Engineering high-k/SiGe interface with ALD oxide for selective GeO_x reduction

Mahmut S. Kavrik[1], Peter Ercius[2], Joanna Cheung[1], Kechao Tang[3], Qingxiao Wang[4], Bernd Fruhberger[5], Moon Kim[4], Yuan Taur[1], Paul C. McIntyre[3] and Andrew C. Kummel[1]*

[1]Materials Science and Engineering, University of California San Diego, La Jolla, California, 92093, United States

[2]National Center for Electron Microscopy, Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California, 94720, United States

[3]Materials Science and Engineering, Stanford University, Stanford, California, 94305, United States

[4]Material Science and Engineering, University of Texas, Dallas, Texas 75080-3021, United States

[5]California Institute for Telecommunications and Information Technology, University of California San Diego, La Jolla, CA 92093, United States

Corresponding Author: Andrew Kummel, akummel@ucsd.edu

KEYWORDS: SiGe CMOS, low-power electronics, high-mobility transistor, high-k dielectrics, Al_2O_3, HfO_2 interface trap charge, atomic layer deposition

Abstract

Suppression of electronic defects induced by GeO_x at the high-k gate oxide/SiGe interface is critical for implementation of high mobility SiGe channels in CMOS technology. Theoretical and experimental studies have shown that a low defect density interface can be formed with an SiO_x-rich interlayer on SiGe. Experimental studies in literature indicates better interface formation with Al_2O_3 in contrast to HfO_2 on SiGe however the mechanism behind this is not well understood. In this study, the mechanism of forming a low defect density interface between Al_2O_3/SiGe is investigated using atomic layer deposited (ALD) Al_2O_3 insertion into or on top of ALD HfO_2 gate oxides. To elucidate the mechanism, correlations are made between the defect density determined by impedance measurements and the chemical and physical structure of the interface determined by high resolution scanning transmission electron microscopy and electron energy loss spectroscopy (STEM – EELS). Compositional analysis reveals an SiO_x rich interlayer for both Al_2O_3/SiGe and HfO_2/SiGe interfaces with insertion of Al_2O_3 into or on top of the HfO_2 oxide. The data is consistent with the Al_2O_3 insertion inducing decomposition of the GeO_x from the interface to form an electrically passive, SiO_x rich interface on SiGe. This mechanism shows that nanolaminate gate oxide chemistry cannot be interpreted as resulting from a simple layer by layer ideal ALD process because the precursor or its reaction products can diffuse though the oxide during growth and react at the semiconductor interface. This result shows that in scaled CMOS, remote oxide ALD (oxide ALD on top of the gate oxide) can be used to suppress electronic defects at gate-oxide semiconductor interfaces by oxygen scavenging.

Introduction

SiGe alloys are employed as stressor layers in mainstream complementary metal–oxide–semiconductor (CMOS) transistors and are being investigated as p-type field effect transistor (FET) channels due to their high mobility[1] and ease of integration into CMOS[2]. Thermally stable HfO_2 gate oxides with high dielectric constants reduce CMOS device power consumption[3-4]. SiGe p-FETs with high-k gate dielectrics which have low defect interfaces can provide better electrostatic control of the channel and higher drive current for low gate bias voltage. Conversely, a high density of interface defects between the high k gate oxide and the SiGe channel degrades device performance metrics such as subthreshold slope and reduces the on/off current ratio[5]. The main challenge for implementing SiGe FETs is the binary atom termination (Si-Ge) of the surface which results in formation of $SiGeO_x$ mixed oxides and dangling bonds on both Si and Ge atoms[6-8]. GeO_x and associated dangling bonds are the main sources of defects producing interface trapped charge (D_{it}), while SiO_x is a stable oxide that forms a nearly defect-free interface according to theoretical calculations[9].

Previously, several techniques such as nitride and sulfur passivation on $Si_{0.7}Ge_{0.3}(001)$ were studied with Al_2O_3 gate oxides and reduction in the interface defect density via suppression of GeO_x formation was reported [10-11]. However similar low defect density interfaces could not be established with HfO_2 gate oxide. This is because oxygen containing species such as excess H_2O, OH, and/or O can diffuse through HfO_2 during atomic layer deposition (ALD), forming GeO_x defects on the SiGe surface; in addition, the nature of HfO_2 allows diffusion of Ge and GeO_x to the top surface of the oxide as a result of reaction with HfO_2 and GeO_2 decomposition[11-14] Recently, HfO_2/SiGe interfaces formed with Al_2O_3-HfO_2 nanolaminate gate dielectric stacks were found to have a low interface state density, and it was hypothesized that the mechanism was

reduction of GeO_x out-diffusion during ALD[15]. Theoretical DFT models of the amorphous $HfO_2/Si_{0.5}Ge_{0.5}(001)$ interface have shown that low-defect interfaces may be formed even before hydrogen passivation with short anneals (<10 ps) when the interface is comprised solely of SiO (silicon monoxide)[16]. Experimental studies have shown a 10× D_{it} reduction at the Al_2O_3/SiGe and HfO_2/SiGe interfaces via selective oxygen scavenging by using an oxygen-scavenging, metallic Al gate[17]; the selectivity is due to the difference in formation enthalpy of GeO_x compare to SiO_x facilitating transfer of oxygen from GeO_x to Al.[6, 17] However, this process induced thicker gate oxides and reduced the maximum capacitance density (C_{max}) resulting from Al oxidation on top of ALD grown gate oxide.

In the present study, formation of low interface defect density HfO_2/SiGe gate stacks using inorganic and organometallic based ALD Al_2O_3 insertion in bilayers and nanolaminates of Al_2O_3 and HfO_2 was investigated with STEM-EELS analysis. It is found that insertion of the highly oxygen reactive trimethyl aluminum (TMA) ALD precursor for Al_2O_3 in HfO_2 containing gate stacks reduces defects consistent with remote selective oxygen scavenging from the interface. This new selective oxygen scavenging technique is most effective when Al_2O_3 layers are uniformly distributed across the HfO_2 in a nanolaminate (NL) structure, but it also is effective when the Al_2O_3 ALD deposition occurs on top of the HfO_2. In Al_2O_3 ALD, during each TMA half cycle, the TMA is dosed in excess; therefore, after the surface hydroxyl groups are eliminated, the TMA is available to reduce additional species either by diffusion through HfO_2 or remotely at the growth surface. This is consistent with GeO_x out diffusing to the top of HfO_2 gate oxides[11, 13, 18], and the TMA remotely reducing the GeO_x during each half cycles[13]. It is found that using the Al_2O_3 ALD prior to HfO_2 deposition is not as effective as using the Al_2O_3 ALD in the nanolaminate, consistent with the suggestion that gettering is most effective after deposition of sufficient oxide to act as an

H_2O barrier to reduce additional formation of GeO_x during water-based ALD. This hypothesis of the gate oxide acting as a H_2O barrier and TMA being a GeO_x reducing agent was supported by experiments with ALD of purely Al_2O_3 gates, because Al_2O_3 is a better diffusion barrier than HfO_2[13]. Oxygen vacancies in HfO_2 is well documented and prevention of Ge out diffusion with Al_2O_3 incorporation into gate oxide was shown[11, 13]. Once the Al_2O_3 gate oxide reaches a critical thickness, an additional 25% increase in oxide thickness results in a 4× reduction in D_{it} and nearly complete elimination of the $SiGeO_x$ interfacial layer. For HfO_2, the selective scavenging process benefits from the difference in formation enthalpy of SiO_x in comparison to GeO_x and reduces the interface trapped charge density by forming Si-rich SiO_x at the interface, consistent with the predictions of the DFT models[16]. This is also consistent with the known ability of TMA to reduce low enthalpy of formation oxides on substrates at the start of ALD, a process known as ALD cleanup[19-20].

Methods

Metal oxide semiconductor capacitor (MOSCAP) devices were fabricated with HfO_2 and Al_2O_3 oxides onto $Si_{0.7}Ge_{0.3}$(100) epitaxially grown on p-type Si (100) (Applied Materials). Samples were degreased by sonication in methanol for one minute followed by 10 second rinses in acetone, IPA, and deionized H_2O. Native oxides were removed by cycling 2.5 times through 1 min 2% HF solution and 1 min DI H_2O, ending with 2% HF. The samples were dried in N_2, passivated in an ammonium sulfide solution (25% $(NH4)_2S$) for 15 minutes, rinsed with water for 30 seconds, and dried with N_2. Subsequently, samples were transferred to the atomic layer deposition (ALD) chamber and oxide structures were grown by using a Beneq TFS200 cross flow reactor at 275ºC using tetrakis dimethyl amino hafnium (TDMAH), $HfCl_4$ and TMA metal precursors and water as an oxidant. Each HfO_2 cycle consisted of 1s of a TDMAH or $HfCl_4$ pulse and a 500ms of H_2O

pulse, each Al_2O_3 cycle consisted of a 1s TMA pulse and a 500ms H_2O pulse. Ar was used as carrier gas for all processes, and 6s purges were employed between each pulse. A set of samples were fabricated in bilayer, tri-layer and nanolaminate (NL) structures formed by Al_2O_3 and HfO_2 in gate stack for MOSCAPs as shown in Fig. S1 and the insets of Fig. 1. Nickel gates (50nm thick, 150um diameter) were deposited with a shadow mask onto the oxide surface using a Denton 502A thermal evaporator in vacuum < 2×10^{-6}. Al back contacts were deposited by sputtering after native oxide removal with Ar plasma at 100W, 5mTorr in Denton Discovery 635 sputtering system. Samples were annealed using an optimized recipe for 30 minutes total (10 min at 300°C, 10 min at 330°C, and 10 min at 350°C) in forming gas (5% H_2, 95%N_2) in a Ulvac MILA-3000 Minilamp annealer at 3 slpm at 1atm. MOSCAP characterization at 300 K was performed with Keysight B1500. Multi-frequency capacitance-voltage (C-V) and conductance-voltage (G-V) measurements were obtained from 2 kHz to 1MHz with 30mV AC signal superimposed on DC gate bias varied from 2V to -2V. For the structural and compositional analysis, TEM specimens (<50 nm) were prepared from MOSCAP devices using a FEI-Scios Ga focus ion beam (FIB). STEM/TEM analysis on high-k/SiGe gate stacks were carried out with FEI Titan 80-300, FEI Metrios and JEOL ARM 200CF. STEM High-angle annular dark field (HAADF) and bright field (BF) along with TEM images were obtained. Compositional analysis was performed with electron energy loss spectroscopy (EELS) both at 80keV and 200keV with JEOL ARM 200CF equipped with Gatan Quantum EELS spectrometer.

