UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Atomic Layer Deposition of Hafnium Dioxide on Sulfur-Passivated Silicon-Germanium Surfaces

Permalink

https://escholarship.org/uc/item/03v557xw

Author

Clemons, Maximillian Samuel

Publication Date

2016

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Atomic Layer Deposition of Hafnium Dioxide on Sulfur-Passivated Silicon-Germanium Surfaces

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Electrical Engineering (Applied Physics)

by

Maximillian Samuel Clemons

Committee in Charge

Professor Andrew Kummel, Chair
Professor Peter Asbeck, Co-Chair
Professor Yeshaiahu Fainman
Professor Yu-Hwa Lo

2016

Copyright

Maximillian Samuel Clemons, 2016

All rights reserved.

The Thesis of Maximillian Samuel Clemons is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Aluminum	Al
Ammonium Sulfide	$(NH_4)_2S$
Ampere	A
Angstrom	Å
Argon	Ar
Atomic Layer Deposition	ALD
Capacitance-Voltage	C-V
Chemical Vapor Deposition	CVD
Complementary Metal Oxide Semiconductor	CMOS
Conductance-Voltage	G-V
Current-Voltage	I-V
Density of Interface Traps	D_{it}
Drain Induced Barrier Lowering	DIBL
Equivalent Oxide Thickness	EOT
Farad	F
Flat Band Voltage	V_{FB}
Hafnium Dioxide	HfO_2
Hafnium Tetrachloride	$HfCl_4$
Hydrofluoric Acid	HF
Metal Oxide Semiconductor Capacitor	MOSCAP
Metal Oxide Semiconductor Field Effect Transistor	MOSFET
Molecular Beam Epitaxy	MBE
Nickel	Ni
Palladium	Pd
Physical Vapor Deposition	PVD
Short Channel Effect	SCE
Siemens	S
Silicon Dioxide	SiO_2
Silicon Germanium	SiGe
Silicon on Insulator	SOI
Titanium	Ti
Titanium Nitride	TiN
Volt	V

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Andrew Kummel, for allowing me the opportunity to be part of his research group. Over the last 12 months my knowledge and experience has grown significantly, and without his guidance, none of that would have been possible. I look forward to pursuing the tremendous opportunities he has opened up for me.

I would also like to thank the other members of Kummel Lab for helping me get to this point in my academic career. Specifically, Kasra Sardashti has worked long and tirelessly with me to help complete this research, and has been a seemingly endless source of semiconductor physics and materials science knowledge.

Additionally, I would like to thank my family for their unwavering support through this long and arduous journey. With my parents and brother cheering me on the whole way, no obstacle seemed insurmountable.

Chapter Three, in part is currently being prepared for submission for publication of the material. K. Sardashti, M. Clemons, A. C. Kummel, "Atomic Layer Deposition of HfO_2 on Sulfur-passivated $Si_{0.7}Ge_{0.3}$ Surfaces".

ABSTRACT OF THE THESIS

Atomic Layer Deposition of Hafnium Dioxide on Sulfur-Passivated Silicon-Germanium Surfaces

by

Maximillian Samuel Clemons

Master of Science in Electrical Engineering (Applied Physics)

University of California, San Diego, 2016

Professor Andrew Kummel, Chair

Professor Peter Asbeck, Co-Chair

To integrate SiGe into future CMOS devices, it is essential to realize reliable strategies to deposit very thin high-k dielectrics on SiGe surfaces with a low density of interfacial defects. HfO_2 was deposited by atomic layer deposition (ALD) using $HfCl_4$ and H_2O precursors. The quality of interfaces was varied by ex-situ surface treatment prior to ALD, including HF clean and HF clean followed by wet ammonium sulfide treatment. Electrical properties of the interfaces were examined by variable frequency capacitance-

voltage (C-V) spectroscopy. Interfaces passivated by sulfur were found to have nearly 2x smaller density of interface traps than HF-treated interfaces, particularly near the edge of the valence band. The effect of Pd/Ti/TiN as a gettering gate electrode on the electrical characteristics of the interfaces were compared with Ni. By using Pd/Ti/TiN gate electrodes, lower equivalent oxide thicknesses (EOT) were achieved, but no significant improvement in the interface quality was observed.

CHAPTER ONE

Introduction

1.1 Preview

Improving CMOS technology requires scaling devices to increase circuit density and speed while reducing power consumption. With traditional planar MOSFET devices nearing the limit of scalability, alternative device architectures are necessary to continue the development of nanoscale electronics[1]. FinFETs and other multi-gate device structures have been offered as possible solutions to this problem, and have become more prevalent in recent years.

Reducing oxide layer thickness is a critical aspect of MOSFET device scaling, but due to quantum mechanical tunneling, current scaling trends cannot continue beyond the fundamental SiO_2 thickness limit of approximately 0.5 nm [2]. Increasing the oxide permittivity using a high-k dielectric material allows for a thicker oxide layer with the same capacitance as SiO_2. High-k dielectric layers can then be scaled, resulting in an equivalent SiO_2 layer thinner than 0.5 nm. Due to the difficulty of growing high-k dielectrics on semiconductors, as well the large interface area required by alternative device structures, interface quality between oxide and semiconductor is a critical issue in ensuring device performance and reliability. Specifically, parameters such as density of interface traps (D_{it}) of the oxide are critically important.

Alternative semiconductor materials can additionally help to maintain or improve performance during device scaling. SiGe is a promising material due to its higher hole mobility compared to Si, the tunability of its mobility and bandgap by varying the Ge content, as well as its easy integration into the existing Si process flows.

1.2 Device Scaling and Short Channel Effects

With planar MOSFETs nearing the limit of scaling due to gate oxide tunneling, short channel effects play a much greater role in device performance. For long-channel MOSFETs, a one-dimensional analytical model can be used to describe channel behavior and electric field effects across the oxide[3]. This is because the source and drain depletion regions are far enough removed from the channel that their effects are negligible over most of the device body.

As channel length scales with smaller devices, the long-channel model becomes invalid. When the source-drain distance becomes comparable to the MOS depletion width, junction depletion effects begin to have significant influence on the channel, and a two-dimensional analytical model is required to accurately describe channel behavior[4]. Electric field effects across the oxide, and in the channel direction, must both be taken into account. By modifying the model to include depletion terms from the source and drain junctions, Short Channel Effects (SCEs) are observed, including a decrease in threshold voltage and degradation of subthreshold slope for short-channel MOSFETs. Applying high drain bias further lowers the potential barrier and threshold voltage, an effect called Drain-Induced Barrier Lowering (DIBL).

Figure 1.1 gives a visual comparison of band bending in long-channel and short-channel MOSFETs using constant surface potential contours. In the long-channel device (L_g = 2 μm), surface potential curves are parallel to the gate for most of the device length, meaning lateral field effects play very little role in channel behavior. At short channel gate lengths (L_g = 0.35 μm), surface potential is no longer constant along the semiconductor-

oxide interface, and there is significant variation moving across the channel from source to drain. Barrier lowering described by the short-channel model can be seen in the figure as a dramatic increase of surface potential adjacent to the drain contact.

1.3 Performance Advantages of Multi-Gate Devices

To address short channel effects and DIBL in MOS devices, alternative device architectures have been developed to improve electrostatic gate control by increasing the number of gates and semiconductor-oxide interface area[5]. FinFETs (double-gate) and tri-gate FETs, two prevalent multi-gate device structures shown in Figure 1.2, have demonstrated electrostatic improvement over single-gate planar FETs[6]. In addition to improving short channel performance degradation, multi-gate FET threshold voltage is set by gate work function instead of substrate doping levels, meaning the device body can be lightly-doped or undoped[7]. Limits on semiconductor doping levels, as well as doping fluctuations due to process variation, have provided challenges in device scaling.

1.4 High-k Dielectrics and EOT

SiO_2 has been the prevailing choice of gate oxide for Si-based MOSFETs. As the native oxide for Si, it exhibits many favorable electrical and material properties: thermal growth on Si with good thickness and uniformity control, low interface defect density and easy defect passivation via annealing, high thermal and chemical stability, and large bandgap for good electrical isolation[8]. Unfortunately, as devices scale and oxide thickness decreases, SiO_2 approaches a fundamental thickness limit. Devices with oxide layers near

this limit experience a high probability of tunneling leakage current and degraded device performance.

