Chapter 1

Introduction

1.1 Background: CMOS Scaling Theory and Limitation

In the last four decades, silicon has been widely used as the fundamental material for manufacturing a large number of semiconductor devices. The abundance, thermal stability and its perfect formation of an insulating silicon dioxide makes silicon the suitable choice for developing complementary metal oxide semiconductor (CMOS) integrated circuits for various kinds of electronic applications.

Moreover, the success of CMOS technology is largely due to the fact that the metal-oxide semiconductor field-effect transistors (MOSFETs) can be scaled to smaller and smaller dimensions, and associated with the device scaling comes the improved device performance. Although the architecture and working principle of the MOSFET have basically remained the same, the physical dimensions have been continually reduced to double the number of transistors on a chip every two years following Moore's Law [1]. The exponential decrease in transistor dimensions with enhanced device performance has resulted in powerful circuit functionality and high speed performance over technology generations [2].

CMOS technologies have experienced aggressive reduction of the feature size down to the sub 100 nm regime in order to deliver increased performance in terms of higher integration densities, increased speed and lower power consumption. However, further down sizing of the feature size to the nanometer regime becomes more challenging with each successive CMOS technology node as physical limits of existing manufacturing processes and materials are reached. Consequently, the real device benefit of scaling is declining as these physical limits are approached, and semiconductor industry requires novel solutions to meet the ever challenging device performance target. For example, the incorporation of new materials, from Cu, low-k backend interconnect level [3,4], to front-end high-k dielectrics, metal gate [5-6], new lithography technology [7] and even the new substrate materials [8-10] are emerging as important ways to continue to improve circuit performance.

1.2 Motivation

The conventional device dimension scaling cannot continue forever [3] as the continuing shrinkage of the transistor dimensions involves a lot of issues, such as reduced short channel effects (SCE) controllability [11], and the increased leakage current. For instance, the thickness of the oxide-based gate insulator has to be reduced with the reducing gate length, and thus results into increased gate tunneling leakage currents [12, 13]. Such oxide thinning effect is crucial since the mobility in scaled devices is significantly degraded due to the presence of high vertical fields. At the same time, an increased pocket dose has to be applied for better SCE control to prevent device punch or threshold voltage roll-off with the reduced gate length. The increased impurities neighboring around the channel actually increase the carrier scattering, and thus degrade carrier mobility and drain current of a MOSFET. Due to these increasing dominance of non ideal effects accompanied with each successive device scaling, the device gain brought by the scaling is actually reduced. Thus, the continuous size reduction of the critical dimensions of the transistors necessitates the adaptation of the new developed process techniques to meet the ever challenging device performance target [14, 15].

Without major change to the current process, strain induced band-structure modification turns out to be the easiest way to increase the carrier mobility and has been the dominant technology since 90nm technology node [15-17]. Proper strain improves carrier mobility and MOSFET drive currents by deforming the crystal structure of silicon. In other words, the modification of the lattice structure by creating tensile or compressive stress in the vicinity of the channel region can produce a corresponding strain to boost electrons and holes mobility. Therefore, the transistors employing the strained silicon technique have larger carrier mobility than transistors made of standard bulk silicon, and the scaled CMOS transistors are allowed to have higher current driving abilities with decent device performance.

Among recent newly developed uniaxial strain techniques [15-18], the so-called stress memorization technique (SMT) approach involves complex process steps including strained cap layer deposition, and amorphorized silicon plastic deformation annealing to generate the required strain for electron mobility enhancement [19]. Since SMT effect happens directly on the transistor main structure and is a complex processes including cap layer deposition, activation annealing, silicon recrystallization, and subsequent cap layer removal. Therefore, it is the theme of this thesis to explore and understand the SMT strain in depth to maximize its benefit for nFET device enhancement.

1.3 Thesis Organization and Overview

There are seven chapters in this dissertation, and is organized as follow:

In Chapter 1, the overview of our study and motivations of this thesis are described.

Chapter 2 gives an introduction of strain silicon theory upon the change of electronic and hole band structures. Furthermore, a brief summary of the evolution of strained silicon technology is outlined. Modern strain silicon fabrication methods such as substrate-induced strain and process-induced strain are also described.

After reviewing the basic concepts of the strain silicon theory, the process and literature overview of SMT are discussed in Chapter 3. The SMT effect and its influence on nFET and pFET device performance are summarized, with an emphasis on the SMT strain formation mechanisms for electron carrier mobility enhancement.

In Chapter 4, we've realized the fabrication of SMT strained nFET in the National Nano Device Laboratories (NDL) using conventional nitride as the SMT strain cap layer. The introduction of the SMT process for the as-fabricated nFET device can improve its transcondutance and short-channel effects (SCE). N-type dopant diffusion retarded by the SMT-strain is explained by the point defect engineering.

