論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Threshold voltage control technology in metal/high-k pFET consisting of high germanium content SiGe channel and fixed charge/oxygen vacancy control in gate stack
著者(和文)	山口晋平
Author(English)	Shimpei Yamaguchi
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10986号, 授与年月日:2018年9月20日, 学位の種別:課程博士, 審査員:若林 整,筒井 一生,渡辺 正裕,大見 俊一郎,角嶋 邦之
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10986号, Conferred date:2018/9/20, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Threshold voltage control technology in metal/high-k pFET consisting of high germanium content SiGe channel and fixed charge/oxygen vacancy control in gate stack

by

Shimpei Yamaguchi

Bachelor of Science, The University of Tokyo (2002)

Master of Science, The University of Tokyo (2004)

Submitted to the

Department of Electronics and Applied Physics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy in Engineering

at the

TOKYO INSTITUTE OF TECHNOLOGY

July 27, 2018

Signature of Author____

Department of Electronics and Applied Physics, July 27, 2018

Certified by ____

Hitoshi Wakabayashi, Professor, Thesis Supervisor

Threshold voltage control technology in metal/high-k pFET consisting of high germanium content SiGe channel and fixed charge/oxygen vacancy control in gate stack

by

Shimpei Yamaguchi

Submitted to the
Department of Electronics and Applied Physics
on July 27, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Engineering

Abstract

In advanced complementary metal oxide semiconductor (CMOS) technology, metal gate electrode and high-k gate dielectric have been introduced from 45nm node to overcome gate leakage issue with conventional silicon dioxide (SiO_2) film and enable further area scaling. One of the challenges associated with metal gate / high-k gate dielectric technology has been a control of threshold voltage of field effect transistor (FET), especially p-type FET (pFET). In gate-first integration, where metal gate / high-k dielectric stack is formed before junction formation, high thermal budget (activation anneal) is applied to gate stack and effective work function (eWF) of pFET typically shifts to mid-gap direction and cannot provide sufficiently low threshold voltage (V_T). To achieve practical pFET V_T, two solutions have been proposed against this problem, one is gate-last or replacement metal gate (RMG) technology where gate stack is formed after high thermal treatment by replacing dummy poly silicon gate and SiO_2 gate dielectric with metal gate and high-k gate dielectric. The metal gate and high-k dielectric stack doesn’t receive high thermal budget, therefore effective work function can be kept high.

Another solution is to implement silicon germanium ($Si_{1-x}Ge_x$) in the channel of pFET. $Si_{1-x}Ge_x$ is typically grown on Si substrate by epitaxy. Thanks to its higher valence band energy, pFET V_T is reduced without modifying gate stack itself.

In first part of this thesis, eWF control technique using gate stack engineering is discussed. The base process is gate-last FinFET devices. Even with gate-last or RMG integration scheme, a recent report highlighted effective work function lowering (shift towards mid-gap) in scaled equivalent oxide thickness (EOT) region (EOT < 10 Å). Therefore, it is still highly important to push the eWF further towards valence band edge to achieve lower pFET V_T at scaled EOT region. Here we identified post sacrificial-Si deposition anneal (called WF setting anneal in this thesis) as key enabler for low pFET V_T at scaled EOT in this work. And we revealed two mechanisms to explain pFET V_T reduction. One is fixed charge generation at the interface between high-k dielectric and metal gate electrode. The intermixed layer created in-between titanium nitride (TiN) electrode and hafnium dioxide (HfO_2) gate dielectric during WSA has negative fixed charges and therefore they reduce pFET V_T. The other mechanism is passivation of oxygen vacancies (positively charged) in the HfO_2 film by supplying oxygen from TiN electrode during WSA. We could achieve approximately 140 mV pFET V_T reduction by optimizing WSA process without compromising device performance and scalability.

In second part of the thesis, $Si_{1-x}Ge_x$ channel devices have been fabricated and pFET V_T reduction was pursued by increasing Ge contents in the channel and reducing process thermal budget concurrently. Although lower process temperature had been identified as process knob to enable low pFET V_T in $Si_{1-x}Ge_x$ channel transistor, there were very few reports on successful integration to realize high performing devices at scaled gate length. Therefore, we also focused on the device scaling and performance boost for high Ge content (Ge > 50%) $Si_{1-x}Ge_x$ pFET. We systematically investigated the impact from thermal budget and Ge content in $Si_{1-x}Ge_x$ on pFET V_T, carrier mobility, and off-state leakage current and found out that lower temperature process is also a key to achieve high mobility and performance on $Si_{1-x}Ge_x$ channel devices. From this viewpoint, we set our focus on so-called implant free (IF) structure which has in-situ boron doped epitaxial layer as extension and source/drain hence doesn't need high temperature activation anneal. Furthermore, strain effects for $Si_{1-x}Ge_x$ channel has been deeply investigated to enhance the device performance. One effect is an interaction between $Si_{1-x}Ge_x$ channel and embedded silicon germanium (eSiGe) stressor. It was found that eSiGe stressor can give similar performance boost for $Si_{1-x}Ge_x$ channel as well despite of the stress relaxation in longitudinal direction during cavity recess. The other effect is channel width dependence and we confirmed that hole mobility is significantly enhanced at narrower channel width thanks to the relaxation of channel strain in transvers direction. By utilizing these strain engineering, we could achieve the best-in-class performance with SiGe channel device with extremely high Ge content (55%). In the last section of $Si_{1-x}Ge_x$ channel discussion, scalability of $Si_{1-x}Ge_x$ channel device is discussed. We could demonstrate decent device performance at very short gate length (approximately 20nm, which is close to state-of-the-art FinFET technology) thanks to shallow junction by IF structure with Ge 45% SiGe channel.

In summary, two independent approaches for pFET V_T reduction in metal gate/high-k transistor were discussed. One is fixed charge and oxygen vacancy control by PSA process in gate-last process. Another is introduction of $Si_{1-x}Ge_x$ channel with high Ge content in gate-first process. Integration of $Si_{1-x}Ge_x$ channel with high Ge content was enabled by so-called IF structure by reducing thermal budget of the flow.

Contents

Chapter 1 Research background and introduction of this work

1.1 Complementary metal oxide semiconductor (CMOS) technology

CMOSFET (Complementary Metal Oxide Semiconductor Field Effect Transistor) is the most commonly used technology for constructing very-large-scale-integration (VLSI) circuits, which was originally invented by Wanlass and Sah (Fairchild) in 1983 [1]. CMOS technology is used in microprocessors, microcontrollers, static RAM, and other digital logic circuits. CMOS technology is also used for several analog circuits such as image sensors (CMOS sensor), data converters, and highly integrated transceivers for many types of communication.

CMOS has pairs of p-type and n-type MOSFETs (pMOSFET/pFET and nMOSFET/nFET, respectively), which are constructed simultaneously on the same Si substrate. A CMOS circuit typically consists of an nFET and pFET connected in series between the power supply terminals, so that there is negligible standby power consumption [2]. Since one transistor of the pair is always off, the series combination draws significant power only momentarily during switching between on and off states. Consequently, CMOS devices do not produce as much waste heat as other components of logic, for example transistor–transistor logic (TTL) and nMOSFET logic, which normally have some standing current even when not changing state. CMOS also allows a high density of logic functions on a chip, as circuits are designed to minimize active power dissipation. It was primarily for this reason that CMOS became the most used technology to be implemented in VLSI chips.

The basic MOS structure consists of a conducting gate electrode (metal or heavily doped poly-Si) on top of a thin layer of SiO_2 grown on a Si substrate or deposited dielectric film, as shown in Figure 1-1 [3].

A typical cross section of modern CMOSFET is shown in Figure 1-2 [3]. On top of p-type Si substrate, nFET (conducting carrier is electron) and pFET (conducting carrier is hole) are fabricated simultaneously. nFET consists of n-type poly-Si gate electrode, gate oxide dielectric, p-type Si channel/well, and n-type source and drain electrode. Likewise, pFET consists of p-type poly-Si gate electrode, gate oxide dielectric, n-type Si channel/well, and p-type source and drain electrode. Metal silicide (typically $NiSi_x$, $TiSi_x$, and $CoSi_x$) is formed on the poly-Si gate and on the source and drain areas. Each transistor is electrically isolated by shallow trench isolation (STI, SiO_2 is filled in the trench).

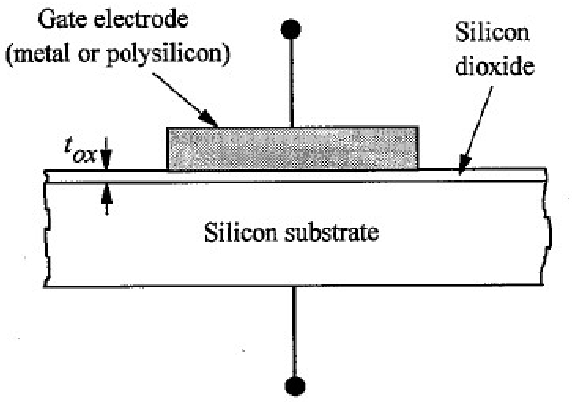


Figure 1-1 A schematic cross section for a MOS structure [3].

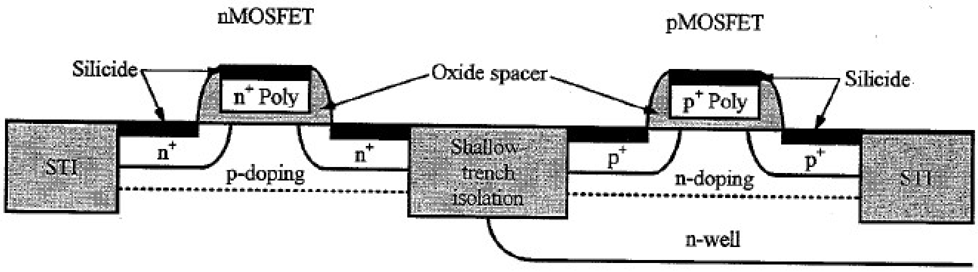


Figure 1-2 A schematic cross section for modern CMOS transistors [3].

1.2 CMOS scaling (Dennard scaling, Moore's law)

CMOS technology evolution in past decades has been supported by transistor scaling which provides density, speed and power improvements simultaneously. Transistor scaling has been enabled by the continuous advancement of lithographic technology. Reducing the transistor channel length leads to so-called short channel effects. The most undesirable short channel effect is a reduction in the gate V_T at which the device turns on, especially at high drain voltages and resultant leakage current increase. For successful advancement of CMOS technology, not only the progress of lithographic technology but also the device architectural optimization has been required to suppress short channel effect.

R. H. Dennard proposed constant-field scaling in 1974 [4], where one can keep short channel effects under control by scaling down the vertical dimensions (gate oxide dielectric thickness, junction depth, etc.) along with the horizontal dimensions, while also proportionally decreasing the applied voltages and increasing the substrate doping concentration (decreasing the depletion width). This is shown in Figure 1-3 [4].

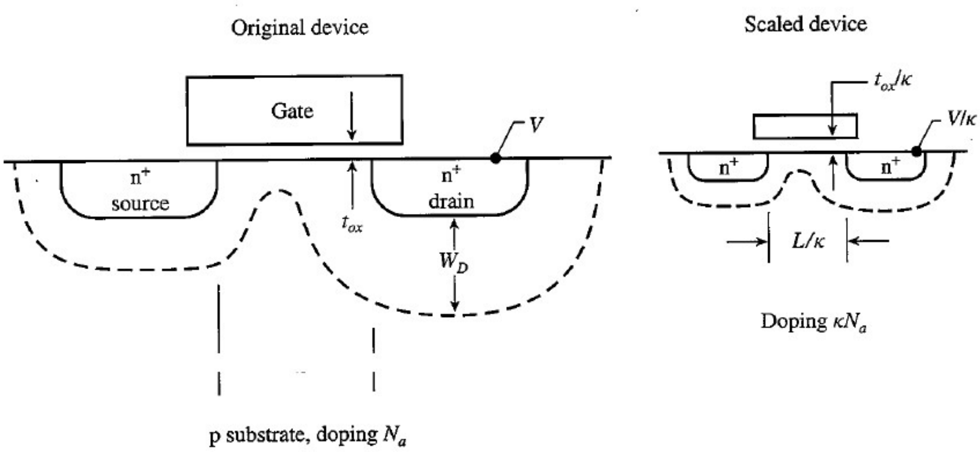


Figure 1-3 Principles of MOSFET constant-field scaling [4].

Table 1-1 shows the scaling rules for various device parameters and circuit performance factors. The doping concentration (N_a, N_d) must be increased by the scaling factor κ in order to scale depletion layer width (W_d) by factor of $1/\kappa$, where W_d is expressed in approximate form as

$$W_D \sim \sqrt{\frac{2\varepsilon_{Si}V_{dd}}{qN_a}} \quad (1\text{-}1)$$

All capacitances scale down by κ, since they are proportional to area and inversely proportional to thickness. The drift current per MOSFET channel width, I_{drift}/W

$$\frac{I_{drift}}{W} = Q_i v = Q_i \mu E \quad (1\text{-}2)$$

is unchanged with scaling, as inversion charge $Q_i = CV$ is unchanged with scaling either. Therefore, the circuit delay, $\tau = CV/I$ scales down by κ. This is the most important conclusion of constant-field scaling: once the device dimensions and the power supply voltage are scaled down, the circuit speeds up by the same factor. Moreover, power dissipation per circuit, which is proportional to VI, is reduced by κ^2, while power density remains unchanged in the scaled-down chip. The constant field scaling described above is known as Dennard's scaling.

Table 1-1 Constant field scaling: MOSFET device and circuit parameters [3, 4].

	MOSFET device/circuit parameters	Multiplicative factor ($\kappa > 1$)
Scaling assumptions	Device dimensions (t_{ox}, L, W, x_j)	$1/\kappa$
	Doping concentration (N_a, N_d)	κ
	Voltage (V)	$1/\kappa$
Device parameters	Electric field (E)	1
Device parameters	Carrier velocity (υ)	1
	Depletion-layer width (W_d)	$1/\kappa$
	Capacitance ($C = \varepsilon A/t$)	$1/\kappa$
	Inversion layer charge density (Q_i)	1
	Current, drift (I)	$1/\kappa$
	Channel resistance (R_{ch})	1
Circuit parameters	Circuit delay time ($\tau = CV/I$)	$1/\kappa$
	Power dissipation per circuit (P ~ VI)	$1/\kappa^2$
	Power-delay product per circuit ($P\tau$)	$1/\kappa^3$
	Circuit density ($\propto 1/A$)	κ^2
	Power density (P/A)	1

As a result of continuous device scaling, one could observe that the number of transistors in a VLSI circuit doubles about every 1.5 year as shown in Figure 1-4 [5], which is known as Moore’s law, which was named after Gordon Moore, the co-founder of Fairchild Semiconductor and Intel [6]. Moore’s law can be demonstrated in actual transistor size the best represented in form of contacted gate pitch (CPP) or SRAM cell size as shown in Figure 1-5 [7]. The dimensions of transistor typically have been scaled down by approximately 0.7x in each node, and SRAM cell size has been scaled down by 0.5x in each node. Moore's prediction proved accurate for several decades and has been used in the semiconductor industry to guide long-term planning and to set targets for research and development.

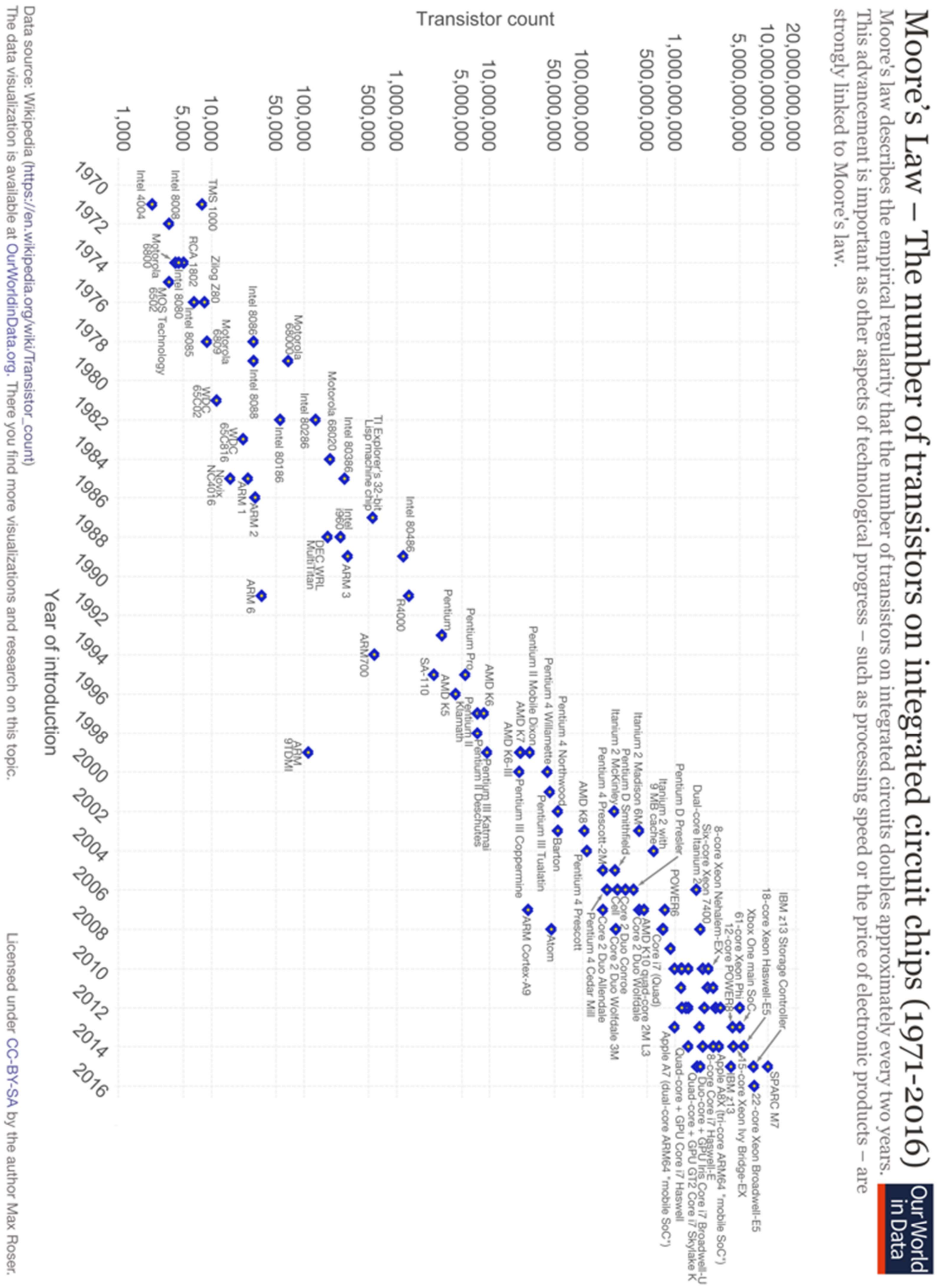


Figure 1-4 Transistor number per chip versus year of introduction demonstrating Moore's law [5].

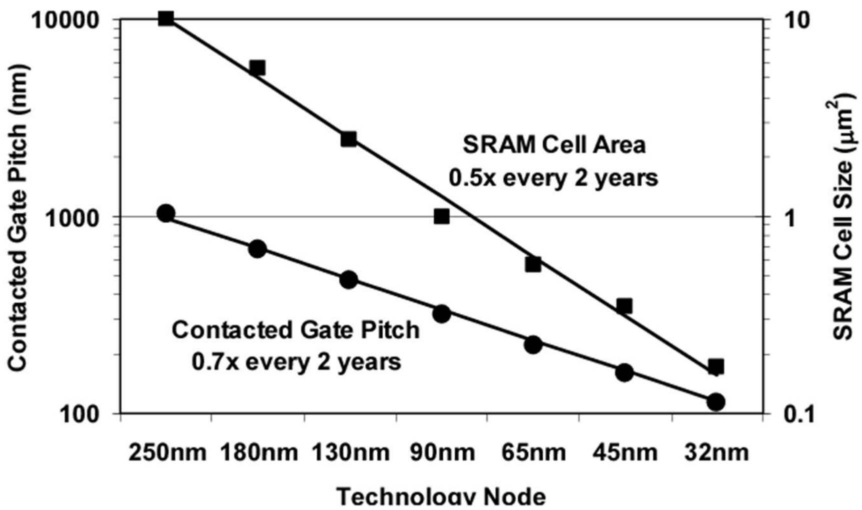


Figure 1-5 Contacted gate pitch (left axis) and SRAM cell size (right axis) versus technology node from 250 nm down to 32 nm node [7].

Moore's law can be rephrased as "costs per transistor is reduced by half about every 1.5 years, as long as wafer processing costs is identical", as the costs per transistor (CPT) can be expressed as CPT = (wafer processing costs) / (number of transistors on the wafer). Historically, wafer-processing costs tended to increase from one generation to the next due to multiple reasons such as introduction of new process, device structures, etc. This offset the increase in number of transistors per area by scaling and slowed down the scaling rate for CPT. Semiconductor industry has been dealing with this issue by migrating to larger wafer sizes, which could sharply reduce wafer processing costs. The net effect was nearly constant with only slight increases in wafer processing costs, as shown in Figure 1-6 (source: Intel). However, after migrating from 200 mm to 300 mm wafer in 130-nm or 90-nm node, the industry hasn't been able to migrate to 450 mm wafer. Therefore, wafer processing cost has kept increasing in every new technology node. Due to this issue, area scaling needs to be accelerated even more than conventional area scaling, to keep CPT scaling. Figure 1-7 is historical trend for costs per unit area (wafer processing costs), area per transistor, and CPT presented by Intel. They accelerated the area scaling rate from 14 nm node technology to overcome the increasing wafer processing costs by introducing so-called hyper-scaling which can realize more shrinkage rate for standard cell than conventional CPP × MxP (MxP: Metal pitch) scaling.

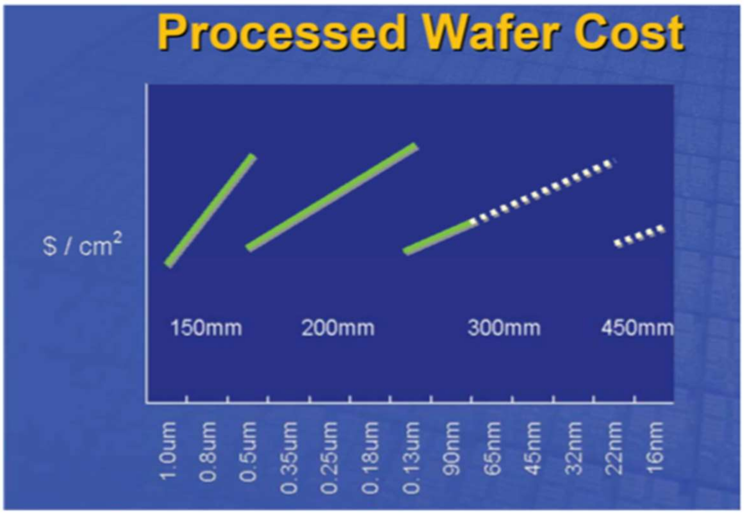


Figure 1-6 Historical trend for wafer processing costs. The cost increase due to new process generation can be offset by introduction of larger wafer sizes (Source: Intel).

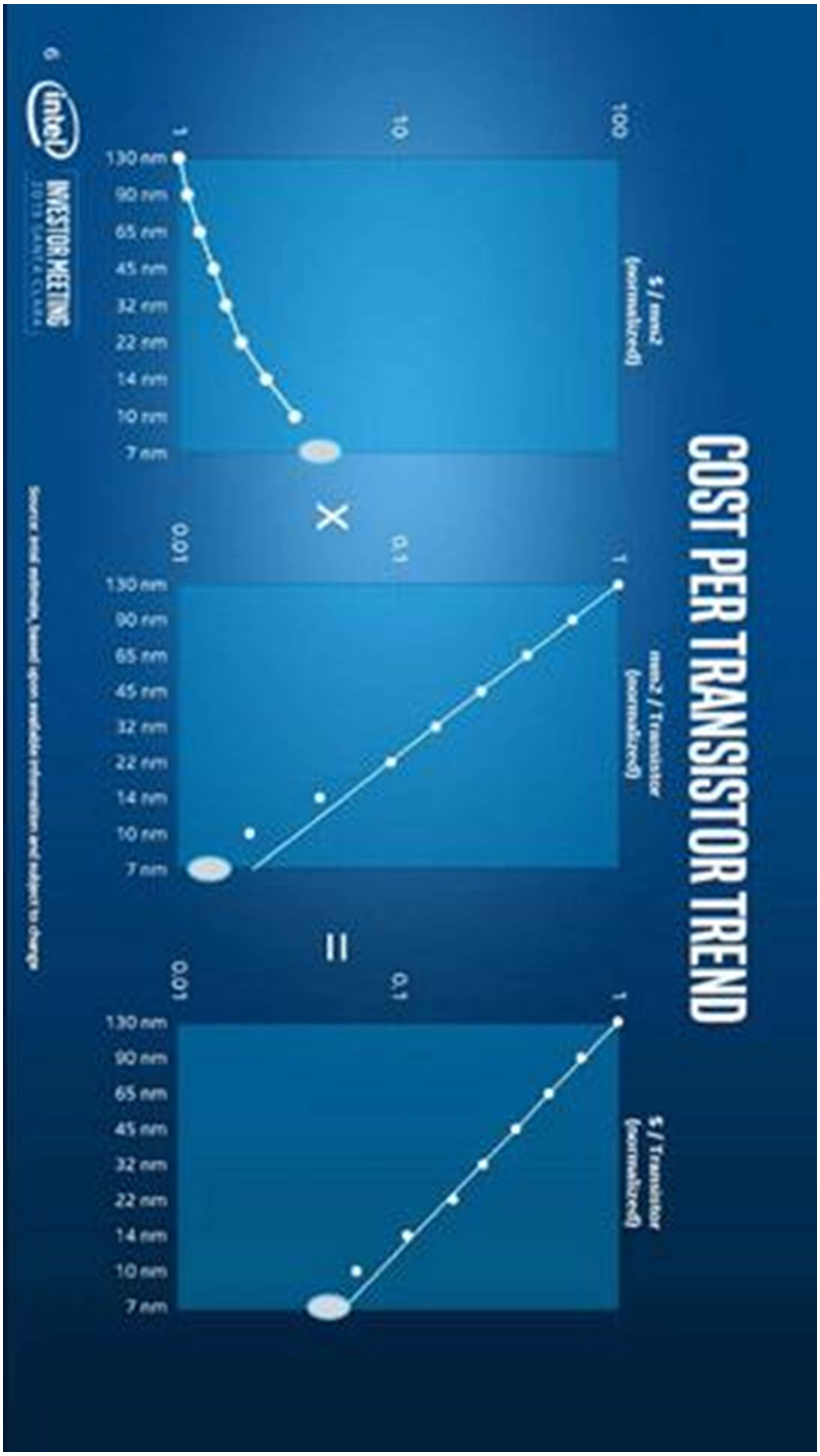


Figure 1-7 Historical trend for costs per unit area (wafer processing costs), area per transistor, and cost per transistor (source: Intel).

1.3 Gate stack scaling, metal gate and high-k dielectric technology

In this section, we will review gate stack scaling (T_{inv} scaling) and introduction of metal gate/high-k dielectric technology. As described in Table 1-1, T_{inv} is scaled down continuously as part of the constant field scaling of the MOSFET. In constant field scaling, supply voltage V_{dd} scales down by factor of κ, therefore T_{inv} is scaled down as well by factor of κ to keep vertical electric field constant. Especially in advanced technology node, T_{inv} needs to be scaled down aggressively to enable gate length scaling and maintain manageable short channel effect.

In earlier technology nodes, SiO_2 and silicon oxynitride (SiON) had been used as gate dielectric material thanks to its decent interfacial property at Si surface (low interfacial trap density) and large band gap. However, since around 90-nm node technology, SiON stopped scaling down further due to increasing gate leakage current, which is shown in Figure 1-8 [8]. From 180-nm node technology, gate leakage current increased exponentially due to direct tunneling current, which is dominant in this thickness regime (below 2 nm). SiON gate dielectric thickness reached its limit in 90 nm node technology and couldn't be scaled down further in 65 nm node due to severe gate leakage current which leads to high power consumption. To overcome this situation, high-k gate dielectric and metal gate electrode have been introduced from 45 nm node technology by Intel [8, 9]. High-k gate dielectric has larger dielectric constant than SiO_2 (dielectric constant 3.9) [10, 11, 12], therefore scaled T_{inv} can be realized even with thicker physical thickness than SiO_2 and gate leakage current can be suppressed as shown in Figure 1-9 [8].

Development efforts have focused on finding a material with a requisitely high dielectric constant that can be easily integrated into a Si-based VLSI manufacturing process. Key considerations include band alignment to Si (which may alter gate leakage current) [13, 14], film morphology [15], thermal stability [16, 17, 18], maintenance of a high mobility of charge carriers in the channel and minimization of electrical defects in the film/interface [19, 20, 21]. In these regards, Hafnium-based dielectric (HfO_2, $HfSiO_x$) is now the most commonly used high-k dielectric material in the industry. Dielectric constant of HfO_2 is ranging approximately from 16 to 20 [10, 11, 12].

For the gate electrode, the industry had been using n+ or p+ poly-Si for nFET and pFET respectively. The problem of using n+/p+ poly-Si gate is depletion layer created in the poly-Si gate when the transistor is turned on, which became considerable portion of total T_{inv} in advanced node [22]. The depletion-layer thickness can be reduced by increase in implant dose so that the poly-Si is degenerated. However, strip of photo resist would be impossible with such implant dose due to severe crusting. Metal gate can completely eliminate depletion layer while avoiding this issue and promote further T_{inv} scaling.

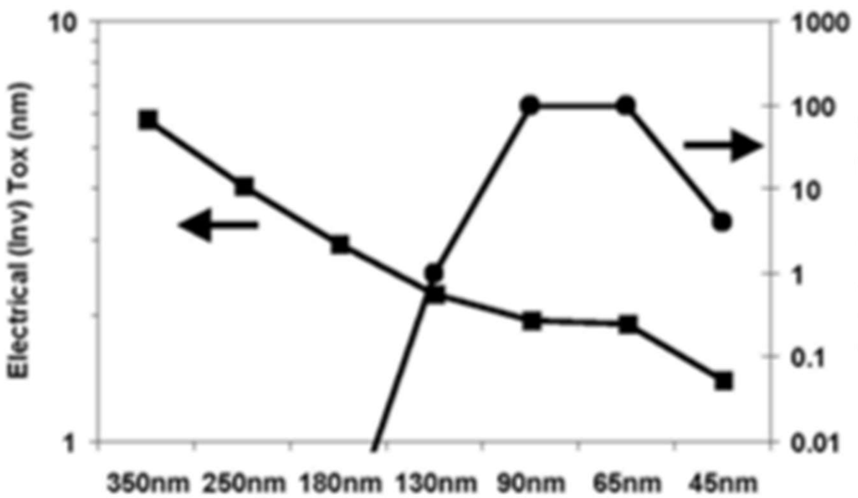


Figure 1-8 Historical trend for T_{inv} and gate leakage current over past several technology nodes from 350 nm node down to 65 nm node [8].

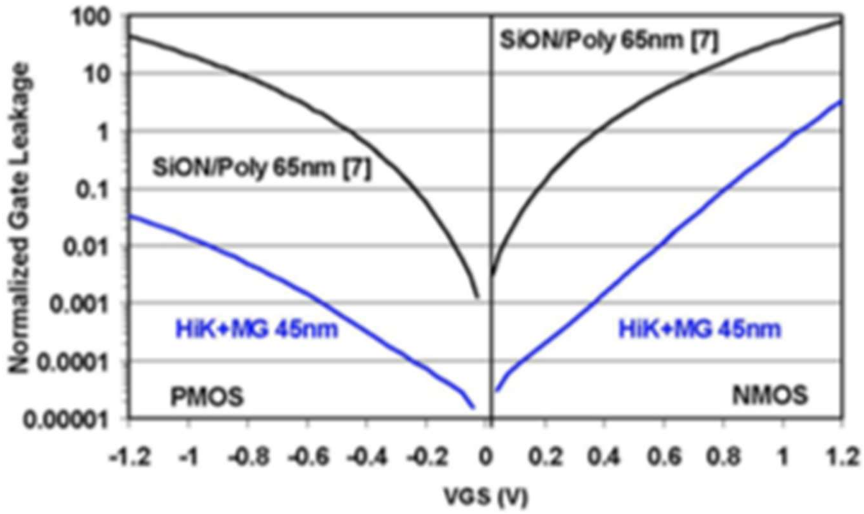


Figure 1-9 Gate leakage current for Poly-Si/SiON gate stack (65 nm node) and high-k/meal gate (45 nm node) showing significant reduction [8].

Among two integration schemes for metal gate/high-k technology (gate-first, gate-last/RMG), gate-first was the mainstream scheme in research phase because of their relatively lesser degree of integration difficulty especially when thin metal gate is inserted between poly-Si and high-k dielectric so that gate patterning RIE (reactive ion etching) is not severely affected (Metal Inserted Poly Silicon Gate: MIPS Gate) [23, 24].

However, gate-first scheme had an issue that pFET V_T gets higher (eWF gets closer to mid-gap) especially at scaled EOT region, compared to conventional p+ poly-Si case (fermi level for p+ poly-Si is only several tens of mV away from valence band edge of Si) [16, 25, 26, 27, 28, 29, 30, 31, 32]. Akiyama *et al*. reported flat band voltage (V_{fb}) roll-off behavior in gate-first metal gate/high-k stack, where eWF of the gate stack shifts toward mid-gap direction and pFET V_T shifts higher with thinner SiO_2-IL (EOT less than 3 nm), as shown in Figure 1-10 [32]. They attributed this roll-off behavior to oxygen vacancy (positively charged) generation in HfO_2 layer (the dissolved oxygen atoms oxidize substrate and regrow the interfacial SiO_2 layer) during high temperature anneal (PDA, 800ºC). When SiO_2-IL is thinner than a certain thickness, oxygen can dissolve from HfO_2 and move to SiO_2-IL/Si interface. However, if SiO_2-IL is thick enough, this reaction can be suppressed. This mechanism can explain V_{fb} shift at thin EOT region. Cartier *et al*. reported significant V_{fb} modulation of various p-type metal gate (Ru, Re, and Pt) on HfO_2 by annealing conditions as shown in Figure 1-11 [26]. Although their vacuum work-function is approximately from 4.9 to 5.1 eV which is close to valence band edge, eWF of p-type metal/HfO_2 stack was modulated up to 0.75 eV (more than half of the Si band gap, 1.12 eV) with various anneal ambient and temperature. Oxidizing ambient (low O_2 partial pressure N_2/O_2 mixture) shifts eWF toward valence band edge (positive V_{fb} shift) and reducing ambient (N_2/H_2 forming gas) acts oppositely and shifts eWF toward conduction band edge (negative V_{fb} shift). They claimed this eWF/V_{fb} behavior is due to modulation of oxygen vacancy concentration in HfO_2. Oxygen in N_2/O_2 mixture ambient can fill the oxygen vacancies and electrically neutralize them. As oxygen vacancies are positively charged and shifts V_{fb} negatively, this reaction shifts back V_{fb} in positive direction (eWF shifts toward valence band edge). On the other hand, reducing ambient (N_2/H_2 forming gas) can create oxygen vacancy in HfO_2, therefore V_{fb} shifts negatively.

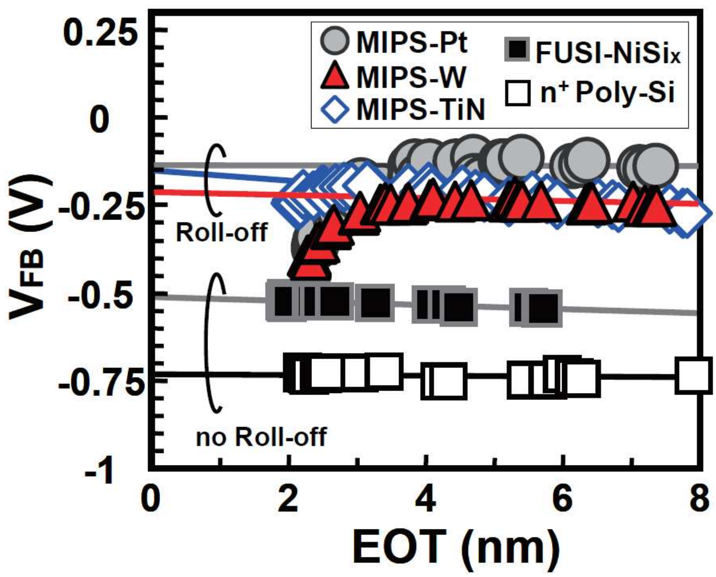


Figure 1-10 V_{fb} as a function of EOT. V_{fb} roll-off behavior was observed for high WF metal [32].