Increasing the permittivity of the oxide is the electrostatic equivalent of reducing the gate oxide thickness. Therefore, by using high-k dielectrics as alternatives to SiO_2, it is possible to generate the same MOS capacitance as an ultra-thin SiO_2 layer while avoiding excessive gate leakage. Equivalent oxide thickness (EOT) can be calculated for high-k dielectric layers to show what the oxide layer thickness would be if an SiO_2 layer were used instead:

$$EOT = t_{high-k}\left(\frac{k_{SiO2}}{k_{high-k}}\right) \tag{1.1}$$

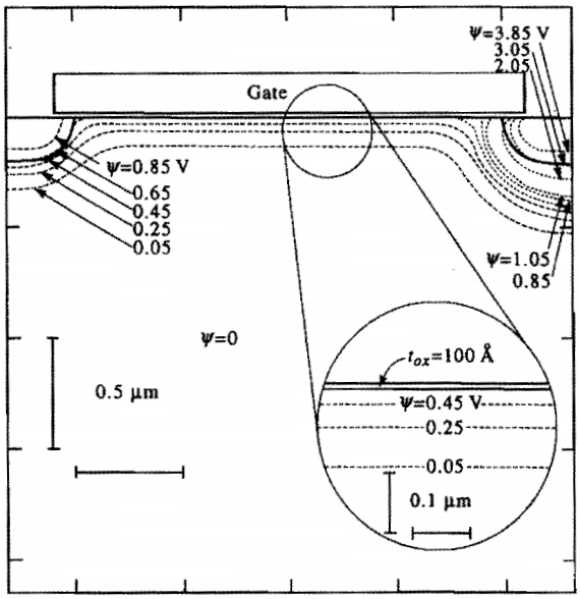
The term $t_{high\text{-}k}$ is the high-k oxide thickness, $k_{high\text{-}k}$ is the high-k dielectric constant, and k_{SiO2} is the dielectric constant of SiO_2 ($k_{SiO2} = 3.9$). EOT values less than the SiO_2 thickness limit indicate that device scaling can continue with alternative high-k materials.

1.5 SiGe

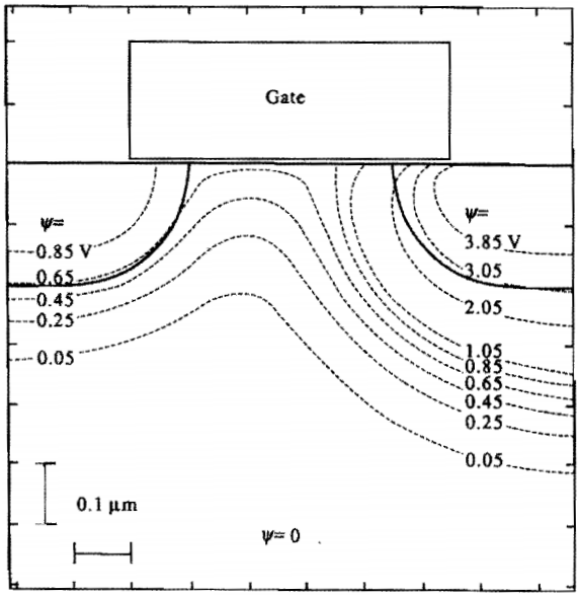
Due to the lower hole mobility than electron mobility of Si (μ_p = 500 cm^2/V-s, μ_n = 1450 cm^2/V-s at room temperature)[9], PMOS devices are often the performance limiting factors in CMOS technology. Ge is favorable as an alternative semiconductor material because it has four times greater hole mobility and two times greater electron mobility than Si. However, because of the low thermal stability of GeO_x, use of thermally grown Ge oxide gate dielectric leads to significant reliability issue[10].

$Si_{1-x}Ge_x$ alloy with high Ge content ($x > 0.5$) has boosted mobility due to strain-induced heavy-hole/light-hole splitting and reduction of in-plane hole effective mass. However, because of the 4.2% lattice mismatch between Si and Ge, only very thin SiGe layers grown on Si can avoid mobility degradation from strain relaxation and dislocations[11]. In order to grow a suitably thick layer and gain the benefit of SiGe with high Ge content, a strain relieved buffer must be used. SiGe with low Ge content ($0.15 \leq x \leq 0.5$), on the other hand, still provides a small mobility improvement over Si, and is favored in research for its ability to be grown directly on Si substrates at usable thicknesses[12].

1.6 Figures

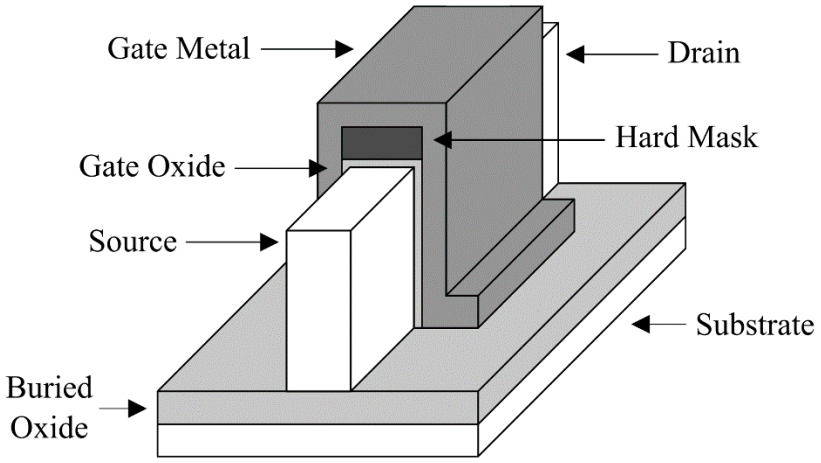


(a)

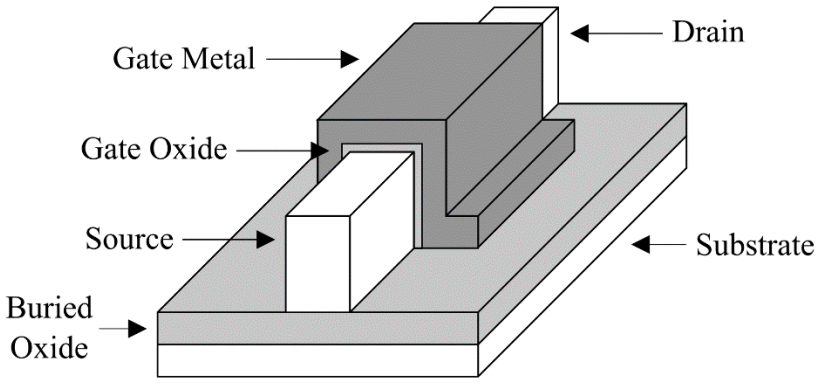


(b)

Figure 1.1 Diagram of a (a) long channel and (b) short channel planar MOSFET. Constant potential contours (dashed lines) show competing vertical and lateral electric field effects at each scale[4].



(a)



(b)

Figure 1.2 Two types of Multi-Gate FETs. A (a) FinFET is considered a double-gate structure due to fin height being sufficiently larger than fin width, as well as having a hard mask added to inhibit the third gate. A (b) Tri-Gate FET utilizes three sides as gates by not including a hard mask, and having fin height and fin width closer in scale.

CHAPTER TWO

Materials and Methods

2.1 Atomic Layer Deposition

Fabrication of device features at nanometer scale, especially thin-film gate oxide layers, requires exceptional control of deposition thickness and uniformity. While many conventional deposition methods such chemical vapor deposition (CVD) and physical vapor deposition (PVD) can achieve high quality thin-film deposition, they do not have the required precision for conformal coating of high aspect ratio electronic structures. Atomic layer deposition (ALD) is considered a subset of CVD, modified to deposit in a cyclic manner for thickness control at the monolayer level[13].

The ALD process is a binary reaction that requires two precursor materials, and a cycle can be broken down into four stages: 1) Precursor 1 dose, 2) First purge, 3) Precursor 2 dose, and 4) Second purge. Figure 2.1 gives a schematic diagram of a single HfO_2 deposition cycle using ALD, with $HfCl_4$ and H_2O as precursors. Similar high-k oxides such as Al_2O_3, ZrO_2, and TiO_2 can also be deposited via ALD with various precursor materials. To analyze HfO_2 ALD specifically, a cycle can be broken down into the following two half reactions (stages 1 and 3 above)[14]:

$$-Hf-OH+HfCl_4 \rightarrow -Hf-O-HfCl_3+HCl\uparrow \tag{2.1}$$

$$-Hf-O-HfCl_3+3H_2O \rightarrow -Hf-O-Hf\left(OH\right)_3+3HCl\uparrow \tag{2.2}$$

Following each cycle, only Hf-O-Hf-OH sites are present, meaning the reaction process can start again as it had in the previous iteration. The key to an ALD reaction is its self-limiting behavior, whereby the finite number of surface sites available saturate after a certain amount of precursor dosing. With both of the surface reactions in the process being self-limiting, sequential growth can continue at the monolayer level independent of precursor flux[15]. Purge steps added between dosing stages further differentiate ALD from conventional CVD. In CVD, precursors react in the gas phase and are subject to decomposition, but adding ALD purge steps separates precursors into surface reactions without self-decomposition[16]. Failure to perform adequate purges can result in unintentional CVD reactions between lingering precursor materials in the chamber, ultimately leading to increased oxide growth rate and nonuniform deposition.

