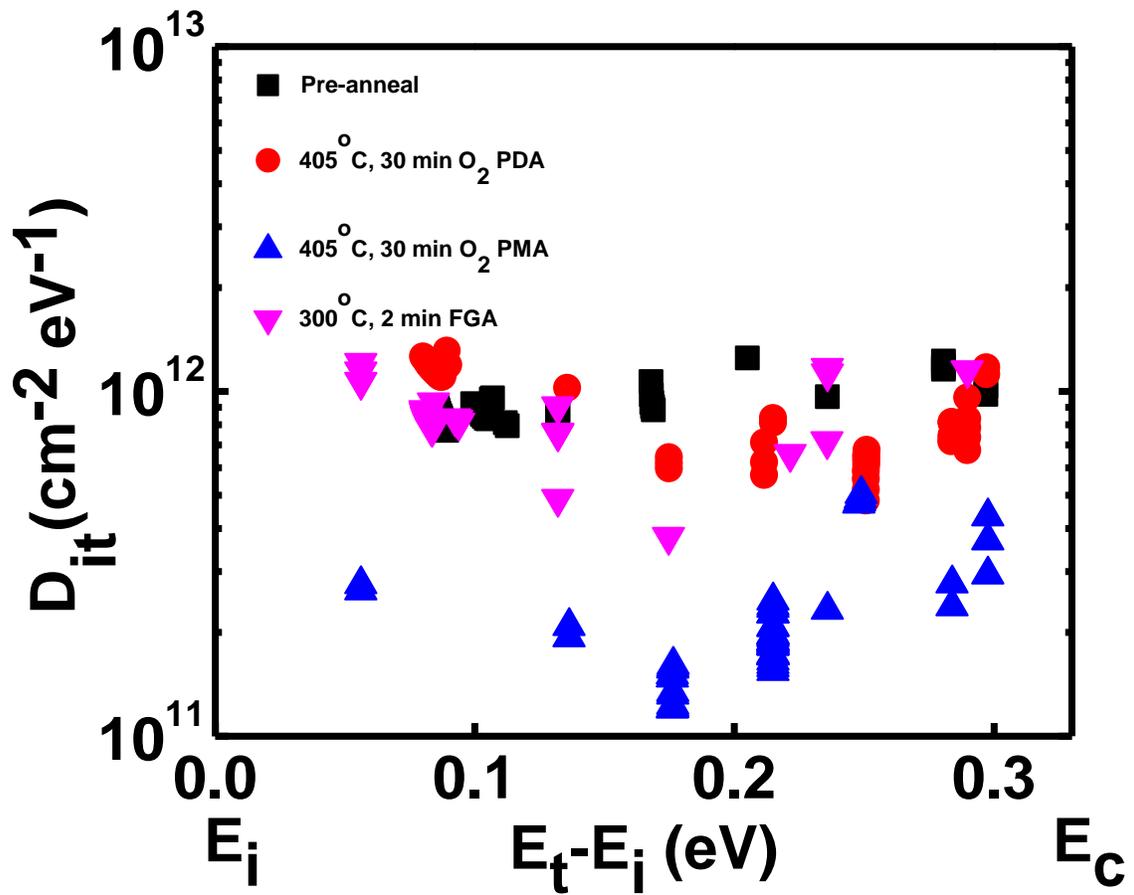
Figure 5.13 D_{it} as a function of energy of the (100) Ge MOS-C on GaAs via an AlAs buffer architecture for various annealing schemes.

5.5.2 Interface Passivation via Different Annealing Schemes for (110) Ge MOS Devices

Fig. 5.14a-l below shows the C-V characteristics of the (110) Ge MOS-C, including pre-anneal C-V curves, as well as curves measured after annealing using the annealing schemes mentioned in Table 5.1. As with the (100) case, all the samples were measured at temperatures from 78 – 300 K. Unlike the (100) case, however, minority carrier was not effectively suppressed except at the higher frequencies of 750 kHz and 1 MHz, despite measuring at the low temperatures of 78 and 150 K. This was also found to be true regardless of annealing scheme. This demonstrates that despite the passivation of D_{it} using the various annealing schemes, there still exists shallow traps with quick response times distributed across the bandgap. Similar to the case for the (100) Ge devices, C_{max} decreases with FGA and O_2 PMA, demonstrating an increase in *EOT* with a significant increase in *EOT* for the case of O_2 PMA. However, unlike the case of the (100) Ge devices, *EOT* *decreases* for the PDA case. More investigation would be necessary to determine why that would be the case. Similar to the (100) case as well is the small frequency dispersion kink in the depletion regime, both apparent for the pre-anneal and the PDA case, as shown in Fig. 5.14a-b and 5.14d-e. Fortunately, the dispersion kink is fully suppressed with FGA and O_2 PMA as shown in Fig. 5.14g-h and Fig. 5.14j-k, demonstrating effective D_{it} passivation with both annealing schemes.
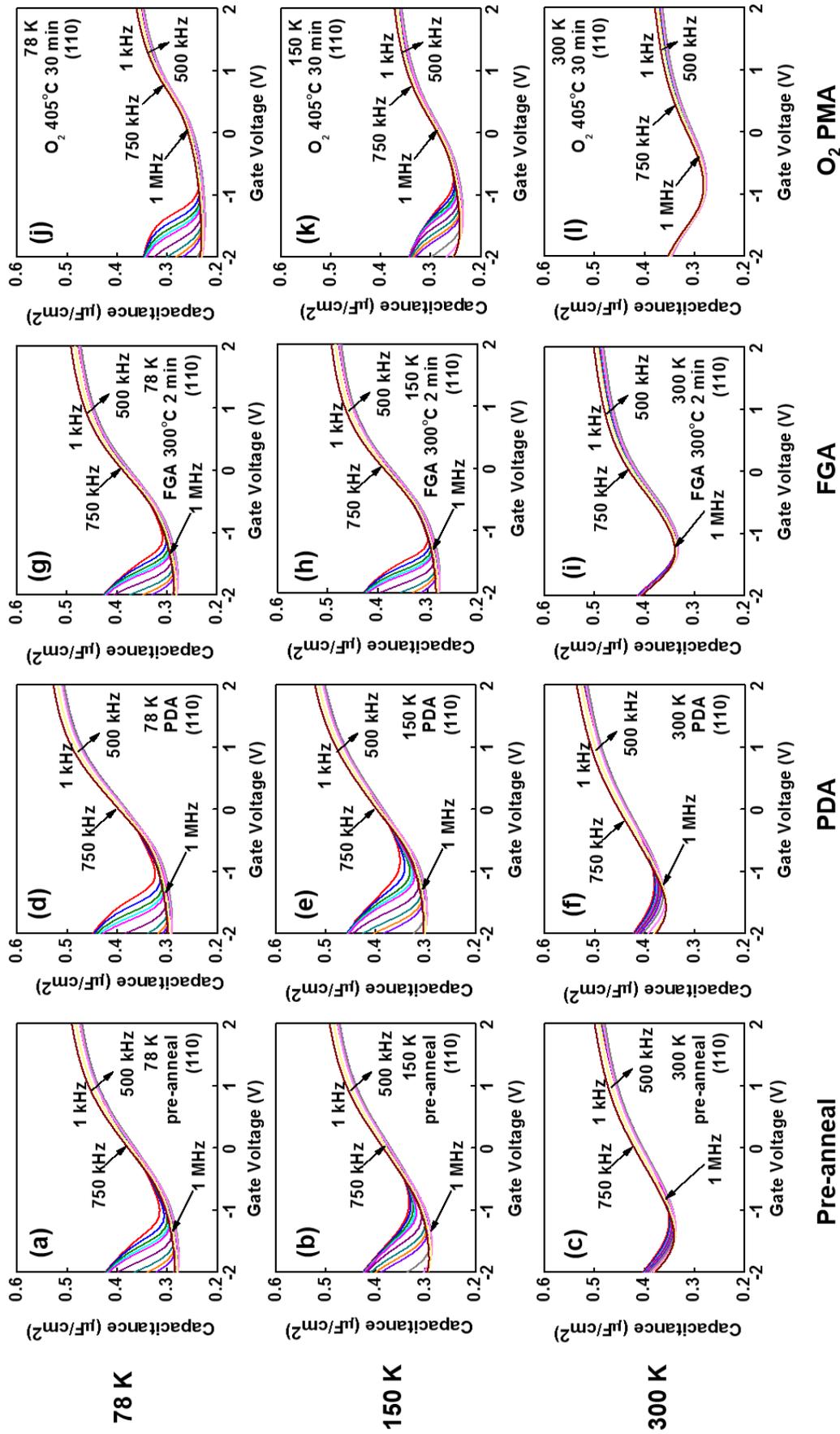
Figure 5.14 C–V characteristics of the (110) Ge MOS-C on GaAs via an AlAs buffer architecture, where (a-c) are C–V curves measured at 78, 150, and 300 K for the pre-anneal case, (d-f) are C–V curves measured at 78, 150, and 300 K for the PDA case, (g-i) are C–V curves measured at 78, 150, and 300 K for the FGA case, and (j-l) are C–V curves measured at 78, 150, and 300 K for the O$_2$ PMA case.