RESULTS AND DISCUSSION

Electrical characterization of the MOSCAP devices with various gate oxide structures after forming gas annealing (FGA) are shown in Fig. 1. The C-V analysis for 45 cycles (~4.5nm) of HfO_2 control sample with a C_{max} of 2.5 uF/cm^2 and corresponding G-V characterization are

presented in Fig.1a and 1f; a peak defect density of (D_{it}) 3.74×10^{12} $eV^{-1}cm^{-2}$ (see Figure 2) is obtained. The full interface state model[21-22] is used to calculate the energy distribution of the D_{it} by fitting multi frequency C-V and G-V curves for each bias point from inversion (2V) to accumulation (-1V). The result peak D_{it} values from the full interface state model are compared with the D_{it} values from conductance method[23] and shown to be in agreement as documented in table 1. Fig. 1b and 1c show the C-V for 5 cycle Al_2O_3 insertion below and above 45 cycles of HfO_2. As expected, Al_2O_3 insertion decreases C_{max} in both cases due to an increase in total oxide thickness and the lower dielectric constant of Al_2O_3 in comparison to HfO_2. However, the magnitude of the dispersive, depletion capacitance D_{it} feature also decreases as shown in Fig. 1b and 1c. Al_2O_3 insertion below or above HfO_2 has a nearly identical effect on interface defect density corresponding to peak D_{it} values of 3.30×10^{12} $eV^{-1}cm^{-2}$ and 3.15×10^{12} $eV^{-1}cm^{-2}$. Note full D_{it} distributions as a function of the thickness are shown below confirming the trends. Because interface trap response involves defects at the semiconductor oxide interface, a change in D_{it} resulting from the addition of 5 ALD cycles of Al_2O_3 on top of a 4.5 nm thick HfO_2 layer is unexpected. Moreover, as shown in Fig 1d, Al_2O_3 insertion both above and below HfO_2 in an Al_2O_3 /HfO_2/Al_2O_3 tri-layer structure further decreases the interface trapped charge density by 10% to 2.53×10^{12} $eV^{-1}cm^{-2}$. Furthermore, as shown in Fig 1e, when Al_2O_3 layers dispersed across the HfO_2 in the nanolaminate structure (NL), the D_{it} decreases further down to 2.22×10^{12} $eV^{-1}cm^{-2}$, 12% lower than the tri-layer with a small increase in C_{max}. The C_{max} reduction due to an increase in total oxide thickness is expected, but D_{it} reduction with more Al_2O_3 incorporation remote from the interface is not.

To better document the effects of bottom vs top Al_2O_3, samples were growth with 10 cycles of Al_2O_3 inserted either below the HfO_2 or above the HfO_2. This caused very large decreases in D_{it} as shown in Fig. 2. Again, the deposition of Al_2O_3, on top of the HfO_2 induced a 57% decrease

in the peak D_{it} as well as a 54% decrease in the integrated D_{it} while the Al_2O_3 inserted below the HfO_2 only induced a 40% decrease in the peak D_{it} as well as a 40% decrease in the integrated D_{it}. The integrated D_{it} are obtained by summation of the defect densities across the band gap in Fig.4; this is equivalent to integration of the areas under the D_{it} curves; these integrated D_{it} are a measure of the total defects induced in the band gaps. urves provides total defects induced in band gap.

For the 5 cycle Al_2O_3 insertion, although absolute defect density quantification is challenging, small decreases in the magnitude of the interface trap-related depletion capacitance feature while maintaining an essentially constant C_{max} suggests a lower defect density at the interface oxide/semiconductor interface for the nanolaminate structure compared to tri-layer sample or the bilayer samples. For each processing condition, 5-10 devices were studied. The 2-3 devices with the most consistent C-V were chosen for further analysis. Therefore, although absolute D_{it} calculations are accurate only within 30%; even 10% changes in D_{it} decay with Al_2O_3 insertion are reliable as confirmed via fabrication multiple sample sets as shown in Fig S8. The D_{it} standard errors of the mean (in /cm^2-eV and percentage) for samples were calculated as shown in table 1. The typical standard error is 3.9%, and changes in D_{it} of 10% are significant in this comparison. The D_{it} standard error analysis shown documents that changes in D_{it} of 10% can reliably be determined. Error bars are not given for each data point since the absolute D_{it} is less precisely determined in part because the uncertainty in C_{ox} is about 10% which can translate into 10-20% uncertainly in absolute D_{it}[24]. Additional details for standard error analysis are shown in Table 1. As shown in Table 1, there is no significant change for V_{fb} as a function of processing conditions. In addition, there is very little dependence of V_{fb} on frequency; however, when there is frequency dependence, the highest frequency C-V is employed to access V_{fb} since it is the least affected by traps.

To investigate the mechanism of D_{it} reduction due to insertion of Al_2O_3 ALD layers, MOSCAP devices with only Al_2O_3 gate oxides of varying thickness were prepared, as shown in

Fig. 3. As expected, C-V and G-V measurements from Al_2O_3 devices after FGA as a function of deposition cycles or oxide thickness show a decrease in C_{max} from 1.3uF/cm^2 to 1.0 uF/cm^2 by increasing the oxide thickness. The depletion capacitance D_{it} feature also decreases in amplitude. Once the ALD-grown oxide thickness increases above a critical value, there is a super-linear decrease in D_{it}. For an increase of 25% in the number ALD cycles (40 to 50 cycles), the D_{it} decreases by > 4× from 2.93×10^{12} to 0.67×10^{12} eV^{-1}cm^{-2}. The lower D_{it} from increasing the number of ALD cycle from 40 to 50 is not the result of a sudden increase in oxide thickness due to an ALD onset delay. The C_{max} for 50 cycles is just 25% lower than for 40 cycles consistent with nearly linear growth. Because the D_{it} originates from defects at the interface with the semiconductor, D_{it} reduction by growing additional Al_2O_3 layers is consistent with chemical modification of the interface by exposure to the TMA-based ALD environment. The C-V curves obtained from MOS capacitors containing Al_2O_3 before FGA shown in Fig. S4 exhibit a similar trend. It is noted that for Si interfaces, Al_2O_3 ALD is reported to reduce defects by a hydrogen passivation mechanism. However, for SiGe, the most important dangling bonds are on Ge which are not readily passivated by hydrogen as shown by the lowest D_{it} devices being ones which eliminate interfacial GeO_x using a gettering gate. This is also consistent with previous reports of extremely low D_{it} for HfO_2/SiGe by deposition of Al gettering metal on top of HfO_2 in which the FGA was performed after Al deposition[17].

The suppression of the interface defects with Al_2O_3 insertion into gate oxide influences is not limited to a single energy but is distributed across the band gap as shown in the D_{it} energy distributions extracted from measured C-V and conductance-voltage (G-V) data using the full interface state model[21] as shown in Fig. 4. Note that, interface defect distribution across the band gap for all the devices in this work is summarized in fig. 4 side by side for better comparison. In addition, the insertion of 1, 3, 5 and 10 monolayers of Al_2O_3 below HfO_2 (Fig. S2) and on top of HfO_2 (Fig. S3) was studied via C-V and G-V and indicate similar trends. Trapped charge energy distributions of the corresponding capacitors shown in Fig. 4a and b are consistent with Fig. 1 and show a decrease of D_{it} with incorporation of Al_2O_3 ALD layers but below and above HfO_2 layers. Inset D_{it} values in Fig. 4 c and d denote the integrated defect density across the band gap and are in good agreement with the behavior observed for the peak D_{it}. As shown in Fig 2 and 4, the shapes of the D_{it} distributions for HfO_2/SiGe and Al_2O_3/SiGe are similar (not identical), but the D_{it} values

vary more than 2x; this is consistent with the source of D_{it} being similar for both oxides (for example GeO_x) but present in greater concentration for HfO_2/SiGe than Al_2O_3/SiGe.

In Al_2O_3 ALD, during each TMA half cycle, the TMA is dosed in excess; therefore, after the surface hydroxyl groups are eliminated, the TMA is available to reduce additional species. However, during each H_2O half cycle, the H_2O is present in excess at the end of the pulse, and some H_2O or H_2O-derived species may diffuse through thin Al_2O_3 gate stack. Also, although the Al_2O_3 is better diffusion barrier in comparison to the HfO_2, it has been reported previously that GeO_x can diffuse through thin Al_2O_3 gate stacks[10]. The Al_2O_3 – SiGe interface is known to include GeO_x species which are source of interface defects[6]. It is hypothesized that above a critical thickness, the Al_2O_3 becomes a good barrier to H_2O; therefore, excess TMA exposure during ALD can scavenge oxygen from GeO_x species and convert it into Ge which might be redeposited on SiGe without formation of new GeO_x from the water pulses.[6] This mechanism is consistent with decreasing D_{it} by insertion of Al_2O_3 into HfO_2-containing gate dielectric stacks or deposition of Al_2O_3 on top of HfO_2 gate dielectrics. However, as noted earlier, Al_2O_3 is a more effective diffusion barrier in comparison to HfO_2 for oxygen containing species; consequently, the D_{it} suppression with HfO_2 dielectrics is not expected to be as great as with Al_2O_3 dielectrics of similar thickness[13]. This suggests that the number of Al_2O_3 layers incorporated into oxide is mainly responsible for the interface defect density reduction observed in these experiments. Similar to $6\times(9HfO_2+1Al_2O_3)$ NL gate dielectric structure, gate oxide without oxidant during Al_2O_3 ALD is performed by only TMA dosing and D_{it} reduction along with C_{max} observed indicating Al_2O_3 layer formation by oxygen scavenging." It is noted that for HfO_2 based gate stacks with incorporated Al_2O_3 layers the TMA may diffuse to the interface to directly reduce the GeO_x to Ge since HfO_2 is a poor diffusion barrier; however, the mechanism likely coexists with the remote scavenging mechanism.