In Chapter 5, we've demonstrated the SMT fabrication process and analyzed the nFET electrical characteristics with different SMT cap nitride processes. Both PE-SiN and LP-SiN were investigated. Basic physical and chemical properties of nitride films, such as thickness, shrinkage, bonding composition change, and stress change upon S/D activation annealing are also characterized. The nFET electrical performance and interfacial quality degradation have been well correlated to the effect of the wafer bowing exerted by the SMT cap nitride. This suggests that wafer bowing induced by strain silicon process should be carefully controlled in order to minimize gate edge damage for high-yield production.

In Chapter 6, the channel backscattering characteristics for strained and non-strained nFETs using the temperature-dependent extraction method are examined for highly scaled nFET. We demonstrate that nFET channel backscattering can be effectively reduced by the uniaxial strain. Furthermore, the strain-reduced conductivity effective mass also increases the thermal injection velocity and ballistic efficiency. Such phenomenon becomes more appreciable as the gate length is reduced. It is thus expected that the performance enhancement induced by the uniaxial strain will be maintained even in the ballistic transport regime with very short gate length.

Finally, conclusions of this dissertation and recommendation for further research are presented in Chapter 7.

Chapter 2

Strain Silicon Technology

2.1 Introduction of Strain Silicon Technology

Historically, the number of transistors on an integrated circuit has doubled approximately every two years according to the well-known Moore's law [1]. To sustain this trend, the device physical dimensions (width, length, junction depth…) are approximately 70% scaled in the subsequent technology node to achieve a half reduction of transistor active area, ultimately resulting in today's CMOS devices with nano meter (nm) dimension. However, the increased channel impurity concentration and the vertical field required to control short channel effect have also been increased to the point where they substantially degrade carrier mobility. Maintaining short channel control without sacrificing drive current performance is one of the most important tasks in further scaling of today's nano-scale transistors.

Strained silicon channels are one of the most important technologies for further silicon CMOS developments. The mobility enhancement obtained by applying appropriate strain provides higher carrier velocity in MOS channels, resulting in higher current drive under a fixed supply voltage and gate oxide thickness [16, 18]. The strained silicon technology was first developed using a virtual substrate [20] to create a strain layer in the CMOS channel region aiming at strain-induced carrier mobility enhancement. Several state-of-the-art and low cost strained silicon technologies have been developed and integrated in advanced CMOS fabrication processes. The physical mechanism of carrier mobility enhancement with strain silicon techniques, and the methods of strain generation in advanced VLSI devices are presented next.

2.1.1 Electron Conductivity Enhancement by Biaxial Tensile Strain

Under low field operation, the current in doped n-type semiconductor is directly proportional to the electric field as guided by the Ohm's law. For n-type silicon with electric field applied, the current flows in parallel to the electric field direction as indicated by Eq 2-1 [21]:

$$J_n \equiv q \cdot n \cdot \mu_n \cdot E \qquad \text{Eq. 2-1}$$

where J_n is electron current density, n is the number of electrons per volume in silicon, q is the elementary charge of a single electron, E is the electric field, and μ_n is the mobility of electrons given by [21]

$$\mu_n \equiv \frac{q \cdot \tau}{m^*} \qquad \text{Eq. 2-2}$$

where τ is the average relaxation time and m^* is the electron transport effective mass.

The effect of strain on the band structure of silicon and on the carrier mobility has been extensively studied [22, 23]. The conduction band of unstrained bulk silicon has six equivalent valleys in E-K space as shown in Fig. 2-1. If there is no strain on silicon, the six equivalent valence bands are equally populated by electrons, so the conductance in either x, y and z directions is the same due to their symmetries.

When applying biaxial tensile strain for silicon in the x-y plane, the conduction band 6-fold valley degeneracy will split into Δ_2, and Δ_4 with the energy of minima of the 4-fold valleys rises with respect to the energy of the 2-fold valleys as shown in Fig. 2-2. As a consequence, the electrons prefer to populate the lower valleys, which are energetically favored. This results to increased electron mobility via reduced averaged electron conductivity mass. Furthermore, electron scattering can also be reduced due to the conduction valleys splitting into two sets of energy levels with the lowered rate of intervalley phonon scattering. Hence, the suppression of intervalley scattering by energy band splitting using strain also contributes to relaxation time increase and the mobility enhancement as predicted in Eq 2-2.