Fig. 5.15a-l below shows the G_p/ω contours of the (110) Ge MOS-C, including pre-anneal and contours measured after annealing using the annealing schemes mentioned in Table 5.1, where the contours followed the same color scale as shown in Fig. 5.15c, 5.15f, 5.15i, and 5.15j. For the pre-anneal case and all three annealing schemes, the samples were measured at temperatures from 78 – 300 K. The G_p/ω contours shown are from measurements at 78, 150, and 300 K with the intention to suppress minority carrier response [8-9, 11, 15] from the Fermi-level tracing to yield a more accurate representation of the devices' *FLE's*. However, unlike the (100) samples, there is still significant minority carrier response, as evidenced in the C-V measurements of the (110) samples. Despite this, the Fermi-level can be traced out at the lower temperatures of 78 and 150 K, as shown in Fig. 5.15a-b, 5.15d-e, 5.15g-h, and 5.15j-k. As with the (100) samples, there is significant broadening of the contours at 78, 150, and 300 K for the pre-anneal and PDA case, as shown in Fig. 5.15a-c and Fig. 5.15d-f, with the broadening worsening for the PDA case. This suggests that for the pre-anneal and PDA case, there is a large presence of D_{it} that results in electron trapping. Furthermore, the Fermi-level trace for both the pre-anneal and PDA case demonstrates arcing at the lower temperatures, which is a sign of poor *FLE* [11]. However, for the FGA and O_2 PMA case, not only do the contours narrow, but the Fermi-level traces become steeper, demonstrating good modulation of the Fermi-level with respect to gate voltage bias due to the passivation of D_{it} [10]. In additionally, the contours for the FGA and O_2 PMA case changes from the lighter blue to the darker blue, demonstrating a decrease in conductance magnitude, which correlates with a decrease in D_{it} with these two annealing schemes.
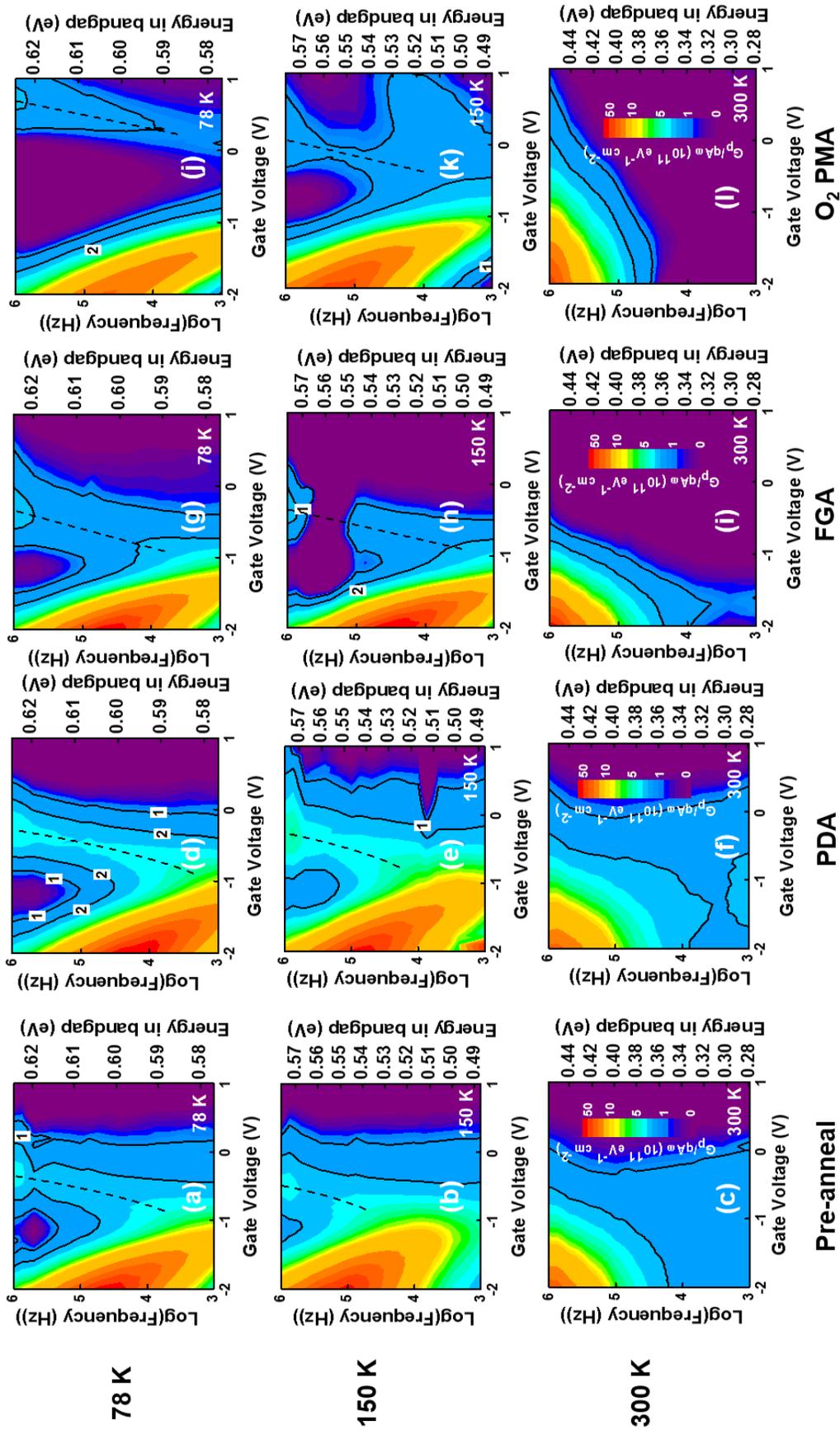
Figure 5.15 G_p/ω contours of the (110) Ge MOS-C on GaAs via an AlAs buffer architecture, where (a-c) are G_p/ω contours measured at 78, 150, and 300 K for the pre-anneal case, (d-f) are G_p/ω contours measured at 78, 150, and 300 K for the PDA case, (g-i) are G_p/ω contours measured at 78, 150, and 300 K for the FGA case, and (j-l) are G_p/ω contours measured at 78, 150, and 300 K for the O$_2$ PMA case.

Table 5.3 summarizes the EOT and N_{ot} values extracted from the (110) Ge MOS C-V measurements for the pre-anneal and all three annealing schemes. As with the (100) samples, EOT is highest for the case of O_2 PMA, correlating well with the significant decrease in C_{max}, as shown in Fig. 5.14j-l. N_{ot} also decreases with annealing and is best for the case of O_2 PMA, with FGA yielding the second best. As mentioned previously, the PDA sample demonstrates the lowest EOT, but also the highest N_{ot}.

Fig. 5.16 shows D_{it} extracted from the measured devices for the pre-anneal case and all annealing schemes mentioned in Table 5.1. D_{it} is highest across the bandgap for the pre-anneal sample, shown as a black square, and is improved using the different annealing schemes, with the best D_{it} distribution being achieved with the O_2 PMA scheme. The second best D_{it} distribution was achieved using FGA, with a minimum D_{it} value of 3.84×10^{11} cm^{-2} eV^{-1}. The increase in D_{it} closer to mid-gap observed across all samples is due to the temperature-dependent supply of minority carriers to the inversion layer which contribute to higher conductance, and thus lead to an overestimation of D_{it} [8, 11, 15].

Table 5.3 EOT and N_{ot} extracted for pre-anneal and the three annealing schemes of the (110) Ge MOS-C devices.