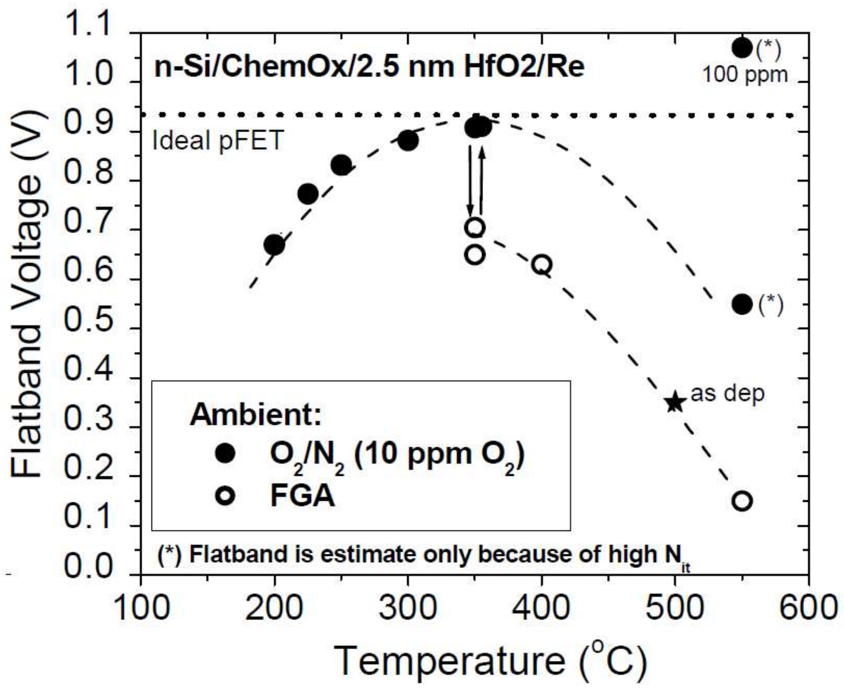


Figure 1-11 V_{fb} as a function of annealing temperature for Re/HfO_2 gate stack [26].

As metal gate/high-k stack cannot avoid receiving high temperature anneal in gate-first integration scheme (junction activation annealing), oxygen vacancy generation in HfO_2 and hence negative V_{fb} shift (higher pFET V_T) is supposed to be inevitable unless oxygen is supplied to HfO_2 in later stage of the processing to electrically neutralize them. Cartier *et al*. proposed lateral oxygenation technique [25]. After gate stack (poly-Si/metal/HfO_2) patterning, oxygen anneal is conducted to supply oxygen to HfO_2 from the edge of the gate (Figure 1-12). This technique worked for short gate length devices as oxygen diffused from both sides of the gate edge can reach to the middle of the gate, then oxygen vacancies are neutralized all along the gate. On the other hand, at long channel gate devices, oxygen cannot reach to the middle of the gate if the gate length is too long. Therefore, this technique has pattern loading, which makes this technique very difficult to be applied to manufacturing.

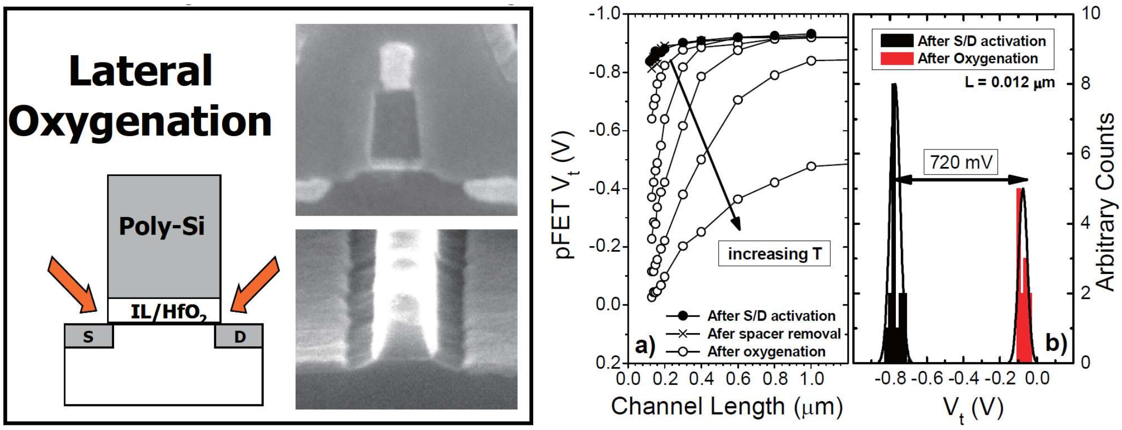


Figure 1-12 (left) Concept of lateral oxygenation. (right) pFET V_T as a function of channel length. Oxygenation anneal can reduce pFET V_T at shorter gate length, but the effect of anneal becomes smaller with longer gate length [25].

1.4 pFET V_T control method

To overcome the high pFET V_T issue at scaled EOT, there have been several approaches to date which have been implemented in manufacturing.

The most commonly accepted technique is to use Al_2O_3 or Al capping layer deposited on top of Hf-based high-k [27, 33, 34, 35, 36]. The first report of Al_2O_3 capping layer was from Cartier *et al.*, where pFET V_T could be reduced by approximately 100 mV with Al_2O_3 deposited on top of $HfSiO_x$ (Poly-Si/Al_2O_3/$HfSiO_x$ gate stack) as shown in Figure 1-13. In Figure 1-13, C-V characteristics for Poly-Si/(Al_2O_3)/$HfSiO_x$ stack showed pFET V_T reduction with Al_2O_3. Tatsumura *et al.* and Ando *et al.* proposed the physical model of explaining pFET V_T modulation by Al_2O_3 cap in ref. [35] and [36], respectively. In their model, Al and Si atoms form electrical dipole moment at HfO_2/SiO_2-IL interface. Although they reported approximately 150~200 mV of pFET V_T reduction with Al_2O_3 capping layer which is encouraging, T_{inv} increase and hole mobility degradation were confirmed at the same time (Figure 1-14). As dielectric constant for Al_2O_3 is approximately 10 and lower than that for HfO_2 (16-20), T_{inv} increase with the additional Al_2O_3 layer is understandable. Regarding hole mobility degradation with Al_2O_3 capping layer, this is in good contrast with La_2O_3 capping layer for nFET application (La_2O_3 cap shifts V_T negatively), where no extrinsic mobility degradation with La_2O_3 cap was observed [35, 36]. Their hypothesis is that the Al-Si dipoles are densely distributed at HfO_2/SiO_2-IL interface, so that hole mobility is degraded due to remote Coulomb scattering. In case of La_2O_3 cap, on the other hand, La-O-Si network is formed at the top of SiO_2-IL (La silicate formation) along the depth direction of the gate stack and creates longer and lower density dipole moments. Therefore, the mobility degradation by remote Coulomb scattering can be minimized. As described above, although the V_T shift amount is significant, Al-based capping layer has inherent disadvantage of performance degradation due to T_{inv} and mobility degradation.

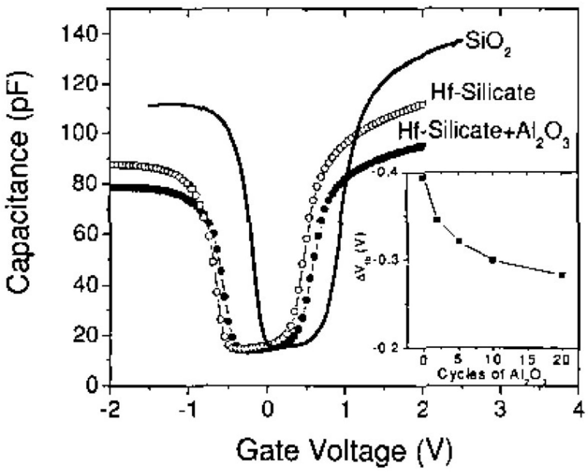


Figure 1-13 C-V characteristics for Poly-Si/(Al_2O_3)/$HfSiO_x$ stack showing pFET V_T reduction with Al_2O_3 [27].

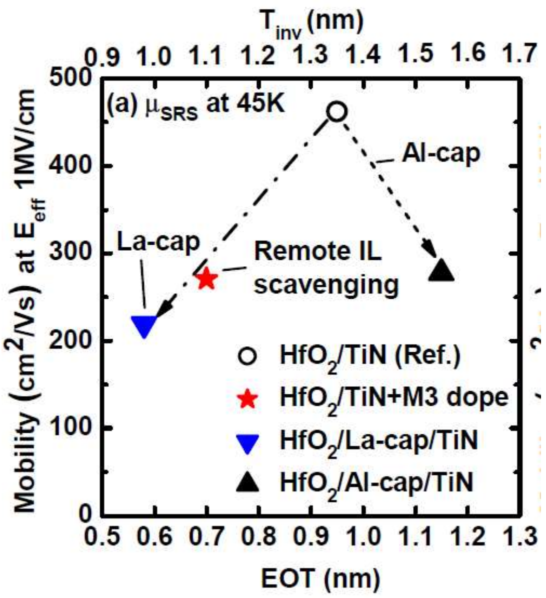


Figure 1-14 Carrier mobility as a function of EOT for La-based cap and Al-based cap [36].

Gate-last (or RMG) integration scheme is another approach to mitigate pFET V_T issue, although it requires significant change in the CMOS fabrication flow [8, 9, 37]. In RMG process flow, final gate stack is formed after junction activation anneal by replacing dummy poly-Si gate and SiO_2 gate dielectric with metal gate and high-k gate dielectric. Relatively lower pFET V_T and eWF closer to valence band edge have been consistently confirmed in many literatures [38, 39, 40, 41, 42], as shown in Figure 1-15. This is supposedly due to absence of high temperature anneal on metal/high-k stack (less amount of oxygen vacancies in HfO_2) in case of RMG, although HfO_2 PDA (typically 700-900ºC RTA) is still needed for densification in gate-last flow as well.

RMG based metal gate/high-k technology was firstly industrialized by Intel Corporation in their 45nm node technology [8, 9] and currently most companies are using RMG based metal gate/high-k technology.

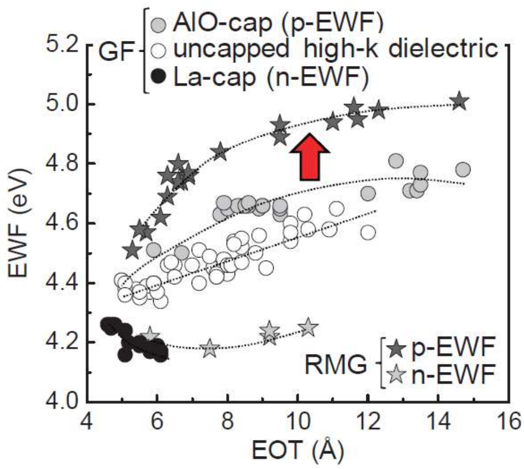


Figure 1-15 eWF as a function of EOT comparing gate-first and gate-last (from [39]).

The other approach is to use epitaxial $Si_{1-x}Ge_x$ film as a channel material instead of conventional Si channel [43, 44, 45, 46] to shift valence band edge energetically higher so that inversion layer can be formed with relatively lower gate voltage. $Si_{1-x}Ge_x$ film is typically epitaxially-grown on Si substrate in pFET active area by masking nFET area with hard mask [47]. The band alignment for strained $Si_{1-x}Ge_x$ grown on bulk Si substrates is shown in Figure 1-16. The valence band edge is shifted higher depending on the Ge content in the $Si_{1-x}Ge_x$ channel ($\sim 0.74x$ eV). The bandgap for strained $Si_{1-x}Ge_x$ grown on bulk Si substrates and for unstrained $Si_{1-x}Ge_x$ is shown in Figure 1-17 [48]. Experimentally measured bandgaps have been fitted to two quadratic equations for the Si-like and Ge-like parts of the band structure as described as below in ref. [48, 49]

$$E_g = 1.155 - 0.43x + 0.0206x^2 \text{ (eV) for } x < 0.85 \quad (1\text{-}3)$$

$$E_g = 2.010 - 1.27x \text{ (eV) for } x > 0.85 \quad (1\text{-}4)$$

Figure 1-18 shows I_d-V_g characteristics for Si and $Si_{1-x}Ge_x$ channel pFET [43], demonstrating pFET V_T reduction with $Si_{1-x}Ge_x$ channel. $Si_{1-x}Ge_x$ channel has been introduced in manufacturing successfully by IBM in their 28nm node technology with a significant effort to overcome its integration challenges like interface quality between gate dielectric and $Si_{1-x}Ge_x$ channel [46].

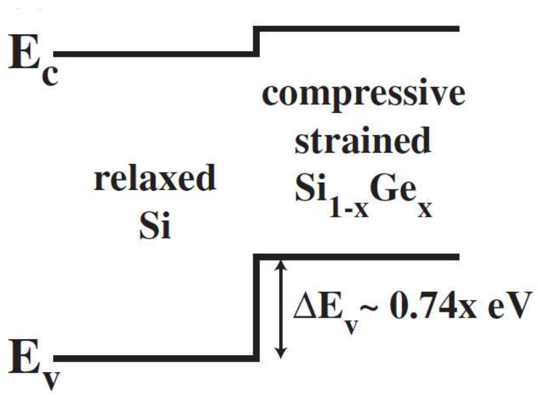


Figure 1-16 The band alignments for a compressively strained-$Si_{1-x}Ge_x$ heterolayer grown on relaxed Si [48].

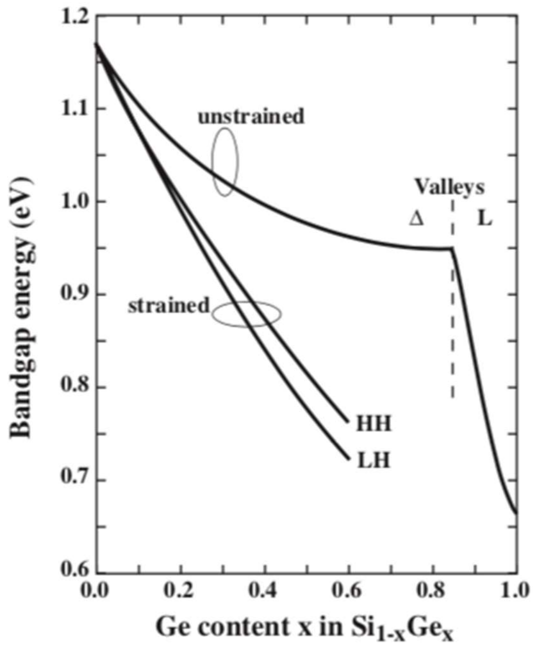


Figure 1-17 Bandgap for strained $Si_{1-x}Ge_x$ grown on bulk Si substrates and for unstrained $Si_{1-x}Ge_x$ [48].

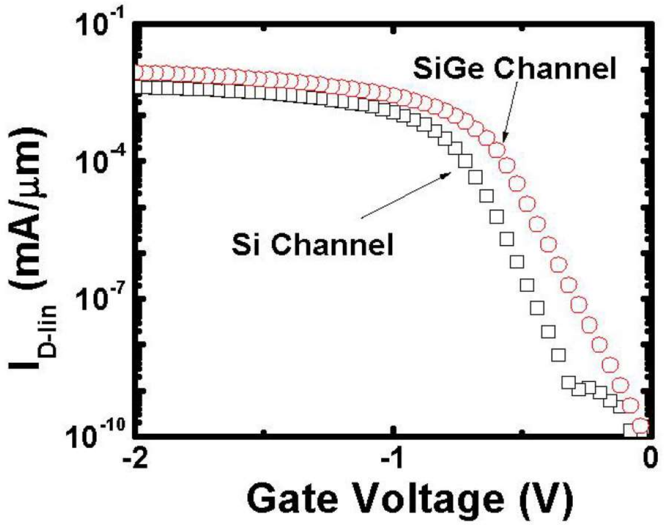


Figure 1-18 I_d-V_g characteristics for Si and $Si_{1-x}Ge_x$ channel [43].

Transistor gate voltage V_g in subthreshold region is expressed by

$$V_g = V_{fb} + \psi_s - \frac{Q_s}{C_{ox}} = V_{fb} + \psi_s + \frac{\sqrt{2\varepsilon_{Si} q N_d \psi_s}}{C_{ox}} \quad (1\text{-}5)$$

where ψ_s is surface potential in Si, Q_s is total charge per unit area in Si, C_{ox} is oxide capacitance per unit area, ε_{Si} is Si permittivity, q is electronic charge, N_d is donor impurity density. In case of Si channel, strong inversion starts where ψ_s reaches to $2\psi_B$ (ψ_B is difference between Fermi level and intrinsic level in Si) as shown in Figure 1-19. If we use $Si_{1-x}Ge_x$ as a channel material, the necessary surface band bending will be reduced by

$$\Delta E_v/q \equiv {}^{1}\!/_{q}\,(E_v^{SiGe} - E_v^{Si})$$

and (1-5) will be

$$V_T = V_{fb} + 2\psi_B - \Delta E_v/q + \frac{\sqrt{2\varepsilon_{Si} N_d (2q\psi_B - \Delta E_v)}}{C_{ox}} \quad (1\text{-}6)$$

V_{fb} is given by

$$V_{fb} = \phi_{ms} - \frac{Q_{ox}}{C_{ox}} \quad (1\text{-}7)$$

$$\phi_{ms} \equiv \phi_m - \phi_s \tag{1-8}$$

where Q_{ox} is effective oxide charge per unit area at the gate dielectric and Si interface. ϕ_m and ϕ_s are work-function of metal gate and semiconductor (channel material), respectively. Using (1-7) and (1-8), (1-6) can be written as

$$V_T = \left(\phi_m - \frac{Q_{ox}}{C_{ox}}\right) - \phi_s + 2\psi_B - \Delta E_v/q + \frac{\sqrt{2\varepsilon_{Si} N_d (2q\psi_B - \Delta E_v)}}{C_{ox}} \tag{1-9}$$

The first term of RHS in (1-9) can be defined as eWF,

$$\phi_m^{EFF} \equiv \phi_m - \frac{Q_{ox}}{C_{ox}}$$

In case of techniques modulating the amount of electrical dipole (Al_2O_3 cap) or fixed charge (RMG, where density of oxygen vacancy is supposedly smaller than gate-first scheme), ϕ_m^{EFF} is modulated, whereas $Si_{1-x}Ge_x$ channel is modulating ΔE_v depending on Ge content.

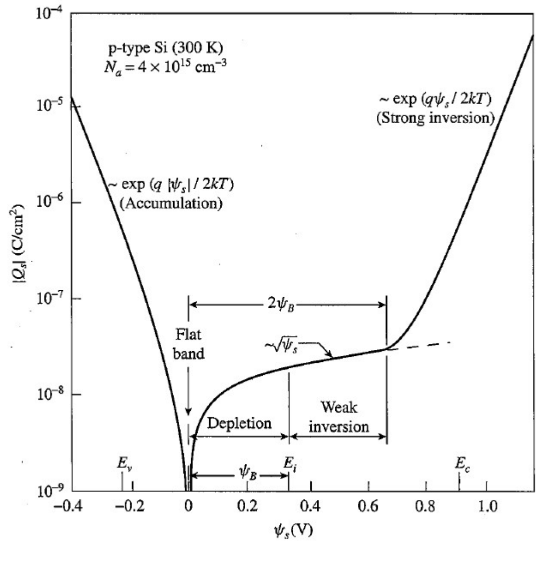


Figure 1-19 Total charge density in Si as a function of surface potential ψ_s for pFET [3].

1.5 The scope of this thesis

In this thesis, two kinds of pFET V_T modulation techniques will be discussed in detail. In Chapter 2, as a first topic, eWF control technique by gate stack engineering in RMG FinFET technology will be discussed. Even with gate-last or RMG integration scheme, recent reports highlighted eWF lowering in scaled EOT region (EOT < 10 Å) [39, 50]. Therefore, it is still highly important to push eWF further towards valence band edge to achieve lower pFET V_T at scaled EOT region. However, the solution for this problem has not been identified clearly yet. In this thesis, we conducted systematic study to investigate the impact from process condition on eWF or device V_T. Based on this result, we propose the practical method to reduce RMG pFET V_T in aggressively scaled EOT region (< 10 Å) by controlling effective oxide charge Q_{ox}.

From Chapter 3 to Chapter 5, as a second topic, device characteristics for $Si_{1-x}Ge_x$ channel pFET will be discussed from various aspects. Although higher Ge content and lower process temperature had been already identified as a method to reduce pFET V_T, there was very limited report on successful integration of high Ge content $Si_{1-x}Ge_x$ channel (x > 0.5) at scaled gate length (20-30 nm, necessary for sub-32nm node technology) at the time when we started research (2010), as shown in

Table 1-2. Therefore, in this thesis, we focused on demonstration of high performing $Si_{1-x}Ge_x$ channel pFET with extremely high Ge content (x = 0.55) at scaled gate length (below 30 nm).

In Chapter 3, impact from thermal budget on various device characteristics such as V_T, carrier mobility, off-state leakage current will be discussed. In Chapter 4, strain effect in $Si_{1-x}Ge_x$ channel pFET will be discussed and device performance boost elements will be shown. In the last section of $Si_{1-x}Ge_x$ channel in Chapter 5, scalability of $Si_{1-x}Ge_x$

channel pFET will be discussed and optimal device structure will be proposed.

Table 1-2 Benchmark for Ge content, gate length, and drive current performance for $Si_{1-x}Ge_x$ channel pFET.

Reference	Year	Structure	Ge content x	Minimum gate length	I_{ON}@I_{OFF}=100 nA/μm
[42]	2010	$Si_{1-x}Ge_x$/Si-sub	Approx. 25%	NA	0.68 mA/μm (V_{dd}=-1V)
[43]	2007	$Si_{1-x}Ge_x$/Si-sub	70%	1 μm	NA
[44]	2007	$Si_{1-x}Ge_x$/Si-sub	25%	80 nm	0.56 mA/μm (V_{dd}=-1.2V)
[45]	2008	$Si_{1-x}Ge_x$/Si-sub	Approx. 50%	1 μm	NA
[51]	2010	$Si_{1-x}Ge_x$/Si-sub	63%	100 μm	NA
[52]	2008	$Si_{1-x}Ge_x$/Si-sub	55%	80 nm	NA
[53]	2010	$Si_{1-x}Ge_x$/sSOI	40%	22 nm	0.3 mA/μm (V_G-V_T=V_d=-1V)
[54]	2009	$Si_{1-x}Ge_x$/Si-sub	23%	100 nm	NA
[55]	2009	$Si_{1-x}Ge_x$/Si-sub	25%	1 μm	NA
[56]	2010	$Si_{1-x}Ge_x$/Si-sub	43%	10 μm	NA
[57]	2006	$Si_{1-x}Ge_x$/Si-sub	30%	> 60 nm	0.71 mA/μm (V_{dd} = -1.2V)
[58]	2010	Strained SGOI	25-35%	20 nm	0.52 mA/μm (V_{dd} = -1V)

Chapter 2 pFET effective work function control techniques in RMG

2.1 Background

In this chapter, we will discuss eWF control technique for RMG FinFET devices. As mentioned in Chapter 1, even with gate-last or RMG integration scheme which provides relatively low pFET V_T without a help from $Si_{1-x}Ge_x$ channel or various capping techniques [38, 59], recent reports highlighted eWF lowering (higher pFET V_T) in aggressively scaled EOT region (EOT < 10 A) [39, 50]. Therefore, it is still highly important to push the eWF further towards valence band edge to achieve lower pFET V_T at scaled EOT region on RMG devices. In this chapter, the eWF control methods are demonstrated for pFinFET RMG devices based on the process flow proposed in ref. [50] and [59], and the details of device fabrication flow will be discussed in Section 2.2. Section 2.3 and 2.4 will address experimental results and proposal for possible pFET eWF control techniques.

2.2 Device fabrication flow

Fabrication process flow (only RMG part is shown which is relevant to the discussion here) is shown in Figure 2-1. We fabricated RMG bulk FinFET device. In this work, we investigated process condition dependence of gate stack properties to explore eWF control methods. In ref. [50], so-called WF setting anneal is done after sacrificial TiN (sac-TiN) and amorphous Si (a-Si) cap layer were deposited on HfO_2 gate dielectric. We followed the same process flow in this work, and we abbreviate this WF setting anneal as WSA in this thesis.

After dummy poly-Si gate and dummy gate oxide (SiO_2) removal, SiO_2 interfacial layer (SiO_2-IL) and HfO_2 are formed as gate dielectric. Post deposition anneal (PDA) is done right after HfO_2 deposition to densify the film [60, 61]. After that, 1st-TiN (TiN1) and a-Si cap layers are deposited on top of HfO_2 gate dielectric. Then, WSA is done at various temperature (T_1, T_2, and T_3, spike RTA). We will discuss the impact from this WSA temperature on the gate stack properties such as V_T, mobility, and reliability in Section 2.4. After WSA, a-Si cap and TiN1 are removed successively by wet etching (a-Si cap removal is done in all cases. TiN1 removal is optional). A-Si cap is removed with ammonium hydroxide (NH_4OH), and TiN1 is removed with SC1 clean. In this experiment, TiN1 wet etch time was varied and the impact on gate stack properties were investigated. The result will be discussed in Section 2.3. After TiN1 removal (or after a-Si removal in case TiN1 is not removed), second-TiN (TiN2) is deposited in case of pFET as p-type WF (pWF) metal. In case of nFET, n-type WF (nWF) metal is deposited. All TiN and nWF metals are deposited with ALD.

The thickness of TiN1 is supposed to be thicker than critical thickness (T_{crit}) discussed in ref. [59]. In ref. [59], T_{crit} is defined as the TiN1 thickness above which eWF is stable over the air exposure time between a-Si layer removal and TiN2 deposition. We confirmed that eWF was very stable over long air exposure time.

After WF metal is deposited, gate trench is filled with CVD-W and W is planarized by CMP. Then standard BEOL process follows.

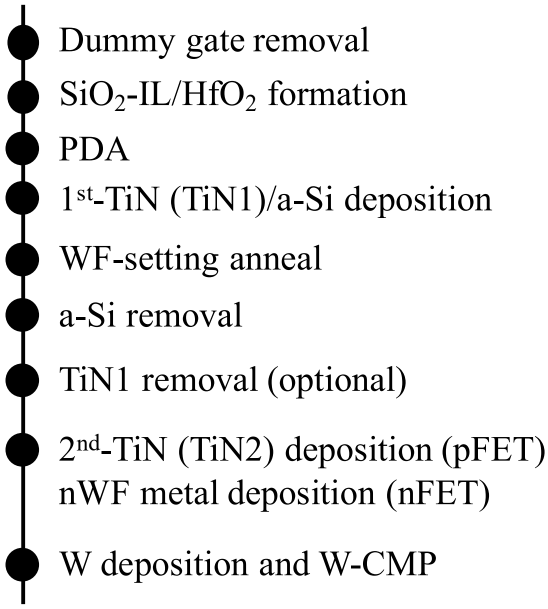


Figure 2-1 RMG process flow for FinFET devices.

2.3 Threshold voltage control by introduction of fixed charge layer

Figure 2-2 shows pFET linear I_D-V_G curve comparison between devices with and without TiN1 removal process. Gate length is 200 nm, and V_D is -50 mV. Devices without TiN1 removal process have TiN1/TiN2 stack as WF metal, while devices with TiN1 removal have only TiN2 as WF metal. The WSA was done at temperature T_2 unless mentioned otherwise. TiN1 removal over etch condition is 38x over etch. Devices without TiN1 removal process showed approximately 90 mV pFET V_T reduction compared to the ones with TiN1 removal process, even though both devices have same TiN as WF metal and eWF is supposed to be identical. This trend is actually opposite to the results of ref. [59], where pFET V_T is lowered by doing TiN1 removal process, though the process details in ref. [59] are not disclosed and their results cannot be directly compared to our results.

Regarding process uniformity of the TiN1 removal process, both devices show comparable within wafer V_T variation (Within-Wafer (WiW) range of V_T is approximately 20 mV for both devices) and there was no uniformity degradation due to TiN1 removal process.

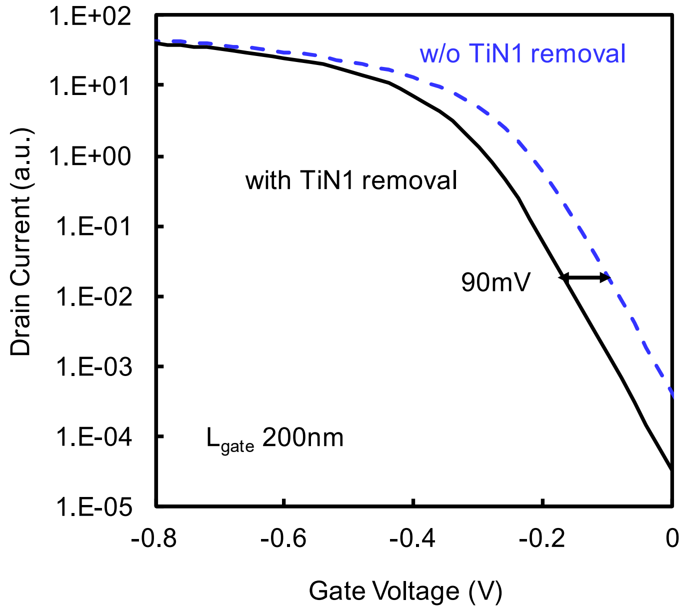


Figure 2-2 pFET linear I_D-V_G curve comparison: devices with TiN1 removal (black) and without TiN1 removal (blue). WSA was done at T_2 temperature. TiN1 removal over etch is 38x.

In order to explore the root cause of this considerable pFET V_T reduction, TiN1 wet etch time has been systematically changed and HfO_2 surface was probed with X-ray photoemission spectroscopy (XPS) to investigate the interface between TiN1 and HfO_2. Figure 2-3 shows summation of Ti 2p and N 1s XPS signal intensity (normalized to the value with the longest over etch time) as a function of TiN1 over etch time ratio with respect to no over etch case (over etch time ratio 1 means no over etch). Even though sufficient amount of over etch (more than ~5x over etch) has been applied, Ti and N signals (well above detection limit) have still been detected and they are reduced very slowly over the etch time (while over etch time is increased from ~5x to ~38x, summation of Ti and N signal is only decreased by a factor of 2). This result strongly indicates that some interfacial layer is formed between TiN1 and HfO_2 (supposedly intermixed layer) by high thermal budget of WSA and this layer should have very slow etching rate to SC1 wet etch chemistry. For convenience, this interfacial layer is called $HfTiON_x$ in later part of the thesis.

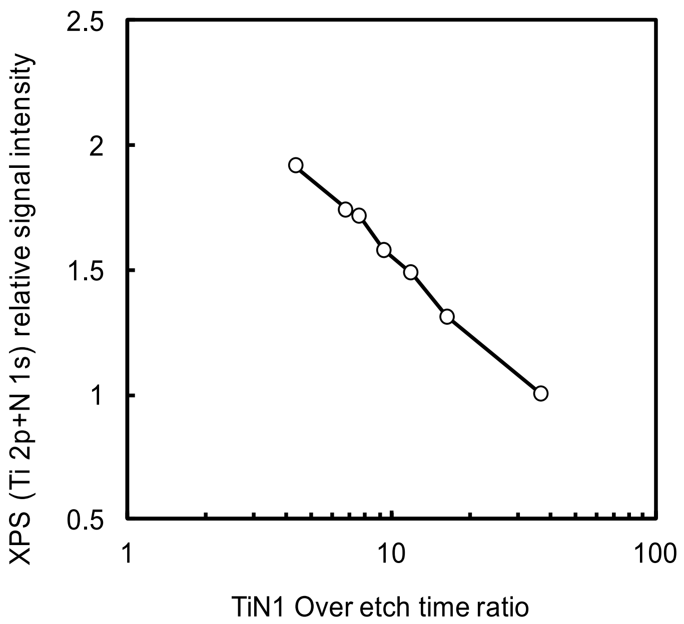


Figure 2-3 XPS Ti 2p + N 1s signal intensity vs. over etch time ratio for TiN1 removal process with respect to no over etch case. XPS has been done after TiN1 wet etch process. WSA was done at T_2 temperature.

To investigate electrical properties of the $HfTiON_x$ interfacial layer, we checked device characteristics as a function of $HfTiON_x$ layer thickness. Devices were fabricated by removing TiN1 layer with various etching time, then nWF metal (nFET) or TiN2 (pFET) was deposited as WF metal. $HfTiON_x$ thickness was calculated based on summation of Ti 2p and N 1s signal intensity measured by XPS. Figure 2-4 shows nFET and pFET V_T shift as a function of $HfTiON_x$ thickness. Gate length of the devices is 24 nm. V_T shift was calculated with respect to V_T at $HfTiON_x$ thickness of 0.28 Å (longest TiN1 over etch time, thinnest $HfTiON_x$ layer thickness). Linear V_T shift was observed as a function of $HfTiON_x$ thickness. Both nFET V_T and pFET V_T show positive shift with thicker $HfTiON_x$, meaning nFET V_T increases and pFET V_T reduces. It should be noted that magnitude of V_T shift is almost same for n- and pFET at a given $HfTiON_x$ thickness. As seen in Figure 2-2, if this $HfTiON_x$ layer is not removed at all (in case of no TiN1 removal process), it gives approximately 90 mV V_T shift.

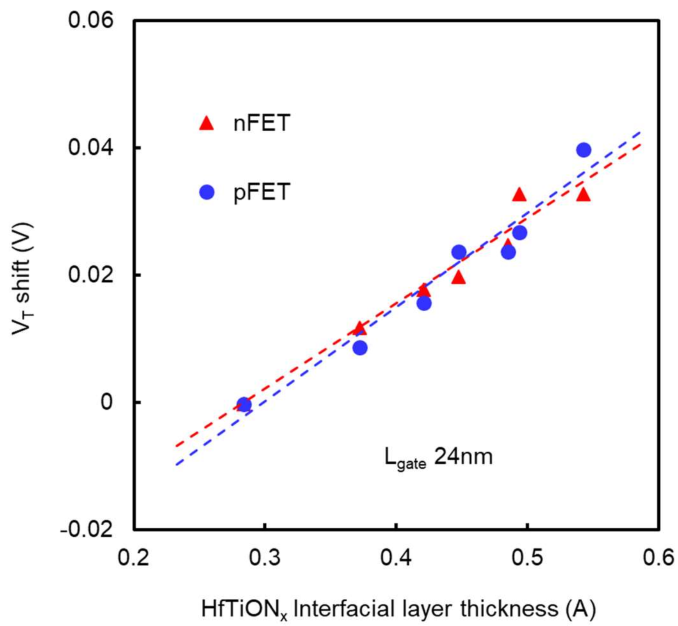


Figure 2-4 n- and pFET V_T shift vs. $HfTiON_x$ interfacial layer thickness. WSA was done at T_2 temperature.

Figure 2-5 shows normalized T_{inv} and J_{ginv} (gate leakage in inversion) trend as a function of $HfTiON_x$ interfacial layer thickness. T_{inv} and J_{ginv} are normalized with respect to the ones at $HfTiON_x$ thickness of 0.28 Å. T_{inv} gets thicker and J_{ginv} becomes smaller with increasing $HfTiON_x$ interfacial layer thickness, which suggests that the interfacial layer is dielectric material, not metal. Therefore, V_T shift observed in Figure 2-4 is supposed to be caused by negative fixed charge in $HfTiON_x$ layer.

Based on the discussion so far, it can be concluded that WSA creates interfacial dielectric layer at TiN1/HfO_2 interface (we call it $HfTiON_x$ for convenience) and this layer has negative fixed charges which increases nFET V_T and reduces pFET V_T (V_T shift towards positive direction) as seen in Figure 2-4.

Fixed charge density (N_f) in $HfTiON_x$ layer can be estimated from the equation

$$N_f = {}^{1}/_{e}\, C_g \times \delta V_T \qquad (2\text{-}1)$$

where e is elemental charge, C_g and δV_T are gate capacitance per unit area and V_T shift, respectively. In this case, N_f was estimated to be approximately 2×10^{12} cm^{-2}.

Figure 2-6 shows normalized electron and hole mobilities with various $HfTiON_x$ interfacial layer thicknesses (0.28 and 0.54 Å). Normalization was done with respect to peak mobility of $HfTiON_x$ 0.28 Å sample. Although mobility at high effective field is all similar regardless of the $HfTiON_x$ interfacial layer thickness, there is slight mobility degradation only at low N_{inv} region with thicker $HfTiON_x$ interfacial layer, especially on nFET (electron mobility). This carrier mobility degradation is the most likely due to additional remote Coulomb scattering from the fixed charges in $HfTiON_x$ interfacial layer, as it only affects lower effective field region. Practically, as we would either remove this $HfTiON_x$ layer from nFET side or remove TiN1/a-Si stack from nFET side before WSA to have this $HfTiON_x$ layer only on pFET when integrated into CMOS process flow, relatively minor mobility degradation on pFET side may not be a concern from device performance point of view.