2.1.2 Hole Conductivity Enhancement by Biaxial Tensile Strain

The strain technique for hole mobility applies the similar principle as that of electron as discussed above. The principle is to split sub-bands in energy and then let carriers to populate in low energy sub-bands. For holes, the valence band structure of silicon is more complex than the conduction band due to three complex band structures as well as valence band warping under strain [23]. For unstrained silicon at room temperature, holes occupy the top two bands; the heavy and light hole bands [24].

Fig 2-3 shows the change of valence bands by using biaxial tensile strain. Besides valance sub-bands splitting, the shape of the sub-bands also changed under strain [16, 25, 26]. At the same time, the hole effective mass becomes highly anisotropic due to band warping, and the energy levels become mixtures of the pure heavy, light, and split-off bands. Thus, the light and heavy hole bands lose their meanings, and holes increasingly occupy the top band at higher strain due to the energy splitting.

The hole mobility improvement is mainly due to the reduced inter-band and intra-band scatterings by lifting of the degeneracy between light-hole and heavy-hole bands and lowering of the spin-off band, not due to the reduction of the effective transportation mass as the electron behavior under strained silicon.

2.1.3 Piezoresistance Coefficient for Strain Silicon

The piezoresistive effect describes the change of electrical resistance of a material due to applied mechanical stress because the electron transportation could be changed due to the deformation of crystal. For the semiconductor materials like single crystal silicon and germanium, this effect is even more pronounced since the applied strain changes not only the carrier transportation but also the carrier population due to the band structure deformation [27, 28]. The relationship between stress and electrical resistance is called as piezoresistance, which can be expressed as [18]

$$\pi = \frac{1}{\sigma} \cdot \frac{\Delta\rho}{\rho} \qquad \text{Eq. 2-3}$$

where π is the piezoresistance, σ is the force, and $\Delta\rho/\rho$ represents the change in resistance.

The piezoresistance effect comes from the change of the band structure. Once the mechanical force is applied, the lattice is distorted, and the lattice symmetry is reduced. This change in real lattice also reflects the change in the reciprocal lattice, and thus the E-K relationship can be varied. Consequently, the shapes of the valleys can thus be changed, and give rise to the change of the effective mass of the carriers therein accompanied with a redistribution of the carriers between valleys.

The corresponding piezoresistance coefficients for silicon and germanium are first measured by Charles Smith [27]. The CMOS device silicon piezoresistance coefficient for (100) and (110) silicon substrate, and the orientations most commonly used are listed in Table 2.1 and Table 2.2. [15-17]. Furthermore, mobility is in proportion to the stress applied at different directions as shown in equation 2-4 [15-17].

$$\frac{\Delta\mu}{\mu} \propto \left[\pi_L \cdot \sigma_L + \pi_W \cdot \sigma_W + \pi_Z \cdot \sigma_Z\right] \qquad \text{Eq. 2-4}$$

where the directions L, W and Z depend on the orientation of the device.

It should be noted that a positive piezoresistance coefficient means the resistance increases with tensile (or positive) stress according to Eq. 2-3, so a compressive stress should be applied in order to improve the carrier mobility. Thus, it is apparent that nFET mobility improves with tensile stress while the pFET mobility improves with compressive stresses. Based on this rule, these relations are further simplified into Table 2.3 indicating either the tensile or compressive strain should be given for mobility improvement.

2.2 Engineering Methods to Introduce Strain in MOSFET Channels.

There are mainly two different methods to introduce strain in the channel region, and can be categorized into two groups: substrate induced strain (namely biaxial strain) and process induced strain (namely uniaxial strain) [18, 20, 21].

Biaxial strain is also referred to as global strain and is introduced by epitaxial growth of Si and SiGe layers (substrate engineering). Such strain can be induced by taking advantage of the lattice mismatch between Si and SiGe as shown in Fig. 2-4 [22-26]. Although biaxial tensile strained silicon has received considerable attention in the last decade as a technique for mobility enhancement, it has been proven to be difficult to implement because of misfit and threading dislocations, Ge up-diffusion, fast diffusion of S/D extensions, and high cost [22-26].

In contrast, uniaxial strain can be more easily implemented by simple structure or process modifications, thus avoiding the complex wafer fabrication, high cost, and defects of biaxial strain. Furthermore, different uniaxial strain techniques are compatible and can be combined together to deliver larger drive currents [15-17, 25-26]. Consequently, uniaxially (or locally) strained devices have emerged as the main stream for carrier mobility enhancement.

2.2.1 Global Strain (Substrate-Induced Strain)

The global strain is generally biaxial, and is induced by the epitaxial growth of $Si_{(1-X)}Ge_X$ and silicon multiple layer structure as shown in Fig. 2-4. In order to reduce the defect density in the strained silicon, a relaxed $Si_{(1-X)}Ge_X$ buffer is required grown on a graded $Si_{(1-X)}Ge_X$ layer. The relaxed layer is usually a graded $Si_{1-X}Ge_X$ layer with increasing “x” value (Ge fraction) from substrate to the channel layer. Because the Ge concentration x increases continuously by about 10 % per μm, the thickness of the graded buffer is several micrometers [20, 24].