Annealing scheme	EOT (nm)	N_{OT} (10^{11} cm^{-2})
Pre-anneal	5.06	9.68
PDA	4.75	11.4
FGA	5.51	7.74
O_2 PMA	6.51	5.17
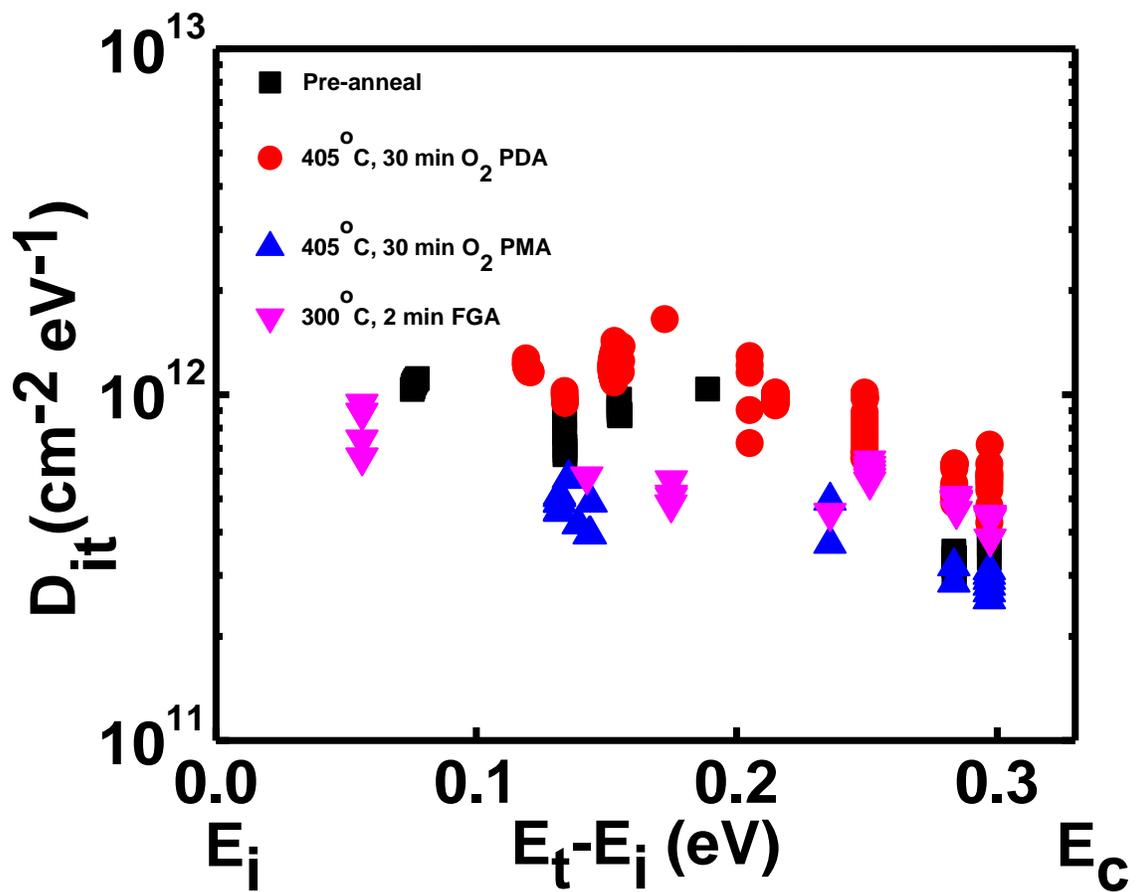
Figure 5.16 *D$_{it}$* as a function of energy of the (110) Ge MOS-C on GaAs via an AlAs buffer architecture for various annealing schemes.

References

[1]   J. Kavalieros *et al.*, "Tri-gate transistor architecture with high-k gate dielectrics, metal gates and strain engineering," in *VLSI Symp. Tech. Dig.*, 2006, pp. 50–51.

[2]   T. Sasada *et al.*, "Surface orientation dependence of interface properties of GeO2/Ge metal-oxide-semiconductor structures fabricated by thermal oxidation," *J. Appl. Phys.*, vol. 106, no. 7, pp. 073716-1–073716-7, Oct. 2009.

[3]   S. Gaudet *et al.*, "Thin film reaction of transition metals with germanium," *J. Vac. Sci. Technol. A.*, vol. 24, no. 3, pp. 474–485, May 2006.

[4]   K. Gallacher, P. Velha, D. J. Paul, I. MacLaren, M. Myronov, and D. R. Leadley, "Ohmic contacts to n-type germanium with low specific contact resistivity," *Appl. Phys. Lett.*, vol. 100, no. 2, p. 022113, Jan. 2012.

[5]   J. Lin, "Low resistance contacts to n-type germanium," Ph.D. dissertation, Dept. Electr. Eng., Stanford Univ., Stanford, CA, USA, Jun. 2013.

[6]   F. Bellenger *et al.*, "Passivation of Ge (100)/GeO2/high-k gate stacks using thermal oxide treatments," *J. Electrochem. Soc.*, vol. 155, no. 2, pp. G33–G38, Feb. 2008.

[7]   D. K. Schroder, Semiconductor Material and Device Characterization, 3rd ed. Hoboken, NJ, USA: Wiley, 2006.

[8]   K. Martens *et al.*, "On the correct extraction of interface trap density of MOS devices with high-mobility semiconductor substrates," IEEE Trans. Electron Devices, vol. 55, no. 2, pp. 547–556, Feb. 2008.

[9]   D. Lin *et al.*, "Enabling the high-performance InGaAs/Ge CMOS: A common gate stack solution," in *IEDM Tech. Dig.*, Dec. 2009, pp. 327–330.

[10]   H. C. Lin *et al*., "The Fermi level efficiency method and its applications on high interface trap density oxide-semiconductor interfaces," *Appl. Phys. Lett.*, vol. 94, no. 15, Apr. 2009.

[11]   P.D. Nguyen *et al.*, "Heteroepitaxial Ge MOS Devices on Si Using Composite AlAs/GaAs Buffer," *IEEE J. Electron Devices Soc.*, vol. 3, no. 4, pp. 341-348, July 2015.

[12]   J. Maserjian, G. Peterson, and C. Svensson, "Saturation capacitance of thin oxide MOS structures and the effective surface state density of silicon," *Solid-State Electron.*, vol. 17, pp. 335–339, Apr. 1974.

[13]   J. R. Brews, "Rapid interface parametrization using a single MOS conductance curve," *Solid State Electron.*, vol. 26, no. 8, pp. 711–716, Aug. 1983.

[14]   E. H. Nicollian and J. R. Brews, *MOS Physics and Technology*. Hoboken, NJ, USA: Wiley, 2003.

[15]   D. Kuzum, "Interface-engineered Ge MOSFETs for future high performance CMOS applications," Ph.D. dissertation, Dept. Electr. Eng., Stanford Univ., Stanford, CA, USA, Dec. 2009.

[16]   C.-W. Cheng, G. Apostolopoulos, and E. A. Fitzgerald, "The effect of interface processing on the distribution of interfacial defect states and the C-V characteristics of III-V metal-oxide-semiconductor field effect transistors," *J. Appl. Phys.*, vol. 102, no. 2, Jan. 2011.

Chapter 6 – Work Function Tuning for Different Crystallographically Oriented Epitaxial Ge MOS Devices

The replacement of polysilicon gates with metal gates for MOSFETs during the 45 nm node has brought about a large interest in metal and work function engineering MOSFETs [1]. For nMOSFETs and pMOSFETs, two types of gate metals are typically used to accommodate for the difference between φ_s (which differs for n- and p-type channels) and φ_m, or, in other words φ_{ms}. φ_{ms} has a direct impact on V_{FB}, which in turn affects threshold voltage V_{TH}. However, the ability to tune V_{FB} and therefore V_{TH} is limited by Fermi-level pinning (FLP). FLP results from the formation of a dipole layer at the interface, which is caused by both interfacial defects [2] as well as metal-induced gap states (MIGS) from metal charge transfer [3]. The influence of the dipole layer depends on the difference between φ_m and the charge neutrality level φ_{CNL} in the oxide, as shown in Fig. 1, where $E_{F,m}$ is the Fermi-level energy of the metal and E_{CNL} is the CNL energy level.
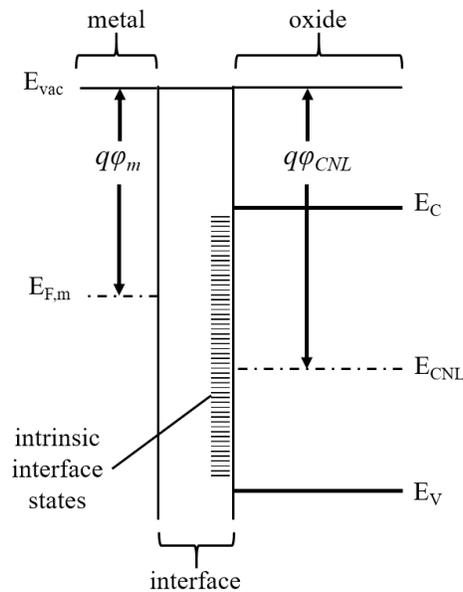
Figure 6.1 Energy band diagram of a metal which is in contact with an oxide.

The interface dipole that forms alters the band alignment, forcing $E_{F,m}$ to bend towards E_{CNL}, thereby altering φ_m and producing $\varphi_{m,eff}$, the effective metal work function. $\varphi_{m,eff}$ can be expressed by the following equation [4]:

$$\varphi_{m,eff} = \varphi_{CNL} + S(\varphi_m - \varphi_{CNL}) \qquad (6.1)$$

where S is known as the slope or pinning parameter. S ranges from zero (full pinning) to unity (no pinning). If MOS-C's implementing different gate metals are fabricated from the same material stack, variations that alter the calculation of V_{FB} between devices such as substrate doping or defect density are now negligible. Thus, the difference in V_{FB}, ΔV_{FB}, that occurs due to the use of different gate metals is now equivalent to the difference in $\varphi_{m,eff}$, $\Delta \varphi_{m,eff}$. Using eq. 6.1, it can be shown that:

$$\Delta V_{FB} = \Delta \varphi_{m,eff} = S\Delta \varphi_m \qquad (6.2)$$

where S is known as the slope parameter or the pinning parameter. An S of zero means full pinning and an S of unity means no pinning. In this chapter, the integration of two different gate metals to form MOS-Cs on the (100) and (110) crystallographically oriented epitaxial Ge heterostructures from *Chapter 4* is studied and the amount of pinning is analyzed in detail.