Fig. 5 shows STEM high-angle annular dark-field (HAADF), bright field (BF) and TEM images of the HfO_2-only, HfO_2 - Al_2O_3 bilayers and nanolaminate gate stacks. White and black arrows in HAADF and BF images indicate the estimated oxide interfacial layer (IL) thickness. The SiGe-HfO_2 interface in Fig. 5a shows a 0.8 nm IL. Insertion of Al_2O_3 between SiGe and HfO_2 increases the apparent IL thickness (Fig. 5b) which is expected because the interface now consists of both Al_2O_3 and $SiGeO_x$. EELS or another spatially-resolved composition profiling method is needed to differentiate between these layers because of the similar atomic mass contrast of these oxides in HAADF-mode imaging. Al_2O_3 grown on HfO_2 and the Al_2O_3 - HfO_2 nano-laminate have thickness similar that shown in Fig 5.

Elemental profiles across the gate stacks were investigated with STEM-EELS analysis as shown in Fig. 6. A multiple linear least square (MLLS) fitting procedure[25] is used to resolve spectroscopic feature overlay issues, especially for Al, Hf and Si. The red dash line intercepts the half max of oxygen peak and is employed to indicate the SiGe surface. Black and green arrow denote Si and Ge composition of the SiGe surface. It should be noted that the determination of interface location in STEM-EELS is subject to interpretation especially due to surface roughness. However even with the surface roughness, the IL should be within few angstroms of the dashed lines, and this metric was confirmed by determining the location of the Si/SiGe interface since these layers are epitaxially grown. Any uncertainty in the exact location of the interface will not impact the trends in the elemental composition across the oxide which indicating suppression of Ge out diffusion with Al_2O_3 incorporation. Fig. 6a and 6b show 200keV EELS analysis of the same structures fabricated with hafnium tetrachloride at 300ºC and tetradimethylamido hafnium (IV) ALD precursors at 275ºC; no significant differences were observed as seen in Fig. 6 a and b consistent with interface defects obtained with multifrequency C-V analysis. Both the $HfCl_4$ and

TDMAH gate stacks have nearly identical elemental distribution across the device and have a Ge tail extending about 2 nm into the HfO_2 layer in contrast to the capacitors in Fig. 6c-f which show a diminished or zero Ge tail. This may be attributed to the Al_2O_3 layer impeding Ge out-diffusion[11, 13]. It is surprising that Al_2O_3 reduces Ge out diffusion even when Al_2O_3 is deposited onto HfO_2 (Fig 6c - 6f). This may be result of TMA reaction products (either AlO_x or AlC_xH_y) diffusion into HfO_2 reaching IL.

Further documentation that even remote Al_2O_3 ALD can control the GeO_x in the interlayer is observed in the composition of the interlayers. For all the gate stacks with Al_2O_3, the amount of Ge and the ratios of Si to Ge at the interface (between full max and 1/2 height of the oxygen peak) are greater than the control HfO_2/SiGe (Fig 6a); exact comparison of the interfacial Si/Ge ratio between the samples with Al_2O_3 (Fig 6b-f) is challenging since the ratio can vary with slight adjustment of the nominal interface position. All the samples have an interlayer above the SiGe as shown by the gap between the Hf and O edges; however, the Si signal falls off most steeply for Si on the nanolaminate samples (Fig S5-7) consistent with these samples having the most abrupt interface.

STEM-EELS at both 80 keV and 200keV is employed to investigate the AlO_x distribution in the ALD HfO_2 layer, while controlling beam induced Al damage due to high energy electrons[26]. For the HfO_2/Al_2O_3/SiGe structures (Figs 6c and 6d), the 200 keV spectrum shows complete diffusion of the Al while the 80 keV spectrum shows only a small retention of Al at the interfaces likely due to the limited signal to noise of the 80 keV spectra; the data is consistent AlO_x diffusing during HfO_2 deposition. For the Al_2O_3/HfO_2/SiGe structures (Figs 6e and 6f), the 200 keV spectrum shows two Al peaks, one above and one below HfO_2 while the 80 keV spectrum shows a just small retention of AlO_x above the HfO_2 layer, consistent with the limited signal to noise of

the 80 keV spectra. Overall both are consistent with diffusion of TMA or its reaction products into the HfO_2 when the Al_2O_3 ALD is performed after HfO_2 ALD. The nanolaminate gate stack (Fig.1e) composition studied with 80 keV EELS shows dispersed Al (Fig. S5).

This unexpected inter diffusion of Al_2O_3 is also shown with EELS raw data in 3D semi-log graph in Fig. 7 ($Ni/Al_2O_3/HfO_2$ /SiGe) and Fig S6 (Ni/HfO_2 / Al_2O_3/Ni). The oxide compositional profile can be seen from the electron energy loss peaks starting after element specific edges; for example, the Si K edge is at 1839 eV (orange arrow) and the Ge L edges at 1217 eV (pink arrow). The blue arrow indicates the $SiGe/HfO_2$ interface region. It is seen that the Ge peak decays earlier than the Si peak as a function of distance from the SiGe surface consistent with a SiO_x rich region at the interface. The black arrow indicates energy loss due to Al (K edge 1560 eV) across the oxide. The peak intensity is significantly lower in comparison to other elements; however, it can be seen in several regions in the HfO_2 and reaches a maximum close to the HfO_2/Ni interface since it is deposited on HfO_2 as shown in 3D data and the region of interest (RIO) in Fig 7 a-b. However, the Al peak can be also seen close to $SiGe/HfO_2$ interface, almost 4 nm away where it was deposited. The existence of Al signal in this region is consistent with AlO_x diffusion through HfO_2. A similar raw data analysis performed for Ni/HfO_2 /Al_2O_3 /SiGe bilayer device also show interdiffusion (Fig S6). However, comparison of EELS analysis for both bilayer structure indicates enhanced Al_2O_3-HfO_2 interdiffusion when Al_2O_3 is deposited prior to HfO_2 ALD.

Fig. 8 shows simultaneously acquired HAADF and BF-STEM images of Ni/Al_2O_3-/SiGe/Si (a-b) and Al/HfO_2/SiGe/Si (c-d) MOSCAP devices along with corresponding C-V graphs (e-f). In contrast to the Ni/HfO_2/SiGe/Si gate stack in Fig. 5, Al/HfO_2/SiGe/Si has a gettering gate which is known to remove oxygen from the interfaces and reduces IL thickness; this is supported

by EELS analysis (Fig. S6) indicating similar O and Hf peak decay profiles at the SiGe interface unlike with Ni gates which show offsets between O and Hf peak at the interface (Fig. 6a-f)[17].

The gate stacks in Fig. 8e and f have two very dissimilar oxides and show very low depletion capacitance feature resulting from charging/discharging of interface traps. Similar integrated D_{it} values are displayed in the inset. However, these two very different gate oxide structures deposited on SiGe show similar, almost abrupt interfaces with the underlying SiGe (Fig S6 and Fig 8g). For $Al/HfO_2/SiGe$ capacitors, it was previously shown that an Al metal gate can scavenge oxygen from the interfacial layer and reduce the interface defect density while also thinning the IL[17]. For the case of the $Ni/Al_2O_3/SiGe$ device, it is most likely that introduction Al_2O_3 ALD in the gate stack fabrication process provides an effect similar to that of an Al gate and scavenges oxygen from oxide/SiGe interface. The mechanisms are similar because TMA is a highly reactive precursor with oxygen which can interact with nearly all oxygen-containing molecules to form Al_2O_3. Insertion of each additional Al_2O_3 layer using TMA precursor can scavenge excess oxygen from the gate oxide or the high-k/SiGe interface. Similar to $6\times(9HfO_2+1Al_2O_3)$ NL gate dielectric structure, gate oxide without oxidant during Al_2O_3 ALD is performed by only TMA dosing and D_{it} reduction along with C_{max} observed indicating Al_2O_3 layer formation by oxygen scavenging. It is hypothesized that TMA exposure scavenges weakly bound oxygen from the interface either by diffusing into the interface as TMA or TMA reaction products (for example monomethyl aluminum) or it decomposes the GeO_x remotely, producing suboxide species that diffuse readily through even thin Al_2O_3 (remote gettering)[10]. As TMA interacts and scavenges oxygen from the interface, it is likely that GeO_x dissociates and donates oxygen to TMA due to the lower Gibbs free energy of formation of GeO_x in comparison to SiO_x.[27] Therefore, TMA can selectively scavenge

oxygen from the interface layer and reduce the interface trap density, while also thinning the IL, which has important benefits for gate stack dimensional scaling.

STEM-EELS compositional analysis for the Ni/Al_2O_3/SiGe/Si device shown in Fig. 8g supports this picture and indicates that a Si-rich interface forms, as shown by the intersection of the red dashed line marking the SiGe surface and half maximum count-rate of oxygen. The black and green arrows denote the Si and Ge composition at the oxide/SiGe interface, respectively. It is clearly seen that GeO_x composition is diminished significantly at the SiGe surface consistent with less diffusion of diffusion of GeO_x into the gate oxide and less diffusion of H_2O through the Al_2O_3 once the Al_2O_3 reaches a critical thickness.