Since the lattice parameter of $Si_{(1-X)}Ge_X$ ($0 \leq x \leq 1$) alloys varies between 0.5431 nm (for pure silicon when $x = 0$) and 0.5657 nm (for pure germanium when $x = 1$), tensile biaxial strain is induced in a silicon layer epitaxially grown on top of the the relaxed SiGe layer. For nFET, such biaxial tension splits the conduction band and can boost electron mobility as predicted by Table 2.1 since tensile strain is preferred in both L and W direction.

However, for hole mobility operated under high vertical electric fields, biaxial strain actually reduces mobility [15, 17, 26]. There are several reasons for this. First, the decrease of the valence band separation due to surface potential confinement is diminished. Moreover, the valence band split is not as high as the conduction band, so a higher Ge fraction is needed to give decent improvement. Besides, the split also saturates at high Ge fraction. Another aspect to view this effect is from Table 2.3, for hole mobility enhancement, a compressive strain is preferred in the L direction while a tensile strain is preferred in the W direction. Therefore, by simply applying a biaxial strain one cannot satisfy both strain preferences.

From manufacturing point of view, growing such a multiple of SiGe and silicon layer structure is rather complicated and costly. Second, improper epitaxial process would generate dislocation loops within these high Ge fraction multi-layers, forming a substrate leakage path. To resolve all these problems, uniaxial strain processes for current semiconductor devices are presented next.

2.2.2 Local Strain (Process-Induced Strain Silicon)

Unlike the global strain created over the whole wafer, the uniaxial or the so-called “local strain” is realized locally especially in the transistor channel by using novel process technologies, and thus it is well known as process-induced strain (PIS). However, electron and hole mobilities respond differently to mechanical stresses as shown in Table 2.3, therefore, different strained channel approaches have to be employed to improve both the electron mobility in nFET and the hole mobility in pFET. The commonly used methods are using SiGe as the S/D for pFET [16, 18, 29] and SiC as the S/D for the nFET [30, 31]. The strained contact etch stop layer technique by means of different nitride film is also used to create stress [32, 33]. Other currently techniques include the Stress Memorization Technique (SMT) [19], and Shallow Trench Isolation Strain [34]. Some further discussions on these process-induced strain techniques widely used nowadays are presented in the section 2.3 below.

2.3 Uniaxial Strain Silicon Techniques for MOSFET Enhancement

Uniaxial strain has already been employed in 90-nm CMOS [15-17] by using typical process schemes including CESL nitride films covering on the nFET and pFET, SiGe regrown at pFET source/drain regions and stress memorization…etc as shown in Fig. 2-5. In all of these cases, the mobilities of carriers are increased by the introduction of tensile strain or compressive strain to enhance electron or hole mobility, respectively.

2.3.1 Embedded SiGe / SiC Source/Drain

Since Ge has larger lattice spacing than Si, incorporation of Ge into the Si lattice causes compressive stress on the nearby atoms. So by using SiGe as the source/drain on both sides of the channel, uni-axial compressive strain can be successfully applied to the channel direction to boost pFET mobility. It's similar to the argument for incorporating SiC as the S/D for nFET [30, 31]. There carbon atoms are smaller than than silicon atoms, so they shrink the lattice structure when they fit into a silicon lattice site. This causes tensile stress on the nFET channel direction [30].

Epitaxially-grown strained SiGe in the S/D regions and extension location is the widely used process technique to induce compressive strain in the channel to produce significant pFET hole mobility improvement [29]. The compressive stress is mainly dependent on the epitaxial SiGe thickness and Ge concentration.

Nevertheless, the boron doped SiGe regrown source/drain regions may contribute to pFET variability in short-channel effect control, due to source/drain regions recess profile local variations such as etch depth and proximity to channel. as a trade-off for the mobility improvement.

The integration of the different compressive stressors into advanced CMOS technologies results in increasing device performance. In this respect, pFET devices with a 45 nm gate length fabricated with recessed source/drains made of silicon germanium (20% germanium) and compressive nitride contact etch stop stressor films can demonstrate huge 200% improvement in hole mobility compared with equivalent transistors with no strain applied [35].