A. Experimental Setup and Pulsing Procedure

For the experimental work in this thesis, ALD was performed using a Beneq TFS-200 continuous flow reactor. In this type of ALD reactor, carrier gas (Ar specifically) flows continuously over samples in the chamber during all four stages of an ALD cycle, precluding the need for a separate purge gas dose during stages 2 and 4. Because of this configuration, deposition in a flow-type reactor is performed at relatively high pressures (1.7 Torr in the reactor) and extensive pumping isn't required to achieve these vacuum levels. Previous work using this system has yielded a large range of reaction temperatures sufficient for depositing Al_2O_3 on SiGe[17], but for HfO_2 deposition, the optimal reaction temperature strongly depends on the Hf precursor chemistry.

At room temperature and atmospheric pressure, $HfCl_4$ is a crystalline solid powder, and therefore requires heating to achieve sublimation and sufficient vapor pressure for dosing. Heating the source and precursor material to manufacturer-specified 190 °C yields 1 Torr of vapor pressure, which is still not adequate given the higher base pressure of the flow-type reactor used in this study. Thus, a bubbling configuration using Ar carrier gas is essential to effective delivery of HfCl4 gas to the sample surface.

Details of the Ar bubbling procedure for HfCl4 container is shown in Figure 2.2. The Beneq ALD system is equipped with three valves adjacent to the hot source. DV-BH1 and DV-PH1 are the inlet and outlet pulse valves, respectively, and DV-BHA1 is an additional pulse valve controlling carrier gas flow from the main gas line to the other two valves. A connecting line between inlet and outlet pulse valves is designed to enable the Ar purge of the HfCl4 dose line. A 200 um orifice has been installed in this connection line to prevent backstreaming of HfCl4 into the Ar feeing line. With carrier gas continuously flowing throughout the entire reaction, flow can only be directed forward into the inlet, out of the outlet, and across the connecting channel from inlet to outlet.

The alternative pulsing procedure (“bubbling”) starts with DV-BHA1 and DV-BH1 pulsed open to inject carrier gas into the vessel. Next, all valves are closed to build vapor pressure and mix the Ar carrier gas with $HfCl_4$ vapor. Finally, all three valves (DV-BHA1, DV-BH1, and DV-PH1) are pulsed simultaneously to dose the Ar-HfCl4 gas mixture into the dosing line. Pulsing just DV-PH1 can be used instead as the final step, but injecting carrier gas at the same time gives a more controlled pressure response.

2.3 Density of Interface Traps Characterization

Trap states at the semiconductor-oxide interface can arise due to structural lattice defects, dangling bonds, and impurities, and can cause significant degradation of device performance. Many research efforts have been focused on calculating and reducing the density of interface traps (D_{it}) by improving the interface quality. Two conventional methods have been used in this thesis to calculate D_{it} from experimental capacitance-voltage (C-V) and conductance-voltage (G-V) measurements.

A. Full Interface State Model

In classical interface theory, single-level trap states can be modeled using a Y equivalent circuit, which consists of capacitance C_T connected to conductances G_n and G_p, associated with the conductance and valence bands, respectively[18]. In order to model trap states at all energy levels, however, this Y equivalent circuit must first be converted to a Δ equivalent circuit model (shown in Figure 2.3a), then the circuit elements can be integrated over trap energy to yield the following complex admittances[19]:

$$C_{Tn} = qD_{it}\tau_n^{-1}\int_0^1 df\,(1-f)\left[\,j\omega f\,(1-f) + f\,\tau_p^{-1} + (1-f)\,f\,\tau_n^{-1}\right]^{-1} \tag{2.3}$$

$$C_{Tp} = qD_{it}\tau_p^{-1}\int_0^1 df\,(1-f)\left[\,j\omega f\,(1-f) + f\,\tau_n^{-1} + (1-f)\,f\,\tau_p^{-1}\right]^{-1} \tag{2.4}$$

$$G_{gr} = qD_{it}\tau_p^{-1}\tau_n^{-1}\int_0^1 df\left[\,j\omega f\,(1-f) + f\,\tau_p^{-1} + (1-f)\,f\,\tau_n^{-1}\right]^{-1} \tag{2.5}$$

where

$$\tau_n = \left(\sigma_n v_{th}^n n_s\right)^{-1} \tag{2.6}$$

$$\tau_p = \left(\sigma_p v_{th}^p p_s\right)^{-1} \tag{2.7}$$

At each frequency, the complex admittances are completely specified by the D_{it}, τ_n, and τ_p terms. Therefore, using a numerical calculation program, measured capacitance and conductance can be fitted to the equivalent model's $C_{tot}(\omega)$ and $G_{tot}(\omega)$ using the previous terms as fitting parameters. In this way, D_{it} can be manually extracted at several bias points over the entire bias range respective to the band edges.

B. Conductance Method

While full interface state model requires a dedicated numerical calculation tool to perform D_{it} extraction, the conductance method makes approximations to produce a simple analytical model in the depletion region. In the depletion region, capture and emission processes at trap levels primarily involve majority carriers, and therefore the only trap level contribution comes from a single admittance element C_{it} in parallel with the depletion capacitance C_d. Using this approximation, the equivalent model can be reduced as shown in Figure 2.3b, with the following descriptions for parallel capacitance and conductance[19]:

$$C_p = C_d + C_{it}\left(\omega\tau\right)^{-1}\tan^{-1}\left(\omega\tau\right) \tag{2.8}$$

$$\frac{G_p}{\omega} = C_{it}\left(2\omega\tau\right)^{-1}\ln\left[1+\left(\omega\tau\right)^2\right] \tag{2.9}$$

Experimental measurements of capacitance (C_m) and conductance (G_m) are modeled as the simple parallel connection shown in Figure 2.3c. Translation to the depletion interface trap state model can then be done, resulting in the following equation for G_p/ω[20]:

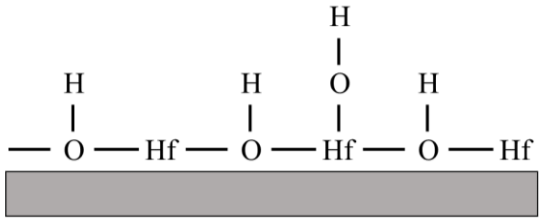
$$\frac{G_p}{\omega} = \frac{\omega C_{ox}^2 G_m}{G_m^2 + \omega^2 \left(C_{ox} - C_m\right)^2} \tag{2.10}$$

Interface trap states are frequency-dependent, as determined by the trap level time constant τ, with resulting energy loss due to changes in occupancy at each level. It follows that maximum energy loss occurs when the modulation frequency applied during measurement is resonant with the interface traps, or when $\omega\tau = 1$. Therefore, D_{it} can be estimated from the peak of the G_p/ω vs. ω plot using the following equation:

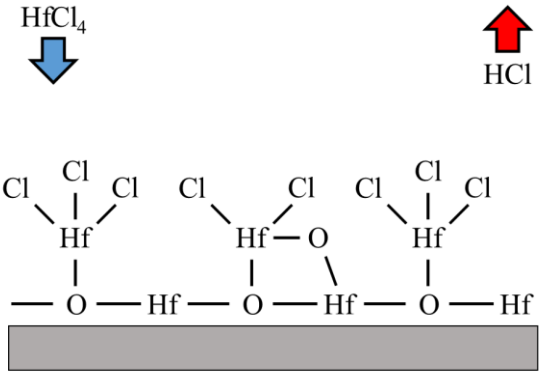
$$D_{it} \approx \frac{2.5}{Aq}\left(\frac{G_p}{\omega}\right)_{\max} \tag{2.11}$$

Using this method provides a quick, non-rigorous way of approximating Dit values in the depletion region, where much of the detrimental effects are seen. When considering total interface quality however, it's necessary to model D_{it} in every region using the full interface state model.