Conclusion

Novel gate oxide structures were investigated which suppress electronic defects at high-k/SiGe interfaces by employing an oxygen scavenging ALD precursor, TMA. The approach utilizes the difference in the heat of formation of SiO_x and GeO_x, achieving lower interface trap densities at the high-k / SiGe interface with just a modest reduction of C_{max}. Although metallic Al remains more effective at oxygen scavenging, it induces a much larger C_{max} reduction, demonstrating the benefit of TMA remote oxygen scavenging. The data is consistent with insertion of Al_2O_3 into the HfO_2 gate oxide, removing a GeO_x component of the interlayer between the channel and the deposited gate dielectric, and suggests that effective oxygen scavenging can be achieved with TMA during ALD. This TMA based oxygen scavenging technique is most effective when the Al_2O_3 layers are uniformly distributed across the HfO_2 in a nanolaminate structure, but it also is effective when the Al_2O_3 ALD deposition occurs on top of the HfO_2. To achieve the effect, during each TMA half cycle of Al_2O_3 ALD, TMA is dosed in excess to provide sufficient TMA for

reduction of additional chemical species after the surface hydroxyl groups are eliminated. In the present study, for the first time, the mechanism for low defect interface formation with Al_2O_3 in contrast to HfO_2 is explained with experimental data. In contrast to previous reports in which the good diffusion barrier properties of Al_2O_3 are considered to be the source of D_{it} improvement, the mechanism for interface defect reduction with Al_2O_3 in the present work is shown to be selective oxygen scavenging by diffusion of TMA reaction products in HfO_2. Furthermore, direct and remote oxygen scavenging using ALD Al_2O_x deposited on HfO_2 gate oxide 4nm away from SiGe surface is demonstrated. By deposition of Al_2O_3 and etching back with atomic layer etching (ALE), it is likely that low D_{it} can be obtained in high aspect ratio structure such as on fins which is not possible with Al deposition for gettering gates.

ASSOCIATED CONTENT: Supplementary materials

The supporting information documenting variation of interface defect density on samples as well as comparison of interface defect calculation methods used in the article and the EELS analysis for few devices are available free of charge on the ACS Publication website at DOI:

ACKNOWLEDGEMENTS: Larry Grissom, Sean Parks of the Nano3 clean room at UCSD are gratefully acknowledged for technical support. Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Chung-Yang Lee, Yu-Chia Liang and Vincent Hou from TSMC is appreciated for the discussions on EELS data analysis. This work was funded by TSMC and performed in part at the San Diego Nanotechnology Infrastructure (SDNI) of UCSD, a member of the National Nanotechnology Coordinated Infrastructure, which is supported by the National Science Foundation (Grant ECCS-1542148)

References

(1) Fischetti, M. V.; Laux, S. E. Band Structure, Deformation Potentials, and Carrier Mobility in Strained Si, Ge, and SiGe Alloys. *Journal of Applied Physics* **1996,** *80* (4), 2234-2252,

(2) Lu, D. Silicon Germanium FinFET Device Physics, Process Integration and Modeling Considerations. *ECS Transactions* **2014,** *64* (6), 337-345

(3) Ceresoli, D.; Vanderbilt, D. Structural and Dielectric Properties of Amorphous h-k Oxides: HfO2, ZrO2 and Their Alloys. *Physical Review B* **2006,** *74* (12), 125108.

(4) 2.0, I. T. R. f. S. International Technology Roadmap for Semiconductors 2.0. **2015**.

(5) Taur, Y.; Ning, T. H. *Fundamentals of Modern VLSI Devices*, Cambridge University Press: 2013.

(6) Zhang, L.; Guo, Y.; Hassan, V. V.; Tang, K.; Foad, M. A.; Woicik, J. C.; Pianetta, P.; Robertson, J.; McIntyre, P. C. Interface Engineering for Atomic Layer Deposited Alumina Gate Dielectric on SiGe Substrates. *ACS Applied Materials & Interfaces* **2016,** *8* (29), 19110-19118,

(7) LeGoues, F. K.; Rosenberg, R.; Nguyen, T.; Himpsel, F.; Meyerson, B. S. Oxidation Studies of SiGe. *Journal of Applied Physics* **1989,** *65* (4), 1724-1728

(8) Cho, M.-H.; Chang, H. S.; Moon, D. W.; Kang, S. K.; Min, B. K.; Ko, D.-H.; Kim, H. S.; McIntyre, P. C.; Lee, J. H.; Ku, J. H.; Lee, N. I. Interfacial Characteristics of HfO2 Films Grown on Strained Si0.7Ge0.3 by Atomic-Layer Deposition. *Applied Physics Letters* **2004,** *84* (7), 1171-1173

(9) Chagarov, E. Density-Functional Theory Molecular Dynamics Simulations of a-HfO2/a-SiO2/SiGe and a-HfO2/a-SiO2/Ge with a-SiO2 and a-SiO Suboxide Interfacial Layers. *Journal of Physical Chemistry*. Vol. 443, 2018, Pg. 644-654

(10) Sardashti, K.; Hu, K.-T.; Tang, K.; Park, S.; Kim, H.; Madisetti, S.; McIntyre, P.; Oktyabrsky, S.; Siddiqui, S.; Sahu, B.; Yoshida, N.; Kachian, J.; Kummel, A. Sulfur Passivation for the Formation of Si-Terminated Al2O3/SiGe(001) Interfaces. *Applied Surface Science* **2016,** *366* (Supplement C), 455-463

(11) Lu, N.; Bai, W.; Ramirez, A.; Mouli, C.; Ritenour, A.; Lee, M. L.; Antoniadis, D.; Kwong, D. L. Ge Diffusion in Ge Metal Oxide Semiconductor with Chemical Vapor Deposition HfO2 Dielectric. *Applied Physics Letters* **2005,** *87* (5), 051922

(12) Zhang, Q.; Wu, N.; Lai, D. M. Y.; Nikolai, Y.; Bera, L. K.; Zhu, C. Germanium Incorporation in HfO2 Dielectric on Germanium Substrate. *Journal of The Electrochemical Society* **2006,** *153* (3), G207-G210,

(13) Ogawa, S.; Asahara, R.; Minoura, Y.; Sako, H.; Kawasaki, N.; Yamada, I.; Miyamoto, T.; Hosoi, T.; Shimura, T.; Watanabe, H. Insights into Thermal Diffusion of Germanium and Oxygen Atoms in HfO2/GeO2/Ge Gate Stacks and Their Suppressed Reaction with Atomically thin AlOx Interlayers. *Journal of Applied Physics* **2015,** *118* (23), 235704

(14) Cho, M. H.; Chang, H. S.; Moon, D. W.; Kang, S. K.; Min, B. K.; Ko, D. H.; Kim, H. S.; McIntyre, P. C.; Lee, J. H.; Ku, J. H.; Lee, N. I. Interfacial characteristics of HfO2 films grown on strained Si0.7Ge0.3 by atomic-layer deposition. *Applied Physics Letters* **2004,** *84* (7), 1171-1173

(15) Kwak, I.; Kavrik, M.; Thomson, E.; Liang, Y.-C.; Ueda, S. T.; Tang, K.; Hou, V.; Lee, C.-Y.; Aoki, T.; Kim, M.; Fruhberger, B.; Taur, Y.; McIntyre, P. C.; Kummel, A. C. HfO2/Al2O3 Nanolaminate on Si0.7Ge0.3(100) Surface by Thermal Atomic Layer Deposition. *Meeting Abstracts* **2018,** *MA2018-02* (31), 1080.

(16) Chagarov, E. A.; Kavrik, M. S.; Fang, Z.; Tsai, W.; Kummel, A. C. Density-Functional Theory Molecular Dynamics Simulations of a-HfO2/a-SiO2/SiGe and a-HfO2/a-SiO2/Ge with a-SiO2 and a-SiO Suboxide Interfacial Layers. *Applied Surface Science* **2018,** *443*, 644-654

(17) Kavrik, M. S.; Thomson, E.; Chagarov, E.; Tang, K.; Ueda, S. T.; Hou, V.; Aoki, T.; Kim, M.; Fruhberger, B.; Taur, Y.; McIntyre, P. C.; Kummel, A. C. Ultralow Defect Density at Sub-0.5 nm HfO2/SiGe Interfaces via Selective Oxygen Scavenging. *ACS Applied Materials & Interfaces* **2018,** *10* (36), 30794-30802

(18) Hua, W.; Lee, M. H.; Chen, P. S.; Maikap, S.; Liu, C. W.; Chen, K. M. Ge Outdiffusion Effect on Flicker Noise in Strained-Si nMOSFETs. *IEEE Electron Device Letters* **2004,** *25* (10), 693-695

(19) Hinkle, C. L.; Sonnet, A. M.; Vogel, E. M.; McDonnell, S.; Hughes, G. J.; Milojevic, M.; Lee, B.; Aguirre-Tostado, F. S.; Choi, K. J.; Kim, H. C.; Kim, J.; Wallace, R. M. GaAs Interfacial Self-Cleaning by Atomic Layer Deposition. *Applied Physics Letters* **2008,** *92* (7), 071901

(20) Milojevic, M.; Contreras-Guerrero, R.; Lopez-Lopez, M.; Kim, J.; Wallace, R. M. Characterization of the "Clean-up" of the Oxidized Ge(100) Surface by Atomic Layer Deposition. *Applied Physics Letters* **2009,** *95* (21), 212902

(21) Chen, H. P.; Yuan, Y.; Yu, B.; Ahn, J.; McIntyre, P. C.; Asbeck, P. M.; Rodwell, M. J. W.; Taur, Y. Interface-State Modeling of Al2O3 InGaAs MOS From Depletion to Inversion. *IEEE Transactions on Electron Devices* **2012,** *59* (9), 2383-2389

(22) Chen, H.-P.; Yuan, Y.; Yu, B.; Chang, C.-S.; Wann, C.; Taur, Y. Re-examination of the Extraction of MOS Interface-State Density by C–V Stretchout and Conductance Methods. *Semiconductor Science and Technology* **2013,** *28* (8), 085008

(23) E. H. Nicollian, J. R. B. *MOS (Metal Oxide Semiconductor) Physics and Technology*, Wiley-Interscience: New York, 1982.