2.3.2 Contact-Etch-Stop-Layer (CESL) Stressor

A highly tensile nitride cap layer used as a contact etch stop layer (CESL) has been widely utilized in advanced MOSFET fabrication technologies owing to the resulting uniaxially tensile mechanical stress created in the nFET channel region [32, 33]. CESL are typically stressed nitride layers deposited after salicidation on top of the transistor’s structure. The strain induced in the channel region depends on the intrinsic stress of the layer, thickness of the layer, and device dimensions. Tensile CESL is deposited over nFET and compressive CESL is deposited for pFET to generate the desired uni-axial strain in the channel according to Table 2.3.

To meet integration requirements for nFET and pFET simultaneously, the so-called dual CESL technique [33] can be applied. In this approach, a highly tensile nitride layer over the entire wafer is firstly deposited, then it is patterned and etched from pFET regions as pFET prefers the compressive strain. Afterwards, a highly compressive nitride layer is deposited. Finally, this compressive layer is patterned and etched from nFET regions. As a result, one single nitride layer with the desired stress property can be formed for both nFET and pFET to boost carrier mobility.

2.3.3 Stress Memorization Techniques (SMT)

In addition to CESL, the channel region may be locally strained through a stress memorization technique resulting in performance improvements for nFET devices [19]. In this approach, the source/drain (S/D) substrate area and polysilicon gate structure are firstly amorphized by S/D and extension implantation. Conventional dopant activation annealing is performed after the deposition of a tensile SMT nitride stressor capping layer. The stress effect is transferred from the SMT nitride cap layer to the channel during the annealing process, then the re-crystallized, reformed S/D and poly-gate layers "memorize" and "retain" the SMT stress. This stress is retained even after the removal of the silicon nitride capping layer for nFET device improvement. The compatibility of uniaxial tensile strain techniques has also been validated by combining tensile CESL nitride and SMT to enhance strained nFET performance relative to unstrained devices [36].

However, SMT process increases nFET drive current and typically degrades the pFET transistor because of increase channel resistance accompanied with p type dopant depletion [37, 38]. The dopant loss happens when the huge amount of hydrogen evolved from SMT stack films. Because these hydrogen atoms can not escape outward when the SMT nitride cap is presented, they react with boron dopant at the S/D regions upon activation annealing [39].

2.4 Summary

Device performance has been mainly improved by geometrical scaling, such as scaling down of gate oxide thickness and gate length. However, gate length scaling is limited by both lithography technology and the accompanied short channel effect degradation. Gate oxide thickness scaling is also limited by the severe gate leakage current through the gate dielectric with thickness of less than 3nm. To overcome the current scaling limitation, strained silicon technologies have been widely studied as promising solutions for current and future device scaling. The mobilities of both electrons and holes can be improved by applying different kinds of stress to induce appropriate strain in the channel to changes their conduction and valence band structures. To conclude, mobility improvement in strained silicon takes place mainly due to

a) The reduction of the carrier conductivity effective mass, and

b) The reduction in the intervalley phonon scattering rates.

Besides, piezoresistance coefficient model correlates the channel stress components with mobility changes and suggests the appropriate stress polarity to boost carrier mobility. The engineering methods hence all follow the similar idea: applying the "proper strain" on the designated direction by modifying the device structure.

Process-induced strain refers to the strain in silicon lattice generated by many of the novel processing steps. Many recent publications on current CMOS technologies present engineering the process-induced strain to improve transistor performance as discussed in section 2.3, and these strain enrineering processes turn out to be the most popular techniques in the current device design technology. However, as the device keeps getting smaller, the increase in mobility may not be in proportion to the increase in current. This is because other effects like velocity saturation or ballistic transport become dominant as the device are highly scaled into nano meter regime as will be discussed in Chapter 6.

Channel	<100>			<110>		
Direction	L	W	Z	L	W	Z
Coefficient	π_{11}	π_{12}	π_{12}	$(\pi_{11}+\pi_{12}+\pi_{44})/2$	$(\pi_{11}+\pi_{12}-\pi_{44})/2$	π_{12}
n-Sub	-102.2	53.4	53.4	-31.2	-17.6	53.4
p-Sub	6.6	-1.1	-1.1	71.8	-66.3	-1.1

Table 2.1: Piezoresistance coefficients on (100) Silicon substrate.

Channel	<100>			<110>		
Direction	L	W	Z	L	W	Z
Coefficient	π_{11}	π_{12}	π_{12}	$(\pi_{11}+\pi_{12}+\pi_{44})/2$	π_{12}	$(\pi_{11}+\pi_{12}-\pi_{44})/2$
n-Sub	-102.2	53.4	53.4	-31.2	53.4	-17.6
p-Sub	6.6	-1.1	-1.1	71.8	-1.1	-66.3

Table 2.2: Piezoresistance coefficients on (110) Silicon substrate.

Here L refers to the channel direction, W is the direction perpendicular to the channel on the surface and Z is the direction perpendicular to both L and Z.