6.1 MBE Growth of (100) Ge/AlAs and (110) Ge/AlAs Heterostructures on GaAs Substrate

The two heterostructures used in this chapter consists of (100) and (110) Ge on AlAs buffer layer integrated on GaAs substrate were grown *in-situ* by solid source MBE utilizing separate III-V and Ge growth chambers connected *via* an ultra-high vacuum transfer chamber, as discussed in *Chapter 5*.

For the (100) oriented epitaxial Ge growth, a 250 nm GaAs nucleation layer was grown on (100) GaAs substrates that were 6° offcut towards the [110] direction, followed by the large bandgap 170 nm AlAs layer, both of which were grown at a growth temperature of ≥ 600 °C. The undoped 270 nm Ge layer growth was grown at 400 °C at a low growth rate of 0.067 Å/s. in a separate Ge chamber. For the (110) oriented epitaxial Ge growth, a 250 nm GaAs epitaxial layer was grown on (110) GaAs substrates. The growth of the 170 nm AlAs layer and undoped 270 nm Ge layer was similar to the (100) oriented sample. More details on the growth of both heterostructures was discussed in *Chapter 4*.

6.2 Fabrication of MOS-Cs on Different Crystallographically Oriented Epitaxial Ge Material Stacks

The N-type MOS-Cs were fabricated on both the (100) and (110) crystallographically oriented epitaxial Ge material stacks. Fabrication began with a standard degrease followed by dipping the samples into dilute (1:10) hydrofluoric acid to remove the native oxide. A native GeO_x interfacial passivating layer was formed via thermal oxidation at 450 °C for 40 minutes, followed by deposition of 4 nm Al_2O_3 gate oxide at 250 °C. Two different gate metals were used: (1) 0.8 nm TiN/100 nm Al for one set of (100) and (110) MOS-C devices and (2) 60 nm Au/40 nm Pt/0.4 nm TiN was used for the other set of (100) and (110) MOS-C devices. 0.8 nm TiN/100 nm Al/10 nm Ti/30 nm Ni was used as the ohmic contact for both sets of devices. All metals were deposited using a Kurt J. Lesker PVD250 electron beam deposition chamber. Following metallization, all devices were annealed at 300 °C for 2 min in a forming gas ambient (N_2:H_2 95%:5%). More details on the MOS-C fabrication process flow and the optimization of the forming gas annealing scheme were discussed in *Chapter 4*. As mentioned in *Chapter 4*, ~3.2 nm was estimated for the

thickness of GeO$_x$ for the (100) sample [5], whereas >3.2 nm was estimated for the thickness of GeO$_x$ for the (110) sample due to the higher thermal oxidation rates for (110) Ge [6].

The composition of the Al metal, which was deposited on HfO$_2$ integrated on Ge/GaAs heterostructures was investigated using a PHI Quantera SXM XPS system with a monochromatic Al-Ka (energy of 1486.7 eV) x-ray source. The Al 2p binding energy spectra was collected with a pass energy of 26 eV and an exit angle of 45u. The binding energy was corrected by adjusting the carbon 1 s CL peak position to 285.0 eV for each sample surface. Curve fitting was performed by CasaXPS 2.3.14 using a Lorentzian convolution with a Shirley-type background.

To evaluate the electrical characteristics of the fabricated MOS-C devices, low- and room-temperature multi-frequency C-V and G-V measurements were performed using an HP4284A precision LCR meter with frequencies ranging from 1 kHz to 1 MHz. Accurate measurements were obtained with the removal of series resistance, as discussed in [7].

6.3 Work Function Tunability of MOS-Cs Implementing (100) and (110) Epitaxial Ge Material Stacks

Fig. 6.2a-c shows the C-V measurements at 78 K, 150 K, and 300 K of the (100) Ge MOS-C's using 0.8 nm TiN/100 nm Al as its gate metal while Fig. 6.2d-f shows the C-V measurements at 78 K, 150 K, and 300 K of the (100) Ge MOS-C's using 60 nm Au/40 nm/0.4 nm TiN as its gate metal. Asides from the large noticeable shift in V_{FB} observed at all temperatures, there is also a difference in C_{max}. This C_{max} difference is not due to oxide thickness variation because both samples underwent the same fabrication process, but rather, due to the oxidation of the Al metal, as will be explained in detail in a later section.
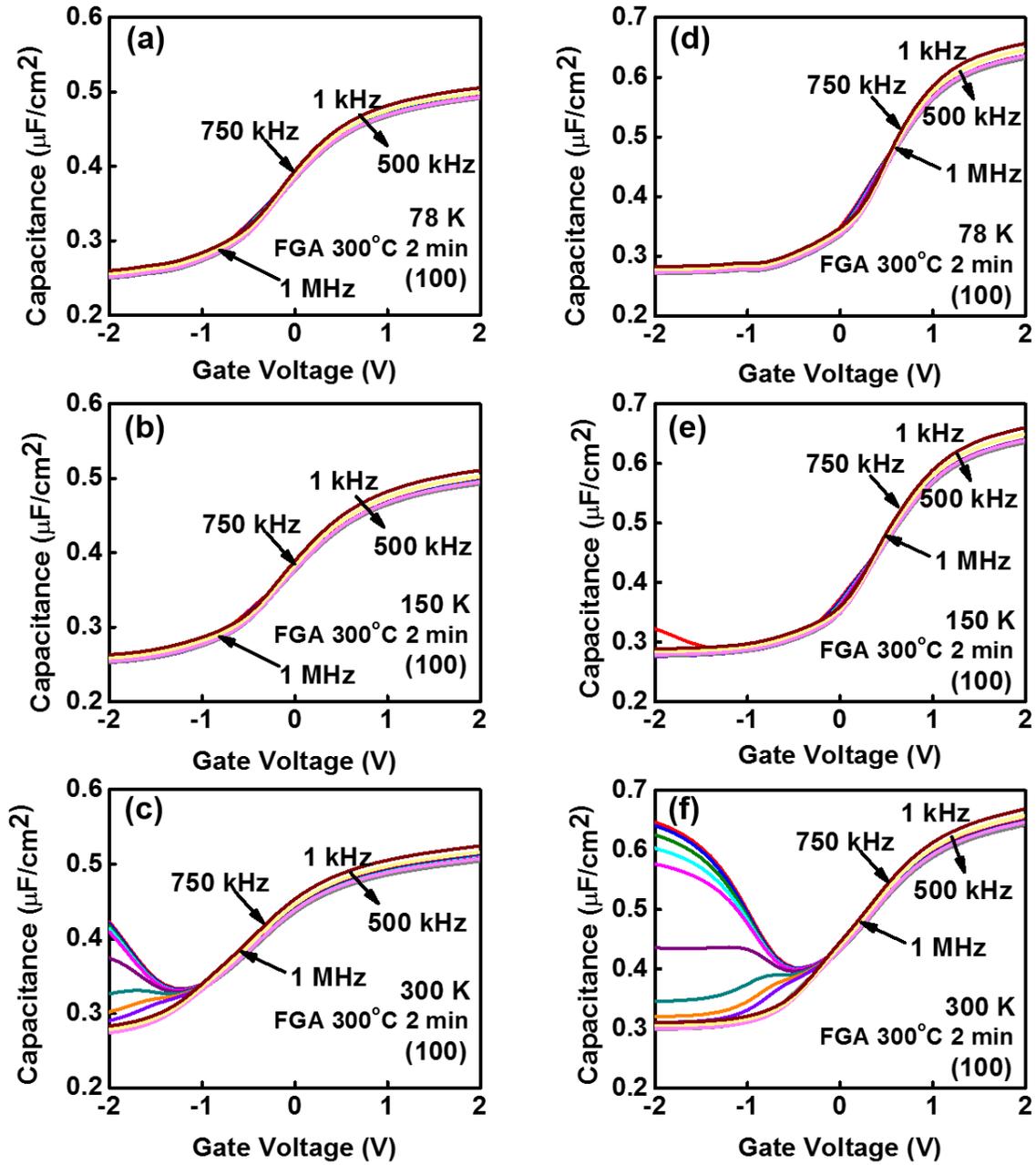
Figure 6.2 C-V characteristics of the (100) Ge MOS-C's on GaAs via an AlAs buffer architecture using two different gate metals: devices with 0.8 nm TiN/100 nm Al measured at (a) 78 K, (b) 150 K, and (c) 300 K. Devices with 60 nm Au/40 nm Pt/0.4 nm TiN measured at (d) 78 K, (e) 150 K, and (f) 300 K.