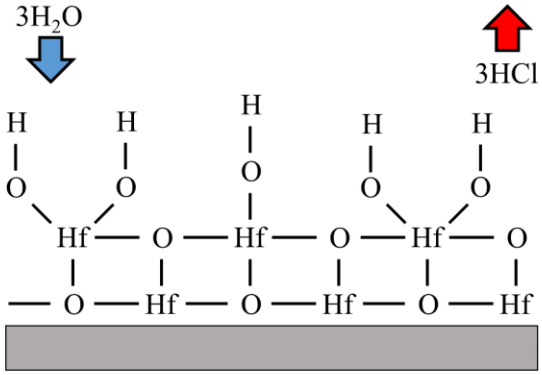
2.5 Figures



(a)



(b)



(c)

Figure 2.1 Schematic of an HfO_2 ALD cycle. At (a) the start of the reaction cycle, Hf atoms are terminated by –OH bonds. With (b) the addition of the $HfCl_4$ precursor, one or two of Hf–Cl bonds break and Hf–O bonds form. The (c) further addition of H_2O as a precursor again removes the rest of Hf–Cl bonds, returning the surface to –OH termination.

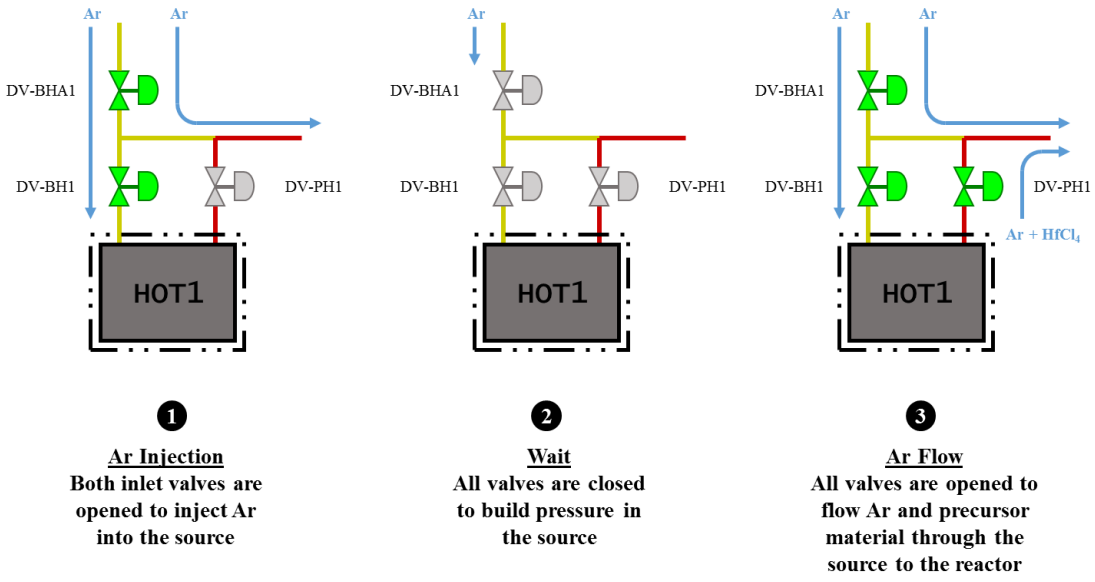


Figure 2.2 $HfCl_4$ hot source Ar bubbling procedure. In order to increase $HfCl_4$ vapor pressure to sufficient levels, additional gas injection into the vessel is required. Open valves are indicated in green and Ar flow is indicated with blue arrows.

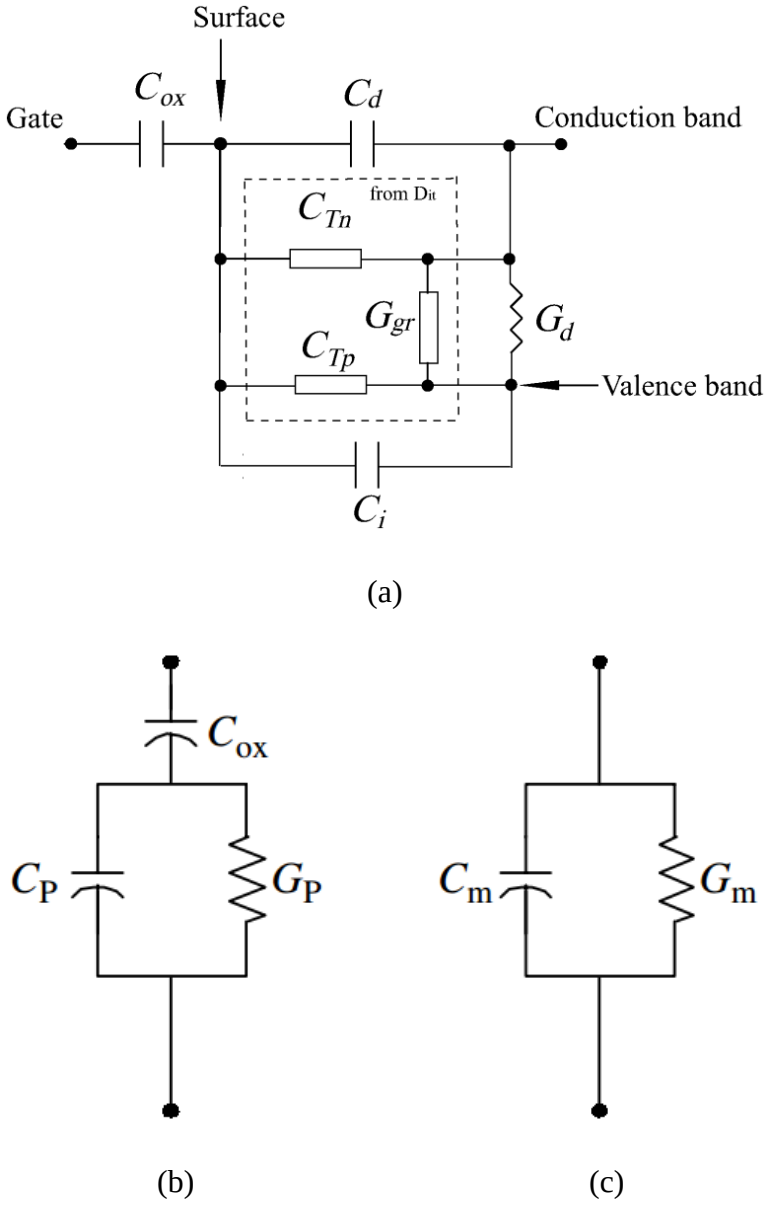


Figure 2.3 Equivalent capacitor circuit models for interface trap characterization. Full interface state modeling can be done using a (a) Δ equivalent circuit model, with complex admittance elements for contributions from the conduction and valence bands, as well as generation-recombination[19]. Approximations can be made to reduce to a (b) simplified model in the depletion region, which can then be equated to the (c) model of measured capacitance and conductance [21].

CHAPTER THREE

Results and Discussion

3.1 Introduction

The present study determined the effect of ex-situ wet sulfur passivation on the electrical properties of HfO_2/SiGe(001) interfaces. Electrical characteristics of the HfO_2 and HfO_2/SiGe interfaces were determined by capacitance-voltage (C-V) and current-voltage (I-V) spectroscopy measurements on MOS capacitors (MOSCAPs). Compared with HF treatment, S-passivation resulted in nearly 2x reduction in the density of interface traps near the edge of the valence band. Moreover, effect of gate electrode composition on the interface quality and Ge diffusion behavior is discussed by comparison between C–V characteristics of Ni/HfO_2/SiGe and Pd/Ti/TiN/HfO_2/SiGe gate stacks.