To conclude this section, we identified that the root cause of the V_T difference between devices with and without TiN1 wet etching is negatively charged dielectric layer which is supposedly formed by intermixing of TiN and HfO_2 layers during WSA. This charged layer can provide 90 mV pFET V_T reduction, although a certain amount of carrier mobility degradation was observed both on n- and pFET. These trade-off relations between various device characteristics are summarized in Table 2-1. To pursue lower pFET V_T, $HfTiON_x$ layer should be kept as thick as possible despite of T_{inv} and mobility degradation.

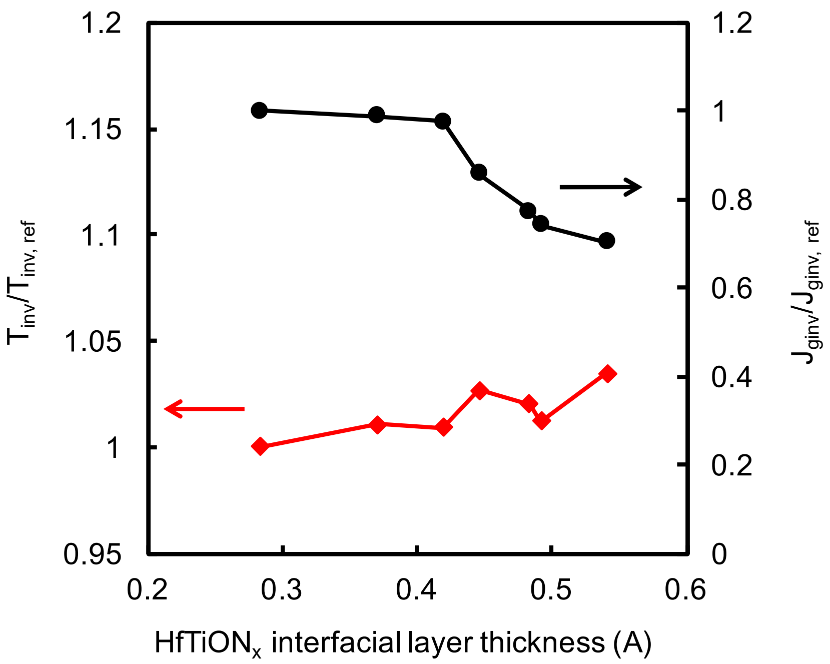
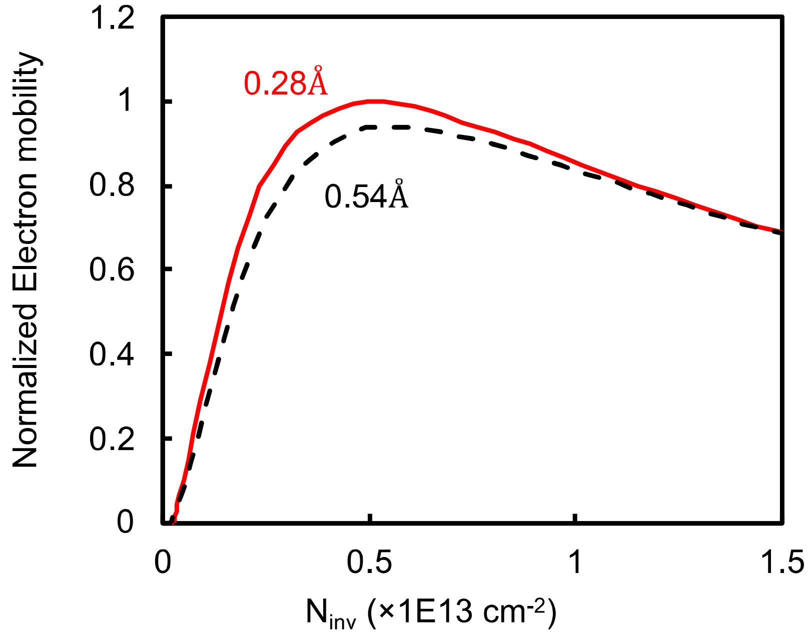


Figure 2-5 Normalized T_{inv} and J_{ginv} vs. $HfTiON_x$ interfacial layer thickness. WSA was done at T_2 temperature.



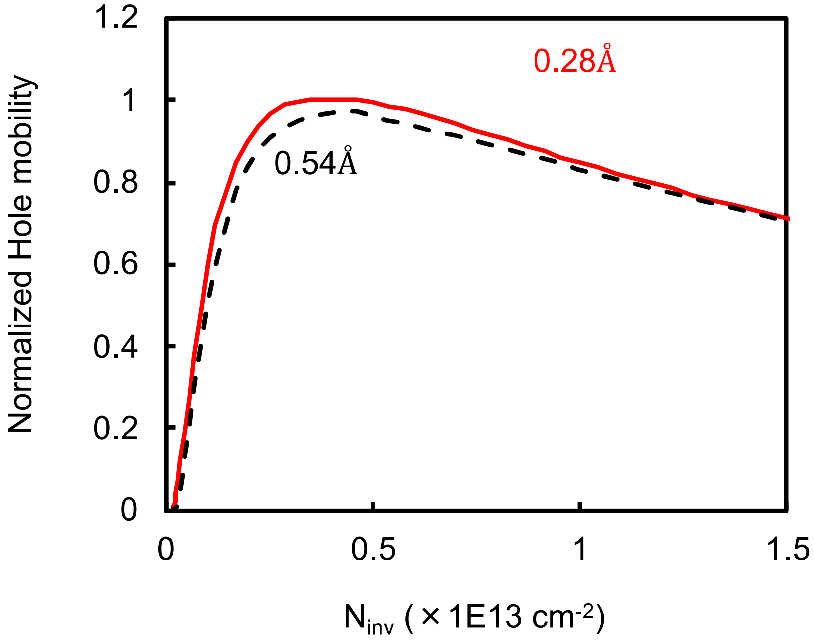


Figure 2-6 Electron mobility (top) and hole mobility (bottom) for various $HfTiON_x$ interfacial layer thickness (0.28, 0.54 Å). WSA was done at T_2 temperature.

Table 2-1 Trade-off of device characteristics by introduction of $HfTiON_x$ fixed change layer.

	$HfTiON_x$ thickness ↑
pFET V_T	lower
T_{inv}	thicker
Gate leakage	lower
Hole mobility	degraded

2.4 Threshold voltage control by oxygen vacancy density modulation

In this section, impact from WSA temperature on gate stack properties will be discussed.

Figure 2-7 shows nFET and pFET relative saturation V_T (relative V_T with regard to V_T at L_{gate} = 2 μm and at T_1 for WSA) as a function of gate length with various WSA temperature ($T_1 < T_2 < T_3$). TiN1 was removed with the longest SC1 time in Figure 2-3 (38x over etching) for all samples. $HfTiON_x$ layer thickness probed with XPS for the samples with T_1 and T_3 temperature is matched to the one with T_2, approximately 0.28 Å. Therefore, the V_T shift by negative fixed charge in $HfTiON_x$ layer can be regarded identical for all samples. By increasing WSA temperature from T_1 to T_3, pFET V_T is consistently reduced up to approximately 50 mV, while nFET V_T is only increased by 20 mV from T_1 to T_2, and nearly no change was observed from T_2 to T_3. Overall, nFET V_T looks less sensitive to WSA temperature. Within wafer V_T variation was similar for all temperatures (not shown).

In pFET, V_T shift happens evenly for all gate lengths (from 24 nm to 2 μm), indicating this V_T shift is not driven by short channel effect (more dopant diffusion due to higher annealing temperature), but driven by gate stack modification. As mentioned earlier, HfTiON$_x$ interfacial layer thickness after TiN1 removal is confirmed to be identical among all WSA conditions, therefore this V_T shift is not due to HfTiON$_x$ layer thickness difference.

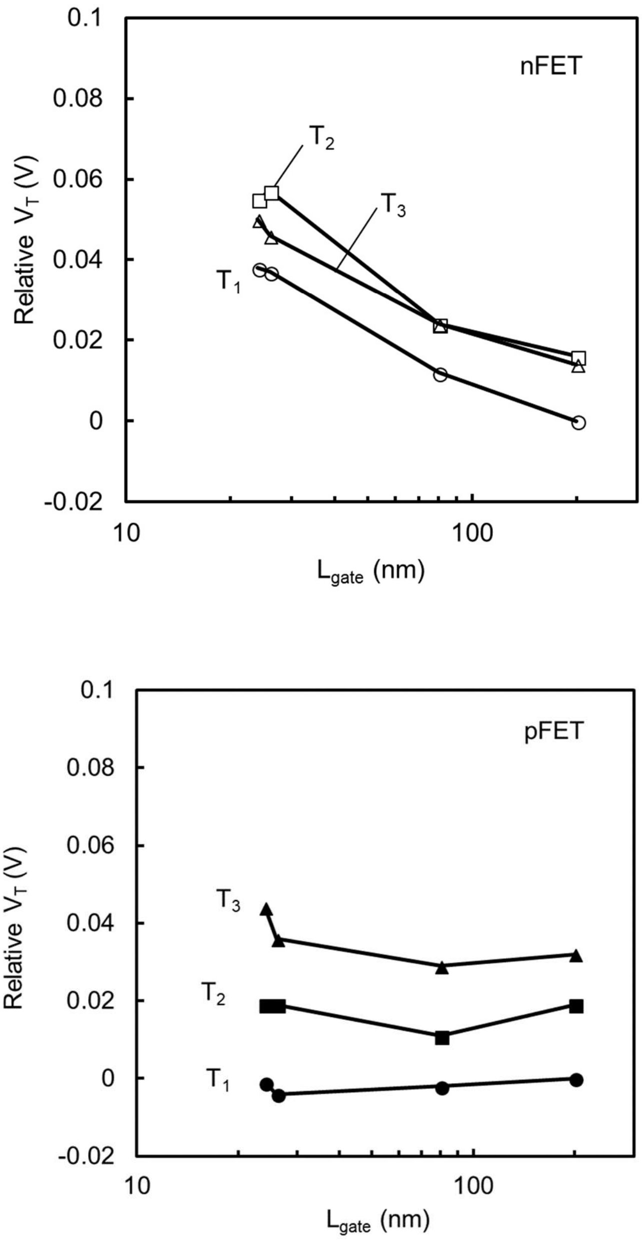


Figure 2-7 nFET (top) and pFET (bottom) Relative V_T vs L_{gate}. T_1 (circle), T_2 (square), T_3 (triangle).

To figure out more what is happening on gate stack by WSA, SiO_2-IL thickness (measured by XPS) and nFET T_{inv} have been checked as a function of anneal temperature. The result is shown in Figure 2-8. As WSA temperature goes higher from T_1 to T_3, SiO_2-IL thickness was confirmed to get thicker, by approximately 0.8 Å. This SiO_2-IL regrowth translates to thicker T_{inv} with higher WSA. Figure 2-8 shows nFET T_{inv}, but similar trend has been confirmed for pFET T_{inv} as well. This SiO_2-IL regrowth is due to oxidation of Si substrate, suggesting oxygen has been supplied from somewhere in the gate stack. Si atom in a-Si cap is isolated from HfO_2/SiO_2-IL by TiN1 layer, therefore Si atom should not be involving in the reaction.

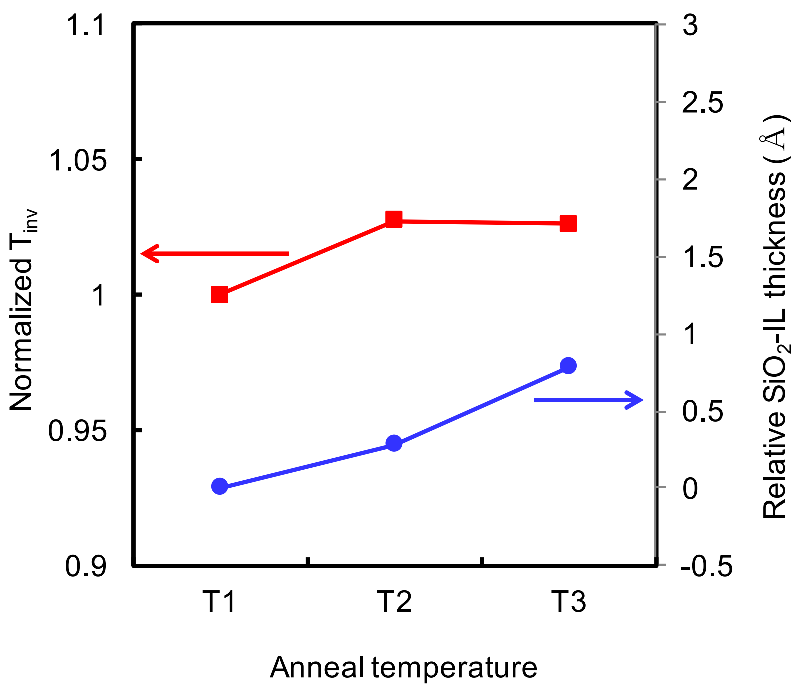


Figure 2-8 Relative SiO_2-IL thickness and nFET T_{inv} trend as a function of WSA temperature.

To understand the mechanism of V_T shift and SiO_2-IL regrowth during WSA, XPS has been carried out on pre and post WSA samples (Si-substrate/SiO_2-IL/HfO_2/TiN1/a-Si) and the results are shown in Figure 2-9. Samples were prepared by sputtering a-Si cap layer close to a-Si/TiN1 interface after a-Si deposition (pre-anneal sample) or after WSA (post-anneal sample). The WSA was done at T_2. Si 2p and Ti 2p XPS spectra are shown for pre- and post WSA. Take-off angle is 90 degrees (normal to the sample surface).

In Figure 2-9 (a), Ti 2p peaks mainly show components from TiN and TiO_x. Peaks corresponding to TiN are at around binding energies of 455 eV and 461 eV, the peak at 455 eV comes from Ti 2p3/2 and the peak at 451 eV is from Ti 2p1/2 due to spin-orbit splitting. Peaks corresponding to TiO_x are at around 458 eV and 464 eV from the above mentioned Ti 2p orbitals. Existence of TiO_x peak suggests surface oxidation of TiN1 layer before a-Si layer is deposited, as no oxygen can be incorporated during deposition itself. This TiO_x peak intensity is clearly reduced after WSA compared to pre-anneal sample both at 458 and 464 eV peaks. On the other hand, SiO_2 peak in Si 2p spectra (Figure 2-9 (b)) shows increase at post WSA.

These XPS results indicate that oxygen atoms diffuse from TiN1 layer towards HfO_2/SiO_2-IL during WSA and some of them oxidize the Si fin and regrow the SiO_2-IL. If there are oxygen vacancies (V_O^{2+}) in HfO_2, some of the diffused oxygen atoms should passivate and electrically neutralize them. As V_O^{2+} is positively charged defect [62], once neutralized, V_T should shift in positive direction. Oxygen vacancies are supposedly created during PDA [32]. Higher WSA temperature should diffuse more oxygen atoms from TiN1, therefore more V_O^{2+} should be passivated and more SiO_2-IL regrowth should happen. This model can explain larger positive pFET V_T shift and larger SiO_2-IL regrowth with higher temperature anneal observed in Figure 2-7 and Figure 2-8, respectively.

According to this model, magnitude of V_T shift should be comparable on nFET as there is no process difference up to this point between n- and pFET. Therefore, relatively insensitive V_T behavior observed in nFET cannot be explained clearly by this model. Difference in work function of metal gate electrode may play a role, but further study is needed for more complete understanding.

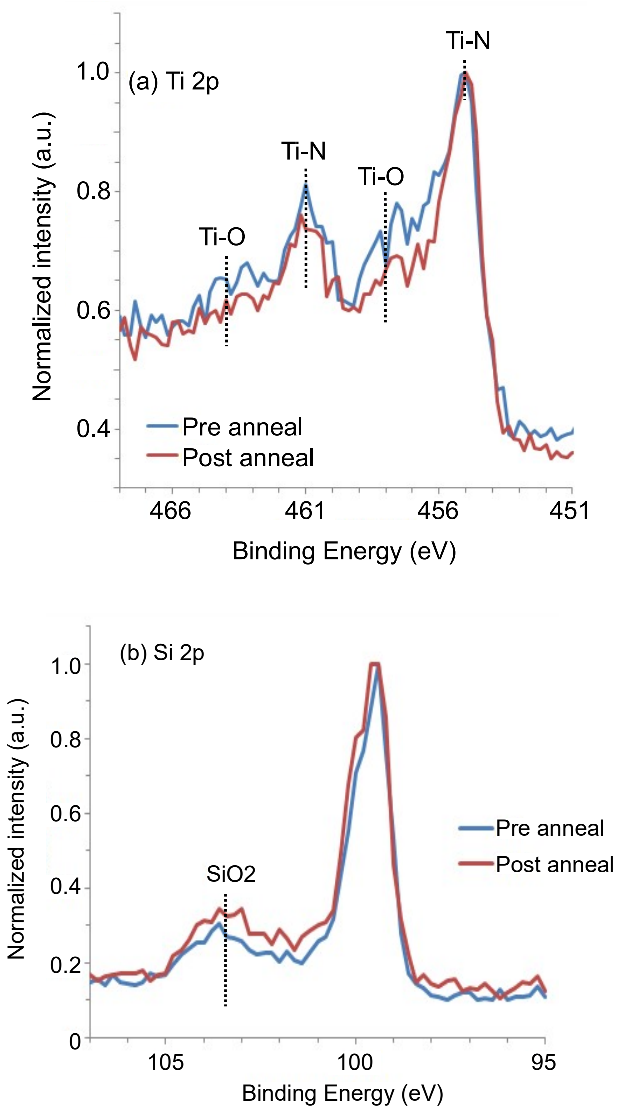


Figure 2-9 AR-XPS results (a) Ti 2p and (b) Si 2pspectra. Pre and post WSA (T_2) are plotted.

Although higher WSA temperature is favorable to pFET V_T reduction, one concern is degradation of short channel effect. Figure 2-10 is summary of drain current, I_{dsat} (at fixed off-current, I_{off_d}) versus overlap capacitance for nFET and pFET at WSA temperature of T_2 and T_3. In pFET, overlap capacitance increase by WSA temperature increase is relatively small (approximately 2%) and not a major issue. In nFET, although overlap capacitance increases significantly with anneal temperature of T_3 (approximately 8%), we could obtain comparable overlap capacitance and I_{dsat} to T_2 by optimizing junction process.

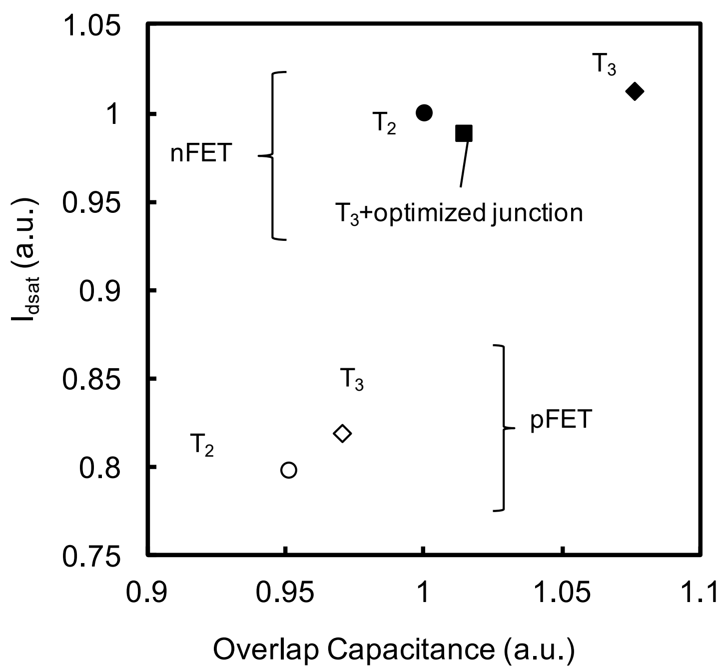


Figure 2-10 I_{dsat} vs. overlap capacitance for n- and pFET with WSA at T_2, T_3, and T_3 with optimized junction process.

Figure 2-11 shows V_T shift of NBTI (Negative Bias Temperature Instability) stress as a function of relative SiO_2-IL thickness (characterized by XPS). SiO_2-IL thickness was varied by temperature of WSA and PDA, while HfO_2 thickness was almost identical for all devices. It is clearly shown that higher WSA temperature significantly improves NBTI by approximately 30% at a given SiO_2-IL thickness.

One possibility is an improvement of SiO_2/Si interface quality by higher temperature anneal. We checked interface trap density (D_{it}) on device A and B in Figure 2-11 (A and B have almost identical SiO_2-IL thickness) and they turned out to be comparable (D_{it} is about 1.1×10^{11} cm^{-2} for both devices). Therefore, NBTI improvement cannot be explained by interface quality improvement. It has been reported that NBTI is more dominantly affected by hole trapping in bulk HfO_2 rather than SiO_2-IL/Si interface degradation when SiO_2-IL is scaled down and direct tunneling from substrate to HfO_2 bulk starts to happen [63]. Based on that, this result suggests that higher WSA helps to reduce hole trap defects in HfO_2. This might be related to passivation of oxygen vacancies, but further study is needed to get more-clear understanding.

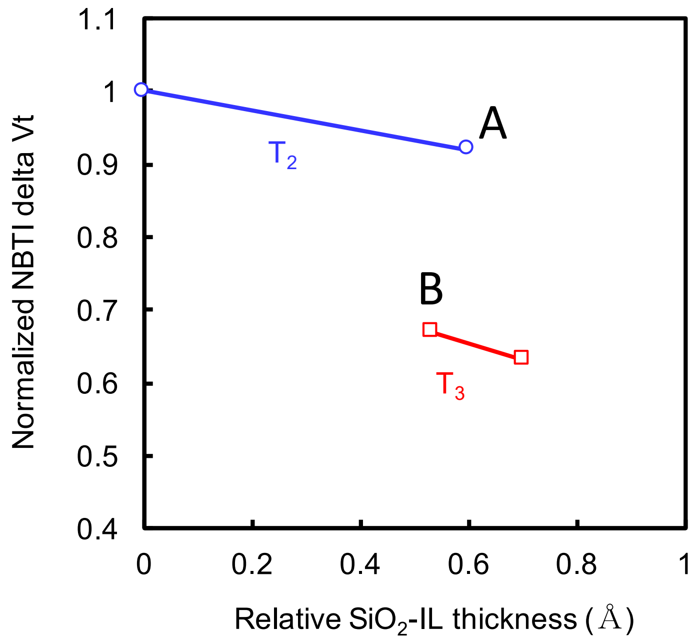


Figure 2-11 NBTI V_T shift as a function of relative SiO_2-IL thickness. WSA temperatures T_2 and T_3 are plotted ($T_2 < T_3$).

2.5 Summary

In this chapter, we have investigated the effect of WSA (high temperature annealing on a-Si/TiN/HfO_2 stack) on gate stack properties to achieve lower pFET V_T with more band-edge work function. The technique of pFET V_T control discussed in this work is illustrated in Figure 2-12. It was found that intermixed layer created in-between TiN and HfO_2 during WSA ($HfTiON_x$) has negative fixed charges and it reduces pFET V_T (positive V_T shift) by about 90 mV. On top of that, it was also found that higher anneal temperature further reduces pFET V_T by about 50 mV (by increasing anneal temperature from T_1 to T_3) while keeping nFET V_T almost unchanged. This could be explained by passivation of oxygen vacancies (positively charged) in HfO_2 with oxygen atoms diffused from TiN1 layer. By combining these effects ($HfTiON_x$ interfacial layer and higher temperature

WSA), one can further push effective work function towards valence band edge by approximately 140 meV and achieve lower pFET V_T.

The pFET V_T reduction is extremely difficult in scaled EOT devices because conventionally used metal gate material cannot provide sufficiently low pFET V_T. This proposed technique has great value as it doesn't require change in work function metal material and integration hurdle is supposedly lower than changing metal material or its thickness. This technique can be applied to multi-V_t integration as well as shown in Figure 2-13. For example, a-Si and TiN1 cap can be removed prior to WSA selectively from certain devices so that WSA modulates pFET V_T for devices which are covered with a-Si and TiN1 cap and V_T difference can be created (flow (A) in Figure 2-13). Alternatively, we can apply WSA for all devices and remove $HfTiON_x$ interfacial layer selectively from certain devices to create V_T difference. In this case, all devices will receive WSA anneal and we can expect NBTI benefit for all devices (flow (B) in Figure 2-13). For multi-V_T integration, multiple work-function metal deposition and patterning [64] is conventional approach and Table 2-2 shows other representative multi-V_T integration methods. The multi-V_T integration method proposed in this work is different from any of those and hence original.

NBTI was significantly improved with higher WSA temperature likely due to reduction of hole trapping sites in HfO_2. PDA is supposed to densify and crystalize HfO_2 film to some extent (depending on HfO_2 thickness and PDA temperature) before depositing TiN1 film, but the effect on reliability is relatively limited compared to WSA. WSA is necessary to improve reliability. Both anneal cause SiO_2-IL regrowth but its thickness is still in acceptable range.

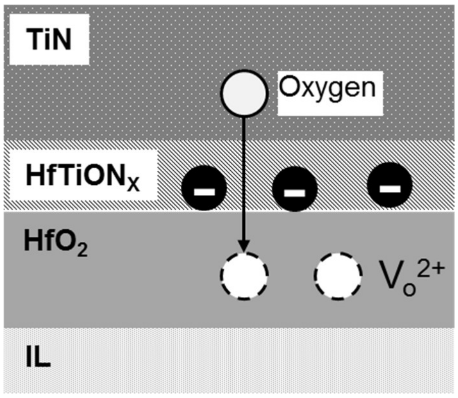


Figure 2-12 Schematic illustration of pFET V_T control techniques discussed in this work. Negative fixed charges in $HfTiON_x$ layer and passivation of oxygen vacancies with diffused oxygen from TiN (elimination of positive charges in HfO_2) reduces pFET V_T.

Figure 2-13 Schematics for possible multi-V_T integration using eWF control technique discussed in this work. (A) a-Si/TiN1 removal prior to WSA, (B) $HfTiON_x$ removal after WSA.

Table 2-2 Representative multi-Vt integration methods [65].

	Method		Principle	N	P
1	Gate Work Function	Separate Metal	Unique Metal Work Function	Ta, Al, Hf	Pt, Ti, Mo
2		Al Diffusion	Al Concentration	O	O
3		Capping Layer	Dielectric Dipole Eng	SrO, La2O3, Lu2O3, Y2O3	Al2O3, TiO2, ZrO2, HfO2
4		SiH4 Soak		O	o
5		Metal Gate Implant	N concentration	O	O
6	Fin	Channel Material	Valence Band (SiGe)	X	O
7		Fin Width	Quantum Confinement	O	O
8		Channel Doping	Depletion Charge	O	O

Chapter 3

Process temperature impact on $Si_{1-x}Ge_x$ channel transistor

3.1 Background

In this chapter, we first discuss fabrication process flow for $Si_{1-x}Ge_x$ channel pFET. Base process is gate-first metal gate/high-k transistor (Si channel) developed in Interuniversity Microelectronics Centre (imec) described in ref. [66, 67, 68]. $Si_{1-x}Ge_x$ channel pFET was fabricated by adding $Si_{1-x}Ge_x$ epitaxy process on top of the base process flow [47, 69]. We also developed so-called implant free (IF-) $Si_{1-x}Ge_x$ pFET which doesn't have ion implanted source/drain and extension [70, 71]. The details of process flow for those various type of transistor structures is discussed below.

3.2 Transistor structures studied in this work

Summary for device structures studied in this work is shown in Table 3-1. As shown in the table, three kinds of device structure were fabricated. All devices are gate-first metal gate/high-k transistors.

First type of device is Si channel as a reference. For Si channel device, extension, halo and source/drain were formed by ion implantation. We fabricated the device with and without embedded in-situ boron doped SiGe (eSiGe:B) stressor to compare the device performance between them. This technology was firstly introduced into manufacturing by Intel Corporation in their 90 nm node technology to boost pFET performance [72]. It was shown that hole inversion mobility was improved by applying uniaxial compressive strain to pFET channel [73].

Second type of device, so-called conventional $Si_{1-x}Ge_x$ channel device also has ion implanted extensions and source/drain [47, 69]. We have fabricated this type of devices with and without eSiGe:B stressor similarly to Si channel case. The interaction between $Si_{1-x}Ge_x$ channel and eSiGe:B stressor was systematically studied and will be discussed in Chapter 4.

The third device, IF-$Si_{1-x}Ge_x$ channel device has very different device structure from the other two, distinct difference is extension structure. While Si channel and conventional $Si_{1-x}Ge_x$ channel devices have extension created by ion implantation, in IF-$Si_{1-x}Ge_x$ channel device, extension is formed by epitaxial growth of in-situ B doped $Si_{1-x}Ge_x$ and raised from the channel (raised SiGe:B extension) [70]. IF-$Si_{1-x}Ge_x$ devices were also fabricated with and without eSiGe:B stressor [71]. We compared device performance and scalability between IF-$Si_{1-x}Ge_x$ with and without eSiGe:B stressor, the result will be discussed in Chapter 5. For IF-$Si_{1-x}Ge_x$ channel device with eSiGe:B, we fabricated the devices with and without source/drain implant to see the impact on device performance and scalability.

Table 3-1 Summary of transistor structures studied in this work.

<table>
<tr><th>Device type</th><th>Channel material</th><th>Extension</th><th>eSiGe:B Stressor</th><th>Source/Drain implant</th><th>Reference</th></tr>
<tr><td rowspan="2">Si channel</td><td rowspan="2">Si</td><td rowspan="2">BF_2 implant</td><td>×</td><td rowspan="2">×</td><td rowspan="2"></td></tr>
<tr><td></td></tr>
<tr><td rowspan="2">Conventional $Si_{1-x}Ge_x$ channel</td><td rowspan="2">$Si_{1-x}Ge_x$</td><td rowspan="2">BF_2 implant</td><td>×</td><td rowspan="2">×</td><td rowspan="2">[70]</td></tr>
<tr><td></td></tr>
<tr><td rowspan="3">IF-$Si_{1-x}Ge_x$ channel</td><td rowspan="3">$Si_{1-x}Ge_x$</td><td rowspan="3">Raised SiGe:B epitaxy</td><td rowspan="2">×</td><td>×</td><td rowspan="3">[71]</td></tr>
<tr><td></td></tr>
<tr><td></td><td></td></tr>
</table>

3.3 Device fabrication flow

Figure 3-1 shows schematics for process flow for three types of devices summarized in Table 3-1. After STI is formed, well is created by ion implantation. Well anneal is done in furnace to cure the damage in the substrate created by ion implantation. Following the well annealing, Si substrate is recessed by HCl chemical etching before growing $Si_{1-x}Ge_x$ channel. Prior to epitaxial growth of $Si_{1-x}Ge_x$ layer, pre-bake treatment is done with H_2 ambient, at 800ºC. This is to remove residual oxygen and moisture (H_2O) from the Si surface and to achieve atomically clean surface. This helps to reduce potential abnormal epitaxial growth. $Si_{1-x}Ge_x$ film with various Ge concentrations (ranging from 25% to 55%) is epitaxially grown, then Si capping layer follows. $Si_{1-x}Ge_x$ layer and Si capping layer thickness are typically 3 to 7 nm and 1 to 3 nm, respectively. Substrate recess, pre-bake and epitaxial growth are processed successively in single chamber without air exposure. After $Si_{1-x}Ge_x$ channel/Si-cap is grown, native oxide on top of Si capping is removed by diluted hydrofluoric acid (DHF) and SiO_2-IL is grown with diluted HCl/O_3 mixture [74]. Thickness of SiO_2-IL is about 0.8 nm. After SiO_2-IL formation, HfO_2 is deposited on top of SiO_2-IL as gate dielectric by ALD method. HfO_2 thickness is 1.8 nm. PDA is done in N2 ambient at 700ºC. After PDA, TiN is deposited by physical vapor deposition (PVD) as metal gate electrode. TiN thickness is 2 nm. After TiN deposition, a-Si deposition by CVD and gate patterning follows. Figure 3-2 shows cross-sectional transmission electron microscopy (XTEM) image of $Si_{1-x}Ge_x$ channel device, taken after HfO_2/TiN and a-Si deposition.

In case of Si channel or conventional $Si_{1-x}Ge_x$ channel device, extension and halo formation by ion implantation is done after the gate patterning. Ion implantation condition for Si channel device is B, 0.7 keV, 7×10^{14} cm^{-2} and F, 10 keV, 2×10^{15} cm^{-2} for extension, As, 40 keV, 3.5×10^{13} cm^{-2} for halo. For conventional $Si_{1-x}Ge_x$ channel device, BF_2, 2 keV, 1×10^{15} cm^{-2} was used as extension implantation. After extension and halo are formed, sidewall spacer is created to offset the channel and source/drain. In case of devices with eSiGe:B stressor, Si substrate is recessed by about 50 nm to create a cavity, then in-situ boron doped $Si_{1-x}Ge_x$ is grown by epitaxy at 650ºC (pre-bake condition is the same as the one for $Si_{1-x}Ge_x$ channel growth, 800ºC). Ge concentration in eSiGe:B stressor is 25%, boron concentration is around 1×10^{20} cm^{-3}. After eSiGe:B growth, source/drain is created by ion implantation, B, 3 keV, 3×10^{15} cm^{-2}.

In case of IF-$Si_{1-x}Ge_x$ channel devices, after the gate patterning, first offset spacer (6 nm) is formed and in-situ boron doped $Si_{1-x}Ge_x$ is epitaxially grown as a raised extension. Ge concentration is 25% and boron concentration is around 1×10^{20} cm^{-3} (same as eSiGe:B stressor). Raised SiGe:B (we call this rSiGe:B in the rest of the thesis) extension thickness is 30 nm. After the rSiGe:B extension formation, second offset spacer is formed. In case of no eSiGe:B stressor, after the second spacer formation, activation anneal is done (this thermal treatment is not meant to activate the dopant but to create a certain amount of overlap between the channel and extension. This will be discussed later). In case eSiGe:B is added for IF-$Si_{1-x}Ge_x$ channel devices, rSiGe:B and Si substrate is recessed (recess depth is 50 nm from gate dielectric – substrate interface) and eSiGe:B is again grown. It should be noted that IF-$Si_{1-x}Ge_x$ channel devices do not receive any halo implantation.

Activation annealing is done either by spike RTA or by laser annealing [66, 68]. Spike RTA temperature is ranging from 950ºC to 1035ºC. Laser annealing temperature is 1150ºC. In case of Si channel and conventional $Si_{1-x}Ge_x$ channel devices, this anneal is literally to activate dopants. On the other hand, in case of IF-$Si_{1-x}Ge_x$ channel devices, this annealing is applied to diffuse boron from rSiGe:B extension and to create a certain amount of overlap between the channel and extension. With this overlap, external resistance (R_{ext}) can be lowered enough to have decent transistor performance.

Then, NiPt silicide is formed to have sufficiently low contact resistance between source/drain material (either Si or $Si_{0.75}Ge_{0.25}$) and substrate contact metal (Ti/TiN) [75, 76]. SiO_2 is deposited as ILD and planarized by CMP. MOL and BEOL process completes the device fabrication flow. Figure 3-3 shows XTEM images for (a) IF-$Si_{1-x}Ge_x$ channel device w/o eSiGe:B and (b) w/ eSiGe:B after NiPt silicidation. It should be noted that rSiGe:B extension is still left next to the first offset spacer in case of devices with eSiGe:B.