(24) Han-Ping, C.; Yu, Y.; Bo, Y.; Chih-Sheng, C.; Clement, W.; Yuan, T. Re-examination of the extraction of MOS interface-state density by C – V stretchout and conductance methods. *Semiconductor Science and Technology* **2013,** *28* (8), 085008.

(25) Leapman, R. D.; Swyt, C. R. Separation of Overlapping Core Edges in Electron Energy Loss Spectra by Multiple-Least-Squares Fitting. *Ultramicroscopy* **1988,** *26* (4), 393-403

(26) Nan, J. Electron Beam Damage in Oxides: a Review. *Reports on Progress in Physics* **2016,** *79* (1), 016501.

(27) Reed, T. B. Free Energy of Formation of Binary Compounds: an Atlas of Charts for High-Temperature Chemical Calculations. *MIT Press [1971]* **1971**.

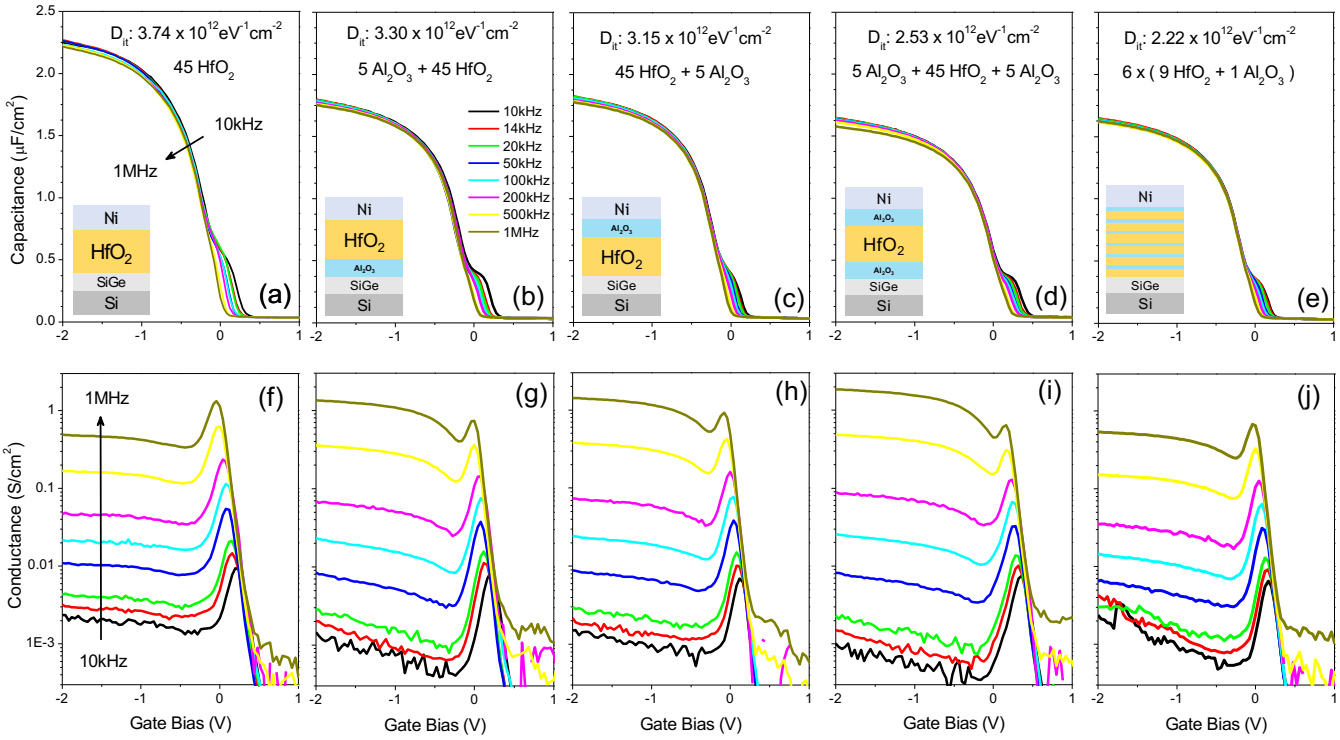

Fig. 1. C-V and G-V graphs of multilayer MOSCAP devices (a-j). Inset drawings indicate device structure for given graph. The inset D_{it} values indicates peak interface defect density value in bandgap obtained with full interface state model. The energy distribution of the defects density and the energy of the peak D_{it} values in band gap are shown in Fig.4. Control device with HfO_2 gate oxide shows highest interface defect density of 3.74 $\times 10^{12} eV^{-1}cm^{-2}$. Al_2O_3 insertion into HfO_2 reduces D_{it} gradually by going from HfO_2-Al_2O_3 bilayer into Al_2O_3-HfO_2-Al_2O_3 tri-layer and finally nanolaminate gate oxides to D_{it}= 2.22 $\times 10^{12}$ $eV^{-1}cm^{-2}$. Al_2O_3 insertion reduces D_{it} regardless of its position shows

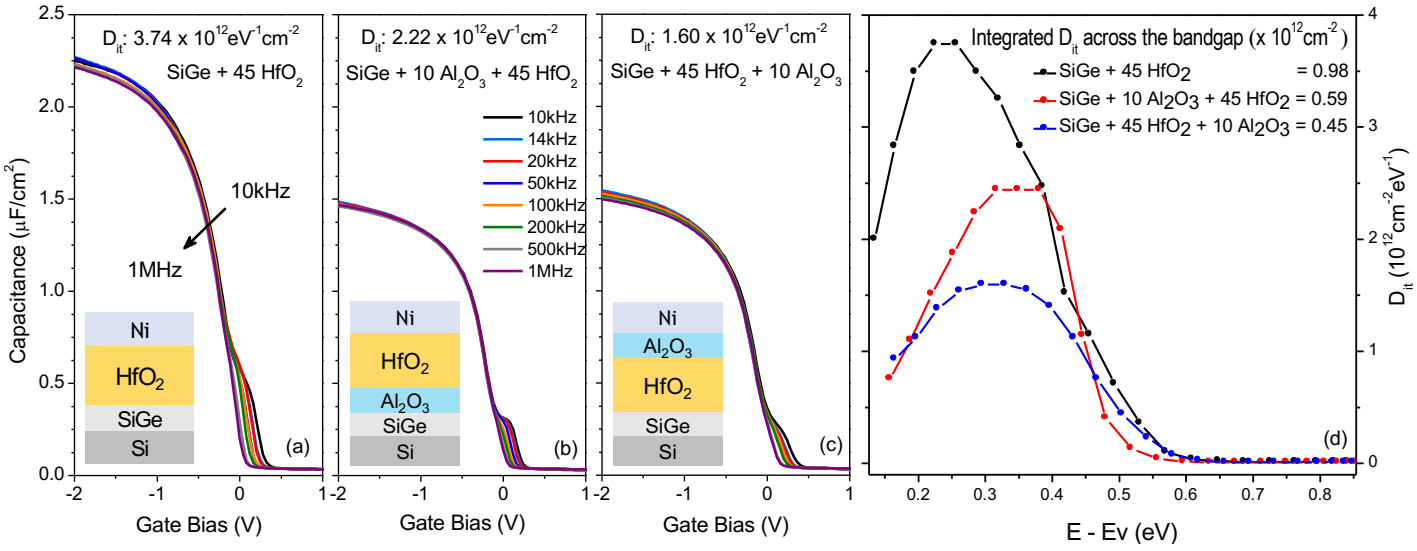

Fig. 2. C-V graphs and Interface defect density distributions across the band gap for MOSCAP devices calculated with the full interface state model. Inset drawings indicate device structure for given graph. The inset D_{it} values indicates peak interface defect density value in bandgap obtained with full interface state model (a) Control device with HfO_2 gate oxide shows highest interface defect density of D_{it} (peak) = 3.74 x10^{12}eV^{-1}cm^{-2}. (b) Interface defects at SiGe oxide interface decrease by insertion of 10 cycles of Al_2O_3 mono layers before HfO_2 D_{it} (peak) = 2.22 x10^{12}eV^{-1}cm^{-2} and (c) after HfO_2 D_{it}(peak) = 1.6 x10^{12}eV^{-1}cm^{-2} gate oxide. (d) Comparison of defect density at SiGe oxide interface across the band gap shows decrease in peak and integrated D_{it} (inset) by insertion of Al_2O_3 layers.