3.2 Experimental Details

A 12nm thick p-type $Si_{0.7}Ge_{0.3}$(100) with doping level of $1x10^{18}$ cm^{-3} (Applied Materials) was grown epitaxially on p-type Si(100) by molecular beam epitaxy (MBE). Prior to ALD, SiGe native oxide was removed by cyclic HF clean using 2% HF solution and DI water at 25 °C, ending with HF dip. For sulfur passivation, HF-treated samples were immersed in 24% $(NH_4)_2S$ solution at 25 °C for various times (15, 30, and 60 min) followed by 30 s of DI H_2O rinse. After surface clean, samples were transferred to the ALD chamber with less than 2 min of air exposure. HfO_2 ALD was performed at 300 °C in a Beneq TFS-200 continuous flow reactor, with Ar as the carrier gas. The chamber base pressure during the ALD process was about 1.7 torr. HfO_2 was deposited by consecutive cycles of 500 ms of $HfCl_4$ and 500 ms of H_2O. After each of $HfCl_4$ and H_2O pulses a 6 s long Ar purge was

employed. After ALD, two different gate electrodes were deposited directly on HfO_2: 1) 50 nm thick Ni gate deposited by thermal evaporation; 2) Pd/Ti/TiN gate, where 10-15 nm of TiN was deposited first using RF magnetron sputtering followed by 10-15 nm of Ti and 50-70 nm of Pd deposited by DC magnetron sputtering. Gate electrode deposition was followed by DC magnetron sputtering of 100 nm thick Al back contact layers on the back sides of the Si substrates. After completion of fabrication, MOSCAPs were annealed in forming gas (5% H_2, 95% N_2) at 250 °C for 15 min.

Capacitance-Voltage (C-V) spectroscopy of the MOSCAPs was performed using an Agilent B-1500 semiconductor analyzer, with AC modulation amplitude of 30 mV, in the gate bias range of -2 to 2 V, at multiple frequencies from 2 KHz to 1 MHz. Using the capacitance and conductance vs. gate voltage at various frequencies, density of interface traps was calculated using the full interface state model with Δ circuit of three complex elements method[19,22]. Gate leakage vs. gate bias was measured in the same bias range.

3.3 Results and Discussion

The effects of surface passivation on the electrical properties of HfO_2/SiGe interfaces were determined by variable frequency C-V spectroscopy. Figure 3.1 displays the C-V results measured from 2 KHz to 1 MHz for SiGe surfaces treated by HF (Figure 3.1a-c) and HF + $(NH_4)_2S$ (noted as HF+S, Figure 3.1d-f). Number of ALD cycles for HfO_2 varied from 60 (a, d) to 45 (b, e) and 40 (c, f). S-passivation was achieved by immersing the sample in 24% $(NH_4)_2S$ solution for 30 min. Decreasing the number of ALD cycles led to larger maximum accumulation capacitance (C_{max}), confirming the scalability of the oxide capacitance for both surface cleaning methods. Moreover, by reducing the oxide

thickness, flat band voltage became smaller, consistent with reduction in negative fixed charge within the oxide. At all oxide thicknesses, the height of the low-frequency "bump" that appears in the C-V curves between 0 and +1 V is smaller for S-passivated samples relative to HF-treated samples, consistent with reduction in density of interface traps by wet sulfur passivation[23,24].

Equivalent oxide thickness (EOT) values, estimated by quasi-static fitting of C-V curves, are shown in Figure 3.2a. By reducing the number of ALD cycles to 40, EOT values as low as of 1.1 nm have been achieved for the HfO_2 directly deposited on $Si_{0.7}Ge_{0.3}$(100) surfaces. S-passivation for all thicknesses resulted in about 0.1 nm larger EOT. This is consistent with either formation of a thicker interfacial layer between HfO_2 and SiGe or higher nucleation density of HfO_2 on the SiGe surface in the presence of sulfur. Reducing the number of ALD cycles from 60 to 45 raised the gate leakage current in accumulation for HF-treated samples by more than two orders of magnitude. Despite higher leakage current for 60 cycles of HfO_2, for thinner HfO_2 layers S-passivation resulted in about 2x lower leakage current relative to HF-treatment, which can be ascribed to higher HfO_2 nucleation density and initial oxide ALD growth rate (Figure 3.2b).

Density of interface traps (D_{it}) as a function of Fermi level relative to the edge of the valence band was calculated using the full interface model for samples with 45 and 40 cycles of HfO_2 ALD as shown in Figure 3.2c. In all cases, the majority of interface traps were concentrated within 0.5 eV above the edge of the valence band. For HF-treated samples, reducing the number of cycles from 45 to 40 resulted in about 100 meV shift to lower energies in part due to the -50 mV shift in V_{FB}. The shift is less significant for S-passivated samples. Moreover, maximum D_{it} value ($D_{it,max}$) near the valence band edge was

significantly smaller for S-passivated vs. HF-treated interfaces: For 45 cycles of HfO_2, $D_{it,max}$ was 1.4 x 10^{13} cm^{-2} eV^{-1} vs. 7 x 10^{12} cm^{-2} eV^{-1} and for 40 cycles of HfO_2 $D_{it,max}$ from 1.3 x 10^{13} cm^{-2} eV^{-1} vs. 8 x 10^{12} cm^{-2} eV^{-1}. The nearly 2x reduction in D_{it} was consistent with sulfur protecting the surface from unfavorable reactions during the ALD process. Therefore, S-passivation after HF cyclic clean improved the HfO_2 insulating properties as well as the quality of the interface between SiGe and HfO_2.

Influence of wet sulfur passivation time on the C-V characteristics of the MOSCAPs was determined by immersing the samples in 24% $(NH_4)_2S$ solution prior to HfO_2 ALD for 15 min, 30 min and 60 min (Figure 3.3). By increasing the S-passivation time, V_{FB} shifted to more negative values consistent with a lower density of negative oxide charge. Devices with 15 and 30 min of S-passivation had identical C-V characteristics except for the low-frequency noise. The noise observed at low frequencies for 60 min passivation is owing to excess leakage in the films (see Figure 3.4) or poor mechanical stability for the electrical probes throughout the measurements.

Average D_{it} near the flat band voltage was calculated for the three devices shown in Fig. 3.3 by the conductance method (Fig. 3.4b)[25]. The samples with 30 min of S-passivation have lower average D_{it} values compared to 15 and 60 min of passivation times, although the difference between 15 and 30 min samples is within the error margins. In addition, increasing the sulfur passivation time to 60 min led to at least an order of magnitude larger leakage current in passivation (Figure 3.4a). However, leakage current in accumulation for the sample with 30 min sulfur passivation was slightly lower than 15 min. Considering all the electrical characteristics, S-passivation longer than 30 min was found to deteriorate the interface quality.

While Ni gates assist H_2 dissociation during the forming gas anneal (FGA) process by formation of atomic H that can potentially passivate the Si and Ge dangling bonds at HfO_2/SiGe interface, it might not lead to significant Ge–O reduction at the interface. Alternatively, Ti and TiN have large oxidation enthalpies and are typically used as oxygen gettering layers in Si-based MOS devices[26]. Therefore, to determine the effect of gate electrode chemistry on the quality of HfO_2/SiGe interfaces, Pd/Ti/TiN stack was tested as the gate electrode on top of HfO_2. Figure 3.5 displays the C-V characteristics of HF-treated samples with 45 cycles of HfO_2 with Ni vs Pd/Ti/TiN gate electrodes, before and after FGA at 250 ºC for 15 min. Before FGA, the sample with Pd/Ti/TiN had higher C_{max} and wider false inversion bump (from 0-1.5 V) than the sample with Ni gate. Extension of false inversion bump beyond 1V before FGA could be due to the damage induced in HfO_2 during TiN, Ti and Pd magnetron sputtering[27]. After FGA, Pd/Ti/TiN sample maintained large C_{max}, but the height and width of D_{it} bump was reduced, consistent with partial reduction of the sputtering damage. Conversely, C-V characteristics of samples with Ni gates underwent no dramatic change after FGA.

Figure 3.6 shows the C-V characteristics of S-passivated samples with 45 cycles of HfO_2 and Ni and Pd/Ti/TiN gate electrodes, before and after FGA at 250 ºC for 15 min. Similar to HF-treated samples, replacing Ni with Pd/Ti/TiN resulted in higher C_{max} and wider false inversion bump (from -0.2-1.2 V) consistent with sputtering damage. Furthermore, the height and width of false inversion bump was significantly reduced by FGA. It should be noted that the frequency dispersion in accumulation in Figure 3.4d is typically due to series resistance effect[25]. Since similar back contact is used for Al samples, this series resistance could be caused by slightly thicker or less conductive TiN layer in the

gate electrode. Comparing Figure 3.5d and 3.6d, it can be concluded that S-passivation results in lower D_{it} than HF-treated samples regardless of the gate electrode chemistry.