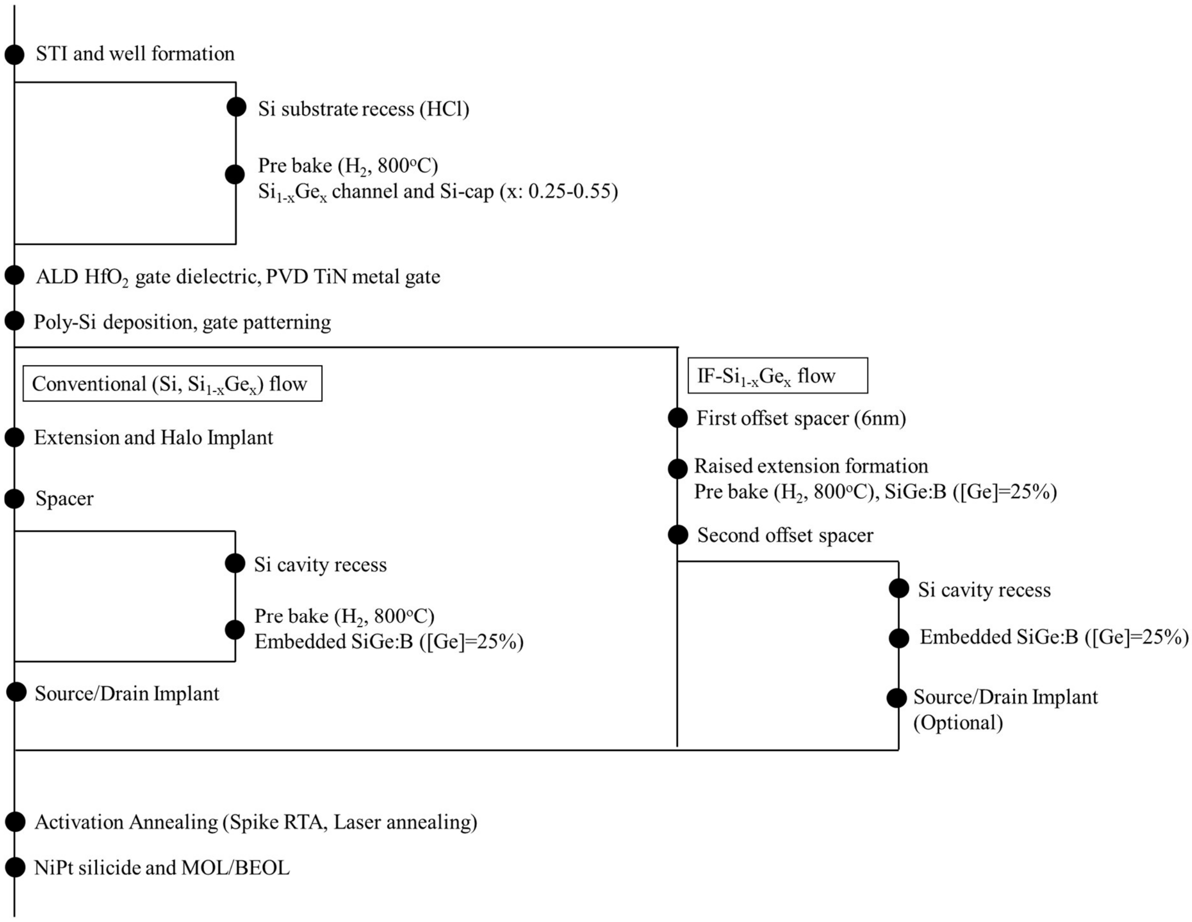


Figure 3-1 Device fabrication flow.

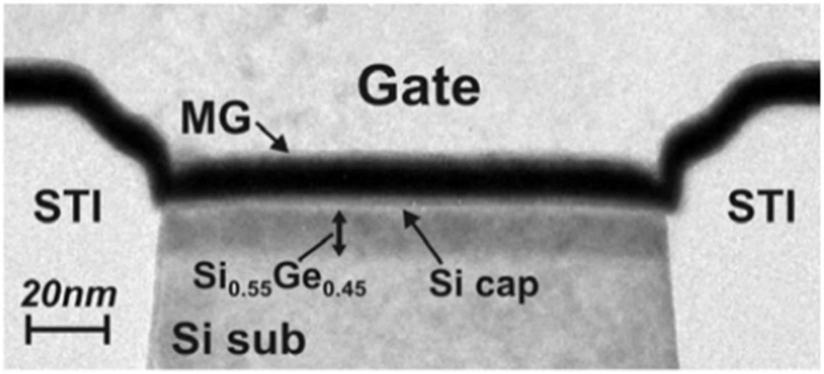


Figure 3-2 XTEM image of $Si_{1-x}Ge_x$ channel transistor (post metal gate and a-Si deposition) [69].

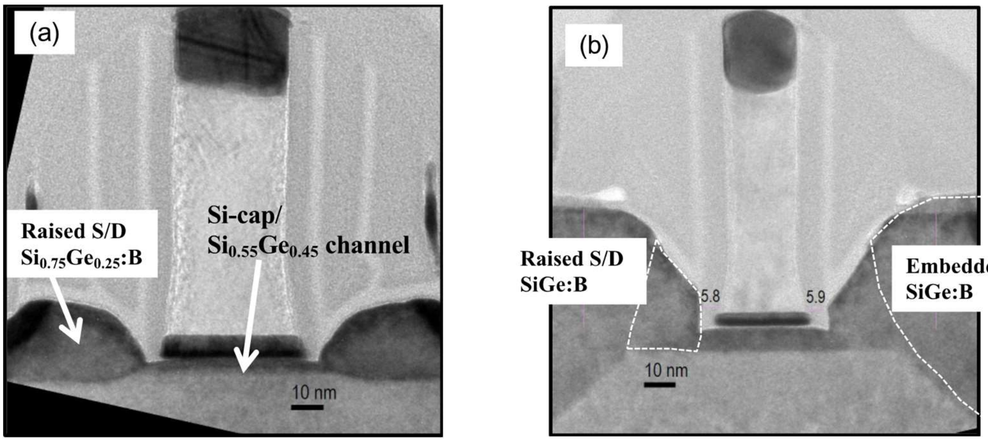


Figure 3-3 XTEM images of IF-$Si_{1-x}Ge_x$ channel devices. (a) IF-$Si_{1-x}Ge_x$ devices w/o eSiGe:B, (b) IF-$Si_{1-x}Ge_x$ devices w/ eSiGe:B.

3.4 Thermal stability of $Si_{1-x}Ge_x$ channel

As $Si_{1-x}Ge_x$ layer is epitaxially grown on Si substrate, it is forced to have identical in-plane lattice constant as Si substrate, despite of $Si_{1-x}Ge_x$ lattice constant being larger than Si [77]. Therefore, $Si_{1-x}Ge_x$ layer is under bi-axial compressive strain and tetragonally distorted due to this lattice mismatch between Si and $Si_{1-x}Ge_x$, as shown in

Figure 3-4 [48]. In Figure 3-4 (a), a_A, a_B, h_A, and h_B are lattice constant of Si, $Si_{1-x}Ge_x$, total layer thickness of Si and $Si_{1-x}Ge_x$, respectively, and in Figure 3-4 (b), $a_{\parallel}$, $a_{A\perp}$, and $a_{B\perp}$ are in-plane lattice constant, perpendicular lattice constant for Si and $Si_{1-x}Ge_x$, respectively. In case of $Si_{1-x}Ge_x$ growth on Si substrate, h_A can be regarded as infinite, and $a_{\parallel}$ and $a_{A\perp}$ are identical to Si lattice parameter a_{Si}.

Lattice constant of Si and Ge are 5.4310 Å and 5.6575 Å, respectively, and mismatch between them is approximately 4.17 %. In case of $Si_{1-x}Ge_x$, the lattice constant ($a_{\mathrm{Si_{1-x}Ge_x}}$) varies almost linearly according to Ge fraction in $Si_{1-x}Ge_x$ layer as predicted by Vergard's law [78], although small variations to this have been measured. An approximation is given by [79]

$$a_{\mathrm{Si}_{1-x}\mathrm{Ge}_x} = 0.5431 + 0.01992x + 0.0002733x^2 \text{ (nm)} \quad (3\text{-}1)$$

As the thickness of the epitaxial $Si_{1-x}Ge_x$ layer is increased, there exists a maximum thickness, called the critical thickness, h_c, above which it costs too much energy to elastically strain additional $Si_{1-x}Ge_x$ in coherence with the Si substrate. Defects appear, in this case misfit dislocations, which act to relieve the strain in the epitaxial film. The epitaxial layer relaxes and the defects interact with the electrical, optical, and thermal properties of the material, typically degrading the device performances. According to Matthews and Blakeslee, h_c, is given by [80]

$$h_c \cong 1.7793x^{-1.2371}\text{nm} \quad (3\text{-}2)$$

This critical thickness, h_c, is plotted in Figure 3-5 [48] and corresponds to the boundary between the stable and metastable regions. Experimentally it was observed that many pseudo-morphic layers could be grown well above the critical thickness values predicted from equation (3-2). This is predominantly related to a kinetic barrier to the relaxation process allowing metastable layers to be grown. Houghton provided a general expression for the extent of plastic strain relaxation, $\Delta\epsilon(t)$, of a strained layer under an arbitrary thermal cycle of temperature T in time, t [81]

$$\Delta\epsilon(t) = 9.4 \times 10^3 N_0 t^2 (\tau_{\mathrm{eff}})^{4.5} \exp\left[\left(-\frac{4.75}{k_B T}\right)\right] \qquad (3\text{-}3)$$

where N_0 is the density of initial nucleation centers for dislocations at $t = 0$, τ_{eff} is the effective stress, k_B is Boltzmann's constant. From (3-3), it is shown that the extent of plastic strain relaxation is thermally activated process, increases with thermal budget (temperature T and time t) applied to strained $Si_{1-x}Ge_x$ layer. The nucleation rate for dislocations $\frac{\mathrm{d}N(t)}{\mathrm{d}t}$, where $N(t)$ is nucleation center density at time t, is determined as

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = 0.7 \times 10^5 N_0 (\tau_{\mathrm{eff}})^{2.5} \exp\left[-\left(\frac{2.5}{k_B T}\right)\right] \qquad (3\text{-}4)$$

This equation shows that the nucleation rate is promoted by temperature of the thermal cycle (and also by effective stress, τ_{eff}). As dislocations in $Si_{1-x}Ge_x$ layer can cause various issues in transistor operation, such as junction leakage current increase due to defects in depletion layer (defects can become recombination generation center), carrier mobility degradation due to additional scattering of the carriers, and V_T shift due to loss of applied compressive strain. Therefore, it is extremely critical to maintain the applied strain in $Si_{1-x}Ge_x$ layer throughout the whole process flow.

In this chapter, to start off the discussion on $Si_{1-x}Ge_x$ channel device characteristics, impact from thermal treatment during device fabrication flow on $Si_{1-x}Ge_x$ channel device (conventional and IF-$Si_{1-x}Ge_x$ channel devices) will be discussed.

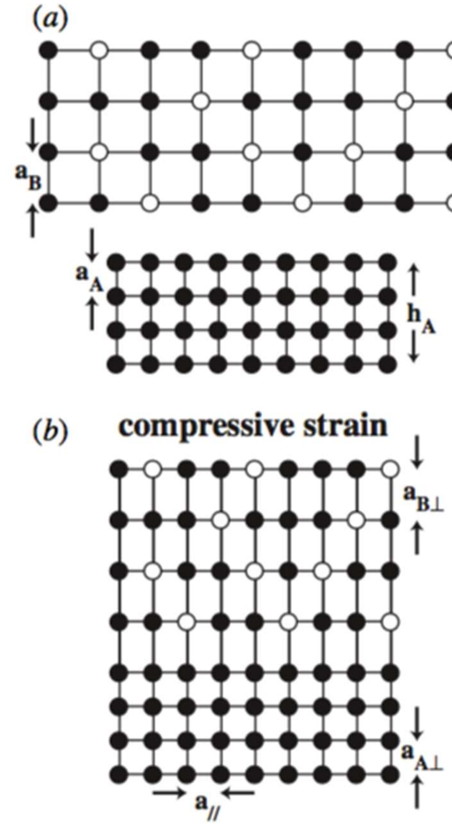


Figure 3-4 (a) A Schematic diagram of the bulk lattice constant of a thin $Si_{1-x}Ge_x$ film (layer B) to be grown on top of a bulk Si layer (layer A). (b) A schematic diagram showing the lattice distortion when $Si_{1-x}Ge_x$ film is grown on Si substrate and being compressively strained [48].

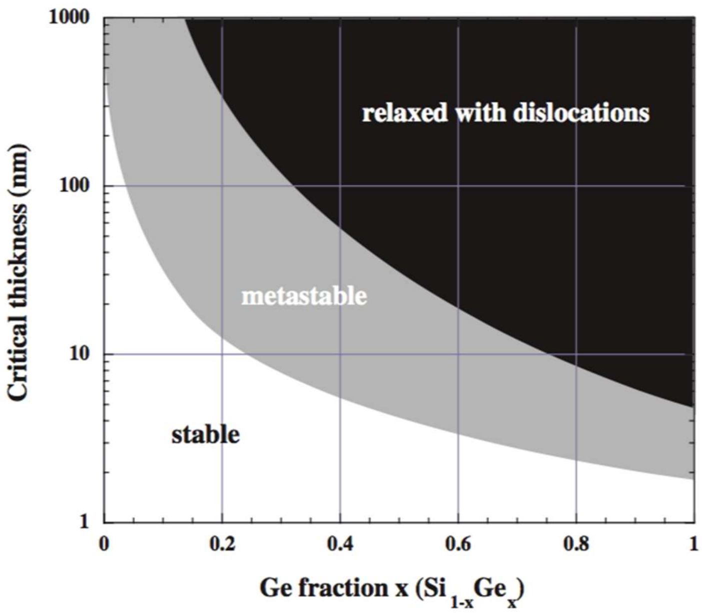


Figure 3-5 Critical thickness of $Si_{1-x}Ge_x$ in $Si_{1-x}Ge_x$/Si-substrate system plotted against Ge fraction for pseudo-morphic $Si_{1-x}Ge_x$ layers grown on bulk (100) Si substrate. A metastable curve is also included for MBE growth at 550ºC [48].

3.5 Threshold voltage of $Si_{1-x}Ge_x$ channel pFET

In this section, V_T of $Si_{1-x}Ge_x$ channel pFET is discussed as a function of process thermal budget. Figure 3-6 shows pFET V_T in linear region as a function of Ge content in $Si_{1-x}Ge_x$ channel (x = 0, 0.25, 0.45, and 0.55). TiN metal gate thickness is 5 nm. Device W/L is 10/10 μm and this is common in all figures in this chapter unless otherwise mentioned. Bias condition is $V_S = V_B = 0$ V, $V_D = -50$ mV. Spike RTA temperature is 1035ºC for all devices. pFET V_T was confirmed to monotonously reduce up to Ge content of 45%. This V_T behavior is also reported in ref. [56, 82], and can be understood from the relationship between bandgap and Ge content in $Si_{1-x}Ge_x$ where the bandgap gets smaller (and higher valence band edge energy) with higher Ge content [48, 83]. Sensitivity of V_T against Ge content estimated from Figure 3-6 is approximately 7.5 mV per Ge%, and this is quite well matched to the theoretical expectation ($\Delta E_v \sim 0.74x$ eV) discussed in Chapter 1. This result clearly shows the advantage of using $Si_{1-x}Ge_x$ channel with higher Ge content. However, no additional reduction was observed at Ge content of 55% as compared to Ge 45%, V_T for $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel devices are comparable. To understand this, V_T for $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel pFETs were compared at lower spike RTA temperature. Figure 3-7 shows linear current (I_{DLIN}) as a function of gate voltage (V_G) for $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel pFETs (conventional $Si_{1-x}Ge_x$ devices) with 1035ºC and 1000ºC spike RTA temperatures. TiN metal gate thickness is 2 nm. Bias condition is the same as the one in Figure 3-6. As seen in Figure 3-6, V_T for $Si_{0.45}Ge_{0.55}$ and $Si_{0.55}Ge_{0.45}$ channel pFETs are comparable with spike RTA 1035ºC. However, interestingly, they show V_T difference at 1000ºC by approximately 150 mV (V_T for $Si_{0.45}Ge_{0.55}$ is 150 mV lower), which is even much larger than the expectation ($\Delta E_v \sim 0.74x$ eV).

The applied compressive strain in $Si_{0.55}Ge_{0.45}$ channel is relatively smaller than $Si_{0.45}Ge_{0.55}$ channel owing to smaller lattice constant mismatch to Si substrate (approximately 1.7 % for $Si_{0.55}Ge_{0.45}$, 2.0 % for $Si_{0.45}Ge_{0.55}$), therefore it is expected that $Si_{0.55}Ge_{0.45}$ channel is more robust against thermal treatment than $Si_{0.45}Ge_{0.55}$ channel in terms of strain relaxation, as discussed in earlier in this chapter (see equation (3-3) and (3-4)). In other words, the amount of strain relaxation when going from 1000ºC to 1035ºC is larger in $Si_{0.45}Ge_{0.55}$ channel than in $Si_{0.55}Ge_{0.45}$. Due to this difference in the amount of strain relaxation, the amount of V_T shift between 1000ºC and 1035ºC spike RTA is larger in $Si_{0.45}Ge_{0.55}$ than $Si_{0.55}Ge_{0.45}$ channel pFET, as can be seen from Figure 1-17.

Other possibility is Ge out-diffusion from $Si_{1-x}Ge_x$ layer. Ge can diffuse out from $Si_{1-x}Ge_x$ layer and the effective Ge content can be reduced from the original value after high temperature processes. The amount of Ge out-diffusion can be larger with higher Ge content $Si_{1-x}Ge_x$, therefore, V_T shift by high temperature process can be larger for higher Ge content $Si_{1-x}Ge_x$ channel.

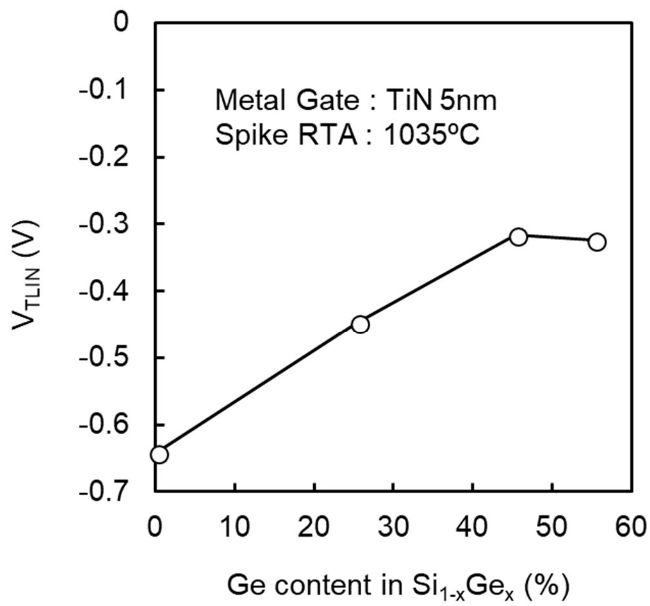


Figure 3-6 pFET linear V_T (V_{TLIN}) as a function of Ge content in $Si_{1-x}Ge_x$ channel. Spike RTA temperature is 1035ºC. V_{TLIN} is confirmed to reduce up to Ge content of 45% but no additional reduction was observed at Ge 55%.

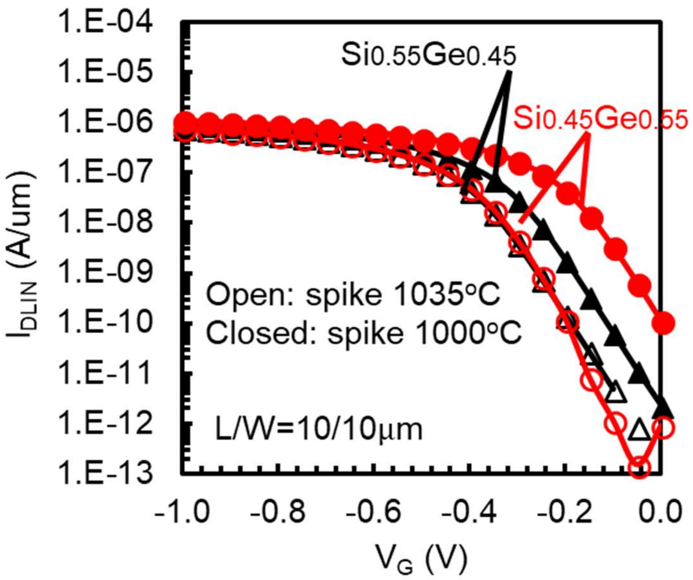


Figure 3-7 Linear current (I_{DLIN}) as a function of gate voltage (V_G) for $Si_{0.55}Ge_{0.45}$ channel and $Si_{0.45}Ge_{0.55}$ channel pFETs. Open markers are with spike RTA temperature of 1035ºC, closed markers are with 1000ºC.

Figure 3-8 shows linear threshold voltage (V_{TLIN}) of $Si_{0.45}Ge_{0.55}$ channel pFET with various activation annealing method and temperature (spike RTA at 1035ºC, 950ºC, and laser annealing at 1150ºC) as another example for the impact from process temperature on V_T of $Si_{1-x}Ge_x$ pFET. Si-cap thickness for devices in Figure 3-8 is different from Figure 3-7. By applying laser annealing for dopant activation [66, 68], it was found that pFET V_{TLIN} could be further reduced compared to 950ºC spike RTA [82]. Despite of high peak temperature of 1150ºC, dwell time of laser annealing is several orders of magnitude shorter than spike RTA [84], therefore, total thermal budget is significantly smaller for laser annealing.

Along with V_T itself, its Within-Wafer (WiW) uniformity (3σ for V_{TLIN}) is also considerably improved by reducing the process temperature as shown in Figure 3-9, where WiW uniformity of V_T for various activation anneal temperatures and methods (spike-RTA, laser anneal) are plotted. As a general trend, V_T WiW uniformity is worse with $Si_{1-x}Ge_x$ channel than Si channel. This is probably due to WiW variation of 1) $Si_{1-x}Ge_x$ layer thickness [56], 2) Si-cap thickness [55], and 3) Ge content [56, 53]. Also, when comparing $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel, $Si_{0.55}Ge_{0.45}$ behaves a little better than $Si_{0.45}Ge_{0.55}$, although the delta between them is not as significant as the one between $Si_{1-x}Ge_x$ and Si channel.

Lower thermal budget annealing generally provides better WiW uniformity for $Si_{1-x}Ge_x$ channel pFET (Ge content 45% and 55%) likely due to suppression of strain relaxation and Ge out-diffusion, which would have a certain WiW variation. For example, in case of $Si_{0.55}Ge_{0.45}$ channel, V_T WiW uniformity goes from 9 mV with spike RTA 1035ºC to 5 mV with 950ºC, and in case of $Si_{0.45}Ge_{0.55}$ channel, WiW uniformity goes from 11 mV with spike RTA 1035ºC to 5 mV with laser annealing. This is another benefit of reducing the thermal budget of process flow in $Si_{1-x}Ge_x$ channel pFET. However, even with this improvement by reducing process temperature, V_T WiW uniformity for Si channel is still much better (less than 2 mV) than $Si_{1-x}Ge_x$ channel.

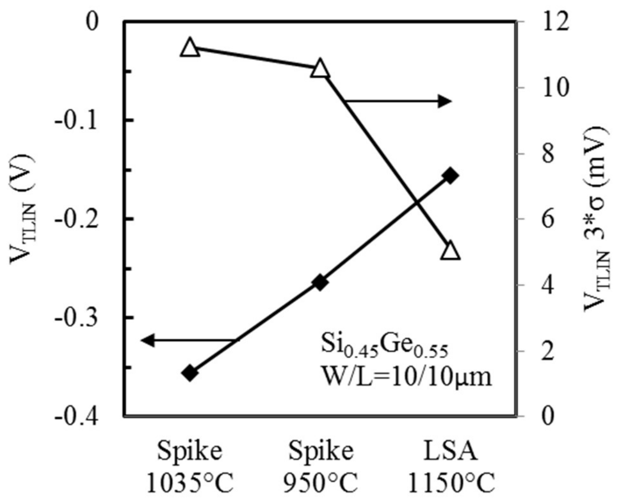


Figure 3-8 Linear threshold voltage (V_{TLIN}) of $Si_{0.45}Ge_{0.55}$ channel pFET with various activation annealing temperature (spike RTA 1035ºC, 950ºC, and laser annealing 1150ºC). Within-Wafer (WiW) variation of V_T is also plotted on right axis.

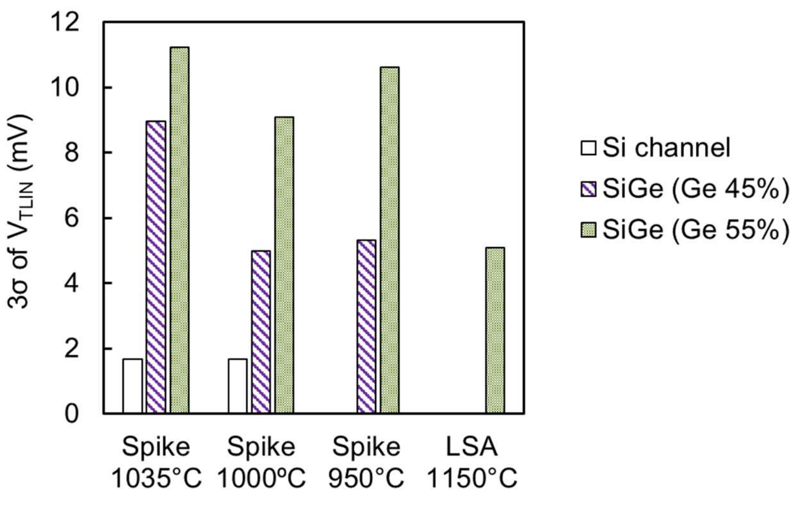


Figure 3-9 WiW uniformity (3σ) of V_T for various activation anneal temperatures and methods (spike-RTA, laser anneal). Lower thermal budget annealing generally provides better WiW uniformity for $Si_{1-x}Ge_x$ channel pFET, however, WiW uniformity for Si channel is still much better.

3.6 High field hole mobility of $Si_{1-x}Ge_x$ channel pFET

Figure 3-10 shows high field inversion layer hole mobility as a function of Ge content in $Si_{1-x}Ge_x$ channel with various activation anneal method / temperatures. Inversion layer mobility was extracted by split C-V method on W/L = 10/10 μm devices. The device structure is conventional $Si_{1-x}Ge_x$ channel pFET. The hole mobility of $Si_{1-x}Ge_x$ channel pFET is much higher than Si channel pFET. For spike RTA 1035°C, hole mobility shows significant increase of about 2.5x from Si channel (39 cm^2 /V-s) to $Si_{0.75}Ge_{0.25}$ channel (106 cm^2/V-s). But there is no further improvement seen with higher Ge content and hole mobility rather decreases with higher Ge content. This is also the case for spike RTA 1000°C, but degradation at higher Ge content is not as significant as spike RTA 1035°C case. On the other hand, if we further lower the spike RTA temperature down to 950°C, significant improvement of hole mobility is seen in $Si_{0.45}Ge_{0.55}$ channel (130 cm^2/V-s) and one can clearly see that hole mobility goes up higher with higher Ge content in the channel with this spike RTA temperature, which is aligned qualitatively to the theoretical calculation in ref [85]. When laser spike annealing is used instead of spike RTA, hole mobility of $Si_{0.45}Ge_{0.55}$ channel shows further improvement and reaches 160 cm^2/V-s. It should be noted that hole mobility of $Si_{0.45}Ge_{0.55}$ is sensitive to process temperature while $Si_{0.55}Ge_{0.45}$ and below shows much less sensitivity.

This behavior can be also explained by strain relaxation in $Si_{1-x}Ge_x$ channel layer and Ge out-diffusion from $Si_{1-x}Ge_x$ layer due to high temperature thermal treatment as discussed in previous section. The strain relaxation by thermal stress would be plastic relaxation and this can generate crystal defects like misfit dislocations and hence degrades carrier mobility [48, 82]. Ge out-diffusion from $Si_{1-x}Ge_x$ channel layer can effectively lower the peak Ge fraction in the channel, which also leads to hole mobility degradation

[82]. Diffused Ge may move towards the interface between SiO_2-IL and Si-capping layer and create unwanted GeO_x which can create an interfacial state inside Si band gap, which is another mechanism for hole mobility degradation [86, 87, 88].

Figure 3-11 shows interfacial trap density (N_{it}) of $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel devices with 950ºC spike RTA, measured by charge pumping method [89, 90]. Device size is W/L = 10/10 μm, gate voltage was pulsed with frequency of 2 MHz. The N_{it} of $Si_{0.45}Ge_{0.55}$ channel is 3.8×10^{11} cm^{-2}, higher than that of $Si_{0.55}Ge_{0.45}$ channel (2×10^{11} cm^{-2}). These values are roughly one order of magnitude higher than conventional thermally grown SiO_2 and Si substrate interface [89]. This suggests that Ge diffuses through Si-capping layer and reaches the interface between SiO_2-IL and Si-capping layer and creates GeO_x, which degrades interfacial trap density (N_{it}) even with 950ºC spike RTA.

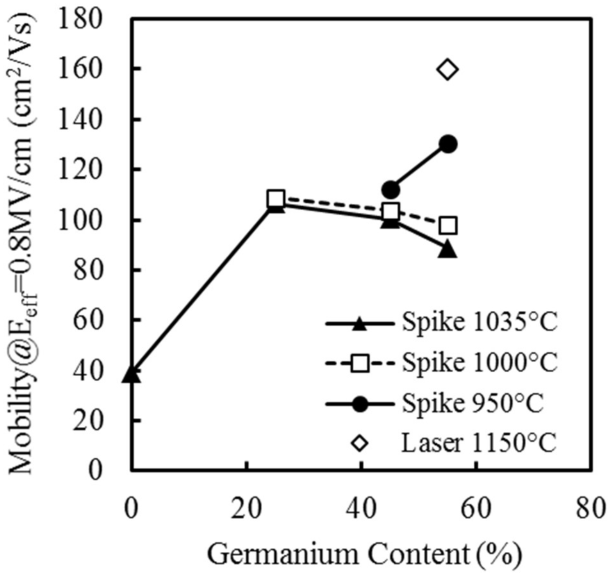


Figure 3-10 High field inversion layer hole mobility (at effective field of 0.8 MV/cm) as a function of Ge content in $Si_{1-x}Ge_x$ channel with various activation annealing method / temperature.

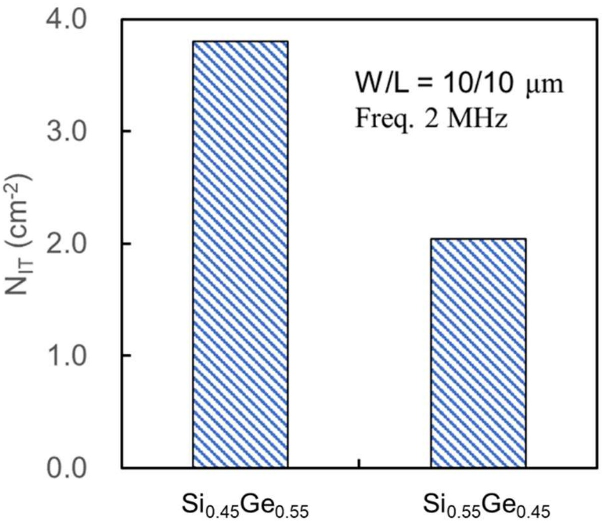


Figure 3-11 Interface trap density (N_{it}) for $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel pFET with spike RTA temperature 950ºC. N_{it} was measured by charge pumping method, frequency of gate bias is 2 MHz.

Figure 3-12 shows T_{inv} for $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel pFET as a function of spike RTA temperature. We observed monotonous T_{inv} reduction as spike RTA temperature goes higher. Si capping thickness and gate stack structure (TiN 2 nm / HfO_2 1.8 nm) are common for all devices. This T_{inv} modulation by spike RTA temperature is considered due to modulation of strength of interfacial layer scavenging. In metal gate/high-k system, phenomena called oxygen scavenging have been reported **[36, 91]**, where SiO_2-IL is significantly reduced during high temperature thermal treatment and hence extremely thin T_{inv} has been achieved. This SiO_2-IL thinning can be realized by doping scavenging element in the metal gate electrode **[36]** or by depositing metal gate alloyed with scavenging element or by adding high-k capping layer on top of HfO_2 **[91]**. The possible mechanism proposed in ref. **[36]** is cascading reactions of oxygen transfer from SiO_2-IL to metal gate. 1) Scavenging metal elements is oxidized by oxygen in HfO_2 film, 2) leaving oxygen vacancies in HfO_2, then 3) the oxygen vacancies are passivated with oxygen atoms decomposed from SiO_2-IL.

The strength of oxygen scavenging by a certain metal element, M, can be measured by magnitude of Gibbs free energy change at a certain temperature, for example 1000 K, (ΔG^o_{1000}) of the following reaction between M and SiO_2

$$\mathrm{Si} + \frac{2}{y}\mathrm{M}_x\mathrm{O}_y \rightarrow \frac{2x}{y}\mathrm{M} + \mathrm{SiO}_2 \qquad (3\text{-}5)$$

If ΔG^o_{1000} for the reaction in (3-5) is negative, this reaction is promoted (substrate is oxidized with oxygen in M_xO_y, SiO_2-IL regrowth) and if positive, the reaction progresses in opposite direction (metal M reduces SiO_2-IL, SiO_2-IL scavenging). ΔG^o_{1000} values for various metal elements are summarized in Table 3-2 **[92]**. Elements like Ti, Zr, Al has positive ΔG^o_{1000}, therefore can be candidates for SiO_2-IL scavenging elements when present in metal gate electrode.

In this study, TiN is used as metal gate electrode and no doping has been done in TiN film, also no alloying based on TiN has been done. Although Ti/N composition in TiN metal gate has not been investigated in this study, if there are excess Ti atoms (Ti-rch composition) or any dangling bonds for Ti atoms in TiN, they can cause oxygen scavenging and form TiO_2, Ti_2O_3, or TiO during high temperature annealing. As this kinetics is thermally activated process, annealing at higher temperature is supposed to promote more oxygen scavenging and hence leaves thinner SiO_2-IL, which can explain the trend in Figure 3-12. This SiO_2-IL thinning could be one reason for hole-mobility degradation with higher temperature annealing (Figure 3-10) due to enhanced remote Coulomb scattering by fixed charge and/or trapped charge in HfO_2 film **[36]**. However, strain relaxation and Ge diffusion should be still the dominant mechanisms for hole-mobility degradation. This can be understood from the fact that the amount of T_{inv} thinning is very similar among $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel devices, though the mobility shows different sensitivity (hole mobility for $Si_{0.45}Ge_{0.55}$ channel is much more sensitive to spike RTA temperature).

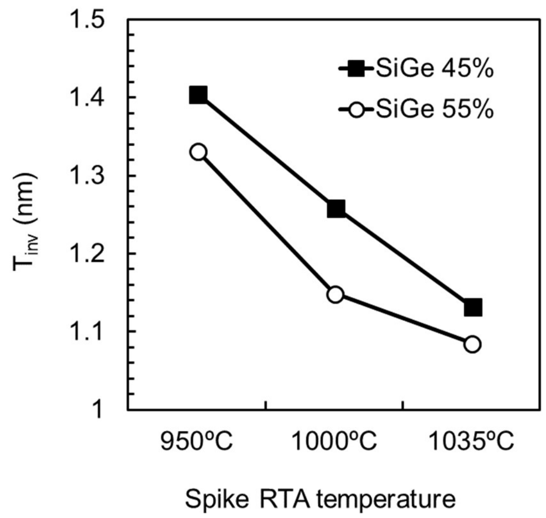


Figure 3-12 T_{inv} for $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel pFETs as a function of spike RTA temperature.

Table 3-2 Gibbs free energy change for the reaction (3-5) [92].

Metal oxide (M_xO_y)	ΔG^o_{1000} per M_xO_y for $Si + \frac{2}{y} M_xO_y \rightarrow \frac{2x}{y} M + SiO_2$ (kcal/mol)
TiO	+17.849
Ti_2O_3	+35.432
TiO_2	+7.527
Ta_2O_5	-52.533
WO_2	-77.126
Al_2O_3	+63.399
La_2O_3	-98.470
ZrO_2	+42.326

3.7 Off-state leakage for $Si_{1-x}Ge_x$ channel pFET

As seen in Section 3.5 and 3.6, lower thermal budget, especially laser annealing is a preferable activation annealing option to implement high Ge content $Si_{1-x}Ge_x$ channel pFET (Ge > 50%) from high field hole mobility and V_T point of view, as the low process temperature can help to keep decent crystallinity and bi-axial compressive strain and abrupt Ge profile in $Si_{1-x}Ge_x$ channel layer. In this section, we discuss the impact on off-state leakage from activation annealing temperature/method.

In Figure 3-13, an off-state current at body terminal (I_{OFF_B}) for $Si_{0.45}Ge_{0.55}$ channel pFETs with various activation anneal temperature/method are plotted as an indicator for junction leakage current. Bias condition is $V_S = V_B = 0$ V, $V_G = V_D = -1$ V. Spike 1035°C gives the lowest I_{OFF_B} and lowering the thermal budget of activation annealing worsens I_{OFF_B}. Laser annealing showed severe degradation of several orders of magnitude. This degradation is supposed to be due to increased residual crystal defects, which are created during halo, extension and source/drain implantation and not fully cured during subsequent activation annealing [93, 94]. The residual defects create energy levels within Si or $Si_{1-x}Ge_x$ bandgap and can act as recombination center if they are located within depletion layer at n/p junction. To make things worse, now the depletion layer is also formed within $Si_{1-x}Ge_x$ layer [48], junction leakage becomes even worse as recombination rate is higher due to narrower band gap for $Si_{1-x}Ge_x$ compared to Si. Therefore, laser annealing which has extremely small thermal budget results in the worst junction leakage current.