Fig 6.3 shows the comparison of 1 MHz curves of the two different gated (100) samples at 300 K, where the red line denotes the samples with 0.8 nm TiN/100 nm Al as its gate metal and the blue line denotes the samples with 60 nm Au/40 nm Pt/0.4 nm TiN as its gate metal. The curves were normalized to their respective C_{ox} values. ΔV_{FB} was measured to be 700 mV. With the work function of Al being 4.28 eV [8] and Pt being 5.65 eV [8] and using eq. 6.2, an S value of 0.511 was calculated. A comparison of the D_{it} values of the two different-gated devices is shown in Fig. 6.4, where D_{it} was calculated using the conductance method accounting for surface potential [9-10], which is summarized through
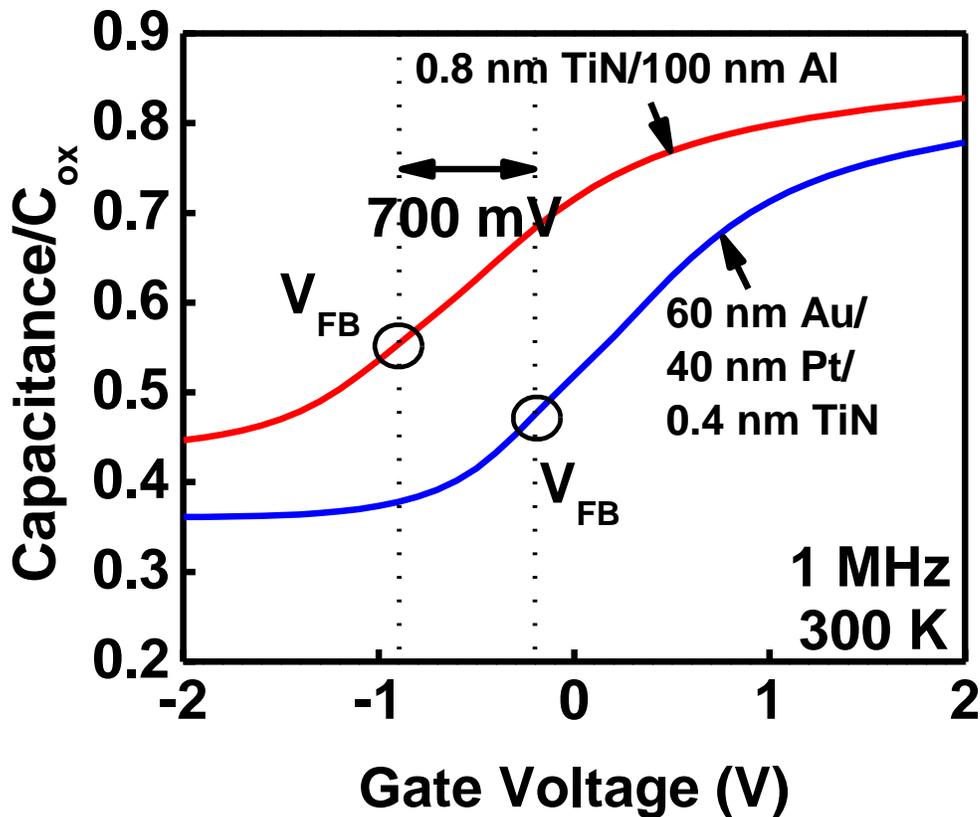
Figure 6.3 Comparison of 1 MHz C-V curves measured at 300 K of (100) Ge MOS-C's on GaAs via an AlAs buffer architecture. Both curves are normalized to their respective C_{ox}.
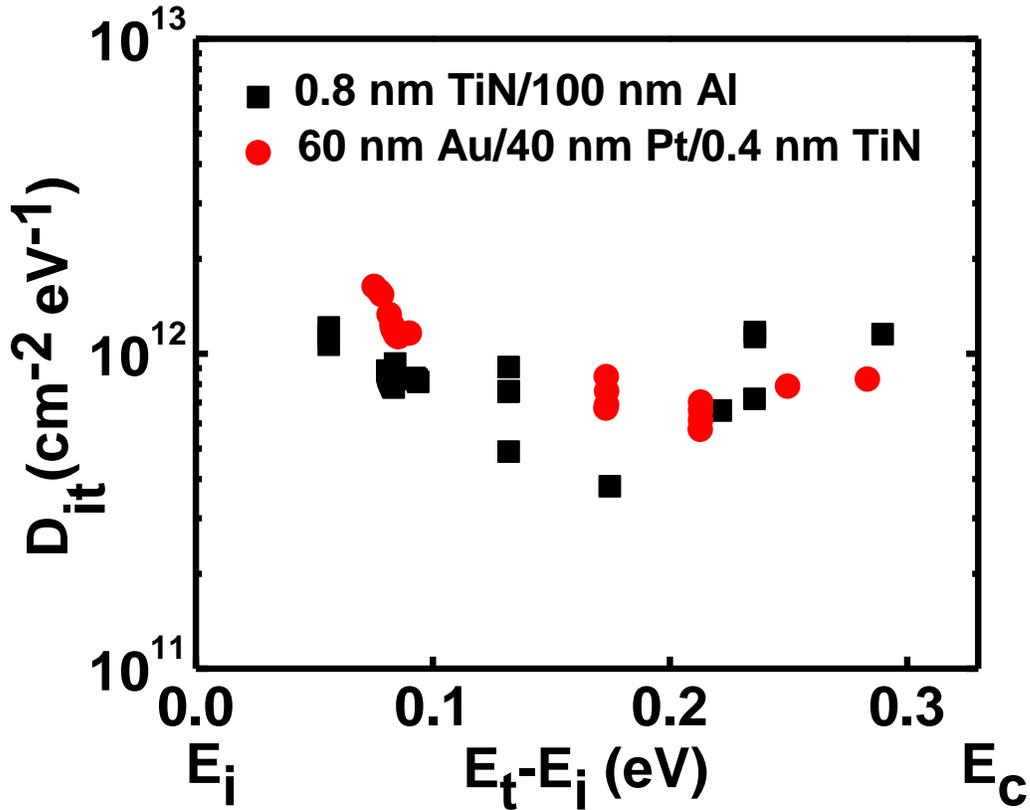
Figure 6.4 D_{it} as a function of energy of two different gated (100) Ge MOS-C's on GaAs via an AlAs buffer architecture.

eq. 3.21. Although the measured D_{it} levels are comparable to those found in some state-of-the-art devices [12], further reduction of D_{it} may allow for further Fermi-level unpinning and therefore, an improvement of S.

Fig. 6.5a-c shows the C-V measurements at 78 K, 150 K, and 300 K of the (100) Ge MOS-C's using 0.8 nm TiN/100 nm Al as its gate metal while Fig. 6.5d-f shows the C-V measurements at 78 K, 150 K, and 300 K of the (100) Ge MOS-C's using 60 nm Au/40 nm/0.4 nm TiN as its gate metal. Similar to the (100) C-V curves shown in Fig. 6.2a-f,
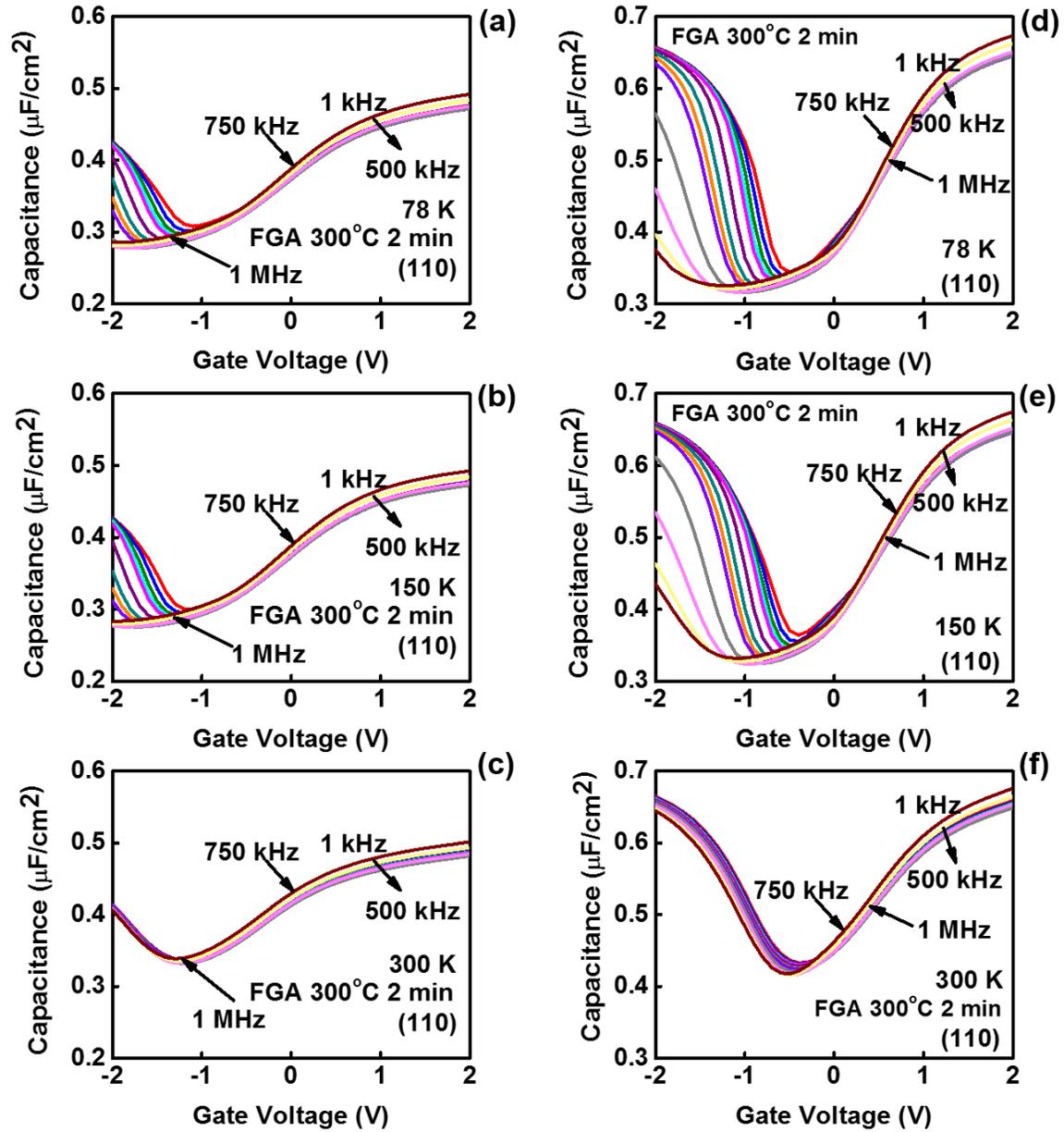
Figure 6.5 C-V characteristics of the (110) Ge MOS-C's on GaAs via an AlAs buffer architecture using two different gate metals: devices with 0.8 nm TiN/100 nm Al measured at (a) 78 K, (b) 150 K, and (c) 300 K. Devices with 60 nm Au/40 nm Pt/0.4 nm TiN measured at (d) 78 K, (e) 150 K, and (f) 300 K.