D_{it} distribution curves for samples with Ni and Pd/Ti/TiN gate electrodes, before and after FGA, are shown in Figure 3.7a and 3.7b. Both HF-treated and S-passivated devices with Pd/Ti/TiN had modest trap densities above the mid-gap (E – EV > 0.55 eV), consistent with the wide D_{it} bumps observed in their low-frequency C-V behavior. The interface trap density above mid-gap was reduced by factors of 2-3 after FGA. However, for HF-treated samples with Pd/Ti/TiN, the D_{it} values above mid-gap are always higher than the D_{it} values for devices with Ni gates. For S-passivated samples after FGA, the D_{it} levels above mid-gap are nearly identical between devices with Ni and Pd/Ti/TiN gates (red and green dotted lines). Nevertheless, Pd/Ti/TiN gates caused larger density of interface traps adjacent to the valence band edge, and as a result, lower interface quality. Comparison between Figure 3.7a and 3.7b confirms lower D_{it} for S-passivated samples relative to HF-treated samples, in agreement with their smaller D_{it} bumps in C-V results. Therefore, D_{it} distribution is found to be more strongly dependent on the surface passivation chemistry vs. gate electrode gettering capabilities.

Figure 3.7c displays the EOT before and after FGA for MOSCAPs with Ni and Pd/Ti/TiN gates that are cleaned by HF and HF+S treatments. Consistent with the qualitative C-V comparison, samples with Pd/Ti/TiN gates before FGA have lower EOT compared to the ones with Ni gates. For HF-treated sample with Pd/Ti/TiN gate electrode, EOT values as low as 1.05 nm are achieved. It should be noted that the EOT of 1.05 nm is smaller than the EOT for Ni gates on 40 cycles of HfO_2 (Figure 3.2a). This reduction in EOT can be attributed to gettering by TiN and Ti layers, that reduces amount of GeO_x

within HfO_2 and increases the equivalent dielectric constant[28,29]. Unlike HF-treated samples, S-passivated samples after FGA have the same EOT regardless of the gate electrode chemistry. This difference in behavior caused by FGA, could be due to the difference in the interface composition between the S-passivated and HF-treated samples. In general, HF-treated samples are expected to have a larger density of GeO_x within or on top of the gate oxide, which would be more readily affected by Ti/TiN gates gettering properties. In contrast, for S-passivated samples, there could be competition between Ge-S bonds at the interface and Ge gettering process by Ti/TiN, therefore no significant influence in the overall dielectric constant of HfO_2 was observed.

3.4 Conclusion

Wet sulfur passivation using ex-situ $(NH_4)_2S$ solution treatment has been applied to $Si_{0.7}Ge_{0.3}$(100) surfaces prior to atomic layer deposition of HfO_2 at 300 ºC. Compared to cyclic HF treatment, which is used to remove the native oxide, S-passivation was found to reduce the density of interface traps near the edge of the valence band by a factor of 2. In addition, by S-passivation accumulation leakage currents as low as 0.2 A/cm^2 were achieved which is about 2-3x smaller than the ones for HF-treated devices. Effect of Pd/Ti/TiN gate electrode as a strong gettering stack was compared to Ni. While using Pd/Ti/TiN led to record EOT value of 1.05 nm for HfO_2 directly deposited on SiGe(001), the D_{it} distribution was found to be stronger function of surface treatment than gate electrode composition for post-deposition annealing temperatures used in this study.

3.5 Acknowledgements

The authors would like to thank Larry Grissom, Ivan Harris, and Sean Parks of UCSD Calit2 for their assistance with ALD and metal deposition at the UCSD Nano3 cleanroom facility. SiGe wafers were provided by Applied Materials Inc. This work is partially supported by a gift from Applied Materials Inc. and Taiwan Semiconductor Manufacturing Co.

Chapter Three, in part is currently being prepared for submission for publication of the material. K. Sardashti, M. Clemons, A. C. Kummel, "Atomic Layer Deposition of HfO_2 on Sulfur-passivated $Si_{0.7}Ge_{0.3}$ Surfaces".

3.6 Figures

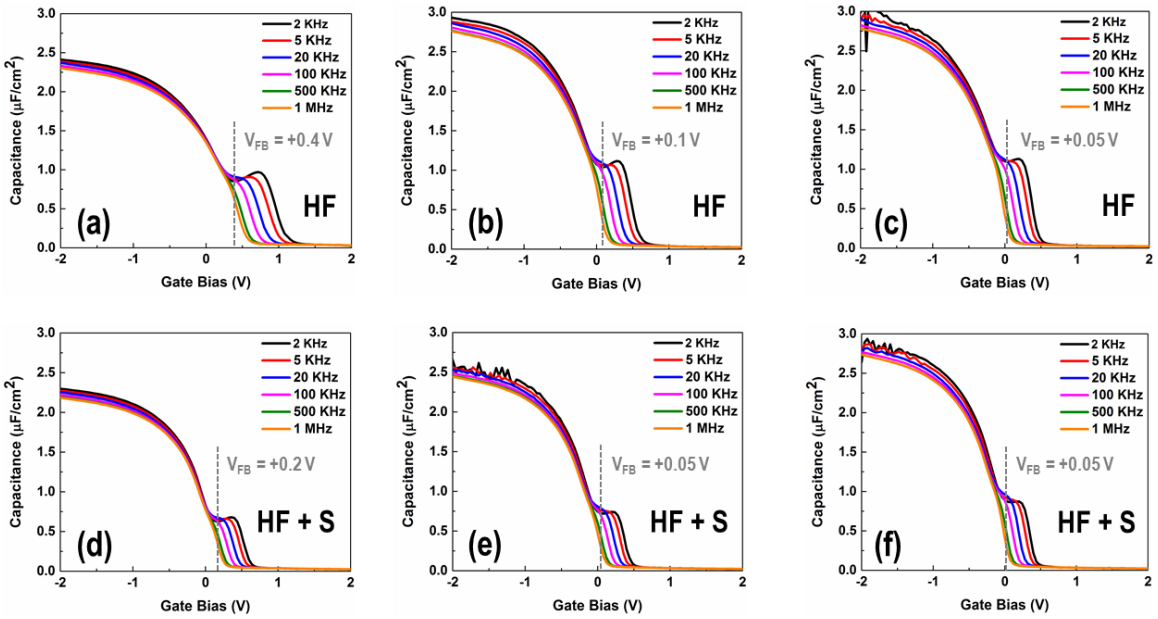


Figure 3.1 Multi-frequency C-V measurements for HfO_2 on $Si_{0.7}Ge_{0.3}$(100): (a, d) 60 cycles, (b, e) 45 cycles and (c, f) 40 cycles of HfO_2 ALD on HF-treated and S-passivated surfaces. For both preparation methods, the C_{ox} increased as the number of ALD cycles was reduced. For each thickness, the false inversion bump associated with D_{it} is about 2x lower for the S-passivated surfaces than the HF-treated surfaces.

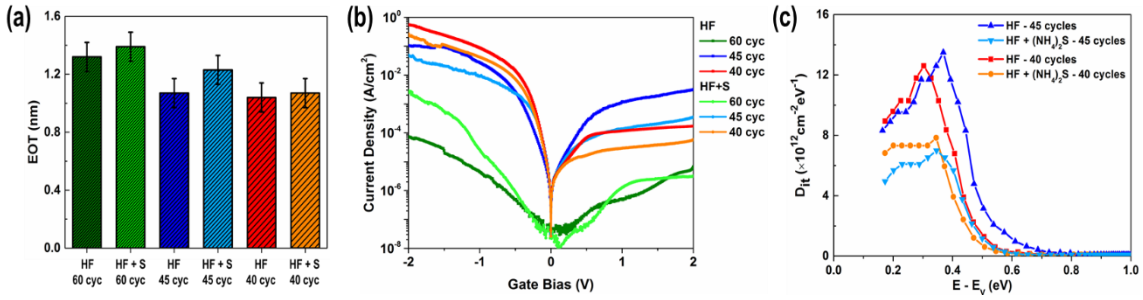


Figure 3.2 Device Characteristics with HF and Sulfur Cleaning from Full State Model: (a) Equivalent oxide thickness (EOT) as a function of number of ALD cycles for HF-treated and S-passivated $Si_{0.7}Ge_{0.3}$(100) samples; (b) Gate leakage characteristics for 60, 45 and 40 cycles of HfO_2 deposited on SiGe prepared with HF-treatment and S-passivation; (c) Density of interface traps vs. Fermi energy level for samples with 45 and 40 cycles of HfO_2 deposited on HF-treated and S-passivated SiGe surfaces. For 40 and 45 cycle films, the sulfur treatments gave lower leakage and lower D_{it}.