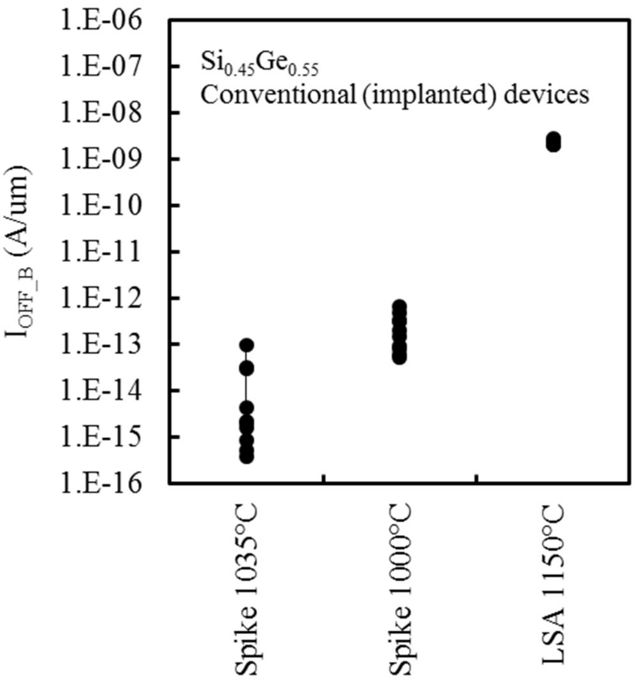


Figure 3-13 Off-state current at bulk terminal (I_{OFF_B}) as an indicator for junction leakage is plotted for various activation annealing method and temperature.

As discussed above, conventional $Si_{1-x}Ge_x$ channel devices, which have implanted extension and source/drain need a certain amount of thermal budget (preferably spike RTA 1000°C or above) to annihilate crystal defects created during ion implantation and to reduce junction leakage current. However, this compromises the high field hole mobility and threshold voltage for $Si_{1-x}Ge_x$ channel pFET, especially $Si_{0.45}Ge_{0.55}$ channel. On the contrary, IF-$Si_{1-x}Ge_x$ channel devices do not need such high temperature activation anneal as they have epitaxially grown rSiGe:B extension and/or eSiGe:B source/drain which already have highly activated dopants (boron ~1×10^{20} cm^{-3}). Therefore, IF-$Si_{1-x}Ge_x$ channel pFET devices have possibility of realizing high hole mobility and low V_T without compromising junction leakage current, although they still do need a certain amount of thermal treatment to diffuse out boron in rSiGe:B extension and to create an overlap between channel and extension as discussed earlier.

In case of nFET (assuming Si channel with conventional implanted S/D), on the other hand, laser-only annealing (no spike RTA in combination with laser annealing) can be applied by optimizing the laser power (high power laser annealing) and showed comparable junction leakage current, as reported in ref. [95]. However, it needs careful assessment to see if the optimized laser annealing condition for nFET can be applied to pFET $Si_{1-x}Ge_x$ channel, as ref. [96] reported on strain relaxation of $Si_{1-x}Ge_x$ layer (x > 45%) even with laser annealing when laser annealing temperature is high.

Even though there should be no implant-related defects in IF-$Si_{1-x}Ge_x$ channel pFETs, there is still a concern on junction leakage current in $Si_{1-x}Ge_x$ channel with higher Ge content, due to smaller band gap E_g than Si as shown in equation (1-3) and (1-4).

However, one can optimize junction leakage current by reducing the thickness of $Si_{1-x}Ge_x$ layer, as shown in Figure 3-14 (conventional $Si_{1-x}Ge_x$ channel device, spike RTA temperature 1000°C). At a given $Si_{1-x}Ge_x$ layer thickness, as Ge content in $Si_{1-x}Ge_x$ channel gets higher, junction leakage current inevitably rises owing to its smaller band gap (equation (1-3) and (1-4), but it can be effectively offset by reducing the $Si_{1-x}Ge_x$ layer thickness. I_{off_B} shows significant reduction by several orders of magnitude by reducing $Si_{1-x}Ge_x$ layer thickness by around 0.5x or so. Although the result shown in Figure 3-14 is from conventional (implanted) $Si_{1-x}Ge_x$ channel devices, reducing channel thickness should work also for IF-$Si_{1-x}Ge_x$ channel devices. In later part of this thesis, all IF-$Si_{1-x}Ge_x$ channel devices have 4-nm-thick $Si_{1-x}Ge_x$ layer.

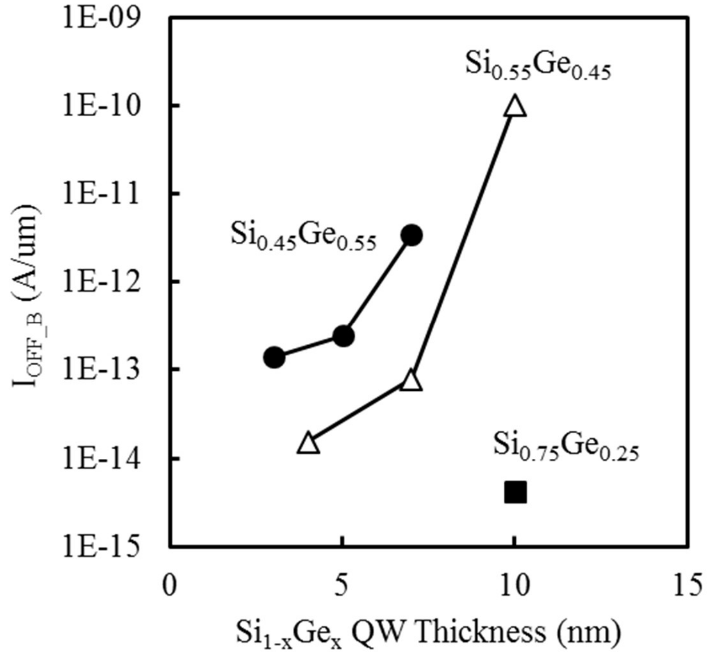


Figure 3-14 I_{OFF_B} as a function of $Si_{1-x}Ge_x$ channel layer thickness for various Ge content in the channel.

3.8 Summary

In this chapter, the impact from process temperature on $Si_{1-x}Ge_x$ channel pFETs has been discussed from various aspects, namely V_T, hole mobility, and junction leakage current and scorecard is summarized in Table 3-3. From $Si_{1-x}Ge_x$ layer integrity point of view, lower thermal budget after $Si_{1-x}Ge_x$ channel growth is preferable to keep biaxial compressive strain in the $Si_{1-x}Ge_x$ channel layer and steep Ge concentration gradient. This helps to keep high hole mobility and low V_T. On the other hand, lower thermal budget at activation anneal can cause junction leakage current issue due to insufficient curing of the defects in case of conventional $Si_{1-x}Ge_x$ channel pFET where extension and source / drain are formed by ion implantation. IF-$Si_{1-x}Ge_x$ channel can overcome this dilemma. This device doesn't need to receive high temperature annealing for dopant activation as the extension and source/drain are formed by in-situ boron doped $Si_{1-x}Ge_x$ epitaxy. Therefore, process thermal budget can be reduced for IF-$Si_{1-x}Ge_x$ channel devices. However, laser annealing is not applicable to IF-$Si_{1-x}Ge_x$ channel devices as this cannot create sufficient gate to extension overlap. Therefore, IF-$Si_{1-x}Ge_x$ with 950ºC spike-RTA is the best option.

In next chapters (Chapter 4 and Chapter 5), device characteristics of $Si_{1-x}Ge_x$ channel pFET will be discussed in details with more focus on IF-$Si_{1-x}Ge_x$ devices.

Table 3-3 Scorecard for process thermal budget selection.

	Spike-RTA				Laser-only annealing	
	1035ºC		950ºC			
	Conv.	IF	Conv.	IF	Conv.	IF
SiGe channel pFET V_t	High		Low		Lower than spike 950ºC	
SiGe channel mobility	Low		High		Higher than spike 950ºC	
SiGe channel junction leakage	Low	Low	High	Low	Very high	Low
nFET (Si-ch) junction leakage	No data from this work but industry standard.				Comparable to Spike with optimized power [95]. Compatibility to SiGe channel is concerned.	
IF-SiGe channel overlap	NA	Good	NA	Good	NA	Poor

Chapter 4

Strain effect in $Si_{1-x}Ge_x$ channel pFET

4.1 Background

In this chapter, we will discuss device characteristics of IF-$Si_{1-x}Ge_x$ channel pFET mainly from the view point of the stress applied to the $Si_{1-x}Ge_x$ channel. As we have seen in Chapter 3, it is extremely critical to maintain the strain applied to $Si_{1-x}Ge_x$ channel from Si substrate to keep high hole mobility and low V_T. The hole mobility of $Si_{1-x}Ge_x$ channel device is modulated by the applied stress via piezoelectric effect [97]. The hole mobility change ($\Delta\mu/\mu$) by the applied stress can be formulated using piezo-resistance coefficient for longitudinal and transverse axis to the transistor current flow (π_L, π_T) and the applied stress for each axis (ΔS_L, ΔS_T) as below [98, 99]

$$\Delta\mu/\mu \approx \pi_L \Delta S_L + \pi_T \Delta S_T \tag{4-1}$$

ΔS_L and ΔS_T are positive in case of tensile stress, negative in case of compressive stress.

Experimentally, the piezo-resistance coefficient can be determined by wafer bending method as show in Figure 4-1 [99]. Uniaxial stress is applied mechanically by bending the wafer and the mobility is extracted by split C-V method under the applied stress. For $Si_{1-x}Ge_x$ channel, J. Pan *et al*. reported piezo-resistance coefficient values for $Si_{0.7}Ge_{0.3}$ channel by this method [99]. Our group (imec, G. Eneman *et al*.) also reported the values for $Si_{0.75}Ge_{0.25}$ and $Si_{0.55}Ge_{0.45}$ channel using same technique [69]. The reported values are summarized in Table 4-1.

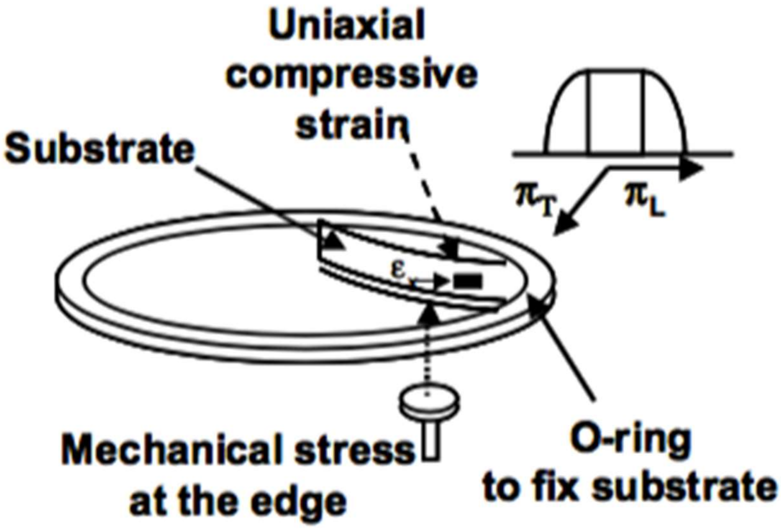


Figure 4-1 Schematic diagram of applying externally compressive uniaxial stress. π_L, π_T are obtained on long channel devices with this uniaxial stress [99].

Table 4-1 Piezo-resistance coefficient (π_L, π_T)for $Si_{1-x}Ge_x$ channel pFET extracted by wafer bending method (Unit: $10^{-11}Pa^{-1}$).

	$Si_{0.7}Ge_{0.3}$	$Si_{0.75}Ge_{0.25}$	$Si_{0.55}Ge_{0.45}$
Longitudinal (π_L)	-125	-86	-101
Transverse (π_T)	+52	+61	+75
Reference	[99]	[69]	[69]

Regardless of Ge content in $Si_{1-x}Ge_x$ channel, the hole mobility of $Si_{1-x}Ge_x$ channel increases when compressive and tensile stress is applied in longitudinal and transverse direction, respectively. Epitaxially grown $Si_{1-x}Ge_x$ channel on Si substrate is under bi-axial compressive stress. While longitudinal stress and transverse stress are counteracting each other (sign of piezo-resistance coefficient is opposite. Longitudinal compressive stress is beneficial to the hole mobility, transverse compressive stress adversely affecting the hole mobility), the net mobility gain is still positive due to absolute value for longitudinal piezo-resistance coefficient being larger than transverse one.

In Section 4.2, interaction between $Si_{1-x}Ge_x$ channel and eSiGe:B stressor will be discussed. In case of Si channel, the mobility boost with eSiGe stressor is well known. We will discuss how the hole mobility of $Si_{1-x}Ge_x$ channel is modulated by eSiGe stressor in comparison to Si channel case. In Section 4.3, as one of the local layout effect (LLE), transistor channel width dependence will be discussed. We observed significant drive current enhancement at narrow width devices, whose strength is dependent on Ge content in the $Si_{1-x}Ge_x$ channel. Elastic strain relaxation at active edge plays a key role in the channel width dependence. In Section 4.4, DC performance of IF-$Si_{1-x}Ge_x$ channel pFET is benchmarked with other literature to conclude this chapter.

4.2 Interaction between $Si_{1-x}Ge_x$ channel and eSiGe:B stressor

Figure 4-2 shows hole mobility as a function of gate length (L_G) for (a) Si channel and (b) $Si_{0.55}Ge_{0.45}$ channel pFETs, with and without eSiGe:B stressor [100]. Short channel mobility down to gate length of 30 nm was extracted and calculated by the second order Y function technique which is described in reference [101]. Both Si channel and $Si_{0.55}Ge_{0.45}$ channel pFETs are conventional structures (implanted extension and source/drain). As a general trend, hole mobility degrades as the gate length gets shorter. This is due to enhanced Coulombic scattering by ionized dopant atoms by halo implant [102]. Although $Si_{0.55}Ge_{0.45}$ channel pFET has almost 2x higher hole mobility compared to Si channel pFET at long channel (L_G = 1 μm), hole mobility for $Si_{0.55}Ge_{0.45}$ degrades more severely at shorter channel length. The resultant hole mobility at gate length of 30 nm is comparable between Si channel and $Si_{0.55}Ge_{0.45}$ channel pFETs. The reason for this phenomenon is not clearly understood but one possibility is that mobility degradation by Coulombic scattering by halo impurity atoms is dominant factor at short channel and the intrinsic mobility improvement seen at long channel may have been smeared.

Looking at the mobility improvement with eSiGe at gate length of 30 nm, they are +60% and +90% improvements for Si channel and $Si_{0.55}Ge_{0.45}$ channel pFETs, respectively. One can say that eSiGe:B stressor gives at least comparable hole mobility boost for $Si_{0.55}Ge_{0.45}$ channel and Si channel pFET. The mobility improvement with eSiGe:B stressor, $\frac{\mu \text{ (with eSiGe)}}{\mu \text{ (without eSiGe)}}$, is plotted as a function of gate length and compared between Si channel and $Si_{0.55}Ge_{0.45}$ channel in Figure 4-3. Hole mobility improvement is larger at shorter gate length (Hole mobility improvement can be hardly seen at gate length of 1 μm), as the stress applied from eSiGe:B at a certain location in the channel reduces over the distance from eSiGe stressor [98].

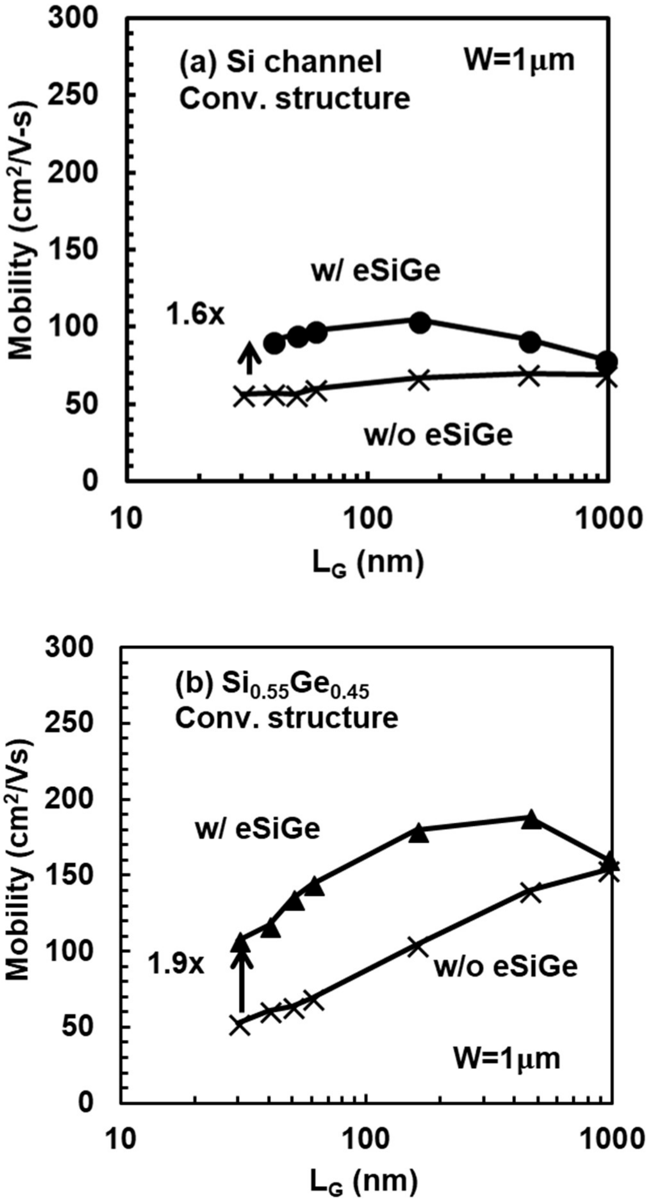


Figure 4-2 Hole mobility as a function of gate length (L_G) for (a) Si channel and (b) $Si_{0.55}Ge_{0.45}$ channel pFETs. Both are conventional structures (implanted extension and source/drain). eSiGe gives larger or at least comparable hole mobility boost for $Si_{0.55}Ge_{0.45}$ channel pFET compared to Si channel pFET.

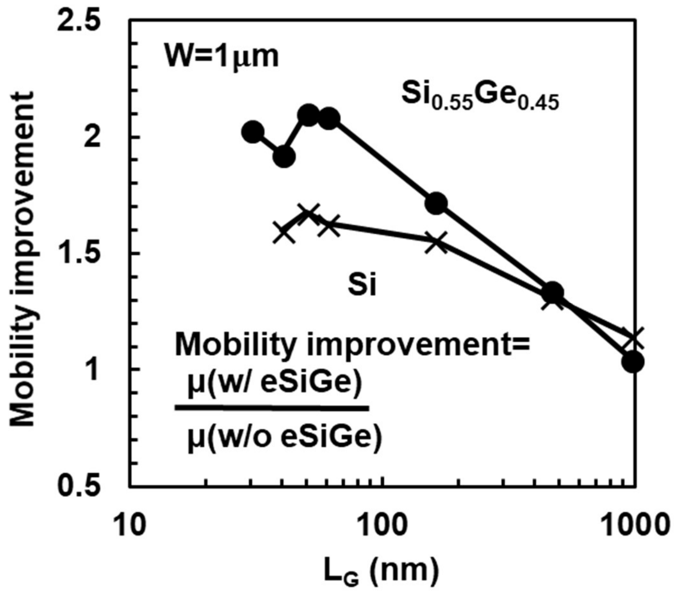
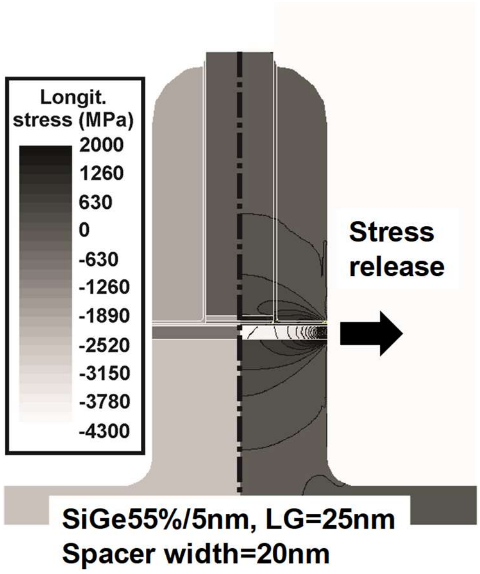


Figure 4-3 Hole mobility improvement with eSiGe:B stressor (fraction of mobility with and without eSiGe) as a function of gate length. $Si_{0.55}Ge_{0.45}$ and Si channel pFETs are compared.

Figure 4-4 (a) is a schematic diagram showing process steps in eSiGe:B stressor formation module (pre cavity RIE, post cavity RIE, and post eSiGe:B growth). The longitudinal stress in $Si_{1-x}Ge_x$ channel is supposed to be elastically relaxed and a certain amount of strain is expected to be lost if not all when cavity RIE is done, because the edge of the $Si_{1-x}Ge_x$ channel becomes open ended and elastic strain relaxation should happen. After eSiGe:B growth, the stress is added back to the $Si_{1-x}Ge_x$ channel.

TCAD simulation was done to track how the stress in $Si_{1-x}Ge_x$ channel evolves at each of these three steps in Figure 4-4 (a) (pre cavity RIE, post cavity RIE, and post eSiGe growth). In this simulation, 5 nm $Si_{0.45}Ge_{0.55}$ channel was used, gate length was varied from 1 μm down to 25 nm. Ge content in eSiGe:B stressor is 25%, cavity depth is 60 nm, and offset spacer width is 20 nm. Figure 4-4 (b) is simulation result showing the stress distribution in the transistor right after the Si cavity RIE. It is clearly seen that the stress at the edge of the $Si_{0.45}Ge_{0.55}$ channel (at the edge of the cavity) is released and dropped to nearly zero, while the stress at the center of the channel is still kept relatively high. This means that the net change of longitudinal stress by cavity RIE and eSiGe growth (ΔS_{xx}) in case of $Si_{1-x}Ge_x$ channel is lower than that of Si channel, especially for $Si_{1-x}Ge_x$ channel with higher Ge content where larger extent of strain relaxation happens. This is illustrated in Figure 4-5, where ΔS_{xx} is plotted as a function of gate length for $Si_{1-x}Ge_x$ channel with various Ge content. In case of Si channel, absolute value for ΔS_{xx} at gate length of 25 nm is approximately 2.0 GPa, however, absolute value of ΔS_{xx} for $Si_{0.45}Ge_{0.55}$ channel is only 1.3 GPa. This is apparently not consistent to the result obtained in Figure 4-2 and Figure 4-3 where hole mobility improvement is at least comparable between $Si_{0.55}Ge_{0.45}$ and Si channel.

(a)



(b)

Figure 4-4 (a) Schematic diagram showing process steps in eSiGe formation module. The longitudinal stress in $Si_{1-x}Ge_x$ channel is supposed to be elastically relaxed when cavity RIE is done. eSiGe adds back the stress. (b) TCAD simulation result showing the elastic stress relaxation in $Si_{1-x}Ge_x$ channel at the edge of cavity.

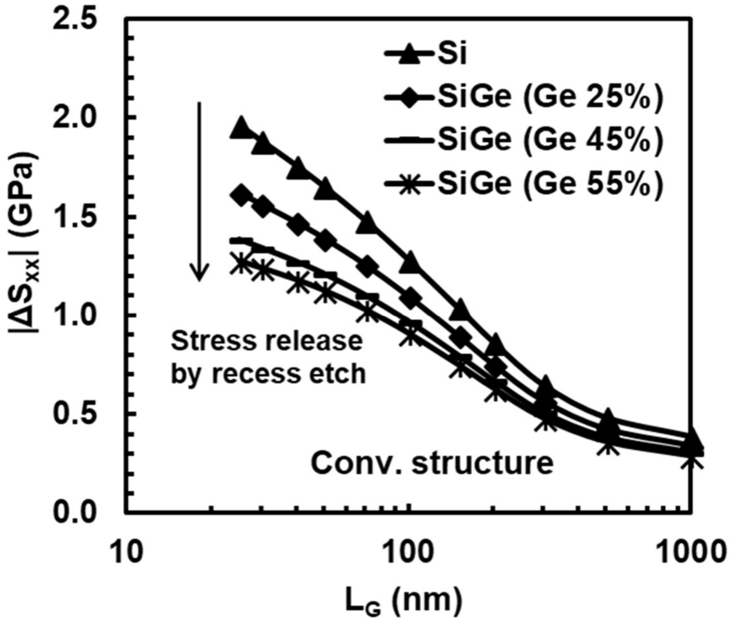


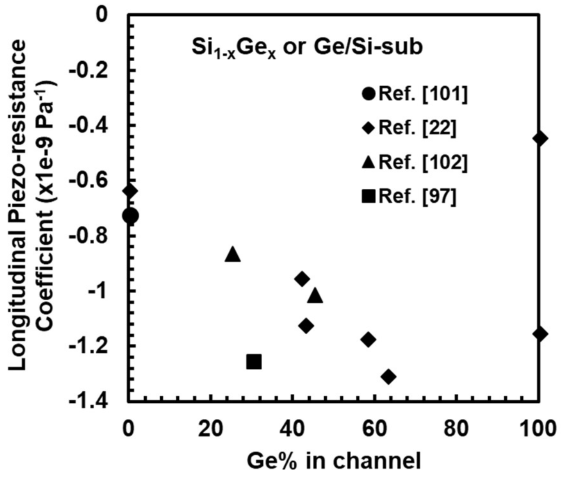
Figure 4-5 Simulated net changes in longitudinal stress (ΔS_{xx}) with eSiGe (stress difference between pre cavity RIE and post eSiGe growth) as a function of gate length for Si and $Si_{1-x}Ge_x$ channel with various Ge content. The net stress gain (ΔS_{xx}) monotonously decreases as Ge content in $Si_{1-x}Ge_x$ channel increases. This is because $Si_{1-x}Ge_x$ channel with higher Ge content loses higher amount of strain during cavity RIE.

One possible explanation for this observation is that $Si_{1-x}Ge_x$ channel may have a larger longitudinal piezo-resistance coefficient than that of Si channel, which can compensate the relatively smaller ΔS_{xx}. A summary for longitudinal piezo-resistance coefficient for unstrained Si, Ge, and strained $Si_{1-x}Ge_x$ (epitaxially grown on Si substrate and hence $Si_{1-x}Ge_x$ is under bi-axial compressive strain) taken from various literatures is shown in Figure 4-6 [51, 99, 103, 104, 105]. In these literatures, longitudinal piezo-resistance coefficient was measured by applying additional uniaxial stress to bi-axially strained Ge and $Si_{1-x}Ge_x$ channel. In Figure 4-6 (a), longitudinal piezo-resistance coefficient is plotted as a function of Ge content in the channel. $Si_{1-x}Ge_x$ channel has larger longitudinal piezo-resistance coefficient than Si channel, but Ge channel doesn't have higher longitudinal coefficient than $Si_{1-x}Ge_x$. Therefore, one can conclude that longitudinal piezo-resistance coefficient is not modulated by Ge contents in the channel. Interestingly, longitudinal piezo-resistance coefficient is apparently rather modulated by initial biaxial compressive strain applied to the channel as shown in Figure 4-6 (b). The amount of biaxial strain is found to strongly correlate with longitudinal piezo-resistance coefficient regardless of Ge content in the channel, as was also pointed out in reference [51]. Further investigation is required to identify the root cause of this phenomenon.

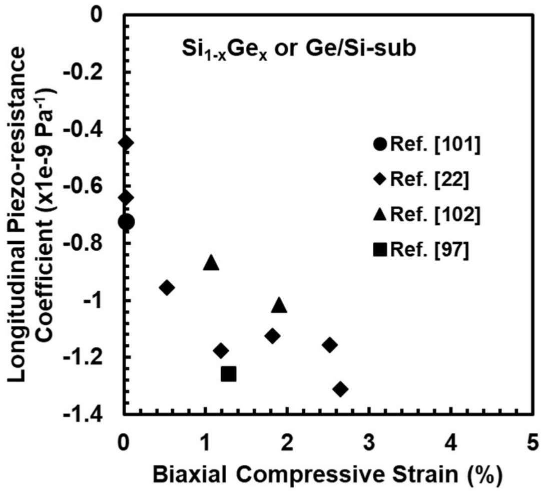
The simulated ΔS_{xx} is approximately 30 – 40 % smaller for $Si_{0.55}Ge_{0.45}$ channel compared to Si channel from Figure 4-5, while longitudinal piezo-resistance coefficient is at least 40 % larger for $Si_{0.55}Ge_{0.45}$ channel than Si channel from our earlier work [69] and also from the literature values shown in Figure 4-6 (b). As a net mobility gain is determined by product of piezo-resistance coefficient and net stress change as shown in equation (4-1), it can explain comparable hole mobility gain with eSiGe stressor between Si and $Si_{0.55}Ge_{0.45}$ channel pFETs.

In case of IF-$Si_{1-x}Ge_x$ channel pFET, we didn't observe additional hole-mobility boost from eSiGe stressor as will be discussed in Chapter 5 and hole mobility for devices with and without eSiGe is comparable. This is because rSiGe:B extension also serves as a stressor and cavity RIE leads to strain relaxation of this stressor as well and net additive stress can be comparable between devices with and without eSiGe stressor at scaled gate length [106]. However, regardless of the presence of eSiGe, IF-$Si_{1-x}Ge_x$ has significant advantage in terms of short channel hole mobility over conventional $Si_{1-x}Ge_x$ channel as shown in Figure 4-7, where hole mobility for conventional $Si_{0.55}Ge_{0.45}$ channel and IF-

$Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel devices are compared. IF-$Si_{1-x}Ge_x$ channel devices showed much higher short channel mobility (L_G less than ~200 nm) than conventional $Si_{1-x}Ge_x$ channel. This is likely because of the absence of Coulomb scattering by ionized impurity atoms from halo (no halo is implanted in IF-$Si_{1-x}Ge_x$ devices). Although IF-$Si_{0.45}Ge_{0.55}$ channel device shows higher hole mobility than that of $Si_{0.55}Ge_{0.45}$ channel at gate length of 1 μm, they are comparable below L_G ~200 nm. The reason for this behavior is not yet well understood.



(a)



(b)

Figure 4-6 Longitudinal piezo-resistance coefficient for Si, Ge and $Si_{1-x}Ge_x$ channel with various Ge content (epitaxially grown on Si substrate) taken from literatures. (a) Plotted as a function of Ge content in the channel. (b) as a function of biaxial compressive strain in the channel. It appears that longitudinal piezo-resistance coefficient is modulated by initial biaxial compressive strain in the channel, not by Ge content in the channel.

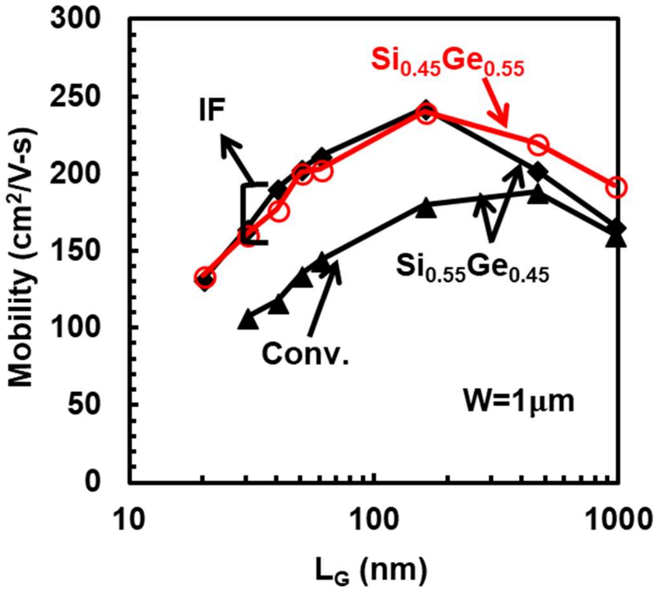


Figure 4-7 Hole mobility as a function of gate length for conventional $Si_{0.55}Ge_{0.45}$ channel and IF-$Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel. All devices have eSiGe:B stressor.

4.3 Channel width dependence for IF-$Si_{1-x}Ge_x$ pFETs due to elastic strain relaxation

Device parameters are becoming more dependent on the layout or local environment of the transistors as new materials (such as eSiGe stressor [72, 98], SiN stress liner [107]) and new integration schemes (such as RMG [8, 9]) are introduced in advanced CMOS technologies [108]. Local layout effect (LLE) includes layout dependency such as n/pFET boundary proximity effect [108, 109], dependence on length of diffusion area (LOD) [110], contacted poly pitch (CPP) [102], channel width [69, 100, 104], and many others. It is critical to understand the physical mechanism of these layout effects and include them in the device models as systematic variability to better predict the actual circuit performance [109]. In this section, we discuss channel width dependence of the drive current for IF-$Si_{1-x}Ge_x$ channel devices, as one of the LLE.

Figure 4-8 shows normalized linear (I_{DLIN}) and saturation (I_{DSAT}) drive current for IF-$Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel pFETs (with eSiGe stressor) at constant gate overdrive ($V_G - V_{TLIN}$ = -0.7 V) as a function of transistor effective channel width (W_{eff}) ranging from 10 μm down to 0.1 μm. Drive current values are normalized with regard to the values at channel width of 10 μm. Significant drive current increase was observed in IF-$Si_{0.45}Ge_{0.55}$ channel pFET, 1.9x and 1.7x increase in linear and saturation current respectively, from channel width of 10 μm down to 0.1 μm), while IF-$Si_{0.55}Ge_{0.45}$ channel pFET showed relatively moderate dependence (1.3x and 1.2x increase from 10 μm to 0.1 μm in linear and saturation current, respectively) as reported in ref [71].

To identify if the source of the drive current increase is either channel resistance (R_{ch}) or external resistance (R_{ext}), R_{ext} of IF-$Si_{0.45}Ge_{0.55}$ channel pFET was plotted as a function of W_{eff} as shown in Figure 4-9. R_{ext} doesn't show any clear W_{eff} dependence, ranging from 100 to 160 ohm-μm. Therefore, this drive current increase can not be explained by R_{ext}. Based on the R_{ext} values in Figure 4-9, R_{ch} has been extacted at channel width of 10 μm and 0.1 μm for IF-$Si_{0.45}Ge_{0.55}$ channel pFET and shown in Figure 4-10. R_{ch} was confirmed to reduce by approximately 60% at 0.1 μm compared to 10 μm, which corresponds to 2.5x increase in hole mobility if no T_{inv} variation at various W_{eff} is assumed.

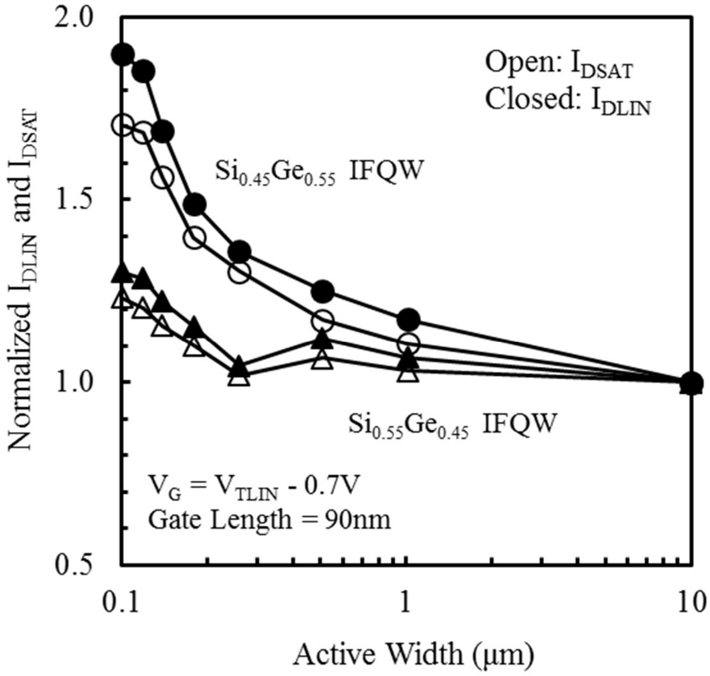


Figure 4-8 Normalized linear (I_{DLIN}) and saturation (I_{DSAT}) drive current at constant gate overdrive ($V_G – V_{TLIN}$ = -0.7V) as a function of transistor effective channel width (W_{eff}). Drive currents are normalized by the current at channel width of 10 μm. V_d = -0.05 V and -1.0 V for linear and saturation current, respectively. IF-$Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel pFETs are compared. Both devices have eSiGe stressor and their gate length is 90 nm.