there is a large noticeable shift in V_{FB} observed at all temperatures as well as a difference in C_{max} metal. As an aside, the (110) epitaxial Ge MOS-C's demonstrate an unusual inversion behavior at all frequencies, including the 1 MHz frequency, which is typically flat for Ge [7, 11], as was demonstrated in **Chapter 5**. Previously in **Chapter 5**, it was believed that some of the minority carriers were suppressed at the lower temperatures of 78 K and 150 K due to the flattening of the 750 kHz and 1 MHz curves, as shown in Fig. 6.5a-b. However, with the large shift in V_{FB} using the Pt gate metal, it is revealed that the minority carriers are actually not suppressed, as shown in Fig. 6.5d-e, although it can be observed that the curves begin to not invert as quickly, demonstrating that the low-temperature does have some minor suppression abilities. As mentioned in **Chapter 5**, the inversion behavior is suspected to be due to shallow traps distributed throughout the bandgap with fast response times, as they are still responsive even at higher frequencies and could also be due the dominance of Ge_2O_3 in the thermally grown GeO_x. However, more in-depth studies would need to be done to conclude both of these facts.

Fig 6.6 shows the comparison of 1 MHz curves of the two different gated (110) samples at 300 K, where the red line denotes the samples with 0.8 nm TiN/100 nm Al as its gate metal and the blue line denotes the samples with 60 nm Au/40 nm Pt/0.4 nm TiN as its gate metal. The curves were normalized to their respective C_{ox} values. ΔV_{FB} was measured to be 770 mV. Again, using eq. 6.2, an S value of 0.564 was calculated, slightly higher than the S value of 0.511, as was calculated for the (100) MOS-C's. Fig. 6.7 shows a comparison of D_{it} values between the different-gated devices, again showing that the D_{it} values are comparable to state-of-the-art devices [12]. Similar to (100), further improvement of D_{it} may help further unpin the Fermi-level and therefore improve S.
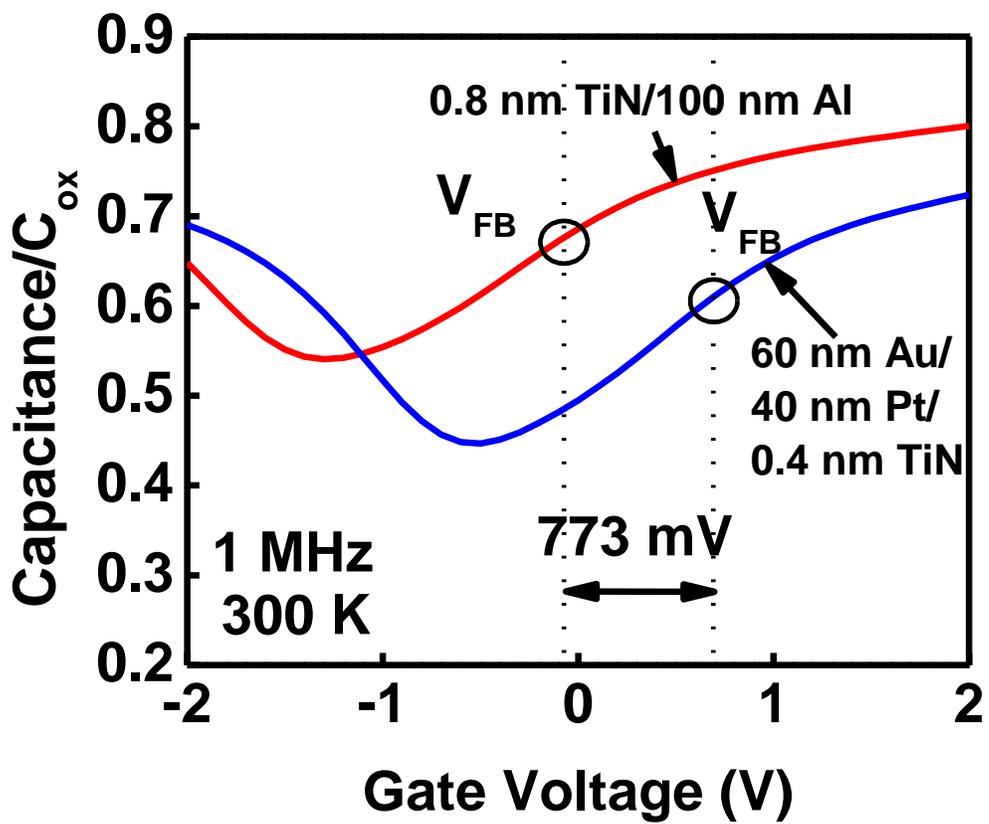
Figure 6.6 Comparison of 1 MHz C-V curves measured at 300 K of (110) Ge MOS-C's on GaAs via an AlAs buffer architecture. Both curves are normalized to their respective C_{ox}.
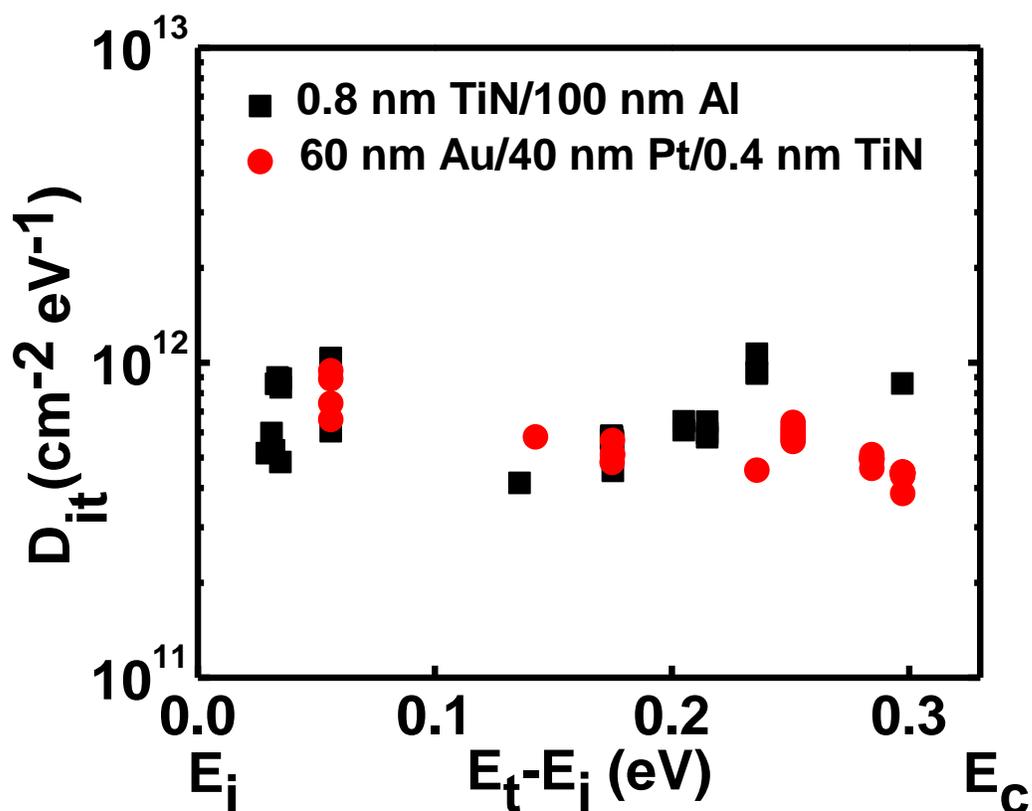
Figure 6.7 *D_{it}* as a function of energy of two different gated (110) Ge MOS-C's on GaAs via an AlAs buffer architecture.