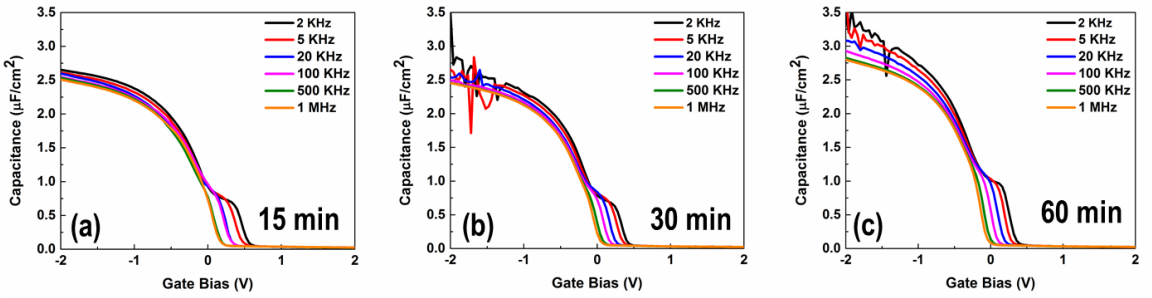


Figure 3.3 Multi-frequency C-V measurements for 45 cycles of HfO_2 ALD on $Si_{0.7}Ge_{0.3}$(100) with various sulfur passivation time in $(NH_4)_2S$ solution: (a) 15 min; (b) 30 min and (c) 60 min. The higher noise for the 30 and 60 min treatments is due to greater leakage or poor mechanical stability.

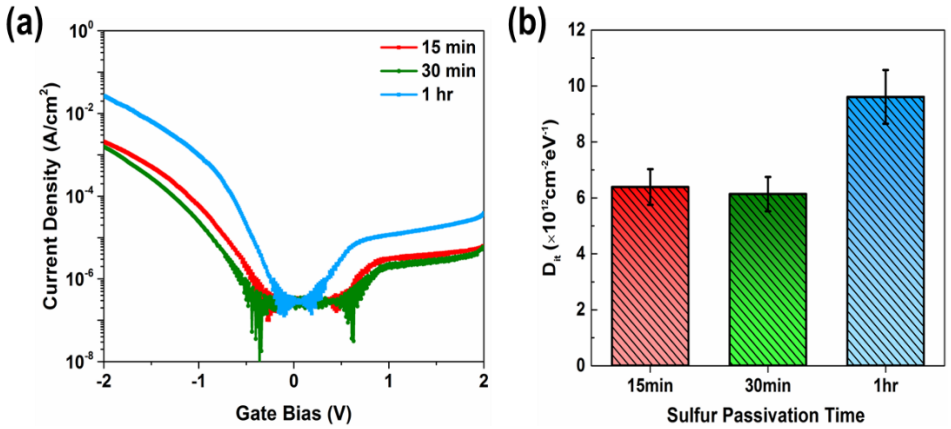


Figure 3.4 (a) Gate leakage characteristics for 45 cycles of HfO_2 deposited on $Si_{0.7}Ge_{0.3}$(100) prepared with various S-passivation times; (b) Average D_{it} calculated by conductance method for devices with 15, 30 and 60 min of wet sulfur passivation followed by 45 cycles of HfO_2 deposition.

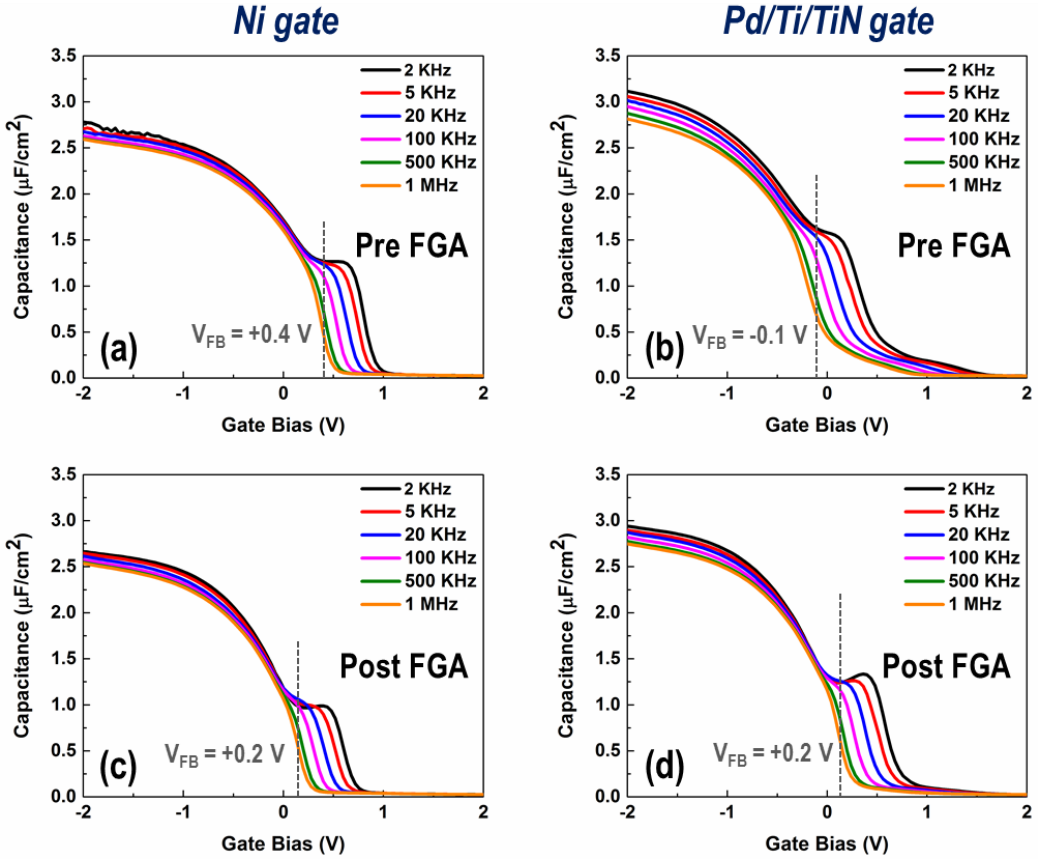


Figure 3.5 Multi-frequency C-V measurements for HfO_2 on HF-treated $Si_{0.7}Ge_{0.3}(100)$ with Ni and Pd/Ti/TiN (a, c) before FGA and (b, d) after 250 ºC for 15 min FGA. Note the Pd/Ti/TiN gate results in C_{ox} being constant before and after FGA while the false inversion bump was reduced by FGA.