Substrate		(100)		(110)	
Channel		[110]	[100]	[110]	[100]
NMOS	L	Tensile	Tensile	Tensile	Tensile
	W	Compressive	Compressive	Compressive	Compressive
	Z	Compressive	Compressive	Tensile	Compressive
PMOS	L	Compressive	Compressive	Compressive	Compressive
	W	Tensile	No significant effect.	No significant effect.	No significant effect.
	Z	No significant effect.	No significant effect.	No significant effect.	No significant effect.

Table 2.3: Preferred stress type for mobility enhancement. Tensile means tensile stress is preferred; compressive means compressive is stress preferred.

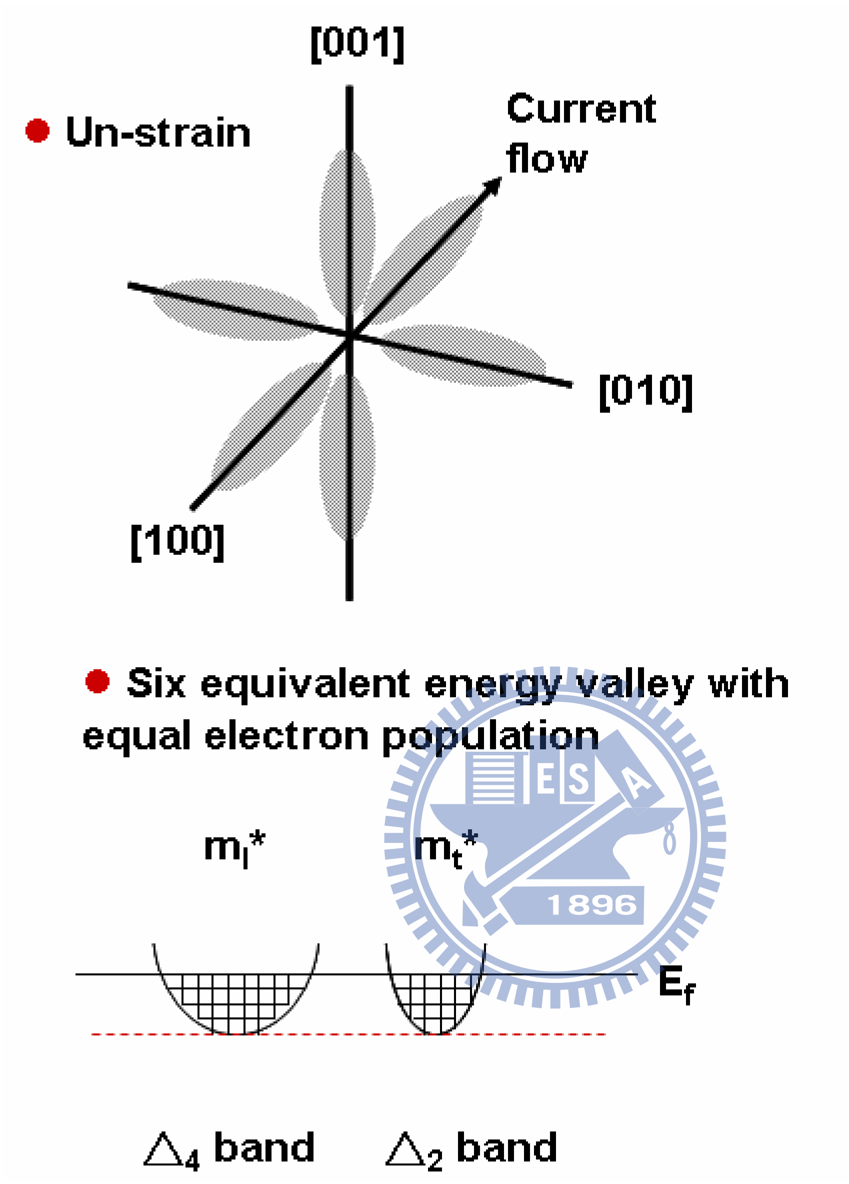


Fig. 2-1: Constant energy surface for un-strained silicon conduction band. The six equivalent valence bands are equally populated by electrons, so the conductance in either x, y and z directions is the same due to their symmetries.