6.4　XPS of Al Metal

XPS measurements were performed to study the composition of the Al metal. Fig. 6.8 shows a 3D overlay of the XPS spectra for the Al 2*p* orbital. There is evidence of oxidation of the Al, as demonstrated by a shoulder peak which is located at a higher binding energy, correlating with the electronegative nature of oxygen. This shoulder peak is prevalent even with increasing sputter time, suggesting that the Al is partially oxidized throughout its entire thickness. This oxidation explains the discrepancy in *C_{max}* for both the (100) and (110) Al-gated samples. The formation of oxidized Al throughout the Al gate metal adds

an extra interfacial layer which therefore increases *EOT* and decreases C_{max}. Further studies would need to be done to understand the nature of why the Al is oxidizing. The smaller intensity peaks to the right of the Al 2*p* peaks are from inelastic collisions.
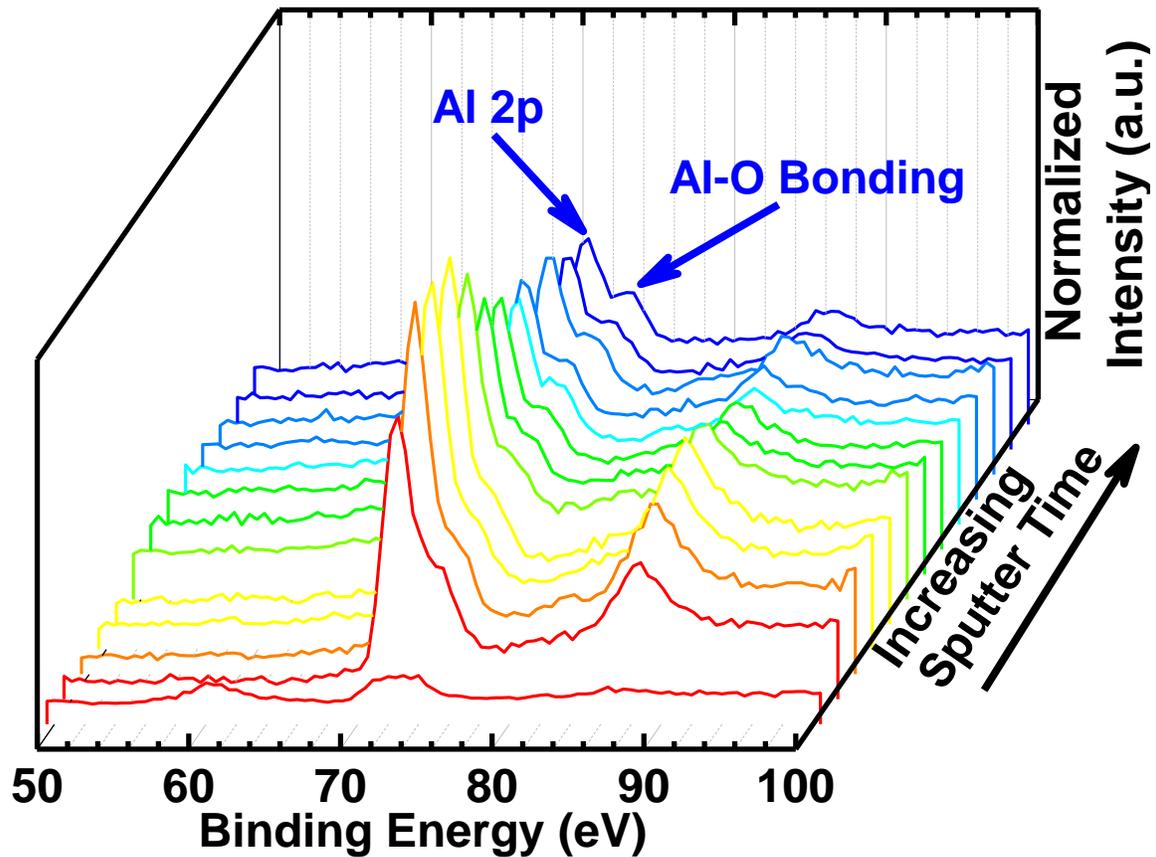
Figure 6.8 XPS analysis on the Al gate metal, demonstrating oxidization of the Al due to the observation of a shoulder peak correlating to an Al-O bond. The Al was deposited on HfO_2 integrated on Ge/GaAs heterostructures.

Reference

[1]   K. Mistry *et al.*, "A 45 nm logic technology with high-k+ metal gate transistors, strained silicon, 9 cu interconnect layers, 193 nm dry patterning, and 100% Pb-free packaging," in *IEDM Tech. Dig.*, 2007, pp. 247–250.

[2]   R. Tung, "Formation of an electric dipole at metal-semiconductor interfaces," *Phys. Rev. B*, vol. 64, pp. 205310-1–205310-15, Nov. 2001.

[3]   V. Heine, "Theory of surface states," Phys. Rev., vol. 138, no. 6A, pp. A1689-A1696, Jun. 1965.

[4]   Y. Yeo, T. –J. King, and C. Hu, "Metal-dielectric band alignment and its implications for metal gate complementary metal-oxide-semiconductor technology," *J. Appl. Phys.*, vol. 92, no. 12, pp. 7266–7271, Dec. 2002.

[5]   F. Bellenger *et al.*, "Passivation of Ge (100)/GeO2/high-k gate stacks using thermal oxide treatments," *J. Electrochem. Soc.*, vol. 155, no. 2, pp. G33–G38, Feb. 2008.

[6]   S. Gaudet *et al.*, "Thin film reaction of transition metals with germanium," *J. Vac. Sci. Technol. A.*, vol. 24, no. 3, pp. 474–485, May 2006.

[7]   D. K. Schroder, Semiconductor Material and Device Characterization, 3rd ed. Hoboken, NJ, USA: Wiley, 2006.

[8]   H. Michaelson, "The work function of the elements and its periodicity," *J. Appl. Phys.*, vol. 48, no. 11, pp. 4729–4733, Nov. 1977.

[9]   J. R. Brews, "Rapid interface parametrization using a single MOS conductance curve," *Solid State Electron.*, vol. 26, no. 8, pp. 711–716, Aug. 1983.

[10]   E. H. Nicollian and J. R. Brews, *MOS Physics and Technology*. Hoboken, NJ, USA: Wiley, 2003.

[11]   K. Martens *et al.*, "On the correct extraction of interface trap density of MOS devices with high-mobility semiconductor substrates," *IEEE Trans. Electron Devices*, vol. 55, no. 2, pp. 547–556, Feb. 2008.

[12]   Q. Xie *et al.*, "Germanium surface passivation and atomic layer deposition of high-k dielectrics—A tutorial review on Ge-based MOS capacitors," *Semicond. Sci. Technol.*, vol. 27, no. 7, Jun. 2012.

Chapter 7 – Conclusions and Future Prospects

7.1. Summary

The material system consisting of epitaxial Ge heterogeneously integrated on Si substrate *via* a composite large bandgap AlAs/GaAs buffer has recently been proposed as a novel solution for future high-speed CMOS logic applications. The low bandgap and high mobility of Ge coupled with the lattice-matched, large bandgap nature of the AlAs/GaAs buffer are promising for the development of high-speed MOSFETs operating within the low-power ($V_{DD} < 0.5$ V) regime. Furthermore, the adoption of the FinFET design utilizing the Ge/AlAs/GaAs/Si heterostructure offers further improvement for I_{ON} while maintaining better gate control and reliable device operation, which becomes especially necessary with the aggressive trend in CMOS scaling. In this research, the material and interfacial properties of Ge/AlAs/GaAs/Si heterostructures were comprehensively investigated. The results are summarized below:

(1) High-quality epitaxial Ge was grown on composite large bandgap AlAs/GaAs buffer heterogeneously integrated on Si substrate *via* solid-source MBE. NMOS capacitors were fabricated from the aforementioned material stack for the first time. Distinct, well-defined Ge, AlAs, and GaAs peaks were observed both from simulated and experimental x-ray rocking curves. In addition, SIMS, EDS elemental mapping, and EDS line profile confirmed negligible interdiffusion of the underlying buffer into the epitaxial Ge, further demonstrating the successful growth of a high-quality epitaxial Ge layer. The fabricated NMOS devices demonstrated excellent electrical characteristics with efficient modulation of the Fermi level from midgap to near the conduction band edge and a peak *FLE* of 27.5% corresponding to the low D_{it} value of 8.55×10^{11} cm^{-2} eV^{-1}. Moreover, 300 K frequency

dispersion was found to be as low as 1.80% per decade and a low frequency-dependent V_{FB} shift of 153 mV was observed, confirming low D_{it} near the conductance band edge. Furthermore, limited C-V stretch-out was observed confirming low D_{it} throughout the upper half of the Ge bandgap. The high-quality growth of epitaxial Ge on AlAs/GaAs buffer heterogeneously integrated on Si substrate and the superior electrical characteristics of the heterogeneously integrated Ge NMOS devices suggest the viability of future high-mobility channel material integration on Si via large bandgap buffer architectures for high-speed, low-voltage, high-performance CMOS logic applications.