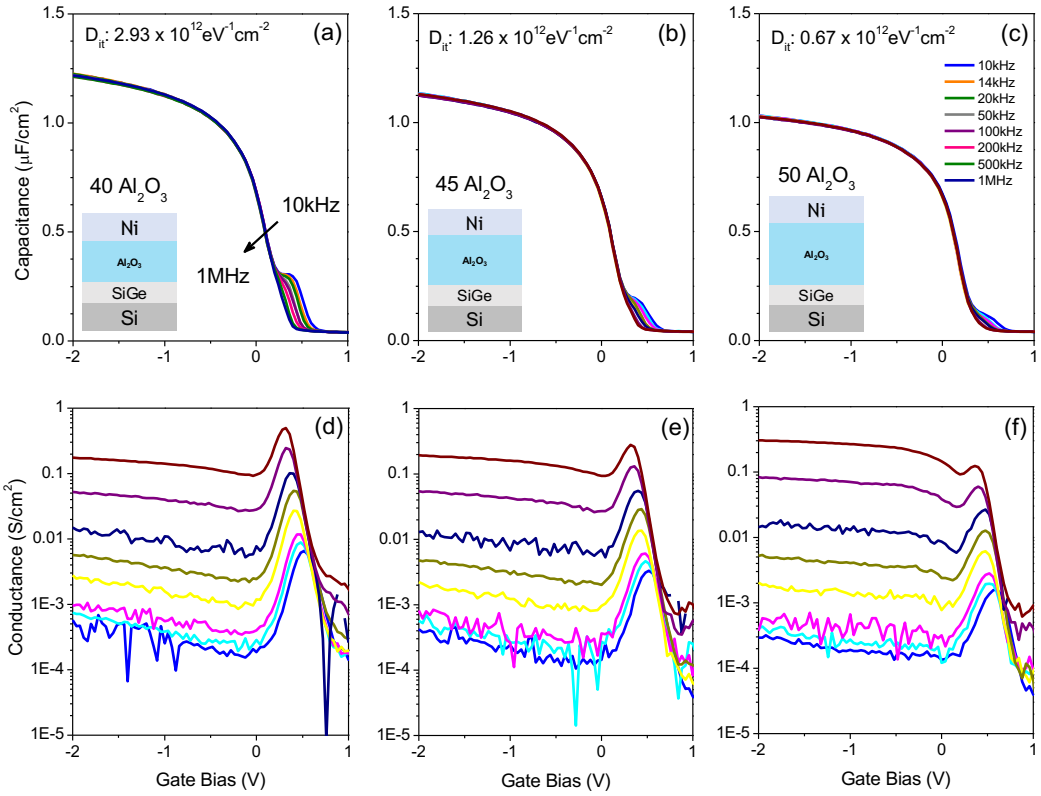

Fig. 3. C-V and G-V graphs of single oxide MOSCAP devices (a-f). Inset drawings indicate device structure for given graph. The inset D_{it} values indicates peak interface defect density value in bandgap obtained with full interface state model. Control device with 40 cycles of Al_2O_3 gate oxide shows highest interface defect density of 2.93 $\times 10^{12}$ $eV^{-1}cm^{-2}$. The small increase of Al_2O_3 thickness reduces D_{it} by > 4×; the lowest interface defect density of 0.67 $\times 10^{12}$ $eV^{-1}cm^{-2}$ is obtained with 50 cycles of Al_2O_3.

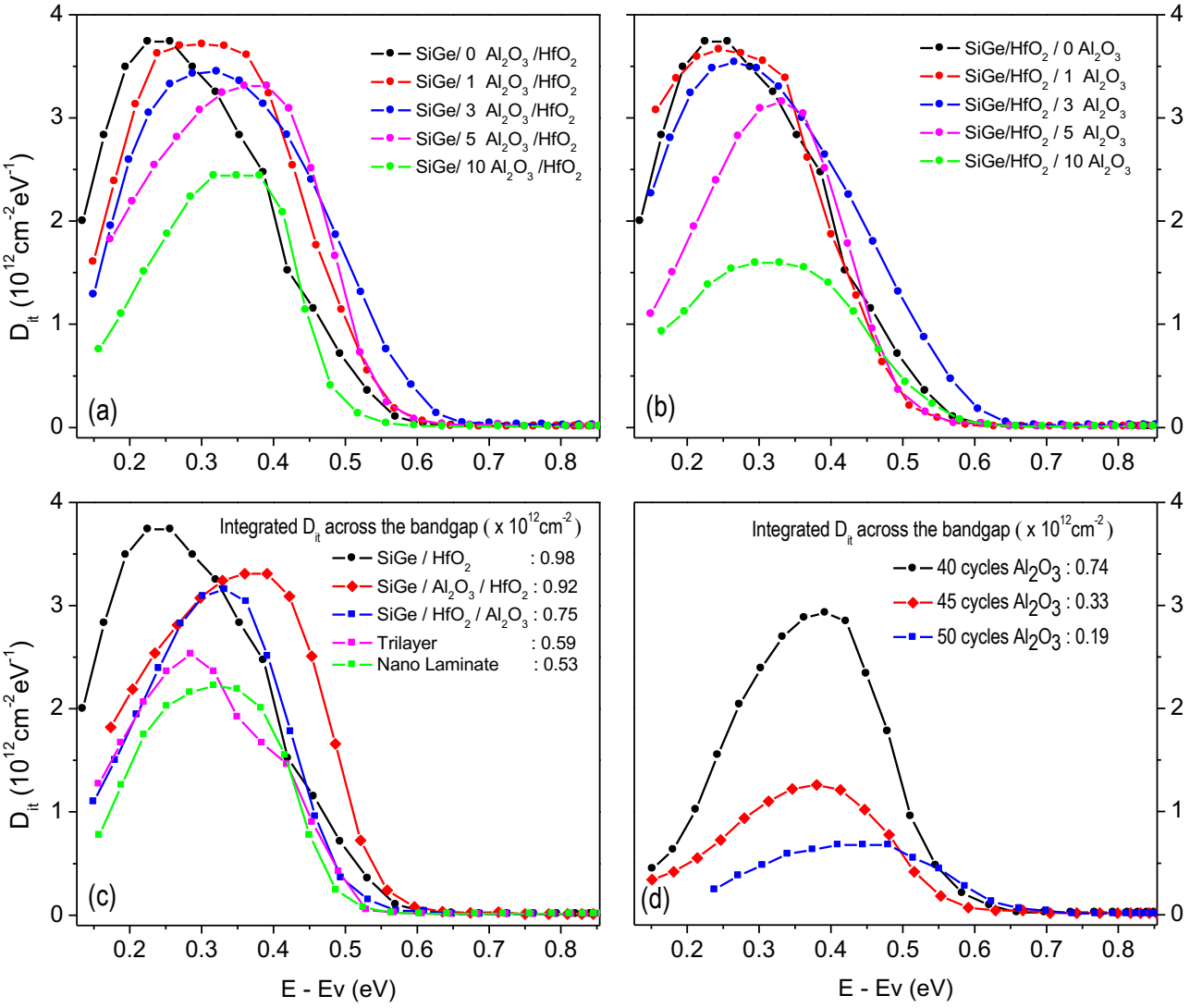

Figure 4. Interface defect density distributions across the band gap for MOSCAP devices calculated with the full interface state model. Interface defects at SiGe- oxide interface decrease by insertion of Al_2O_3 layers before HfO_2 (a) and after HfO_2 (b) gate oxide. (c) Comparison of interface defects variation at SiGe- oxide interface by insertion of Al_2O_3 layers into HfO_2 gate oxide. (d) Interface defect density decreases by increase in Al_2O_3 thickness. For 50 cycles of Al_2O_3 , peak D_{it} reduces to 6×10^{11} $eV^{-1}cm^{-2}$ and integrated defects across the bandgap is as low as 0.19×10^{11} $eV^{-1}cm^{-2}$.