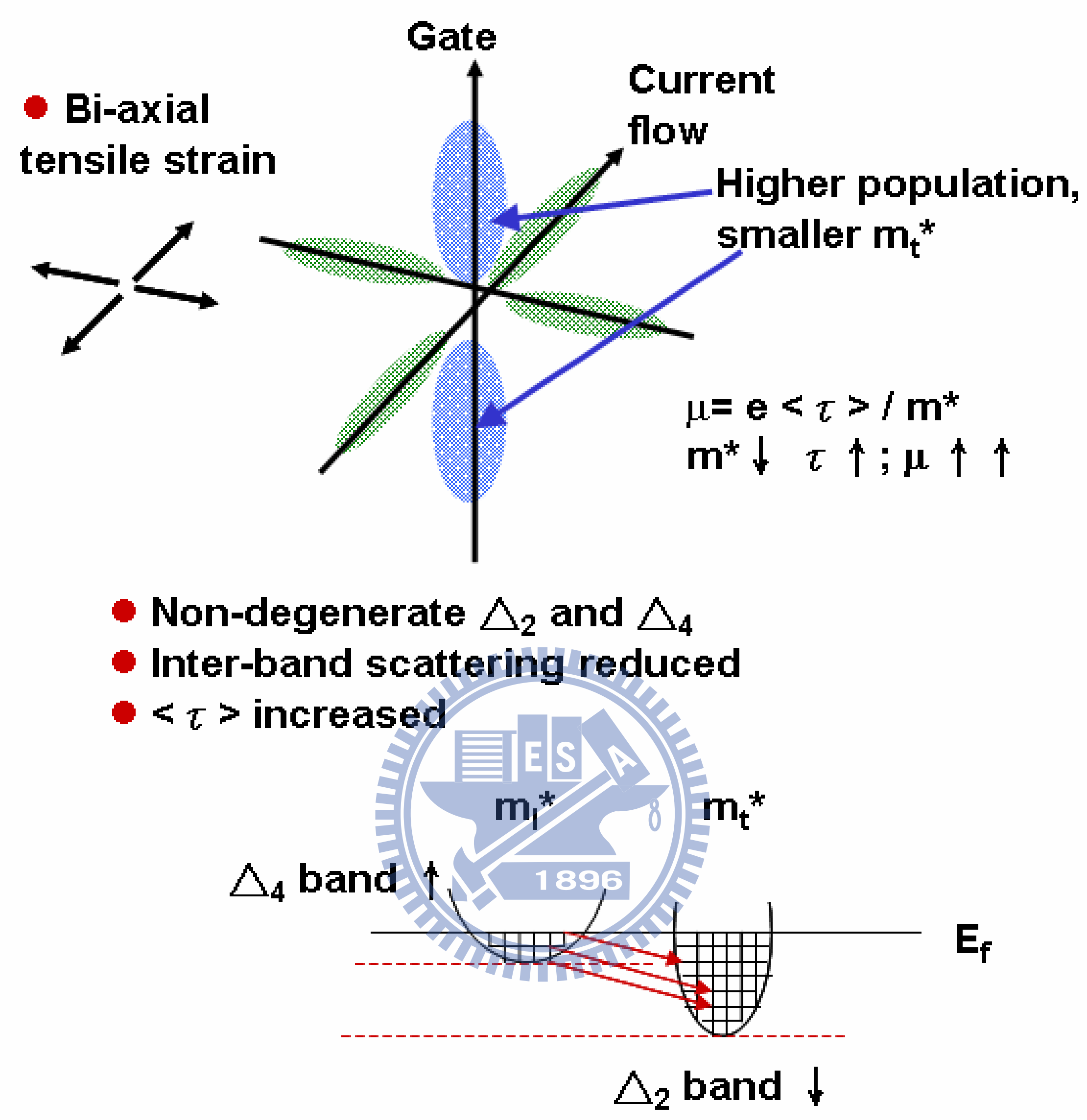
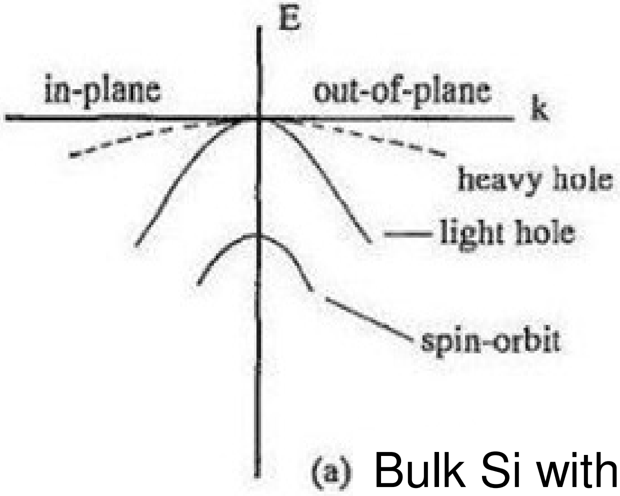
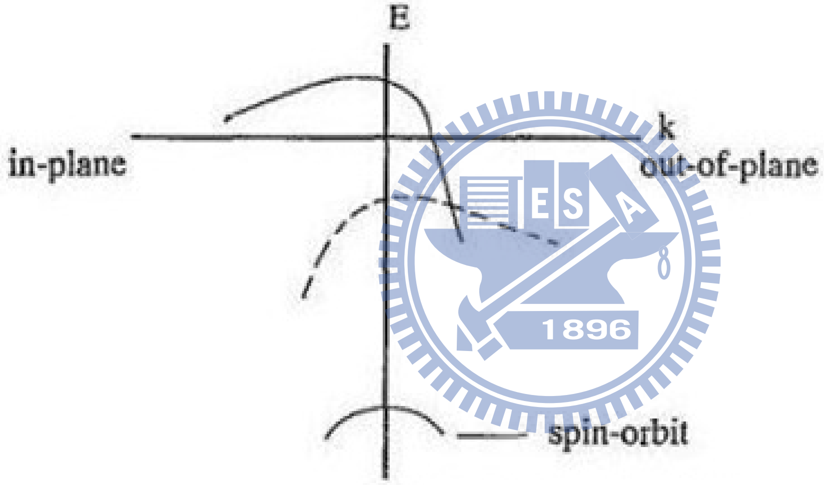