(2) (100) and (110) crystallographically oriented Ge heterogeneously integrated on GaAs substrates *via* a large-bandgap AlAs buffer was grown by solid-source MBE. NMOS capacitors were fabricated from the aforementioned material stack, where several annealing schemes were applied to determine their passivation role on the critical oxide/Ge heterointerface of both (100) and (110) crystallographically oriented epitaxial Ge. It was found that: (a) PDA in O_2 ambient at 405 °C for thirty minutes provided the worst passivation, resulting in an increase in *EOT* while not significantly improving D_{it} or N_{ot} and, as for the case of (110) samples, further degrading these two metrics; (b) post-metallization anneal in O_2 at 405 °C for thirty minutes provided the best passivation, yielding excellent D_{it} distribution and low N_{ot}, but significantly increasing *EOT*; and (c) FGA at 300 °C for two minutes provided the second best D_{it} distribution behind O_2 post-metallization anneal, with minimum D_{it} values of 3.79×10^{11} cm^{-2} eV^{-1} and 3.84×10^{11} cm^{-2} eV^{-1} measured for (100) and (110) devices respectively, low N_{ot} values of 6.15×10^{11} cm^{-2} and 7.74×10^{11} cm^{-2} for (100) and (110) devices respectively, and moderate *EOT* values of 5.45 nm and 5.51 nm for (100) and (110) devices respectively. Thus, FGA for

300 °C, 2 min provided the optimal passivation scheme for the oxide/Ge heterointerface of both (100) and (110) epitaxial Ge. Such an optimal passivation scheme does not significantly increase *EOT* to allow for better channel inversion capabilities by the transistor gate metal, while still mitigating defects and oxide traps that would deleteriously degrade channel current for transistor applications.

However, an unusual inversion behavior at all measured frequencies was observed for the (110) sample, regardless of annealing methods or low-temperature measurements, which would otherwise suppress minority carrier response normally responsible for such inversion behavior. Two possible root causes were suggested, including: (a) the dominance of Ge_2O_3 species for (110) oriented Ge in contrast to the preferable GeO_2 species for (100) oriented Ge, as demonstrated through XPS analysis of the thermally grown GeO_x and (b) shallow traps close to the band edges with very quick response times. However, the interface engineering for both (100) and (110) epitaxial Ge integrated on AlAs/GaAs provides a critical first step in the realization of the usage of these material stacks for high-speed, low-voltage, high-performance CMOS logic applications.

(3) Work function tuning was performed on NMOS capacitors fabricated from (100) and from (110) crystallographically oriented Ge heterogeneously integrated on GaAs substrates *via* a large-bandgap AlAs buffer. Two gate metal stacks (0.8 nm/100 nm Al and 60 nm Au/40 nm Pt/0.4 nm TiN) were implemented on the aforementioned MOS capacitors to demonstrate the extent of V_{FB} shift and thereby, demonstrating the role of Fermi-level pinning on devices implemented from the aforementioned material stacks. An *S* value of 0.511 and 0.564 was calculated for (100) MOS-C's and (110) MOS-C's, respectively, demonstrating good Fermi-level unpinning, which is in good correlation with the low D_{it}

observed in both devices. Furthermore, the difference in C_{max} between the different-gated devices was explained by XPS analysis of the Al metal, where partial oxidization of the Al metal was observed, thereby adding an additional interfacial layer that would therefore decrease C_{max}. Such information would aid in the further *EOT* scaling of these devices by avoiding the usage of Al as the gate metal. Nevertheless, the low Fermi-level pinning observed in the MOS-C's allow for easy V_{TH} tunability for transistor applications, which is critical for the demonstration of low power, high-performance CMOS logic based on epitaxial Ge integrated on AlAs/GaAs.

Thus, this research demonstrates the feasibility of the Ge/AlAs/GaAs/Si material system for applications in low power, high-performance CMOS logic. The superior structural characteristics and electrical characteristics from MOS-C's from the aforementioned material stack, as well as further interface passivation engineering and demonstrated ability to easily tune V_{FB} demonstrates the viability of future high-mobility channel material integration on Si via large bandgap buffer architectures for high-speed, low-voltage, high-performance CMOS logic applications.

7.2. Prospects for Future Research

In order to leverage the research results discussed above and fully exploit the advantages of Ge/AlAs/GaAs/Si-based electronic devices, additional research must be done in the following areas:

(1) In **Chapter 4**, it was demonstrated that MOS-C's fabricated from epitaxial Ge heterogeneously integrated on Si substrate *via* AlAs/GaAs buffer architecture demonstrated an unusually high unintentional doping level of 1.24×10^{18} cm^{-3}, which thereby negatively shifts the V_{FB} of all MOS-C's based on the Ge/AlAs/GaAs material

stack, as observed in C-V measurements from ***Chapter 4***, ***Chapter 5***, and ***Chapter 6***. For transistor applications, this high V_{FB} is undesirable as it makes it more difficult for the gate metal to invert the channel by requiring increased gate voltage to deplete the extra charges as a result of the high doping. In order to fully realize low-power devices implementing the Ge/AlAs/GaAs/Si material stack, it is important to determine the root cause of the unintentional doping and thereby reduce doping as much as possible for future growths implementing the aforementioned material stack.

(2) Further *EOT* scaling and reduction of D_{it} will be necessary to fully unlock the potential usage of Ge as an alternative material to Si, thereby driving device current required for high-performance electronics. To help scale *EOT*, the implementation of high-κ gate dielectrics such as HfO_2 and reduction in passivation layer thickness will be necessary. However, a balance must be achieved in regards to overall D_{it} levels, which must stay competitive with the D_{it} levels as observed in Si-based devices. Implementation of plasma oxidation, where a relatively thick layer of HfO_2 helps reduce gate leakage, but the usage of very thin Al_2O_3 and GeO_x to maintain a competitive *EOT*, may be a possible solution to both scale *EOT* as well as demonstrate low D_{it} values. Furthermore, low D_{it} values will aid in further Fermi-level unpinning, allow for better tunability of V_{TH} using different work function metals.

(3) The root cause of the anomalous inversion behavior observed in MOS-C's implementing (110) epitaxial Ge on AlAs/GaAs would need to be analyzed. Performing deep-level transient spectroscopy on both (100) and (110) epitaxial Ge on AlAs/GaAs material stacks will help determine the existence of shallow traps that may have short response times. Furthermore, the root cause of why Ge_2O_3 is the dominant oxide species

for epitaxial (110) Ge rather than the preferable GeO_2 should also be analyzed. Demonstration of similar MOS-C's behavior for epitaxial (100) and epitaxial (110), as is expected and demonstrated in different oriented Ge bulk devices, will ensure the applicability of (110) surfaces based on Ge/AlAs/GaAs material stacks for FinFET sidewalls, whose orientation is (110).

(4) Lastly, the primary objective of the research presented in this thesis is to demonstrate transistor devices implementing epitaxial Ge heterogeneously integrated on Si substrate *via* composite large bandgap AlAs/GaAs buffers. In addition to addressing the issues mentioned in points 1-3, the most critical obstacle lies in the development of a proper device fabrication process flow for MOSFET devices. Once points 1-3 are addressed and a MOSFET devices can be demonstrated with competitive performance as bulk Ge MOSFETs presented in the literature, the demonstration of FinFET devices will be the next critical step. The realization of a FinFET device implementing epitaxial Ge heterogeneously integrated on Si substrate *via* large bandgap AlAs/GaAs buffers will ultimately demonstrate the feasibility of future high-mobility channel material integration on Si via large bandgap buffer architectures for high-speed, low-voltage, high-performance CMOS logic applications.

Appendix

List of Publications

JOURNAL PAPERS:

1. **P. D. Nguyen**, M. Clavel, P. S. Goley, J.-S. Liu, N. Allen, L. J. Guido and M. K. Hudait, "Heteroepitaxial Ge MOS Devices on Si Using Composite AlAs/GaAs Buffer," *IEEE J. Elec. Dev. Soc.*, vol. 3, no. 4, pp. 341-348, Jul. 2015.

2. D. Maurya, Y. Zhou, B. Chen, M. –G. Kang, **P. Nguyen**, M. K. Hudait, "Functionally Graded Lead-free Ferroelectric Thin Films: Role and Origin of Internal Electric Field and Modulated Electrical Response," *ACS Appl. Mater. Interf.*, vol. 62, no. 10, pp. 22458-22468, Sep. 2015.

3. **P. D. Nguyen** *et al.*, "Effect of Surface Passivation on Crystallographically Oriented Epitaxial Ge Metal-oxide-semiconductor devices," *ACS Appl. Mater. Interf.*, (To be submitted).

4. **P. D. Nguyen** *et al*., "Work Function Tuning on Crystallographically Oriented Epitaxial Ge Metal-oxide-semiconductor devices," *J. Appl. Phys*., (To be submitted).