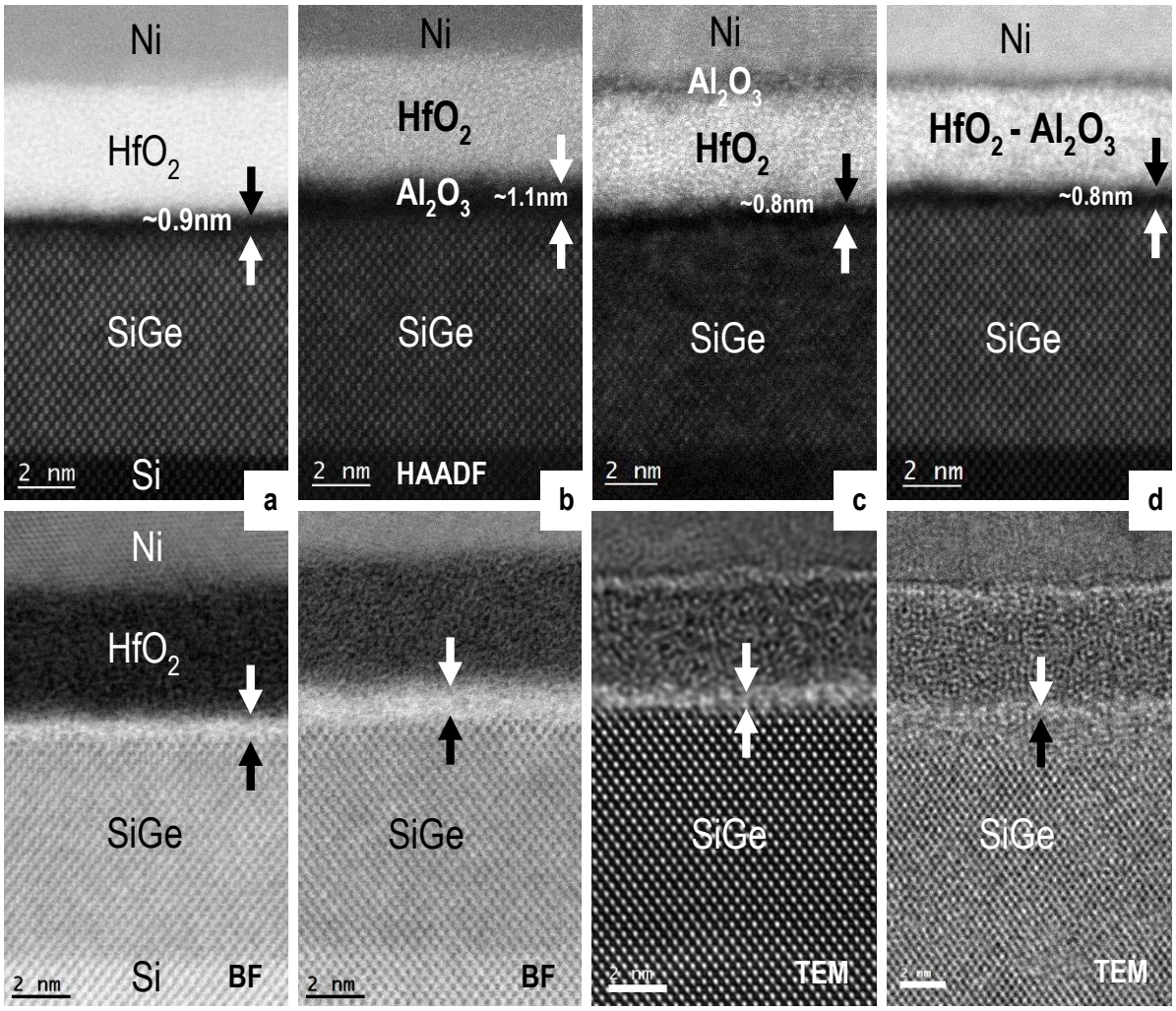

Figure 5. STEM HAADF (high-angle annular dark-field), BF (bright field) images of (a) control HfO_2 , (b) $HfO_2/Al_2O_3/SiGe$ bilayer, (c) $Al_2O_3/HfO_2/SiGe$ bilayer, (d) and Al_2O_3 – HfO_2 Nanolaminate MOSCAPs. In these images, oxide structures and regions are defined according to z contrast. The interfacial layer between SiGe and oxide indicated with black and white arrows on corresponding STEM – HAADF, STEM - BF and TEM image. Note in (b) the interlayer consists of both $SiGeO_x$ and Al_2O_3, so it appears thicker than the control device in (a). In comparison to control device of HfO_2/SiGe, bilayer (c) and NL (d) shows thinner interface consistent with remote Al_2O_3 insertion reducing IL.

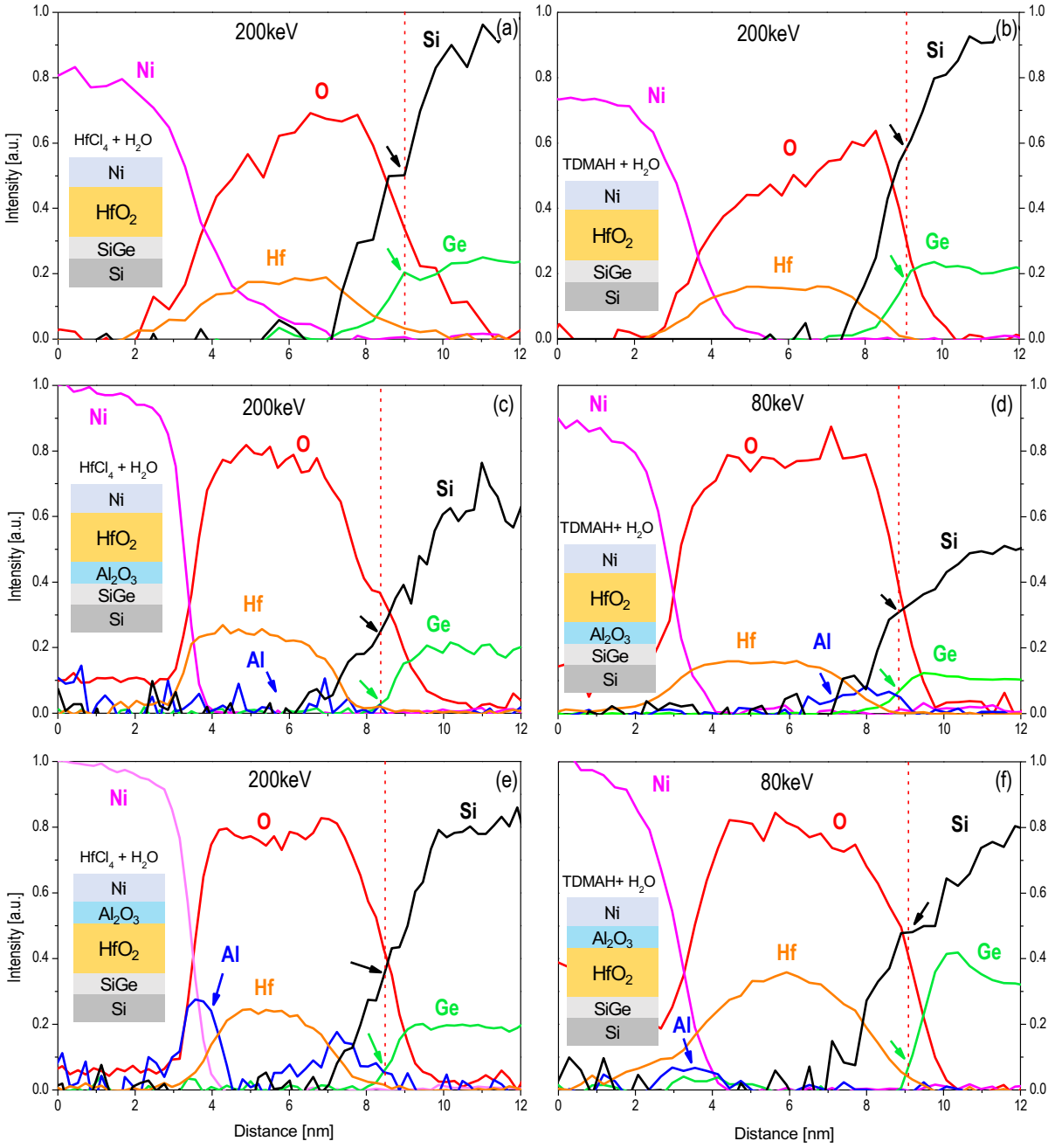

Figure 6. STEM- EELS compositional analysis of MOSCAP devices. EELS experiment performed at 80keV and 200keV as indicated. The inset drawings illustrate corresponding gate stack structure along with the ALD chemistry above it. The compositions of the elements are averages area of ~6 x 0.2 nm parallel to sample surface. The red dashed line intercepts the half peak values of the O signals and indicate the SiGe - HfO_2 interface. Black and green arrows denote Si and Ge composition on the SiGe surfaces respectively. Blue arrow indicates the Al composition in the oxide. AlO_x-HfO_2 interdiffusion is seen for bilayer samples regardless of the initial structure and confirmed with raw data analysis in Fig 6 and Fig. S5. This interdiffusion is prominent for EELS analysis at 80keV. In comparison to a-b, devices in c-f show lower Ge/Si ratio at the intersection with red dashed lines indicates Si rich interface formation with Al_2O_3 incorporation into HfO_2. Ni interdiffusion is seen in devices a and b and Al_2O_3 insertion into HfO_2 impedes the Ni diffusion as seen in c-f.