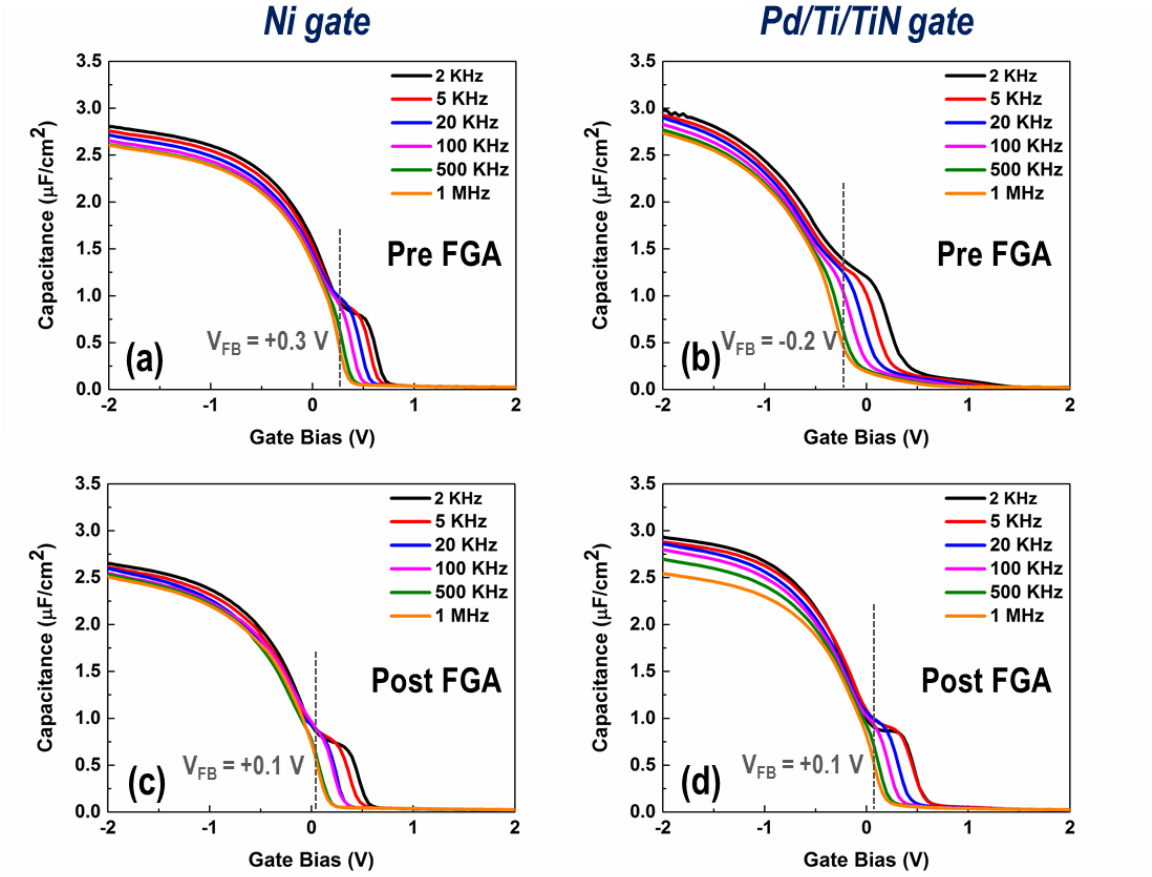


Figure 3.6 Multi-frequency C-V measurements for HfO_2 on S-passivated $Si_{0.7}Ge_{0.3}(100)$ with Ni and Pd/Ti/TiN: (a, c) before FGA and (b, d) after 250 ºC for 15 min FGA S-passivation was carried out by immersing the sample in $(NH_4)_2S$ for 15 min. Note for both treatments, C_{ox} is constant before and after FGA but the Pd/Ti/TiN gate results in the false inversion bump was reduced by FGA.

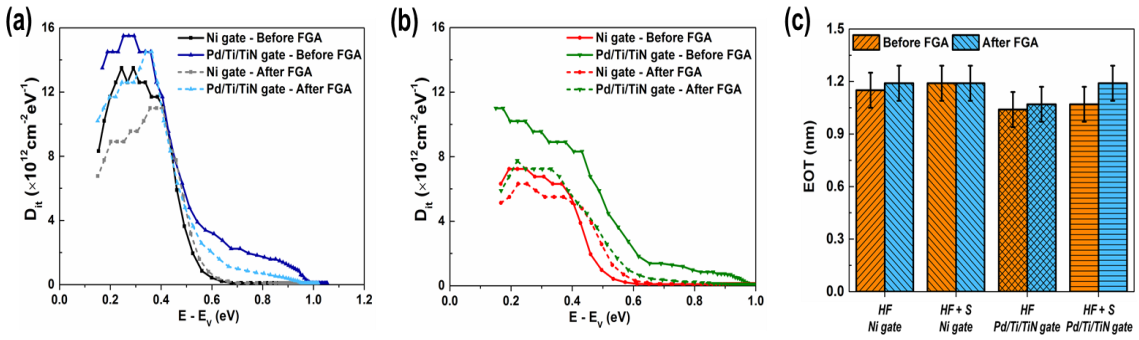


Figure 3.7 Device Characteristics with Ni vs Pd/Ti/TiN Gates from Full State Model (a) D_{it} vs. energy curves for samples with 45 cycles of HfO_2 deposited on HF-treated $Si_{0.7}Ge_{0.3}(100)$ surfaces with and without FGA at 250 ºC for 15 min; (b) D_{it} vs. energy curves for samples with 45 cycles of HfO_2 deposited on S-passivated SiGe surfaces with and without FGA at 250 ºC for 15 min; (c) EOT before (orange columns) and after FGA (blue columns) for HF-treated and S-passivated samples with Ni and Pd/Ti/TiN gates.

REFERENCES

[1] H. Riel, L. E. Wernersson, M. W. Hong, and J. A. del Alamo, Mrs Bull **39** (8), 668 (2014).

[2] B Majkusiak, J Walczak, D Flandre, AN Nazarov, and PLF Hemment, Science and Technology of Semiconductor-On-Insulator Structures and Devices Operating in a Harsh Environment **185**, 309 (2005).

[3] J. R. Brews, Solid State Electron **21** (2), 345 (1978).

[4] Yuan Taur and Tak H. Ning, *Fundamentals of modern VLSI devices*. (Cambridge University Press, Cambrige, UK ; New York, 1998), pp.xxii.

[5] Jean-Pierre Colinge, *FinFETs and other multi-gate transistors*. (Springer, New York, 2008), pp.xiii.

[6] M. Radosavljevic, G. Dewey, D. Basu, J. Boardman, B. Chu-Kung, J. M. Fastenau, S. Kabehie, J. Kavalieros, V. Le, W. K. Liu, D. Lubyshev, M. Metz, K. Millard, N. Mukherjee, L. Pan, R. Pillarisetty, W. Rachmady, U. Shah, H. W. Then, and R. Chau, 2011 Ieee International Electron Devices Meeting (Iedm) (2011).

[7] Yuan Taur, IEEE Electron Device Letters **21** (5), 245 (2000).

[8] Michel Houssa, *High-K gate dielectrics*. (Institute of Physics, Bristol ; Philadelphia, 2004), pp.xi.

[9] S. M. Sze and Kwok Kwok Ng, *Physics of semiconductor devices*, 3rd ed. (Wiley-Interscience, Hoboken, N.J., 2007), pp.x.

[10] Y. Fukuda, T. Ueno, S. Hirono, and S. Hashimoto, Jpn J Appl Phys 1 **44** (9b), 6981 (2005).

[11] K. Ismail, J. O. Chu, and B. S. Meyerson, Applied Physics Letters **64** (23), 3124 (1994).

[12] G. Hock, E. Kohn, C. Rosenblad, H. von Kanel, H. J. Herzog, and U. Konig, Applied Physics Letters **76** (26), 3920 (2000).

[13] M. Leskela and M. Ritala, Angew Chem Int Edit **42** (45), 5548 (2003).

[14] M. A. Alam and M. L. Green, J Appl Phys **94** (5), 3403 (2003).

[15] S. M. George, Chem Rev **110** (1), 111 (2010).

[16] N. Pinna and M. Knez, *Atomic Layer Deposition of Nanostructured Materials*. (Wiley, 2012).

[17] Kasra Sardashti, Kai-Ting Hu, Kechao Tang, Sangwook Park, Hyonwoong Kim, Shailesh Madisetti, Paul McIntyre, Serge Oktyabrsky, S Siddiqui, and B Sahu, Applied Surface Science (2016).

[18] Y. Taur, H. P. Chen, Q. Xie, J. Ahn, P. C. McIntyre, D. Lin, A. Vais, and D. Veksler, Ieee T Electron Dev **62** (3), 813 (2015).

[19] H. P. Chen, Y. Yuan, B. Yu, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, Ieee T Electron Dev **59** (9), 2383 (2012).

[20] R. Engel-Herbert, Y. Hwang, and S. Stemmer, J Appl Phys **108** (12) (2010).

[21] Dieter K. Schroder, *Semiconductor material and device characterization*, 3rd ed. (IEEE Press ; Wiley, Piscataway, NJ Hoboken, N.J., 2006), pp.xv.

[22] Y. Yuan, B. Yu, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, Ieee T Electron Dev **59** (8), 2100 (2012).

[23] M. Houssa, D. Nelis, D. Hellin, G. Pourtois, T. Conard, K. Paredis, K. Vanormelingen, A. Vantomme, M. K. Van Bael, J. Mullens, M. Caymax, M. Meuris, and M. M. Heyns, Applied Physics Letters **90** (22) (2007).

[24] K. Martens, C. O. Chui, G. Brammertz, B. De Jaeger, D. Kuzum, M. Meuris, M. M. Heyns, T. Krishnamohan, K. Saraswat, H. E. Maes, and G. Groeseneken, Ieee T Electron Dev **55** (2), 547 (2008).

[25] E. H. Nicollian and J. R. Brews, *MOS (metal oxide semiconductor) physics and technology*. (Wiley, New York, 1982), pp.xv.

[26] I. S. Park, Y. C. Jung, M. Lee, S. Seong, and J. Ahn, Iop Conf Ser-Mat Sci **54** (2014).

[27] K. Mitani and T. Kawano, Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers **34** (9a), 4649 (1995).

[28] C. Caperaa, G. Baud, J. P. Besse, P. Bondot, P. Fessler, and M. Jacquet, Mater Res Bull **24** (11), 1361 (1989).

[29] B. G. Segda, C. Caperaa, M. Jacquet, and J. P. Besse, Nucl Instrum Meth B **266** (21), 4829 (2008).