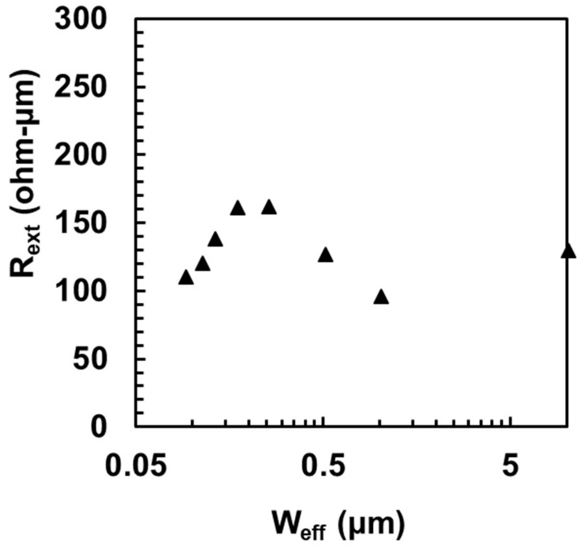


Figure 4-9 External resistance (R_{ext}) for IF-$Si_{0.45}Ge_{0.55}$ channel pFET as a function of channel width (W_{eff}). R_{ext} is ranging from 100 to 160 Ohm-µm without clear W_{eff} dependence.

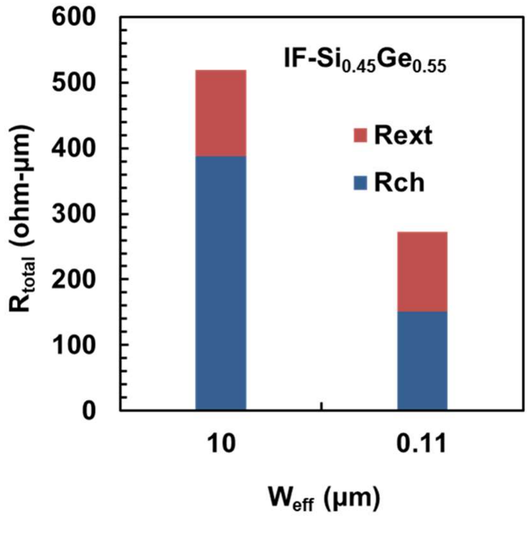
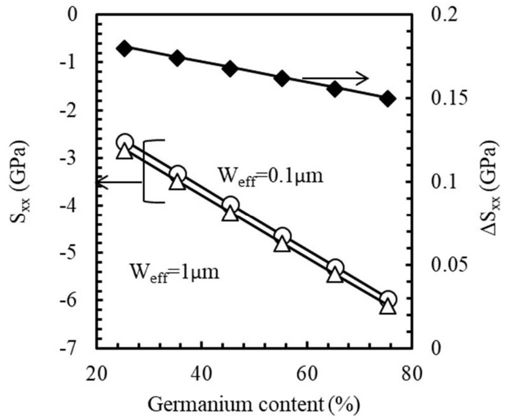


Figure 4-10 On-resistance (R_{total}) and its breakdown into R_{ext} and R_{ch} at channel width of 10 µm and 0.1 µm for IF-$Si_{0.45}Ge_{0.55}$ channel pFET. R_{ch} is reduced by approximately 60%.

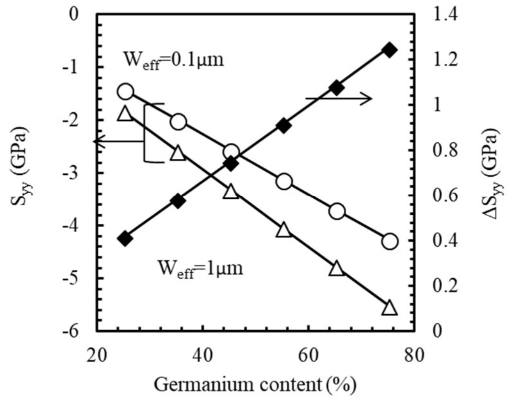
To understand how hole mobility is modulated with device channel width, TCAD stress simulation was carried out for IF-$Si_{1-x}Ge_x$ channel (with eSiGe stressor) with various Ge content in the channel (from 25% to 75%). The stress values for each axis (S_{xx}, S_{yy}, and S_{zz}, where x is longitudinal, y is transverse, and z is vertical to the current flow) at channel width of 1 μm and 0.1 μm are calculated and plotted in Figure 4-11. The amount of stress change between 1 μm and 0.1 μm (ΔS_{xx}, ΔS_{yy}, and ΔS_{zz}) are also plotted together. All stress values are averaged along the transistor channel width direction. As explained earlier, $Si_{1-x}Ge_x$ channel is under compressive stress from Si substrate (biaxial) and additionally from rSiGe extension and eSiGe stressor (uniaxial). It is found that ΔS_{xx} and ΔS_{zz} are relatively smaller compared to ΔS_{yy} and are not that sensitive to Ge content in $Si_{1-x}Ge_x$ channel when device channel width gets narrowed. S_{yy} relaxes significantly at narrow channel width and is very sensitive to Ge content in $Si_{1-x}Ge_x$ channel. Although $Si_{1-x}Ge_x$ at middle of the channel area is still biaxially strained, $Si_{1-x}Ge_x$ layer close to STI edge is elastically relaxing in transverse direction. Therefore as the channel width gets narrower, the transverse strain starts to relax in entire channel region. The elastic strain relaxation happens over several hundreds nano meters range, as can be seen from nano beam diffraction (NBD) result [69]. And the magnitude of the relaxation becomes larger at higher Ge content because the initial strain is higher. As this strain relaxation happens along transverse direction, S_{xx} and S_{zz} are not that affected as shown in Figure 4-11.

Corresponding hole mobility change from active channel width of 1 μm to 0.1 μm is plotted as a function of Ge content in IF-$Si_{1-x}Ge_x$ channel in Figure 4-12. Hole mobility was calculated using piezo-resistance coefficient of $Si_{0.55}Ge_{0.45}$ channel reported in ref. [69] for the whole Ge concentration range for simplicity. ΔS_{xx}, ΔS_{yy}, and ΔS_{zz} from Figure 4-11 are also plotted. It is clearly shown that hole mobility change is mainly driven by

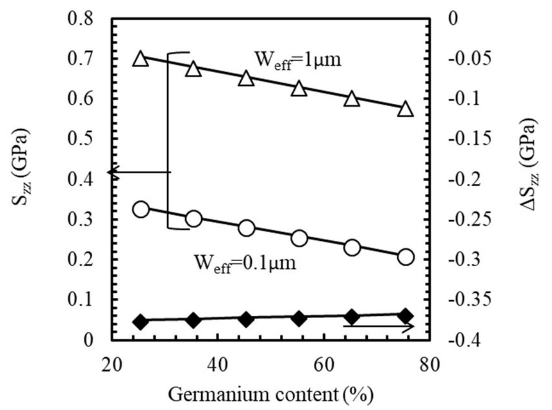
ΔS_{yy}. Although there is rough quantitative matching for IF-$Si_{0.55}Ge_{0.45}$ case between experiment shown in Figure 4-8 and TCAD simulation (assuming I_{DLIN} improvement is coming from mobility improvement), IF-$Si_{0.45}Ge_{0.55}$ showed significantly higher drive current / R_{ch} improvement than TCAD simulation. In case of IF-$Si_{0.45}Ge_{0.55}$ pFET case, hole mobility increase estimated from experiment was approximately 2.5x, which is much higher than 1.4 - 1.5x increase calculated in Figure 4-12. There is no clear explanation for this mismatch, but possible reasons are either higher transverse piezo-resistance coefficient for $Si_{0.45}Ge_{0.55}$ channel than the one for $Si_{0.55}Ge_{0.45}$ channel (We observed higher longitudinal piezo-resistance coefficient for $Si_{1-x}Ge_x$ channel under higher bi-axial strain in Section 4.2, however, is it not yet clear whether this also applies to transverse piezo-resistance coefficient or not) and/or thicker effective $Si_{1-x}Ge_x$ thickness due to Ge diffusion, which also enhances the active width dependence due to larger elastic stress relaxation [69].



(a)



(b)



(c)

Figure 4-11 (a) Longitudinal, (b) transverse, and (c) vertical stress (S_{xx}, S_{yy}, and S_{zz}) and its delta (ΔS_{xx}, ΔS_{yy}, and ΔS_{zz}) between channel width 1 μm and 0.1 μm, as a function of Ge content in IF-$Si_{1-x}Ge_x$ channel.

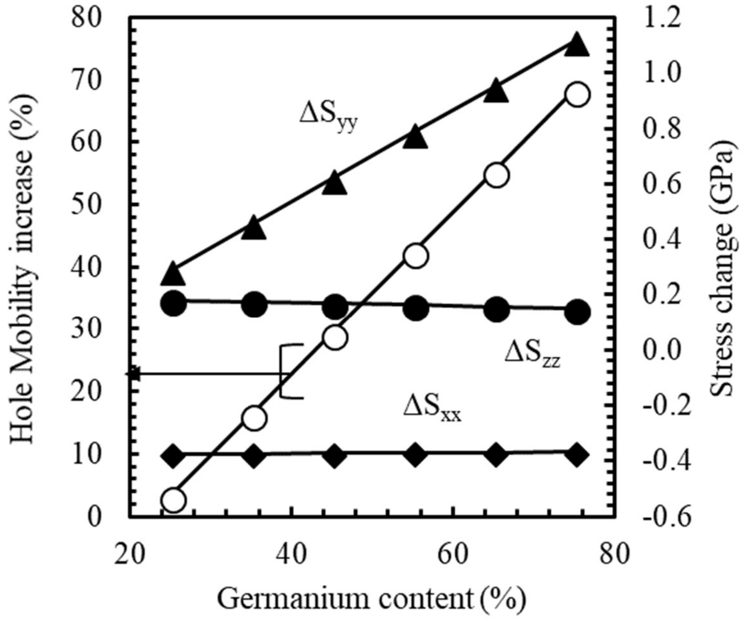


Figure 4-12 Hole mobility change of IF-$Si_{1-x}Ge_x$ channel pFET (with eSiGe stressor) between channel width of 1 µm and 0.1 µm calculated from ΔS_{xx}, ΔS_{yy}, and ΔS_{zz} (plotted together) in Figure 4-11 and piezo-resistance coefficient in ref [69] ($Si_{0.55}Ge_{0.45}$).

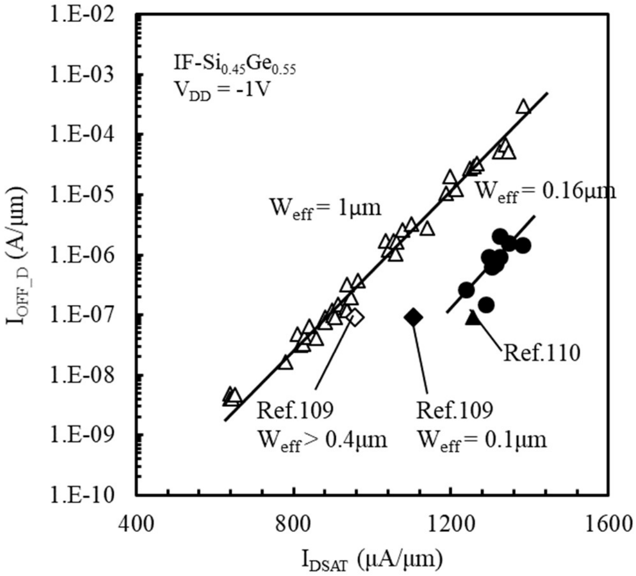
4.4 DC performance benchmark for IF-$Si_{0.45}Ge_{0.55}$ channel pFETs with eSiGe stressor

In previous section, we discussed key performance elements for $Si_{1-x}Ge_x$ channel pFET. One is longitudinal stress enhancement with rSiGe extension and/or eSiGe stressor, and another is stress relaxation in transverse direction at narrow channel width. In this section, DC performance of IF-$Si_{0.45}Ge_{0.55}$ channel pFET with eSiGe stressor is benchmarked with other reports in literature.

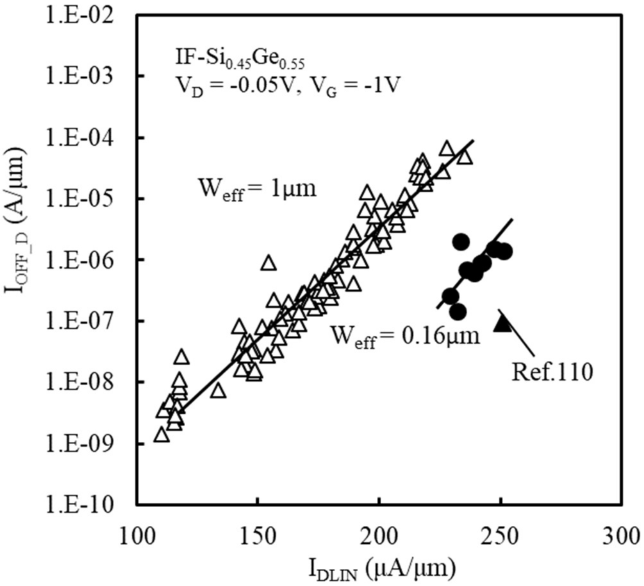
I_{DSAT} - I_{OFF_D} and I_{DLIN} - I_{OFF_D} for IF-$Si_{0.45}Ge_{0.55}$ with eSiGe devices are plotted in Figure 4-13, together with literature data for $Si_{1-x}Ge_x$-based planar devices [111, 112]. Each plot has data from channel width of 1 µm and 0.16 µm. It shows an excellent I_{DSAT}

of 1.28 mA/μm and I_{DLIN} of 231 μA/μm at 160 nA/μm off-current at channel width of 0.16 μm, which is one of the best among all $Si_{1-x}Ge_x$-based planar devices reported to date. Both I_{DSAT} and I_{DLIN} show about 30% improvements from 1 μm to 0.16 μm channel width.

I_D - V_G curve (linear, saturation regime) for the device is shown in Figure 4-14. Despite of the excellent DC performance of the devices, it should be noted that short channel control has a room for improvement, as shown in drain induced barrier lowering (DIBL), sub-threshold swing in saturation (SS_{sat}), and relatively larger gate length of 30 nm. Chapter 5 will address the scalability of $Si_{1-x}Ge_x$ channel devices [111, 112].



(a)



(b)

Figure 4-13 (a) I_{DSAT} vs I_{OFF_D} and (b) I_{DLIN} vs I_{OFF_D} for IF-$Si_{0.45}Ge_{0.55}$ channel pFET at V_D = -1 V and -0.05 V, respectively. Each plot shows channel width of 1 µm and 0.16 µm.

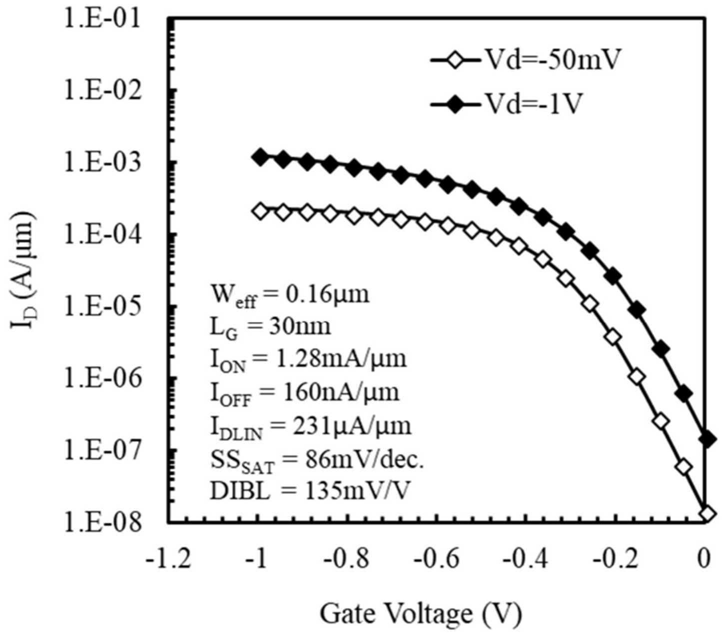


Figure 4-14 I_D - V_G curve for IF-$Si_{0.45}Ge_{0.55}$ channel pFET with eSiGe stressor. Inset shows summary of the device characteristics.

4.5 Summary

In this chapter, we discussed device characteristics of IF-$Si_{1-x}Ge_x$ channel pFET mainly from the view point of the stress applied to the $Si_{1-x}Ge_x$ channel and demonstrated high performing $Si_{1-x}Ge_x$ channel pFET with high Ge content at scaled gate length.

In Section 4.2, interaction between $Si_{1-x}Ge_x$ channel and eSiGe stressor has been investigated. It was shown that even with elastic strain relaxation of $Si_{1-x}Ge_x$ channel during cavity RIE to form eSiGe stressor, $Si_{1-x}Ge_x$ channel pFET shows at least comparable hole mobility enhancement with eSiGe stressor to Si channel. This can be explained by larger absolute values of longitudinal piezo-resistance coefficient for $Si_{1-x}Ge_x$ material under the biaxial compressive strain.

In Section 4.3, active channel width dependence of drive current for IF-$Si_{1-x}Ge_x$ channel pFET with eSiGe stressor was discussed. It was shown that transistor drive current was increased by up to 1.9/1.6x from channel width of 10 μm to 0.1 μm in linear and saturation regime, respectively. This corresponds to 60% R_{ch} reduction (2.5x hole mobility enhancement if no T_{inv} change) in case of IF-$Si_{0.45}Ge_{0.55}$ channel pFET. TCAD simulation has shown that this hole mobility enhancement is due to elastic strain relaxation along transverse direction, although there is still a quantitative gap between the simulation result and the experimental result, which needs to be clarified in future work.

Thanks to the hole mobility enhancement by eSiGe stressor and transverse strain relaxation at narrow channel width, IF-$Si_{0.45}Ge_{0.55}$ channel pFET with eSiGe showed an excellent DC performance which is one of the best $Si_{1-x}Ge_x$ based planar pFET to date, which was described in Section 4.4.

Chapter 5 Scalability of $Si_{1-x}Ge_x$ channel pFETs

5.1 Background

In Chapter 4, DC performance of IF−$Si_{0.45}Ge_{0.55}$ channel pFET has been benchmarked and shown to have best-in-class performance as planar $Si_{1-x}Ge_x$ channel devices. Despite of the decent DC performance, relatively degraded short channel effect (DIBL, SS_{sat}) was raised as a concern in terms of gate length scalability.

In this chapter, we first review short channel effects for conventional and IF-$Si_{1-x}Ge_x$ channel device to demonstrate the superiority of IF-$Si_{1-x}Ge_x$ devices in terms of short channel control. Then we review device performance of IF-$Si_{0.55}Ge_{0.45}$ channel devices with and without eSiGe stressor both from DC performance and from short channel effect / gate length scaling point of view. Projected AC performance comparison between IF-$Si_{1-x}Ge_x$ with and without eSiGe will be also discussed in detail.

5.2 Short channel control comparison between conventional and IF-$Si_{1-x}Ge_x$ channel pFET

To begin with this chapter, we will first discuss short channel control comparison between conventional $Si_{1-x}Ge_x$ and IF-$Si_{1-x}Ge_x$ channel pFETs devices. The conventional $Si_{1-x}Ge_x$ channel pFETs received 1000°C spike RTA, while 950°C spike RTA was applied to IF-$Si_{1-x}Ge_x$ channel pFETs. Figure 5-1 shows Drain Induced Barrier Lowering (DIBL) as a function of gate length for the conventional $Si_{0.45}Ge_{0.55}$, IF-$Si_{0.55}Ge_{0.45}$, IF-$Si_{0.45}Ge_{0.55}$ channel pFETs. All devices have eSiGe stressor. At given Ge concentration (Ge 55%), compared to the conventional devices, IF-$Si_{1-x}Ge_x$ channel device shows much improved short channel control (smaller DIBL) thanks to shallower junction depth enabled by rSiGe extension. When comparing among IF-$Si_{1-x}Ge_x$ channel devices, one can notice that

$Si_{0.55}Ge_{0.45}$ has smaller DIBL than that of $Si_{0.45}Ge_{0.55}$. This trend is supported by larger gate overlap length (L_{OV}) for $Si_{0.45}Ge_{0.55}$ as shown in inset in Figure 5-1. L_{OV} was calculated by substituting electrical gate length from physical gate length (then divided by two). L_{OV} is 9.5 nm and 7 nm for IF-$Si_{0.45}Ge_{0.55}$ and $Si_{0.55}Ge_{0.45}$ channel, respectively. The detailed methodology for L_{OV} extraction is discussed in ref. [113]. One possible mechanism of this result would be enhanced boron diffusion for $Si_{0.45}Ge_{0.55}$ channel compared to $Si_{0.55}Ge_{0.45}$. Boron diffusion has been reported to be retarded in $Si_{1-x}Ge_x$ and Ge compared to Si, and higher Ge content in $Si_{1-x}Ge_x$ suppresses more boron diffusion [114, 115]. Therefore, our experimental result apparently cannot be explained by this mechanism. However, if $Si_{0.45}Ge_{0.55}$ layer has more crystal defects than $Si_{0.55}Ge_{0.45}$ (it is highly likely because of thinner critical layer thickness for $Si_{0.45}Ge_{0.55}$), boron diffusion might have been enhanced due to these defects like transient enhanced diffusion (TED) [116]. Another possible mechanism is degradation of short channel control due to larger effective $Si_{1-x}Ge_x$ layer thickness for $Si_{0.45}Ge_{0.55}$ than $Si_{0.55}Ge_{0.45}$. As Ge diffusion is supposed to be larger in $Si_{0.45}Ge_{0.55}$ channel than in $Si_{0.55}Ge_{0.45}$, final $Si_{1-x}Ge_x$ thickness after all the thermal processes can be thicker for $Si_{0.45}Ge_{0.55}$ channel. The detailed mechanism of the observed difference in short channel control is still unknown and needs further investigation.

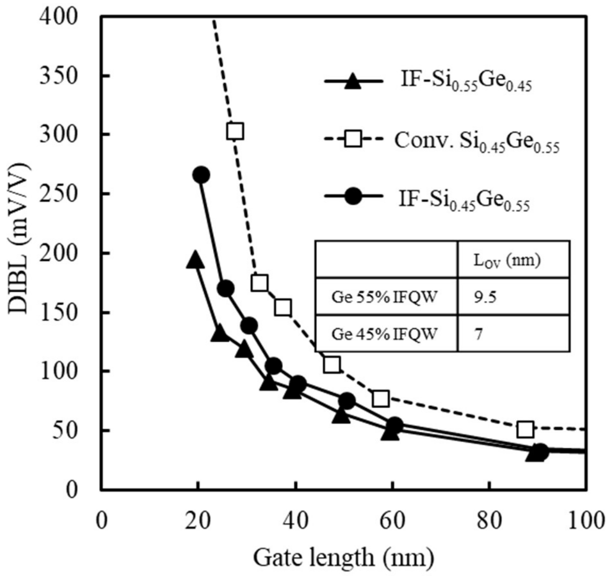


Figure 5-1 DIBL as a function of gate length for conventional devices ($Si_{0.45}Ge_{0.55}$) and IF devices ($Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$). Inset shows overlap length for IF devices.

Saturation drive current (I_{DSAT}) as a function of overlap capacitance is summarized in Figure 5-2 for the same set of devices discussed in Figure 5-1. I_{DSAT} is taken at I_{OFF_D} = 100 nA/µm and at constant gate overdrive V_G - V_{TLIN} = -0.7 V. Device channel width is 1 µm. Overlap capacitance values are well consistent to DIBL trend in Figure 5-1. When comparing IF- and conventional $Si_{1-x}Ge_x$ channel devices with same Ge content, IF-$Si_{1-x}Ge_x$ devices have smaller overlap capacitance and higher drive current than conventional devices at the same time. Higher drive current should be thanks to higher short channel hole mobility for IF devices than conventional devices as seen in Figure 4-7. The higher short channel mobility for IF-$Si_{1-x}Ge_x$ channel is realized by elimination of halo implant (reduced Coulomb scattering) and higher quality of $Si_{1-x}Ge_x$ layer with lower process temperature.

When comparing IF-$Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$ channel devices, $Si_{0.45}Ge_{0.55}$ is less performing than $Si_{0.55}Ge_{0.45}$ (again, device channel width is 1 μm. Considering that short channel mobility (please see Figure 4-7) and external resistance are comparable (179 and 188 Ohm-μm for $Si_{0.55}Ge_{0.45}$ and $Si_{0.45}Ge_{0.55}$, respectively), this performance difference is because of the degraded short channel control compared to $Si_{0.55}Ge_{0.45}$ channel. But annealing condition can be further optimized to adjust overlap length for $Si_{0.45}Ge_{0.55}$ channel in the future.

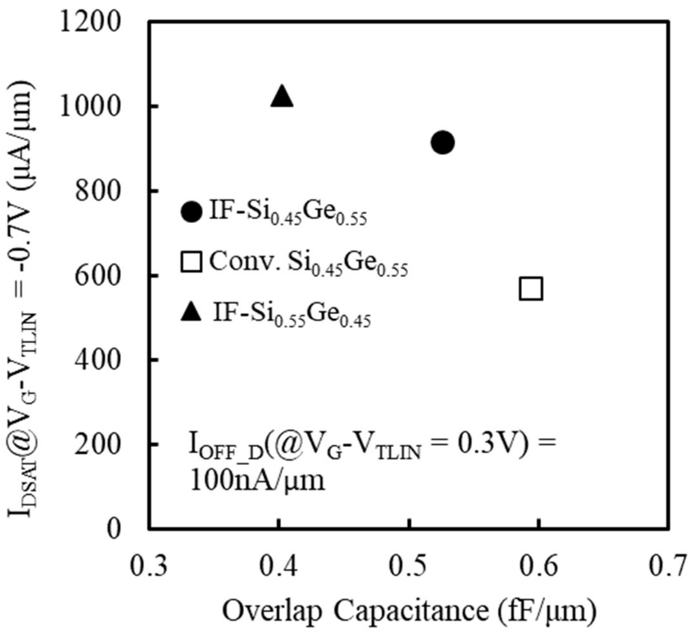


Figure 5-2 I_{DSAT} ($V_G - V_{TLIN}$ = -0.7 V) as a function of overlap capacitance. I_{DSAT} is taken at I_{OFF_D} ($V_G - V_{TLIN}$ = 0.3 V) is 100 nA/μm. Device channel width is 1 μm for all devices.

5.3 Short channel control comparison between IF-$Si_{1-x}Ge_x$ channel pFET with and without eSiGe stressor

In the last section, we confirmed superior performance and scalability of IF-$Si_{1-x}Ge_x$ channel devices over conventional $Si_{1-x}Ge_x$ channel devices. In this section, we compare performance and scalability of IF-$Si_{1-x}Ge_x$ channel devices with and without eSiGe stressor. For convenience, we simply call them IF-rSiGe and IF-eSiGe devices in later part of the thesis. For this comparison, $Si_{0.55}Ge_{0.45}$ channel (thickness is 4 nm) was used. TEM images for studied devices are shown in Figure 3-3 ((a) IF-rSiGe device, (b) IF-eSiGe device). Thickness of the rSiGe layer is 50 nm for rSiGe devices and 30 nm for IF-eSiGe devices. Both devices received spike RTA 950ºC.

Figure 5-3 shows I_{DSAT} - I_{OFF_D} curve for IF-$Si_{0.55}Ge_{0.45}$ rSiGe and eSiGe devices (V_{DD} = -1 V). IF-eSiGe devices have two flavors, namely one without S/D implant and with additional S/D implant. IF-eSiGe devices have about 8% higher I_{DSAT} than IF-rSiGe devices at fixed I_{OFF_D} of 100 nA/µm (1040/1036 µA/µm for IF-eSiGe devices with/without S/D implant, 965 µA/µm for IF-rSiGe devices). Comparable drive current between devices with and without S/D implant is likely because lower external resistance with S/D implant is offset by higher channel resistance due to longer gate length at target I_{OFF_D} (100 nA/µm). Some outliers seen on IF-rSiGe devices are due to high junction leakage (S/D junction can be located very close to or within $Si_{0.55}Ge_{0.45}$ channel layer, therefore any crystal defect can cause very high leakage. This needs to be solved by less defective epitaxial process).

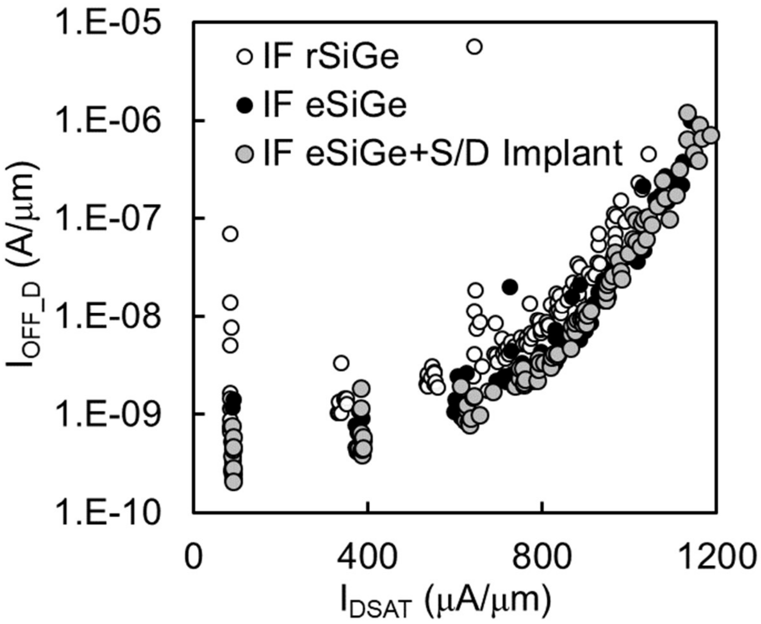


Figure 5-3 I_{DSAT}-I_{OFF_D} for IF-$Si_{0.55}Ge_{0.45}$ channel devices with and without eSiGe ("IF-eSiGe devices" and "IF-rSiGe devices"). IF-eSiGe devices have S/D implant split. V_{DD} = -1 V.

There has been intensive discussion about performance comparison between IF-rSiGe and IF-eSiGe devices [106, 117]. In ref. [106], it was pointed out by TCAD simulation that IF-eSiGe devices have higher longitudinal compressive stress in the channel and higher hole mobility than IF-rSiGe devices at gate length above 20 nm, but the stress for IF-eSiGe devices becomes lower than that for IF-rSiGe devices at gate length below 20 nm in case of $Si_{0.55}Ge_{0.45}$ channel (Figure 5-4). This is because compressive strain in the channel applied from rSiGe extension is fully maintained in case of IF-rSiGe devices regardless of the gate length, while it is partially lost in case of IF-eSiGe devices due to the cavity RIE which is similar phenomenon discussed in Chapter 4 and this strain loss becomes larger at shorter gate length [100, 106, 117]. In this study, it was found that IF-eSiGe devices has slightly higher (+3%) short channel mobility than

IF-rSiGe devices as shown in Figure 5-5 (a). The short channel mobility was extracted and calculated by the same technique as in Figure 4-2 [101]. The gate length at I_{OFF_D} 100 nA/µm is close to 20 nm and the mobility is not very different as TCAD simulation predicted. This small hole mobility difference cannot explain the observed performance difference.

Figure 5-5 (b) shows inversion C-V characteristics of IF-$Si_{0.55}Ge_{0.45}$ rSiGe and eSiGe devices. It was found that their T_{inv} values are considerably different and it can account for I_{DSAT} difference in Figure 5-3 (8% difference, IF-rSiGe devices 1.41 nm, IF-eSiGe devices 1.30 nm) although they started from same gate stack. This T_{inv} difference should be because of different amount of interfacial layer scavenging in two kinds of devices which is driven by thermal budget difference between them. In case of IF-eSiGe devices, there is one more $Si_{0.75}Ge_{0.25}$ epitaxial growth step for embedded stressor compared to IF-rSiGe devices. Therefore, IF-eSiGe devices sees additional thermal budget of pre-epitaxy bake (800ºC) and embedded $Si_{0.75}Ge_{0.25}$ epitaxial growth (650ºC). Oxygen scavenging is reportedly promoted by high temperature anneal, like 1000ºC spike RTA [36, 91, 118] and also we observed T_{inv} modulation by junction activation anneal temperature in Chapter 3. Although the additional thermal budget in this case has lower peak temperature than these cases in literatures, its process time is longer (soak anneal compared to spike RTA in case of activation anneal). Therefore it is supposed that oxygen scavenging is promoted with this additional anneal, leaving thinner interfacial layer for IF-eSiGe devices.

From the discussion above, the performance difference in $I_{DSAT} - I_{OFF\_D}$ is mainly attributed to T_{inv} difference.

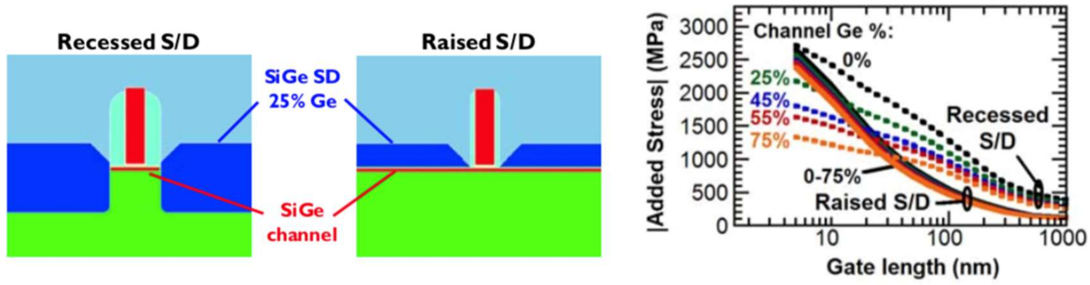


Figure 5-4 (from ref. [106]) TCAD simulation result for added longitudinal stress in the $Si_{1-x}Ge_x$ channel (x: 0, 0.25, 0.45, 0.55, and 0.75) for IF-eSiGe ("Recessed S/D" in the figure) and IF-rSiGe ("Raised S/D"). At relatively longer gate length, IF-eSiGe devices are supposed to have more compressive stress. There is a turn-around point below which IF-rSiGe devices have more compressive stress than IF-eSiGe devices due to stress relaxation during eSiGe cavity RIE.