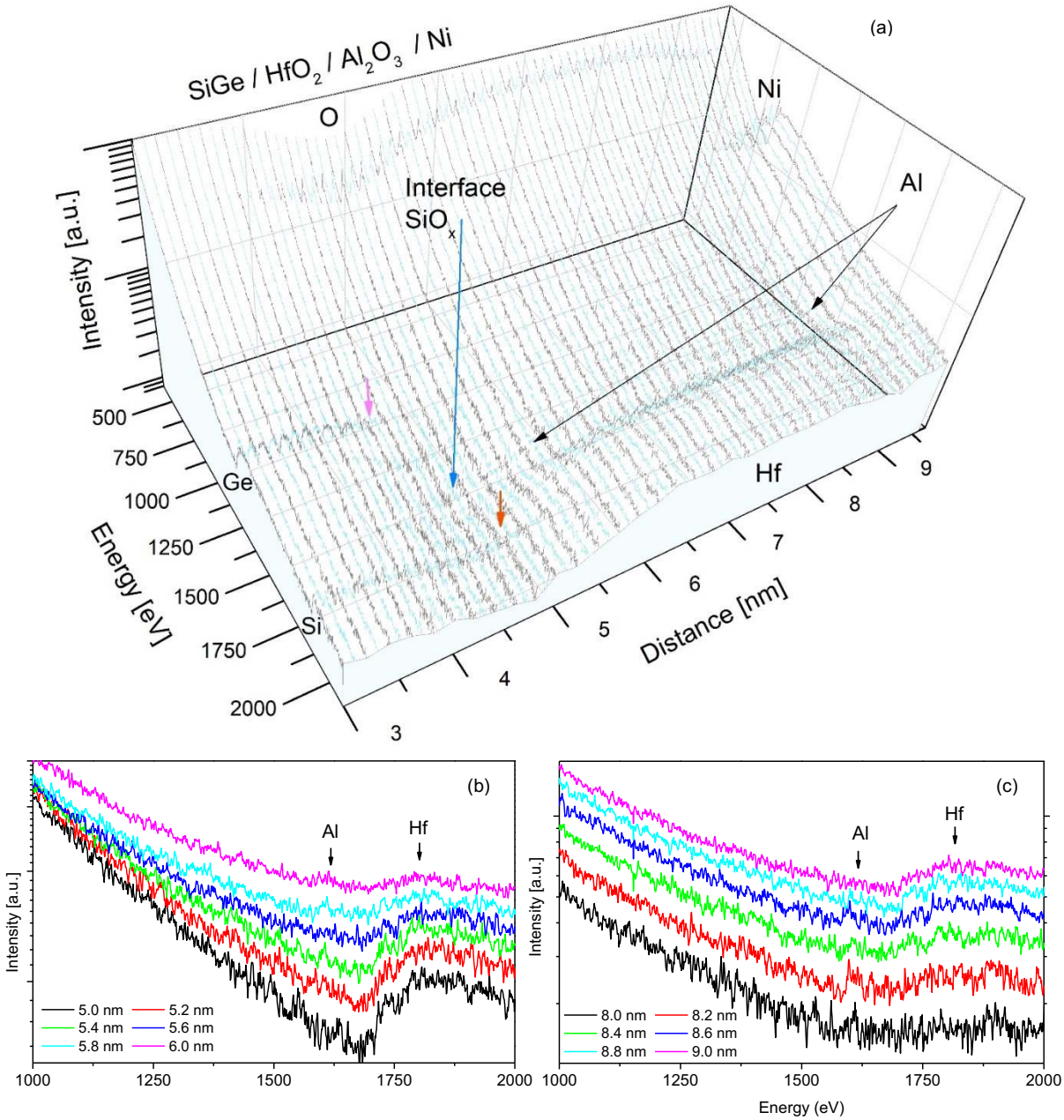

Figure 7. STEM- EELS compositional analysis of Ni/Al_2O_3/HfO_2 /SiGe MOSCAP device. Raw EELS data taken at 200 keV from sample in Fig. 6e is shown in a 3D semi-log graph (a) with the energy axis indicating the electron energy loss and corresponding intensity in arbitrary units. The axis labeled with distance indicates location of the electron beam on sample. The colored consecutive black and light blue lines indicate electron energy loss for the given location on samples and two colors chosen to enhance the image contrast. Each data line projects energy loss averaged from areas of 5 × 0.2nm parallel to the sample surface. The peaks appear on the graphs corresponds to Si K edge (1839 eV), Ge L edge 1217 eV, Hf M edge 1662 eV, O K edge 532 eV, Al K edge 1560 eV, Ni L edge 885 eV. The blue arrow indicates SiO_x interface formation between SiGe and HfO_2. Pink and red arrows indicate the Ge and Si compositions on SiGe surface. The Ge signal decays earlier than Si as it approaches the HfO_2 layer. Black arrows denote Al composition across the oxide. Al_2O_3 insertion onto HfO_2 in bilayer structure forms intermixing by Al diffusion. To increase the visibility of Al peak and inter diffusion, semi-log 2D graph of raw EELS data with 1-2K energy loss range is presented in graph b (SiGe/HfO_2 interface region) and c (HfO_2/Ni interface region) Two graphs prepared with offsets introduced between each curve to improve visibility.

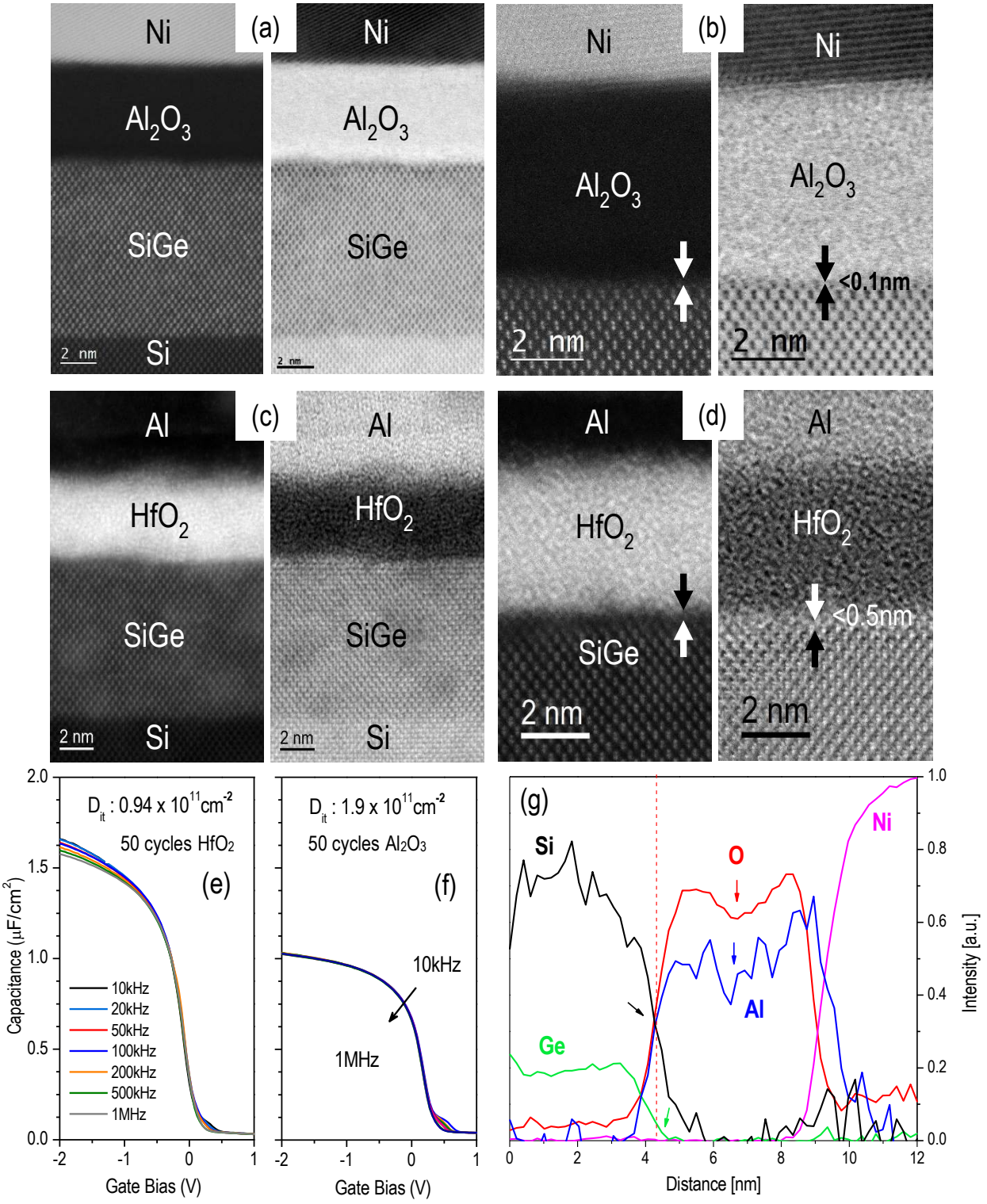

Figure 8. Comparison of Ni/Al_2O_3/SiGe/Si and Al/HfO_2/SiGe/Si MOSCAPs. Note the Al is a gettering gate metal. High resolution STEM HAADF and BF images at 80 keV of devices with (a-b) Ni/Al_2O_3 and (c-d) Al/HfO_2 gate oxides; both have 50 ALD cycles. Nearly abrupt interfaces are observed in both bright field and dark field imaging for both devices. e-f) C-V graphs of Al/HfO_2 and Ni/Al_2O_3 devices with very low depletion capacitance are shown, insets denote integrated D_{it} values across the band gap extracted with full interface state model. EELS elemental composition of Ni/Al_2O_3/SiGe/Si is shown in g. The regions of Al_2O_3 gate oxide defined with Z contrast in a-b are in good agreement with the EELS spectra in g. The red dashed line intercepts the half peak values of the O, Si and Ge signals and delineate the SiGe-HfO_2 interface. Black and green arrows denote Si and Ge composition on SiGe surface; a Si rich interfaces observed. The blue and red arrows indicates the electron beam damaged (80keV) region of Al_2O_3. A similar effect also observed for samples studied at 2000 keV. Ni gate metal overlaying with Al indicating Ni and Al intermix. Note at the SiGe interface, the Al and O profiles are nearly identical consistent with a near zero $SiGeO_x$ interlayer.

MOSCAP devices in SiGe/…/Ni	C_{ox} $\mu F/cm^2$	V_{fb} (V)	D_{it} C.M. $cm^{-2}eV^{-1}$	D_{it} F.I.S.M $cm^{-2}eV^{-1}$	Standard errors of the mean	Standard errors in percentage
45 HfO_2	2.5	-0.02	3.57	3.74	0.266	7.1%
$5Al_2O_3/45HfO_2$	1.96	-0.03	2.43	3.30	0.074	2.2%
45 $HfO_2/5Al_2O_3$	2.02	-0.01	2.35	3.15	0.10	3.1%
$5Al_2O_3/45HfO_2/5Al_2O_3$	1.74	0.12	2.18	2.53	0.124	4.9%
$6\times(9HfO_2+1Al_2O_3)$	1.8	-0.05	2.02	2.22	0.05	2.2%

Table 1. Comparison of D_{it} values obtained with conductance and full interface state model.

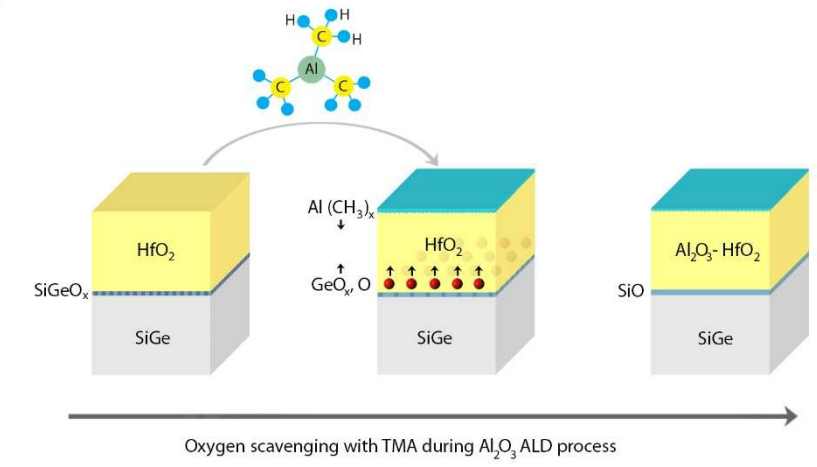

Table of contents