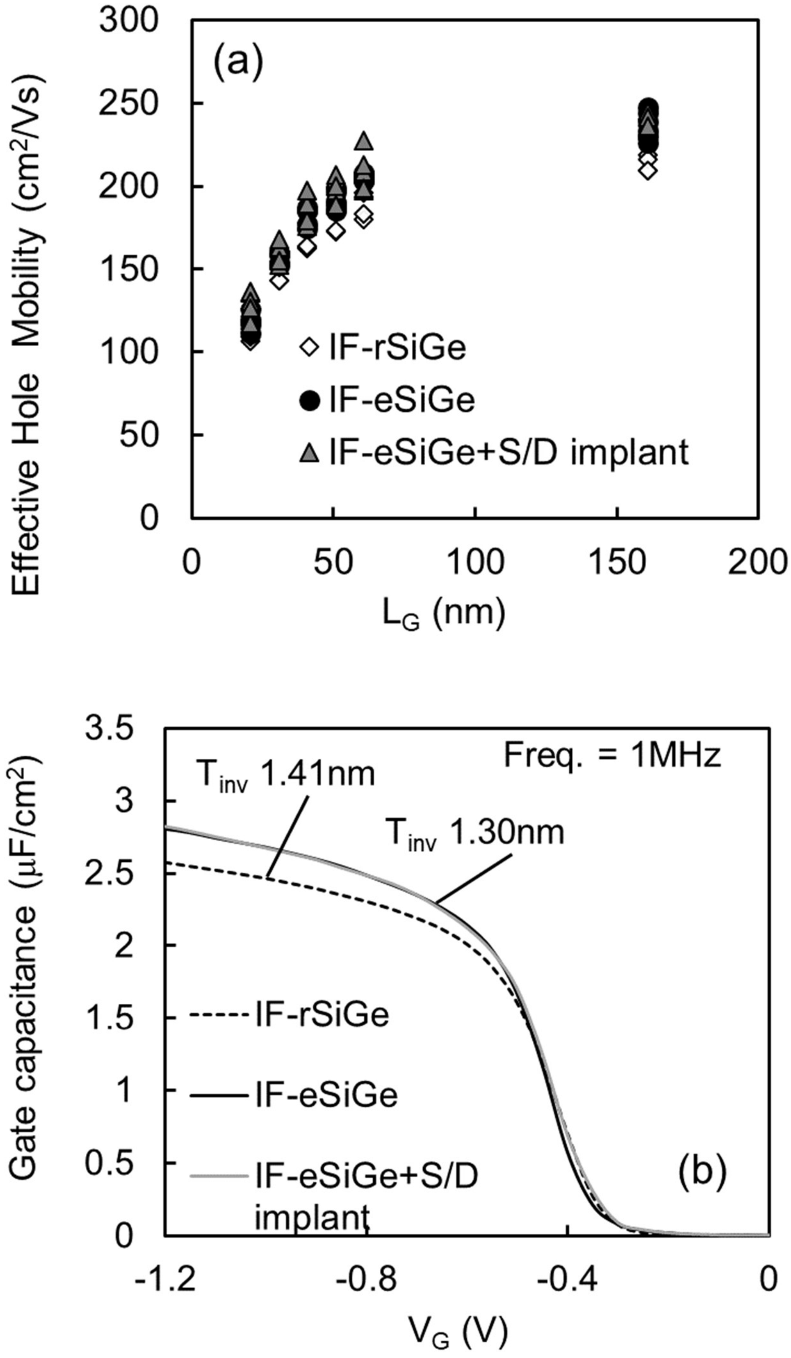
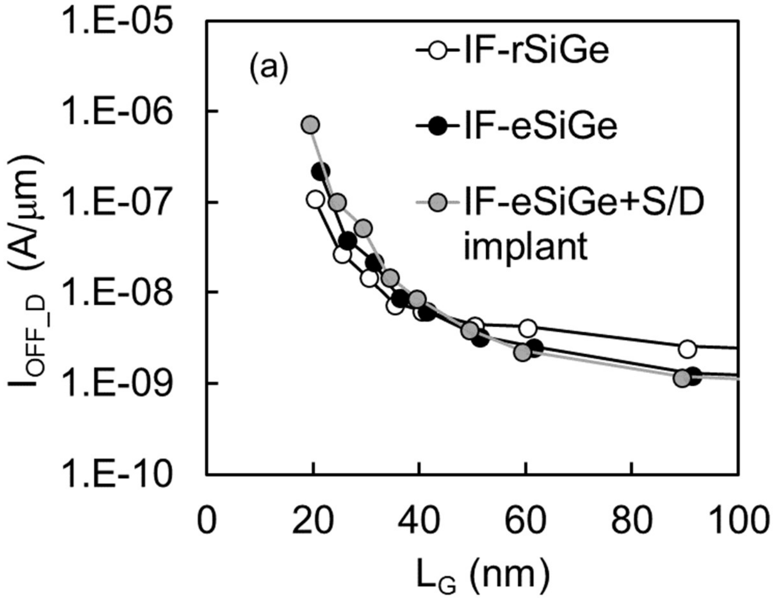


Figure 5-5 (a) Effective hole mobility for IF-rSiGe and IF-eSiGe devices (w/ and w/o S/D implant) as a function of gate length. (b) Inversion C-V curves of same devices.

Figure 5-6 (a) shows I_{OFF_D} as a function of gate length. In short channel region (L_G < 40 nm), IF-rSiGe device has lower I_{OFF_D} than IF-eSiGe device. However, it is opposite at longer gate length (L_G > 40 nm) and there is a turnover at around L_G of 40 nm. Figure 5-6 (b) shows I_D-V_G curves of IF-eSiGe and IF-rSiGe devices at L_G of 0.21 μm. Junction leakage is clearly dominating off-state leakage in IF-rSiGe device. As boron diffuses several nm from rSiGe extension, junction of IF-rSiGe device should be located very close to $Si_{0.55}Ge_{0.45}$ channel layer. On the other hand, in case of IF-eSiGe device, junction is formed in Si or embedded $Si_{0.75}Ge_{0.25}$ S/D, and is significantly deeper. As they have larger band gap than $Si_{0.55}Ge_{0.45}$ [48], junction leakage for IF-eSiGe device is smaller than that of IF-rSiGe device. In shorter gate length (L_G < 40 nm), sub-threshold leakage becomes dominant and IF-rSiGe device has smaller I_{OFF_D} as IF-rSiGe device has better short channel control than IF-eSiGe device thanks to its shallower junction. Additional S/D implant causes excess boron diffusion and leads to further I_{OFF_D} degradation.



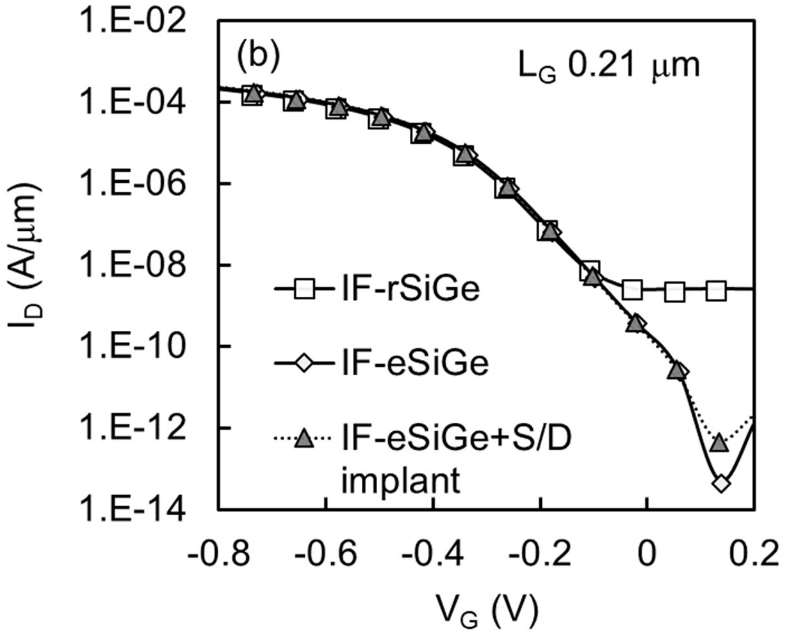


Figure 5-6 (a) I_{OFF_D} as a function of gate length for IF-rSiGe and IF-eSiGe devices (w/ and w/o S/D implant). (b) I_D-V_G curve for same devices at L_G = 0.21 μm.

Superior short channel control of IF-rSiGe devices is also illustrated in DIBL and sub-threshold swing in saturation (SS_{sat}) as a function of gate length in Figure 5-7 (a) and (b), respectively. IF-rSiGe device has smaller DIBL and SS_{sat} than IF-eSiGe device at a given gate length, and additional S/D implant further degrades them. I_{OFF_D}, DIBL and SS_{sat} consistently show superior short channel control with IF-rSiGe devices compared to IF-eSiGe devices.

This trend is also confirmed by gate to drain leakage current, I_{gd}, normalized by area gate leakage current, I_g, (I_{gd}/I_g) shown in Figure 5-7 (c). I_{gd}/I_g is an indicator for direct gate overlap length as I_{gd} is supposed to be proportional to overlap length, and one can compare I_{gd}/I_g between wafers with different gate leakage current levels. From Figure 5-7 (c), it is indicated that the direct overlap length is approximately 60% larger for IF-eSiGe device than IF-rSiGe device, which is consistent to DIBL and SS_{sat} trend.

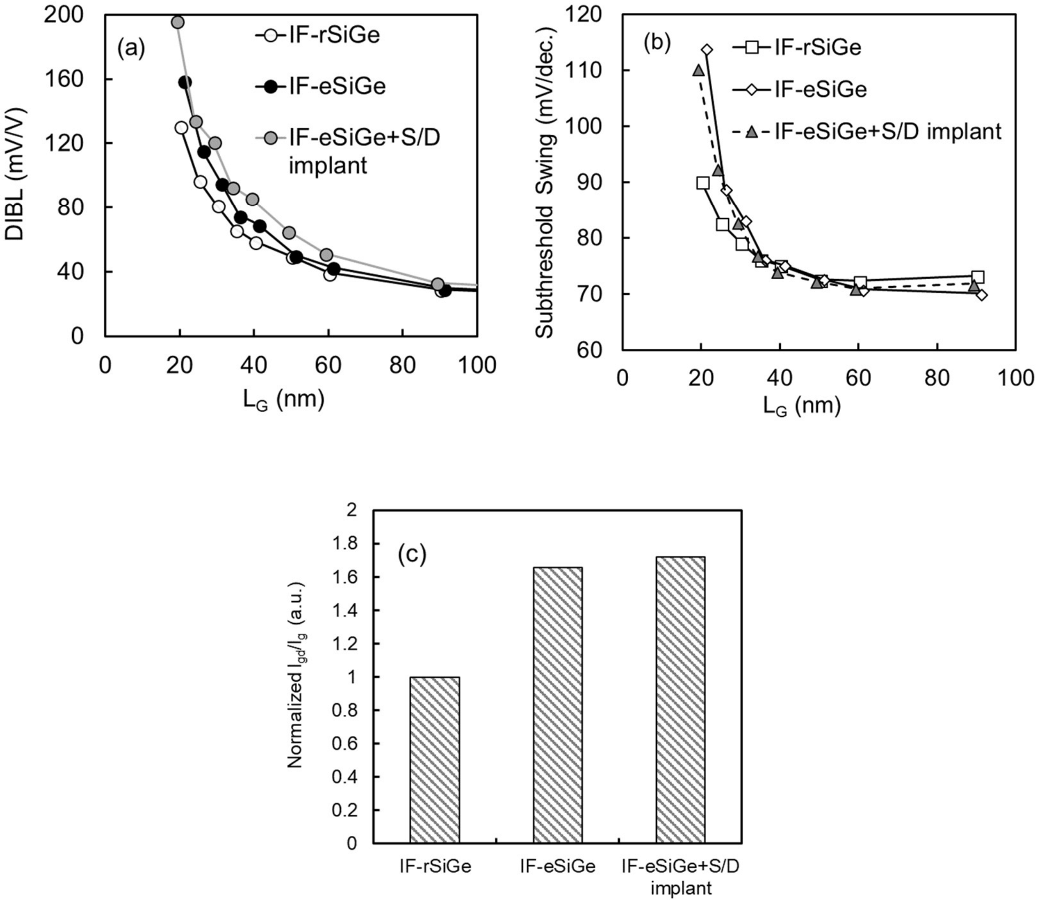


Figure 5-7 (a) DIBL and (b) sub-threshold swing for IF-rSiGe and IF-eSiGe devices (w/ and w/o S/D implant) as a function of gate length. (c) I_{gd}/I_g comparison for same devices. I_{gd}/I_g values are normalized with respect to the value for IF-rSiGe device.

Figure 5-8 (a) shows I_{DSAT} (V_{DD} = -1 V) vs. $L_{G,MIN}$ for IF-rSiGe and IF-eSiGe devices where $L_{G,MIN}$ is defined as L_G at I_{OFF_D} of 100 nA/µm. Although IF-eSiGe devices have higher I_{DSAT} than IF-rSiGe devices thanks to thinner T_{inv} as seen in Figure 5-3, IF-rSiGe devices have much shorter $L_{G,MIN}$ thanks to shallower junction ($L_{G,MIN}$ is approximately 4 nm shorter for IF-rSiGe devices than IF-eSiGe devices) and smaller overlap length which is indicated in Figure 5-7 (c). As shown in Figure 5-3, additional S/D implant didn't improve I_{DSAT} and only degrades $L_{G,MIN}$ by approximately 1.5 nm compared to IF-eSiGe devices without S/D implant. Based on this result, intrinsic delay $C_{inv}V_{DD}/I_{DSAT}$ was calculated and plotted as a function of $L_{G,MIN}$ in Figure 5-8 (b). Inversion layer capacitance C_{inv} is calculated as

$$C_{inv} = \varepsilon_0 \varepsilon_{SiO2} \frac{(L_G \cdot W)}{T_{inv}} \quad (5\text{-}1)$$

where ε_0 and ε_{SiO2} are permittivity of vacuum and dielectric constant of SiO_2, respectively. Thanks to the shorter $L_{G,MIN}$, IF-rSiGe devices show the smallest intrinsic delay among all studied devices.

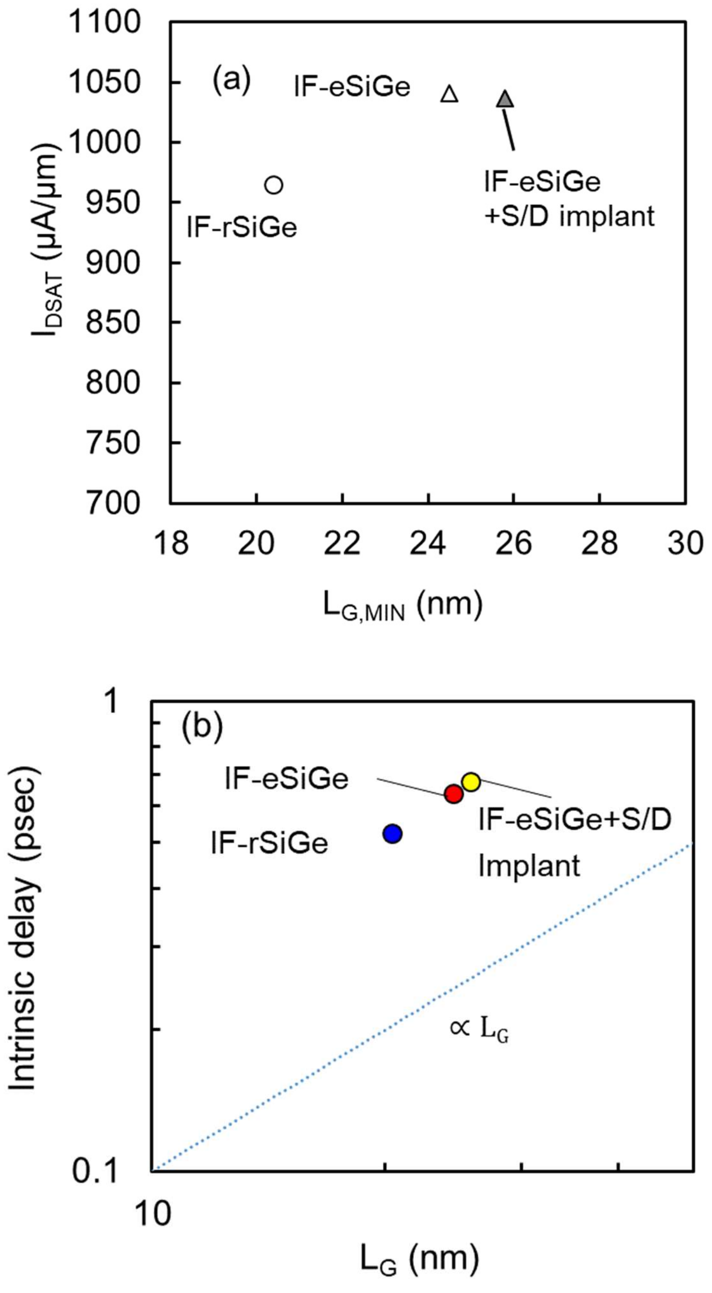


Figure 5-8 (a) I_{DSAT} vs. $L_{G,MIN}$ (L_G at fixed I_{OFF_D} of 100 nA/μm) for IF-rSiGe and IF-eSiGe devices (w/ and w/o S/D implant). (b) Calculated pFET intrinsic delay for same devices.

5.4 Summary

In this chapter, we investigated gate length scalability of various kinds of $Si_{1-x}Ge_x$ channel pFET devices, namely, conventional $Si_{1-x}Ge_x$ channel devices, IF-$Si_{1-x}Ge_x$ channel devices with eSiGe stressor (IF-eSiGe), and IF-$Si_{1-x}Ge_x$ channel without eSiGe stressor (IF-rSiGe). It was demonstrated that IF-$Si_{1-x}Ge_x$ channel devices have higher drive current and better short channel control at the same time, compared to conventional $Si_{1-x}Ge_x$ channel devices (both have eSiGe). This is thanks to higher short channel mobility (no Coulomb scattering due to halo implant) and shallower junction depth (raised SiGe extension) for IF-$Si_{1-x}Ge_x$ channel devices. Further gate length scaling was confirmed with IF-rSiGe devices (further reduction of gate overlap length), with a cost of I_{DSAT} degradation. As a result, IF-rSiGe devices showed the smallest intrinsic delay.

To enable further device scaling based on IF-rSiGe device structure, thinning of $Si_{1-x}Ge_x$ channel thickness can be pursued in future as it can suppress short channel effect further and also may help to reduce junction leakage. However, it should be noted that thinning of $Si_{1-x}Ge_x$ channel may cause V_T increase [56] and performance degradation due to additional scattering by crystal defects at $Si_{1-x}Ge_x$/Si substrate interface [56].

Chapter 6 Conclusions and Future Direction

6.1 Conclusions

In this thesis, two kinds of pFET V_T modulation techniques were discussed in detail. One is eWF control technique by gate stack engineering in RMG FinFET technology. The other is introduction of $Si_{1-x}Ge_x$ channel epitaxially grown on Si substrate.

In first part of the thesis (Chapter 2), eWF control technique by gate stack engineering in RMG FinFET technology was discussed, and we identified the detailed physical mechanism of pFET V_T modulation by WSA (WF setting anneal) and proposed practical eWF control technique to achieve low pFET V_T at scaled EOT. We revealed that intermixed layer created in-between TiN and HfO_2 during WSA ($HfTiON_x$) has negative fixed charges and it reduces pFET V_T by about 90mV. On top of that, we also found that higher WSA temperature further reduces pFET V_T by about 50mV thanks to passivation of oxygen vacancies (positively charged) in HfO_2 with oxygen atoms diffused from TiN layer. By combining these effects ($HfTiON_x$ interfacial layer and neutralization of oxygen vacancies), one can further push eWF towards valence band edge by approximately 140 meV.

In second part of the thesis (Chapter 3~ Chapter 5), we pursued to increase Ge content in $Si_{1-x}Ge_x$ channel to achieve low pFET V_T. Especially, we demonstrated for the first time successful integration of high performing $Si_{1-x}Ge_x$ channel pFET with extremely high Ge content ($x = 0.55$) at competitive gate length ($L_G \sim 20$ nm). In Chapter 3, the impact from process temperature on $Si_{1-x}Ge_x$ channel pFETs has been discussed. Low process temperature was identified as a key to realize high hole-mobility and low pFET V_T by keeping high biaxial compressive strain in high Ge content ($x > 0.5$) $Si_{1-x}Ge_x$ channel layer and steep Ge concentration gradient. From this point of view, IF-$Si_{1-x}Ge_x$

channel device has been identified as potential "Go-To" device structure, because this device doesn't need high temperature activation anneal to cure implant damage. In Chapter 4, key performance elements for IF-$Si_{1-x}Ge_x$ channel pFETs were discussed. One is eSiGe stressor, and the other is narrow channel effect by transverse strain relaxation. $Si_{1-x}Ge_x$ channel pFET showed at least comparable hole mobility enhancement with eSiGe stressor to Si channel even with strain relaxation of $Si_{1-x}Ge_x$ channel during cavity RIE. This can be explained by larger absolute values of longitudinal piezo-resistance coefficient for $Si_{1-x}Ge_x$ material under the biaxial compressive strain. It was also shown that IF-$Si_{1-x}Ge_x$ pFET drive current was increased significantly with channel width scaling. TCAD simulation has shown that this hole-mobility enhancement is due to elastic strain relaxation along transverse direction, although there is still a quantitative gap between the simulation result and the experimental result, which needs to be clarified in future work. Thanks to the hole-mobility enhancement by eSiGe stressor and transverse strain relaxation at narrow channel width, IF-$Si_{0.45}Ge_{0.55}$ channel pFET showed an excellent DC performance, which is one of the best $Si_{1-x}Ge_x$ based planar pFET to date. In Chapter 5, gate length scalability of $Si_{1-x}Ge_x$ channel pFET has been compared. It was demonstrated that IF-rSiGe device (IF-$Si_{1-x}Ge_x$ channel without eSiGe) has the best scalability ($L_{G,MIN}$) and the smallest intrinsic delay despite of I_{DSAT} degradation.

6.2 Future direction

In this work, physical mechanism of eWF modulation by WSA was revealed. However, there are some observations which require further study in the future. One is to clarify the mechanism of negative fixed charge generation in intermixed layer between HfO_2 and TiN ($HfTiON_x$) film. One possibility is formation of electrical dipole moment.

In this work, we couldn't investigate the detail of this fixed charge, such as activation energy of these charged defects. It needs temperature dependence of NBTI to extract the activation energy.

NBTI improvement by WSA is another interesting observation which needs more understanding. As discussed in the thesis, NBTI improvement by WSA is supposedly not due to better interfacial quality between SiO_2-IL and Si-sub, but due to bulk trap reduction in HfO_2 film (D_{it} and physical thickness of SiO_2-IL was confirmed comparable). To evaluate bulk trap in HfO_2 film and prove the hypothesis, more characterization needs to be done such as hysteresis analysis and pulsed I-V measurement.

For $Si_{1-x}Ge_x$ channel part, further reduction of thermal budget is possible direction to further improve device V_t and hole mobility. In general, pFET V_t tends to go higher at scaled EOT, so even lower eWF may be required in future technology nodes. Even with 950ºC spike RTA, there were some indications of crystal defects in case of $Si_{0.45}Ge_{0.55}$ channel, such as faster boron diffusion than in $Si_{0.55}Ge_{0.45}$ (as discussed in the thesis, presence of Ge fraction should retard boron diffusion). Therefore, lower thermal budget not only improves V_t and mobility, but also improves short channel control, which was an issue with $Si_{0.45}Ge_{0.55}$ channel compared to $Si_{0.55}Ge_{0.45}$ channel. There is concern about not enough gate to extension overlap in case of lower thermal budget. This needs to be addressed by reducing offset spacer thickness (in this work, we used 6 nm spacer), however, thin offset spacer would be a significant risk for yield due to possible poor encapsulation of the gate stack. Another knob for further $Si_{1-x}Ge_x$ channel device scaling is thinning of $Si_{1-x}Ge_x$ channel thickness as an analogous to SOI thickness in SOI device. This may cause V_T increase due to quantum confinement effect and performance degradation due to additional scattering by crystal defects at $Si_{1-x}Ge_x$/Si substrate interface, so careful optimization of the device structure and selection of process thermal budget would be required to fully harness the potential of $Si_{1-x}Ge_x$ channel devices.

Lastly, in future technology nodes, Si-Fin would be replaced with $Si_{1-x}Ge_x$-Fin in pFET and metal gate integration scheme should remain RMG. Therefore, two technologies discussed in this thesis ($Si_{1-x}Ge_x$ channel device, RMG process optimization for pFET VT reduction) and their combination should become extremely important in future. In such case, compatibility of WSA anneal with $Si_{1-x}Ge_x$ channel (thermal stability) would be a key.

References

[1] F. M. Wanlass, *et al.*, "Fairchild Application Note 77. "CMOS, the Ideal Logic Family"," *Archived 2015/01/09 at the Wayback Machine,* 1983.

[2] Fairchild, "CMOS, the Ideal Logic Family," *Application Note 77 (archived at the Wayback Machine, 2015/1/9),* 1983.

[3] Y. Taur and T. H. Ning, Fundamentals of Modern VLSI Devices, CAMBRIDGE UNIVERSITY PRESS, 2013.

[4] R. H. Dennard, *et al.*, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid State Circuits,* pp. SC-9, 1974.

[5] M. Roser and H. Ritchie, "Technological Progress I. Empirical View," [Online]. Available: https://ourworldindata.org/technological-progress.

[6] G. E. Moore, *et al.*, "Cramming more components onto integrated circuits," *Electronics,* vol. 38, no. 8, p. 114, 1965.

[7] S. Natarajan, *et al.*, "A 32nm Logic Technology Featuring 2nd-Generation High-k + Melta-Gate Transistors, Enhanced Channel Strain and 0.171um2 SRAM Cell Size in a 291Mb Array," *Technical Digest of IEDM,* p. 1, 2008.

[8] K. Mistry, *et al.*, "A 45nm Logic Technology with High-k + Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," *Technical Digest of IEDM,* p. 247, 2007.

[9] C. Auth, *et al.*, "45nm High-k + Metal Gate Strain-Enhanced Transistors," *Technical Digest of VLSI Symposium on Technology,* p. 128, 2008.

[10] J. F. Conley, *et al.*, "Electrical Properties and Reliability of HfO_2 Deposited via ALD using $Hf(NO_3)_4$ Precursor," *IEEE International Integrated Reliability Workshop Final Report,* p. 108, 2002.

[11] M. Franta, *et al.*, "Microstructure of HfO_2 and $Hf_xSi_{1-x}O_y$ Dielectric Films Prepared on Si for Advanced CMOS Application," *International Conference on Advanced Semiconductor Devices and Microsystems,* p. 47, 2006.

[12] T. S. Boscke, *et al.*, "Tetragonal Phase Stabilization by Doping as an Enabler of Thermally Stable HfO_2 based MIM and MIS Capacitors for sub 50nm Deep Trench DRAM," *Technical Digest of IEDM,* p. 1, 2006.

[13] Y.-C. Yeo, *et al.*, "Metal-dielectric band alignment and its implications for metal gate complementary metal-oxide-semiconductor technology," *Journal of applied physics,* vol. 92, no. 12, p. 7266, 2002.

[14] H. Y. Yu, *et al.*, "Energy gap and band alignment for $(HfO_2)_x(Al_2O_3)_{1-x}$ on (100) Si," *Applied physics letters,* vol. 81, no. 2, p. 376, 2002.

[15] M.-Y. Ho, *et al.*, "Morphology and crystallization kinetics in HfO_2 thin films grown by atomic layer deposition," *Journal of applied physics,* vol. 93, no. 3, p. 1477, 2003.

[16] B. H. Lee, *et al.*, "Thermal stability and electrical characteristics of ultrathin hafnium oxide gate dielectric reoxidized with rapid thermal annealing," *Applied physics letters,* vol. 76, no. 14, p. 1926, 2000.

[17] M.-H. Cho, *et al.*, "Thermal stability and structural characteristics of HfO_2 films on Si (100) grown by atomic-layer deposition," *Applied physics letters,* vol. 81, no. 3, p. 472, 2002.

[18] G. He, *et al.*, "Effect of postdeposition annealing on the thermal stability and structural characteristics of sputtered HfO_2 films on Si (100)," *Surface Science,* vol. 576, no. 1-3, p. 67, 2005.

[19] J. H. Sim, *et al.*, "Effects of ALD HfO_2 thickness on charge trapping and mobility," *Microelectronic Engineering,* vol. 80, p. 218, 2005.

[20] F. Lime, *et al.*, "Carrier mobility in advanced CMOS devices with metal gate and HfO_2 gate dielectric," *Solid-State Electronics,* vol. 47, no. 10, p. 1617, 2003.

[21] M. A. Negara, *et al.*, "The influence of HfO_2 film thickness on the interface state density and low field mobility of *n*-channel HfO_2/TiN gate MOSFETs," *Microelectronic Engineering,* vol. 84, no. 9-10, p. 1874, 2007.

[22] S.-H. Lo, *et al.*, "Modeling and characterization of quantization, polysilicon depletion, and direct tunneling effects in MOSFETs with ultrathin oxides," *IBM Journal of Research Development,* vol. 43, no. 3, p. 327, 1999.

[23] H.-S. Jung, *et al.*, "Integration Friendly Dual Metal Gate Technology Using Dual Thickness Metal Inserted Poly-Si Stacks (DT-MIPS)," *Technical Digest of VLSI Symposium on Technology,* p. 196, 2007.

[24] S. K. Han, *et al.*, "Highly Manufacturable Single Metal Gate Process Using Ultra-Thin Metal Inserted Poly-Si Stack (UT-MIPS)," *Technical Digest of IEDM,* p. 1, 2006.

[25] E. Cartier, *et al.*, "*p*FET V_t control with HfO_2/TiN/poly-Si gate stack using a lateral oxygenation process," *Technical Digest of VLSI Symposium on Technology,* p. 42, 2009.

[26] E. Cartier, *et al.*, "Role of Oxygen Vacancies in V_{FB}/V_t stability of *p*FET metals on

HfO_2," *Technical Digest of VLSI Symposium on Technology,* p. 230, 2005.

[27] E. Cartier, *et al.*, "Systematic study of pFET V_t with Hf-based gate stacks with poly-Si and FUSI gate," *Technical Digest of VLSI Symposium on Technology,* p. 44, 2004.

[28] Y. Liang, *et al.*, "Interface dipole and effective work function of Re in Re/HfO_2/SiO_x/*n*-Si gate stack," *Applied Physics Letters,* vol. 88, pp. 072907-1, 2006.

[29] J. K. Schaeffer, *et al.*, "Contributions to the effective work function of platinum on hafnium dioxide," *Applied Physics Letters,* vol. 85, p. 1826, 2004.

[30] H. Takahashi, *et al.*, "Ti-capping technique as a breakthrough for achieving low threshold voltage, high mobility, and high reliability of *p*MOSFET with metal gate and high-k dielectriccs technologies," *Technical Digest of IEDM,* p. 1, 2009.

[31] K. Akiyama, *et al.*, "Roles of oxygen vacancy in HfO_2/ultra-thin SiO_2 gate stacks – Comprehensive understanding of V_{FB} roll-off -," *Technical Digest of VLSI Symposium on Technology,* p. 80, 2008.

[32] K. Akiyama, *et al.*, "V_{fb} Roll-off in HfO_2 Gate Stack after High Temperature Annealing Process - A Crucial Role of Out-diffused Oxygen from HfO_2 to Si -," *Technical Digest of VLSI Symposium on Technology,* p. 72, 2007.

[33] Y. Yamamoto, *et al.*, "Study of La-Induced Flat Band voltage Shift in Metal/$HfLaO_x$/SiO_2/Si Capacitors," *Japanese Journal of Applied Physics,* vol. 46, no. 11, p. 7251, 2007.

[34] K. Iwamoto, *et al.*, "Experimental evidence for the flatband voltage shift of high-k metal-oxide-semiconductor devices due to the dipole formation at the high-

k/SiO_2 interface," *Applied Physics Letters,* vol. 92, p. 132907, 2008.

[35] K. Tatsumura, *et al.*, "Intrinsic correlation between mobility reduction and V_t shift due to interface dipole modulation in HfSiON/SiO_2 stack by La or Al addition," *Technical Digest of IEDM,* p. 1, 2008.

[36] T. Ando, *et al.*, "Understanding mobility mechanism in extremely scaled HfO_2 (EOT 0.42nm) using remote interfacial layer scavenging technique and V_t-tuning dipoles with gate-first process," *Technical Digest of IEDM,* p. 423, 2009.

[37] P. Packan, *et al.*, "High Performance 32nm Logic Technology Featuring 2nd Generation High-k + Metal Gate Transistor," *Technical Digest of IEDM,* p. 659, 2009.

[38] S. Yamaguchi, *et al.*, "High performance dual metal gate CMOS with high mobility and low threshold voltage applicable to bulk CMOS technology," *Technical Digest of VLSI Symposium on Technology,* p. 152, 2006.

[39] A. Veloso, *et al.*, "Gate-last vs. gate-first technology for aggressively scaled EOT logic/RF CMOS," *Technical Digest of VLSI Symposium on Technology,* p. 34, 2011.

[40] C. L. Hinkle, *et al.*, "Dipole controlled Metal Gate with Hybrid Low Resistivity Cladding for Gate-Last CMOS with Low V_t," *Technical Digest of VLSI Symposium on Technology,* p. 183, 2010.

[41] C. M. Lai, *et al.*, "A Novel "Hybrid" High-k/Metal Gate Process for 28nm High Performance CMOSFETs," *Technical Digest of IEDM,* p. 655, 2009.

[42] C. S. Park, *et al.*, "Alternative approaches for high-k/metal gate CMPS: low temperature process (gate last) and SiGe channel," *Technical Digest of Symposium*

on VLSI-TSA, p. 80, 2010.

[43] S. Suthram, *et al.*, "High performance pMOSFETs using Si/$Si_{1-x}Ge_x$/Si quantum wells with High-k/metal gate stacks and additive uniaxial strain for 22 nm technology node," *Technical Digest of IEDM,* p. 727, 2007.

[44] H. R. Harris, *et al.*, "Band-Engineered Low PMOS V_T with High-K/Metal Gates Featured in a Dual Channel CMOS Integration Scheme," *Technical Digest of VLSI Symposium on Technology,* p. 154, 2007.

[45] J. Huang, *et al.*, "Mechanisms Limiting EOT Scaling and Gate Leakage Currents of High-k/Metal Gate Stacks Directly on SiGe and a Method to Enable sub-1nm EOT," *Technical Digest of VLSI Symposium on Technology,* p. 82, 2008.

[46] S. Krishnan, *et al.*, "A manufacturable dual channel (Si and SiGe) high-K metal gate CMOS technology with multiple oxides for high performance and low power applications," *Technical Digest of IEDM,* p. 634, 2011.

[47] L. Witters, *et al.*, "8Å T_{inv} Gate-First Dual Channel Technology Achieving Low-V_t High Performance CMOS," *Technical Digest of VLSI Symposium on Technology,* p. 181, 2010.

[48] D. J. Paul, "Si/SiGe heterostructures: from material and physics to devices and circuits," *Semiconductor Science and Technology,* vol. 19, p. R75, 2004.

[49] J. Weber, "Near-band-gap photoluminescence of Si-Ge alloys," *Physical Review B,* vol. 40, no. 8, p. 5683, 1989.

[50] T. Ando, *et al.*, "Origins of effective work function roll-off behavior for high-k last replacement metal gate stacks," *IEEE Electron Device Letter,* vol. 34, no. 6, p. 729, 2013.

[51] L. Gomez, *et al.*, "Enhanced Hole Mobility in High Ge Content Asymmetrically Strained-SiGe *p*-MOSFETs," *IEEE Electron Device Letters,* vol. 31, no. 8, p. 782, 2010.

[52] L. Gomez, *et al.*, "Hole Velocity Enhancement in Sub-100 nm Gate Length Strained-SiGe," *IEEE International SOI Conference Proceedings,* p. 163, 2008.

[53] L. Hutin, *et al.*, "Dual Strained Channel Co-Integration into CMOS, RO and SRAM cells on FDSOI down to 17nm Gate Length," *Technical Digest of IEDM,* p. 253, 2010.

[54] M. A. Olayiwola, *et al.*, "Performance Enhancements in Scaled Strained-SiGe *p*MOSFETs with $HfSiO_x$/TiSiN Gate Stacks," *IEEE Transactions on Electron Devices,* vol. 56, no. 10, p. 2277, 2009.

[55] J. Oh, *et al.*, "Additive Mobility Enhancement and Off-State Current Reduction in SiGe Channel *p*MOSFETs with Optimized Si Cap and High-k Metal Gate Stacks," *Technical Digest of Symposium on VLSI-TSA,* p. 22, 2009.

[56] D. C. Gilmer, *et al.*, "Strained SiGe Channels for Band-Edge PMOS Threshold Voltages With Metal Gates and High-k Dielectrics," *IEEE Transaction on Electron Devices,* vol. 57, no. 4, p. 898, 2010.

[57] P. W. Liu, *et al.*, "Superior Current Enhancement in SiGe Channel *p*-MOSFETs Fabricated on (110) Surface," *Technical Digest of VLSI Symposium on Technology,* p. 1, 2006.

[58] L. Hutin, *et al.*, "20nm Gate Length Trigate *p*FETs on Strained SGOI for High Performance CMOS," *Technical Digest of VLSI Symposium on Technology,* p. 37, 2010.

[59] T. Ando, *et al.*, "Simple Gate Metal Anneal (SIGMA) Stack for FinFET Replacement Metal Gate toward 14nm and beyond," *Technical Digest of VLSI Symposium on Technology,* p. 1, 2014.

[60] S. Ohmi, *et al.*, "Investigation of PDA process to improve electrical characteristics of HfO_xN_y High-k dielectric formed by ECR plasma oxidation of HfN," *Technical Digest of International Symposium on Semiconductor Manufacturing,* p. 1, 2007.

[61] M. Takahashi, *et al.*, "Gate-First Processed FUSI/HfO_2/$HfSiO_x$/Si MOSFETs with EOT=0.5 nm – Interfacial Layer Formation by Cycle-by-Cycle Deposition and annealing," *Technical Digest of IEDM,* p. 523, 2007.

[62] K. Shiraishi, *et al.*, "Physics in Fermi Level Pinning at the Poly-Si/Hf-based High-k Oxide Interface," *Technical Digest of VLSI Symposium on Technology,* p. 108, 2004.

[63] M. Cho, *et al.*, "Insight Into N/PBTI Mechanisms in Sub-1-nm-EOT Devices," *IEEE Transactions on Electron Devices,* vol. 59, no. 8, p. 2042, 2012.

[64] M. Togo, *et al.*, "Multiple Workfunction High Performance FinFETs for Ultra-low Voltage Operation," *Technical Digest of VLSI Symposium on Technology,* p. 81, 2018.

[65] D. Ha, "The Evolution of Logic Transistors toward Low Power and High Performance IOT Applications," *IEDM Tutorial,* 2017.

[66] C. Ortolland, *et al.*, "Optimized Ultra-Low Thermal Budget Process Flow for Advanced High-K / Metal Gate First CMOS Using Laser-Annealing Technology," *Technical Digest of VLSI Symposium on Technology,* p. 38, 2009.

[67] H.-Y. Yu, *et al.*, "Integrate LaO_x-capping layer into metal gated CMOS devices

using a gate-first approach for sub-45nm technology node and the device reliability thereof," *Technical Digest of 9th International Conference on Solid State and Integrated Circuit Technology,* p. 1260, 2008.

[68] C. Ortolland, *et al*., "Laser-annealed junctions with advanced CMOS gate stacks for 32nm Node: Perspectives on device performance and manufacturability," *Technical Digest of VLSI Symposium on Technology,* p. 186, 2008.

[69] G. Eneman, *et al*., "High-Mobility $Si_{1-x}Ge_x$-Channel PFETs: Layout Dependence and Enhanced Scalability, Demonstrating 90% Performance Boost at Narrow Width", Technical Digest of VLSI Symposium," *Technical Digest of VLSI Symposium on Technology,* p. 41, 2010.

[70] G. Hellings, *et al*., "Implant-Free SiGe Quantum Well *p*FET: A novel, highly scalable and low thermal budget device, featuring raised source/drain and high-mobility channel," *Technical Digest of IEDM,* p. 241, 2010.

[71] J. Mitard, *et al*., "1mA/µm-I_{ON} Strained SiGe45%-IFQW *p*FETs With Raised and Embedded S/D," *Technical Digest of VLSI Symposium on Technology,* p. 134, 2011.

[72] T. Ghani, *et al*., "A 90nm High Volume Manufacturing Logic Technology Featuring Novel 45nm Gate Length Strained Silicon CMOS Transistors," *Technical Digest of IEDM,* p. 978, 2003.

[73] M. D. Giles, *et al*., "Understanding Stress Enhanced Performance in Intel 90nm CMOS Technology," *Technical Digest of VLSI Symposium on Technology,* p. 118, 2004.

[74] M. M. Heyns, *et al*., "Cost-effective cleaning and high-quality thin gate oxides,"

IBM Journal of Research and Development, vol. 43, no. 3, p. 339, 1999.

[75] P. S. Lee, *et al.*, "New salicidation technology with Ni (Pt) alloy for MOSFETs," *IEEE Electron Device Letters,* vol. 22, no. 12, p. 568, 2001.

[76] Z. Zhang, *et al.*, "Ultra Low Contact Resistivities for CMOS Beyond 10-nm Node," *IEEE Electron Device Letters,* vol. 34, no. 6, p. 723, 2013.

[77] J. P. Dismukes, *et al.*, "Lattice parameter and density in germanium-silicon alloys," *Journal of Physical Chemistry,* vol. 68, no. 10, p. 3021, 1964.

[78] L. Vergard, "Die Konstitution der Mischkristalle und die Raumfullung der Atome," *Zeitschrift fur Physik,* vol. 5, no. 1, p. 17, 1921.

[79] V. T. Bublik, "Calculation of the Binding Energy of Ge-Si Solid Solution," *Phys. Status Solidi,* vol. 65, no. 2, p. K79, 1974.

[80] J. W. Matthews, *et al.*, "Defects in Epitaxial Multilayers," *Journal of Crystal Growth,* vol. 32, no. 2, p. 265, 1976.

[81] D. C. Houghton, "Strain relaxation kinetics in $Si_{1-x}Ge_x$/Si heterostructures," *Journal of Applied Physics,* vol. 70, no. 4, p. 2136, 1991.

[82] S.-H. Lee, *et al.*, "Vth Variation and Strain Control of High Ge% Thin SiGe Channels by Milisecond Anneal Realizing High Performance pMOSFET beyond 16nm node," *Technical Digest of VLSI Syposium on Technology,* p. 74, 2009.

[83] R. People, "Indirect band gap of coherently strained Ge_xSi_{1-x} bulk alloys on <001> silicon substrates," *Physical Review B,* vol. 33, no. 2, p. 1405, 1986.

[84] A. Shima, *et al.*, "Ultra-shallow junction formation by non-melt laser spike annealing for 50-nm gate CMOS," *Technical Digest of VLSI Symposium on Technology,* p. 174, 2004.

[85] M. V. Fischetti, *et al*., "Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys," *Journal of Applied Physics,* vol. 80, p. 2234, 1996.

[86] J. Mitard, *et al*., "Impact of EOT scaling down to 0.85nm on 70nm Ge-pFETs technology with STI," *Technical Digest of VLSI Symposium on Technology,* p. 82, 2009.

[87] Y. Kamata, *et al*., "High-k/Ge *p*- and *n*-MISFETs with Strontium Germanide Interlater for EOT Scalable CMIS Application," *Technical Digest of VLSI Symposium on Technology,* p. 212, 2010.

[88] R. Zhang, *et al*., "Aggressive EOT scaling of Ge *p*MOSFETs With $HfO_2/AlO_x/GeO_x$ Gate-Stacks Fabricated by Ozone Postoxidation," *IEEE Electron Devicee Letters,* vol. 37, no. 7, p. 831, 2016.

[89] G. Groeseneken, *et al*., "A Reliable Approach to Charge-Pumping Measurements in MOS Transistors," *IEEE Transactions on Electron Devices,* Vols. ED-31, no. 1, p. 42, 1984.

[90] G. Groeseneken, *et al*., "Charge Pumping of Single Interface Traps in Submicron MOSFETs," *Technical Digest of ESSDERC,* p. 609, 1994.

[91] K. Choi, *et al*., "Extremely Scaled Gate-First High-k/Metal Gate Stack with EOT of 0.55 nm Using Novel Interfacial Layer Scavenging Techniques for 22nm Technology Node and Beyond," *Technical Digest of VLSI Symposium on Technology,* p. 138, 2009.

[92] K. J. Hubbard, *et al*., "Thermodynamic stability of binary oxides in contact with silicon," *Journal of material research,* vol. 11, no. 11, p. 2757, 1996.

[93] A. Sagara, *et al.*, "Thermal Behavior of Residual Defects in Low-Dose Arsenic- and Boron-Implanted Silicon After High-Temperature Rapid Thermal Annealing," *IEEE Transactions on Semiconductor Manufacturing,* vol. 28, no. 1, p. 92, 2015.

[94] J. Krugener, "Electrical and Structural Analysis of Crystal Defects After High-Temperature Rapid Thermal Annealing," *IEEE Journal of Photovoltaics,* vol. 5, no. 1, p. 166, 2015.

[95] S. Severi, *et al.*, "Optimization of Sub-Melt Laser Anneal: Performance and Reliability," *Technical Digest of IEDM,* p. 1, 2006.

[96] E. Rosseel, *et al.*, "Impact of laser anneal thermal budget on the quality of thin SiGe channels with a high Ge content," *18th International Conference on Advanced Thermal Processing of Semiconductors (RTP),* p. 62, 2010.

[97] C. S. Smith, "Piezoresistance Effect in Germanium and Silicon," *Physical Review,* vol. 94, no. 1, p. 42, 1954.

[98] S. E. Thompson, *et al.*, "A 90-nm Logic Technology Featuring Strained-Silicon," *IEEE Transactions on Electron Devices,* vol. 51, no. 11, p. 1790, 2004.

[99] J. W. Pan, *et al.*, "Mobility and Strain Effects on <110>/(110) SiGe channel *p*MOSFETs for High Current Enhancement," *Technical Digest of IEDM,* p. 1, 2006.

[100] S. Yamaguchi, *et al.*, "High performance $Si_{.45}Ge_{.55}$ Implant Free Quantum Well FET featuring low temperature process, eSiGe stressor and transversal strain relaxation," *Technical Digest of IEDM,* p. 829, 2011.

[101] D. Fleury, *et al.*, "New Y-function-based methodology for accurate extraction of electrical parameters on nano-scaled MOSFETs," *IEEE Conference on*

Microelectronic Test Structures, p. 160, 2008.

[102] J.-S. Goo, *et al*., "Scalability of Strained-Si *n*MOSFETs Down to 25 nm Gate Length," *IEEE Electron Device Letters,* vol. 24, no. 5, p. 351, 2003.

[103] S. Mayuzumi, *et al*., "Channel-Stress Study on Gate-Size Effects for Damascene-Gate *p*MOSFETs with Top-Cut Compressive Stress Liner and eSiGe," *Technical Digest of VLSI Symposium on Technology,* p. 126, 2008.

[104] G. Eneman, *et al*., "Layout Scaling of $Si_{1-x}Ge_x$-Channel *p*FETs," *IEEE Transactions on Electron Devices,* vol. 58, no. 8, p. 2544, 2011.

[105] O. Weber, *et al*., "Examination of additive mobility enhancements for uniaxial stress combined with biaxial strained Si, biaxial strained SiGe and Ge Channel MOSFETs," *Technical Digest of IEDM,* p. 719, 2007.

[106] G. Eneman, *et al*., "Stress techniques in advanced transistor architectures: bulk FinFETs and implant-free quantum well transistors," *ECS Transactions,* vol. 45, no. 3, p. 235, 2012.

[107] R. Gehres, *et al*., "High Volum Manufacturing Ramp In 90nm Dual Stress Liner Technology," *The 17th Annual SEMI/IEEE ASMC 2006 Conference,* p. 411, 2006.

[108] F. Sato, *et al*., "Process and Local Layout Effect interaction on a high performance planar 20nm CMOS," *Technical Digest of VLSI Symposium on Technology,* p. 116, 2013.

[109] M. Hamaguchi, *et al*., "New Layout Dependency in High-K/Metal Gate MOSFETs," *Technical Digest of IEDM,* p. 579, 2011.

[110] C. Ndiaye, *et al*., "Layout Dependent Effect: Impact on device performance and reliability in recent CMOS nodes," *IEEE International Integrated Reliability*

Workshop, p. 24, 2016.

[111] A. Khakifirooz, *et al*., "High-Performance Partially Depleted SOI PFETs With In Situ Doped SiGe Raised Source/Drain and Implant-Free Extension," *IEEE Transactions on Electron Devices,* vol. 32, no. 3, p. 267, 2011.

[112] A. Khakifirooz, *et al*., "Strain Engineered Extremely Thin SOI (ETSOI) for High-Performance CMOS," *Technical Digest of VLSI Symposium on Technology,* p. 117, 2012.

[113] S. Severi, *et al*., "Accurate Channel Length Extraction by Split C-V Measurements on Short-Channel MOSFETs," *IEEE Electron Device Letters,* vol. 27, no. 7, p. 615, 2006.

[114] K. Rajendran, *et al*., "Measurement and Simulation of Boron Diffusion in Strained $Si_{1-x}Ge_x$ Epitaxial Layers," *IEEE Transactions on Electron Devices,* vol. 48, no. 9, p. 2022, 2001.

[115] K. Rajendran, *et al*., "Simulation of boron diffusion in strained $Si_{1-x}Ge_x$ epitaxial layers," *International Conference on Simulation Semiconductor Processes and Devices,* p. 206, 2000.

[116] S. Kamohara, *et al*., "TED model including the dissolution of extended defects," *International Conference on Simulation of Semiconductor Processes and Devices,* p. 85, 1996.

[117] J. Mitard, *et al*., "85 nm-wide 1.5 mA/µm-I_{ON} IFQW SiGe-*p*FET: raised vs embedded $Si_{0.75}Ge_{0.25}$ S/D benchmarking and in-depth hole transport study," *Technical Digest of VLSI Symposium on Technology,* p. 163, 2012.

[118] J. Huang, *et al*., "Gate first high-k/metal gate stacks with zero SiO_x interface

achieving EOT = 0.59 nm for 16 nm application," *Technical Digest of VLSI Symposium on Technology,* p. 34, 2009.

[119] F. Arnaud, *et al*., "32nm General Purpose Bulk CMOS Technology for High Performance Applications at Low Voltage," *Technical Digest of IEDM,* p. 1, 2008.

[120] B. Davari, *et al*., "A new planarization technique, using a combination of RIE and chemical mechanical polish (CMP)," *Technical Digest of IEDM,* p. 61, 1989.

[121] A. Hikavyy, *et al*., "An investigation of growth and properties of Si capping layers used in advanced SiGe/Ge based pFET transistors," *Technical Digest of International Silicon-Germanium Technology and Device Meeting,* p. 1, 2012.

[122] B. H. Lee, *et al*., "Gate stack technology for nanoscale devices," *Materials Today,* vol. 9, no. 6, p. 32, 2006.

List of Publications and Presentations

International Journal (peer reviewed)

1) S. Yamaguchi, L. Witters, J. Mitard, G. Eneman, G. Hellings, A. Hikavyy, R. Loo, and N. Horiguchi, "Scalability Comparison between Raised- and Embedded-SiGe Source/Drain Structures for $Si_{0.55}Ge_{0.45}$ Implant Free Quantum Well *p*FET," *Microelectronics Reliability*, vol. 83, p. 157, 2018.
2) Hsien-Ching Lo, Jianwei Peng, Chloe Yang, Suresh Uppal, Yi Qi, Hui Zhan, Yan Ping Shen, Xiaobo Chen, Jianghu Yang, Baofu Zhu, Shashidhar Shintri, Shimpei Yamaguchi, Talapady Bhat, Wei Hong, Yong Jun Shi, Suresh Regonda, Dongil Choi, Owen Hu, Manoj Joshi, and Srikanth Samavedam, "Trade-Off between Gate Oxide Integrity and Transistor Performance for FinFET Technology," *ECS Journal of Solid State Science and Technology*, vol. 6, no. 8, N137-N141, 2017.
3) Shimpei Yamaguchi, Zeynel Bayindir, Xiaoli He, Suresh Uppal, Purushothaman Srinivasan, Chloe Yong, Dongil Choi, Manoj Joshi, Hyuck Soo Yang, Owen Hu, Srikanth Samavedam, and D. K. Sohn, "Effective Work-Function Control Technique Applicable to p-type FinFET High-k/Metal Gate Devices," *Microelectronics Reliability*, vol. 74, p. 80, 2017.
4) T. Ohashi, K. Suda, S. Ishihara, N. Sawamoto, S. Yamaguchi, K. Matsuura, K. Kakushima, N. Sugii, A. Nishiyama, Y. Kataoka, K. Natori, K. Tsutsui, H. Iwai, A. Ogura, and H. Wakabayashi, "Multi-layered MoS_2 film formed by high-temperature sputtering for enhancement-mode nMOSFETs," *Japanese Jourrnal of Applied Physics*, vol. 54, no. 4, 04DN08, 2015.
5) Shimpei Yamaguchi, Liesbeth Johanna Witters, Jérôme Mitard, Geert Eneman, Geert Hellings, Andriy Hikavyy, Roger Loo, and Naoto Horiguchi, "High-Performance

$Si_{0.45}Ge_{0.55}$ Implant-Free Quantum Well pFET With Enhanced Mobility by Low-Temperature Process and Transverse Strain Relaxation," *IEEE Transaction on Electron Devices*, vol. 61, no. 12, p. 3985-3990, 2014.

6) Geert Eneman, Shinpei Yamaguchi, Claude Ortolland, Shinji Takeoka, Masaharu Kobayashi, Liesbeth Witters, Andriy Hikavyy, Jérôme Mitard, Roger Loo, and Thomas Hoffmann, "Layout Scaling of $Si_{1-x}Ge_x$-Channel *p*FETs," *IEEE Transactions on Electron Devices*, vol. 58, no. 8, p. 2544-2550, 2011.

7) Takashi Ando, Tomoyuki Hirano, Kaori Tai, Shinpei Yamaguchi, Shinichi Yoshida, Hayato Iwamoto, Shingo Kadomura, and Heiji Watanabe, "Band-Edge High-Performance Metal-Gate/High-k *n*MOSFET Using HfSi/HfO_2 Stack," *IEEE Transaction on Electron Devices*, vol. 56, no. 12, p. 3223-3227, 2009.

8) Satoru Mayuzumi, Shinya Yamakawa, Yasushi Tateshita, Tomoyuki Hirano, Masashi Nakata, Shinpei Yamaguchi, Kaori Tai, Hitoshi Wakabayashi, Masanori Tsukamoto, and Naoki Nagashima, "High-Performance Metal/High- k *n*- and *p*-MOSFETs With Top-Cut Dual Stress Liners Using Gate-Last Damascene Process on (100) Substrates," *IEEE Transaction on Electron Devices*, vol. 56, no. 4, p. 620-626, 2009.

International Conferences (peer-reviewed)

9) M. Togo, R. Asra, P. Balasubramaniam, X. Zhang, H. Yu, S. Yamaguchi, E. Geiss, H. S. Yang, B. Cohen, H.-C. Lo, O. Hu, H. Lazar, O. Kwon, D. Brunett, J. Versaggi, E. Banghart, M. K. Hassan, E. Bazizi, L. Pantisano, J. G. Lee, S. B. Samavedam, and D. K. Sohn, "Multiple Workfunction High Performance FinFETs for Ultra-low Voltage Operation," *Technical Digest of VLSI Symposium on Technology*, p. 81, 2018.

10) Y. Okada, S. Yamaguchi, T. Ohashi, I. Muneta, K. Kakushima, K. Tsutsui, and H. Wakabayashi, "Resistivity Reduction of Low-Carrier-Density Sputtered-MoS_2 Film using Fluorine Gas," *Extended Abstract the 17th International Workshop on Junction Technology*, S4-3, 2017.

11) J. Singh, A. Bousquet, J. Ciavatti, K. Sundaram, J. S. Wong, K. W. Chew, A. Bandyopadhyay, S. Li, A. Bellaouar, S. M. Pandey, B. Zhu, A. Martin, C. Kyono, J.-S. Goo, H. S. Yang, A. Mehta, X. Zhang, O. Hu, S. Mahajan, E. Geiss, S. Yamaguchi S. Mittal, R. Asra, P. Balasubramaniam, J. Watts, D. Harame, R. M. Todi, S. B. Samavedam, and D. K. Sohn, "14nm FinFET Technology for Analog and RF Applications," *Technical Digest of VLSI Symposium on Technology*, p. 140, 2017.

12) Xiaoli He, Dina Triyoso, Suresh Uppal, Bianzhu Fu, Xing Zhang, Shimpei Yamaguchi, Chloe Yong, Bingwu Liu, Manoj Joshi, and Srikanth Samavedam, "Room Temperature Aging Effect Improvement for Device Stability and Manufacturability of FinFET Technologies", *ECS Transactions*, vol. 80, no. 1, p. 373, 2017.

13) S. Ishihara, K. Suda, Y. Hibino, N. Sawamoto, T. Ohashi, S. Yamaguchi, K. Matsuura, H. Machida, M. Ishikawa, H. Sudoh, H. Wakabayashi, and A. Ogura, "Improving Crystalline Quality of Sputtering Deposited MoS_2 Thin Film by Post-Deposition

Sulfurization Annealing Using (t-$C_4H_9)_2S_2$," *International Conference on Solid State Devices and Materials*, PS-8-13, 2015.

14) S. Ishihara, K. Suda, Y. Hibino, N. Sawamoto, T. Ohashi, S. Yamaguchi, K. Matsuura, H. Machida, M. Ishikawa, H. Sudoh, H. Wakabayashi, and A. Ogura, "Evaluation of Sputtering Deposited 2-Dimensional MoS_2 Film by Raman Spectroscopy," *MRS Proceedings* 1781, p. 11-16, 2015.

15) T. Ohashi, K. Suda, S. Ishihara, N. Sawamoto, S. Yamaguchi, K. Matsuura, K. Kakushima, N. Sugii, A. Nishiyama, Y. Kataoka, K. Natori, K. Tsutsui, H. Iwai, A. Ogura, and H. Wakabayashi, "Multi-Layered MoS_2 Thin Film Formed by High-Temperature Sputtering for Enhancement-Mode *n*MOSFETs," *International Conference on Solid State Devices and Materials*, P-9-1, 2014.

16) G. Eneman, L. Witters, N. Collaert, J. Mitard, G. Hellings, S. Yamaguchi, A. De Keersgieter, A Hikavyy, B. Vincent, P. Favia, H. Bender, A. Veloso, T. Chiarella, M. Togo, R. Loo, K. De Meyer, A. Mercha, N. Horiguchi, and A. Thean, "Stress Techniques in Advanced Transistor Architectures: Bulk FinFETs and Implant-Free Quantum Well Transistors," *ECS Transaction,* vol. 45, no. 3, p. 235-246, 2012.

17) G. Eneman, G. Hellings, J. Mitard, L. Witters, S. Yamaguchi, M. Garcia Bardon, P. Christie, C. Ortolland, A. Hikavyy, P. Favia, M. Bargallo Gonzalez, E. Simoen, F. Crupi, M. Kobayashi, J. Franco, S. Takeoka, R. Krom, H. Bender, R. Loo, C. Claeys, K. De Meyer, and T. Hoffmann," $Si_{1-x}Ge_x$-Channel PFETs: Scalability, Layout Considerations and Compatibility with Other Stress Techniques," *ECS Transaction*, vol. 35, no. 3, p. 493-503, 2011.

18) S. Yamaguchi, L. Witters, J. Mitard, G. Eneman, G. Hellings, M. Fukuda, A. Hikavyy, R. Loo, A. Veloso, Y. Crabbe, E. Rohr, O. Richard, H. Bender, S. Takeoka, Wei-E

Wang, L.A. Ragnarsson, and N. Horiguchi, "High Performance $Si_{.45}Ge_{.55}$ Implant Free Quantum Well FET Featuring Low Temperature Process, eSiGe Stressor and Transversal Strain Relaxation," *Technical Digest of IEDM*, p. 829-832, 2011.

19) L. Witters, J. Mitard, A. Veloso, A. Hikavyy, J. Franco, T. Kauerauf, M. Cho, T. Schram, F. Sebai, S. Yamaguchi, S. Takeoka, M. Fukuda, W.-E. Wang, B. Duriez, G. Eneman, R. Loo, K. Kellens, H. Tielens, P. Favia, E. Rohr, G. Hellings, H. Bender, P. Roussel, Y. Crabbe, S. Brus, G. Mannaert, S. Kubicek, K. Devriendt, K. De Meyer, L. -Å. Ragnarsson, A. Steegen, and N. Horiguchi, "Dual-channel technology with cap-free single metal gate for high performance CMOS in gate-first and gate-last integration," *Technical Digest of IEDM*, 28.6, 2011.

20) A. Veloso, L. -Å. Ragnarsson, M. J. Cho, K. Devriendt, K. Kellens, F. Sebaai, S. Suhard, S. Brus, Y. Crabbe, T. Schram, E. Röhr, V. Paraschiv, G. Eneman, T. Kauerauf, M. Dehan, S. -H. Hong, S. Yamaguchi, S. Takeoka, Y. Higuchi, H. Tielens, A. Van Ammel, P. Favia, H. Bender, A. Franquet, T. Conard, X. Li, K. -L. Pey, H. Struyf, P. Mertens, P. P. Absil, N. Horiguchi, and T. Hoffmann, "Gate-last vs. gate-first technology for aggressively scaled EOT logic/RF CMOS," *Technical Digest of VLSI Symposium on Technology*, p. 34-35, 2011.

21) J. Mitard, L. Witters, G. Hellings, R. Krom, J. Franco, G. Eneman, A. Hikavyy, B. Vincent, R. Loo, P. Favia, H. Dekkers, E. Altamirano Sanchez, A. Vanderheyden, D. Vanhaeren, P. Eyben, S. Takeoka, S. Yamaguchi, M. J. H. Van Dal, W. -E Wang, S. -H Hong, W. Vandervorst, K. De Meyer, S. Biesemans, P. Absil, N. Horiguchi, and T. Hoffmann, "1mA/um-I_{ON} strained SiGe45%-IFQW *p*FETs with raised and embedded S/D," *Technical Digest of VLSI Symposium on Technology,* p. 134-135, 2011.

22) J. Mitard, L. Witters, M. Garcia Bardon, P. Christie, J. Franco, A. Mercha, P.

Magnone, M. Alioto, F. Crupi, L. -Å. Ragnarsson, A. Hikavyy, B. Vincent, T. Chiarella, R. Loo, J. Tseng, S. Yamaguchi, S. Takeoka, W-E. Wang, P. Absil, and T. Hoffmann, "High-mobility 0.85nm-EOT $Si_{0.45}Ge_{0.55}$-*p*FETs: Delivering high performance at scaled V_{DD}," *Technical Digest of IEDM*, 10.6, 2010.

23) G. Eneman, S. Yamaguchi, C. Ortolland, S. Takeoka, L. Witters, T. Chiarella, P. Favia, A. Hikavyy, J. Mitard, M. Kobayashi, R. Krom, H. Bender, J. Tseng, W. -E. Wang, W. Vandervorst, R. Loo, P. P. Absil, S. Biesemans, and T. Hoffmann, "High-mobility $Si_{1-x}Ge_x$-channel PFETs: Layout dependence and enhanced scalability, demonstrating 90% performance boost at narrow widths," *Technical Digest of VLSI Symposium on Technology*, p. 41-42, 2010.

24) L. Witters, S. Takeoka, S. Yamaguchi, A. Hikavyy, D. Shamiryan, M. Cho, T. Chiarella, L. -Å. Ragnarsson, R. Loo, C. Kerner, Y. Crabbe, J. Franco, J. Tseng, W. E. Wang, E. Rohr, T. Schram, O. Richard, H. Bender, S. Biesemans, P. Absil, and T. Hoffmann, "8Å T_{inv} gate-first dual channel technology achieving low-V_t high performance CMOS," *Technical Digest of VLSI Symposium on Technology*, p. 181-182, 2010.

25) J. Wang, Y. Tateshita, S. Yamakawa, K. Nagano, T. Hirano, Y. Kikuchi, Y. Miyanami, S. Yamaguchi, K. Tai, R. Yamamoto, S. Kanda, T. Kimura, K. Kugimiya, M. Tsukamoto, H. Wakabayashi, Y. Tagawa, H. Iwamoto, T. Ohno, M. Saito, S. Kadomura, and N. Nagashima, "Novel Channel-Stress Enhancement Technology with eSiGe S/D and Recessed Channel on Damascene Gate Process," *Technical Digest of VLSI Symposium on Technology*, p. 46-47, 2007.

26) S. Mayuzumi, J. Wang, S. Yamakawa, Y. Tateshita, T. Hirano, M. Nakata, S. Yamaguchi, Y. Yamamoto, Y. Miyanami, I. Oshiyama, K. Tanaka, K. Tai, K. Ogawa,

K. Kugimiya, Y. Nagahama, Y. Hagimoto, R. Yamamoto, S. Kanda, K. Nagano, H. Wakabayashi, Y. Tagawa, M. Tsukamoto, H. Iwamoto, M. Saito, S. Kadomura, and N. Nagashima, "Extreme High-Performance n- and p-MOSFETs Boosted by Dual-Metal/High-k Gate Damascene Process using Top-Cut Dual Stress Liners on (100) Substrates," *Technical Digest of IEDM*, p. 293-296, 2007.

27) T. Ando, T. Hirano, K. Tai, S. Yamaguchi, K. Tanaka, I. Oshiyama, M. Nakata, K. Watanabe, R. Yamamoto, S. Kanda, Y. Tateshita, H. Wakabayashi, Y. Tagawa, M. Tsukamoto, H. Iwamoto, M. Saito, S. Toyoda, H. Kumigashira, M. Oshima, N. Nagashima, and S. Kadomura, "High Performance and High Reliability Dual Metal CMOS Gate Stacks Using Novel High-k Bi-layer Control Technique," *Technical Digest of Symposium on VLSI-TSA* p. 1-2, 2007.

28) Y. Tateshita, J. Wang, K. Nagano, T. Hirano, Y. Miyanami, T. Ikuta, T. Kataoka, Y. Kikuchi, S. Yamaguchi, T. Ando, K. Tai, R. Matsumoto, S. Fujita, C. Yamane, R. Yamamoto, S. Kanda, K. Kugimiya, T. Kimura, T. Ohchi, Y. Yamamoto, Y. Nagahama, Y. Hagimoto, H. Wakabayashi, Y. Tagawa, M. Tsukamoto, H. Iwamoto, M. Saito, S. Kadomura, and N. Nagashima, "High-Performance and Low-Power CMOS Device Technologies Featuring Metal/High-k Gate Stacks with Uniaxial Strained Silicon Channels on (100) and (110) Substrates," *Technical Digest of IEDM*, p. 1-4, 2006.

29) K. Tai, T. Hirano, S. Yamaguchi, T. Ando, S. Hiyama, J. Wang, Y. Nagahama, T. Kato, M. Yamanaka, S. Terauchi, S. Kanda, R. Yamamoto, Y. Tateshita, Y. Tagawa, H. Iwamoto, M. Saito, N. Nagashima, and S. Kadomura, "High Performance pMOSFET with ALD-TiN/HfO_2 Gate Stack on (110) Substrate by Low Temperature Process," *ESSDERC*, p. 121-124, 2006.

30) S. Yamaguchi, K. Tai, T. Hirano, T. Ando, S. Hiyama, J. Wang, Y. Hagimoto, Y. Nagahama, T. Kato, K. Nagano, M. Yamanaka, S. Terauchi, S. Kanda, R. Yamamoto, Y. Tateshita, Y. Tagawa, H. Iwamoto, M. Saito, N. Nagashima, and S. Kadomura "High Performance Dual Metal Gate CMOS with High Mobility and Low Threshold Voltage Applicable to Bulk CMOS Technology," *Technical Digest of Symposium on VLSI Technology,* p. 192-193, 2006.

31) T. Ando, T. Hirano, K. Tai, S. Yamaguchi, T. Kato, Y. Hagimoto, K. Watanabe, R. Yamamoto, S. Kanda, K. Nagano, S. Terauchi, Y. Tateshita, Y. Tagawa, M. Saito, H. Iwamoto, S. Yoshida, H. Watanabe, N. Nagashima, and S. Kadomura, "Sub-1nm EOT $HfSi_x/HfO_2$ Gate Stack Using Novel Si Extrusion Process for High Performance Application," *Technical Digest of Symposium on VLSI Technology*, p. 166-167, 2006.

32) T. Hirano, T. Ando, K. Tai, S. Yamaguchi, T. Kato, S. Hiyama, Y. Hagimoto, S. Takesako, N. Yamagishi, K. Watanabe, R. Yamamoto, S. Kanda, S. Terauchi, Y. Tateshita, Y. Tagawa, H. Iwamoto, M. Saito, S. Kadomura, and N. Nagashima, "High performance nMOSFET with $HfSi_x/HfO_2$ gate stack by low temperature process," *Technical Digest of IEDM,* p. 890-893, 2005.

International/Domestic Conferences and Symposiums (without review)

33) T. Ohashi, K. Suda, S. Ishihara, N. Sawamoto, S. Yamaguchi, K. Matsuura, K. Kakushima, N. Sugii, A. Nishiyama, Y. Kataoka, K. Natori, K. Tsutsui, H. Iwai, A. Ogura, and H. Wakabayashi, “MoS_2 film thinning on high-temperature sputtering for enhancement-mode nMOSFETs”, IEEE EDS Mini-Colloquium: WIMNACT 45, Suzukakedai Campus, Tokyo Institute of Technology, Japan.

34) T. Ohashi, S. Yamaguchi, K. Matsuura and H. Wakabayashi, “Sputtered MoS_2 Film for Future High-Performance Nanoelectronic Devices,” The 7th Thailand-Japan International Academic Conference, Material Engineering and Technology 2, 2014.

35) 松浦賢太朗，大橋匠，山口晋平，須田耕平，石原聖也，澤本直美，角嶋邦之，杉井信之，西山彰，片岡好則，名取研二，筒井一生，岩井洋，小椋厚志，若林整，“大面積 MoS_2 膜形成に向けた Mo の流加プロセスの検討，”第 62 回応用物理学会春季学術講演会，13p-A23-1，2015.

36) 大橋匠，山口晋平，松浦賢太朗，須田耕平，石原聖也，澤本直美，角嶋邦之，杉井信之，西山彰，片岡好則，名取研二，筒井一生，岩井洋，小椋厚志，若林整，“スパッタ堆積 MoS_2 膜の下地平坦化による電気特性向上，”第 62 回応用物理学会春季学術講演会，12p-A29-10，2015.

37) 石原聖也，須田耕平，澤本直美，大橋匠，山口晋平，松浦賢太朗，若林整，小椋厚志，“ラマン分光法による高温スパッタ堆積 MoS_2 膜の評価，”第 75 回応用物理学会秋季学術講演会，18p-B3-5，2014.

38) 大橋匠，山口晋平，松浦賢太朗，須田耕平，石原聖也，澤本直美，角嶋邦之，杉井信之，西山彰，片岡好則，名取研二，筒井一生，岩井洋，小椋厚志，若林整，“高温スパッタリング法における MoS_2 薄膜化と電気特性，”第 75 回応用物理学会秋季学術講演会，18p-A16-14，2014.

39) 松浦賢太朗，大橋匠，山口晋平，須田耕平，石原聖也，澤本直美，角嶋邦之，杉井信之，西山彰，片岡好則，名取研二，筒井一生，岩井洋，小椋厚志，若林整，“高温スパッタリング法による MoS_2 膜の形成と電気特性，”第 75 回応用物理学会秋季学術講演会，18p-A16-13，2014.

40) 若林整、舘下八洲志、王俊利、長野香、平野智之、宮波勇樹、生田哲也、片岡豊峰、菊池善明、山口晋平、安藤崇志、田井香織、松本良輔、藤田繁、山根千種、山本亮、神田さおり、釘宮克尚、木村忠之、大地朋和、山本雄一、長濱嘉彦、萩本賢哉、田川幸雄、塚本雅則、岩元勇人、齋藤正樹、門村新吾、長島直樹，“{100}及び{110}基板上の埋め込み SiGe 膜による歪み Si チャネルを有する Metal/High-k ゲート電極 MOSFETs，” 第 71 回 半導体集積回路技術シンポジウム，東京農工大学小金井キャンパス, 2007.

41) 山口晋平、田井香織、平野智之、安藤崇志、檜山進、Junli Wang、萩本賢哉、長濱嘉彦、加藤孝義、長野香、山中真由美、寺内佐苗、神田さおり、山本亮、舘下八州志、田川幸雄、岩元勇人、斉藤正樹、長島直樹、門村新吾，“バルク CMOS 向け低閾値電圧・高移動度デュアルメタルゲートトランジスタの開発，”応用物理学会シリコンテクノロジー分科会第 83 回研究集会，東京大学生産技術研究所, 2006

42) 山口晋平、田井香織、平野智之、安藤崇志、檜山進、Junli Wang、萩本賢哉、長濱嘉彦、加藤孝義、長野香、山中真由美、寺内佐苗、神田さおり、山本亮、舘下八州志、田川幸雄、岩元勇人、斉藤正樹、長島直樹、門村新吾，“バルク CMOS 向け低閾値電圧・高移動度デュアルメタルゲートトランジスタの開発，” 電子情報通信学会 SDM/ICD 共催研究会，北海道大学百年記念会館, 2006.

43) 山口晋平，田井香織，平野智之，岩元勇人，“MOCVD-W/WN を用いたメタ

ルゲートトランジスタの電気特性, ”第 66 回応用物理学界秋季学術講演会, 10p-ZK-15，徳島大学, 2005.

Acknowledgements

The support and encouragement of many individuals contributed to the successful completion of this work.

My deepest appreciation goes to Professor Hitoshi Wakabayashi who offered continuing support and constant encouragement. I was greatly fortunate to have him as my supervisor and learn from him, leading to realize the important and essential fundamental of my future career.

I would like to thank Professor Kazuo Tsutsui, Associate Professor Masahiro Watanabe, Associate Professor Shun Ohmi and Associate Professor Kuniyuki Kakushima for reading through this thesis and providing a lot of thoughtful advices.

I deeply appreciate all the co-workers in imec in Belgium where I conducted the research on $Si_{1-x}Ge_x$ channel devices: Drs. Liesbeth Witters, Jerome Mitard, Geert Eneman, G. Hellings, Andriy Hikavyy, Roger Loo, Anabela Veloso, Lars-Åke Ragnarsson, and Naoto Horiguchi. Discussions with those folks have been inspiring.

I also would like to thank my colleagues in Globalfoundries in Malta, U.S. for their kind support: Drs. Zeynel Bayindir, Xiaoli He, Suresh Uppal, Purushothaman Srinivasan, Manoj Joshi, Hyuck Soo Yang, Owen Hu, Srikanth Samavedam, and D. K. Shon.

Last but not least, I would like to thank my family members, my father Masatoshi, my mother Taiko, and my brother Takafumi, for their warm encouragements throughout this study. I would like to close this acknowledgement by expressing my deepest gratitude to my wife, Hanae, and my son, Kanta. I wouldn't have completed this work without them and their support and encouragement.

July 27, 2018

Shimpei Yamaguchi