Fig. 2-2: Constant energy surface for strained silicon conduction band. Biaxial tensile strain in the x-y plane induced silicon conduction band splitting for electrons under z-direction confinement, the conduction band 6-fold valley degeneracy will split into Δ_2, and Δ_4 with the energy of minima of the 4-fold valleys rises with respect to the energy of the 2-fold valleys.



(a) Bulk Si without strain.



(b) Si with bi-axial strain.

Fig. 2-3: Simplified hole carrier valence band diagrams. The change of valence bands by using biaxial tensile strain is shown in Fig. 2-3b. Besides valance sub-bands splitting, the shape of the sub-bands also changed under strain [16, 25, 26]. Furthermore, the hole mobility improvement is contributed by the reduced inter-band and intra-band scatterings.

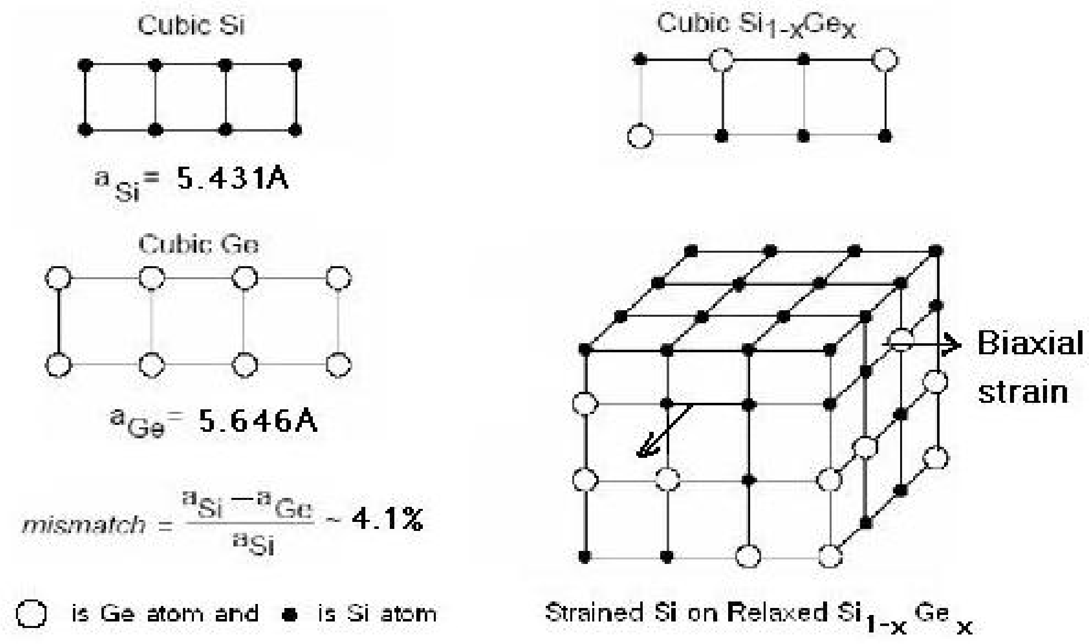


Fig. 2-4: Biaxial tensile strain generation by Si epitaxy growth on SiGe. This technology generates biaxial tensile strain in a silicon epitaxy layer on relaxed SiGe. Because the silicon epitaxy has to match the lattice structure of the SiGe substrate, it is forced to have the same lattice constant as the SiGe substrate. Since the lattice constant of Si is less than SiGe, so the Si epitaxy layer will be forced to "expend" on the surface plane of SiGe substrate and will introduce a biaxial tensile strain for the Si layer.

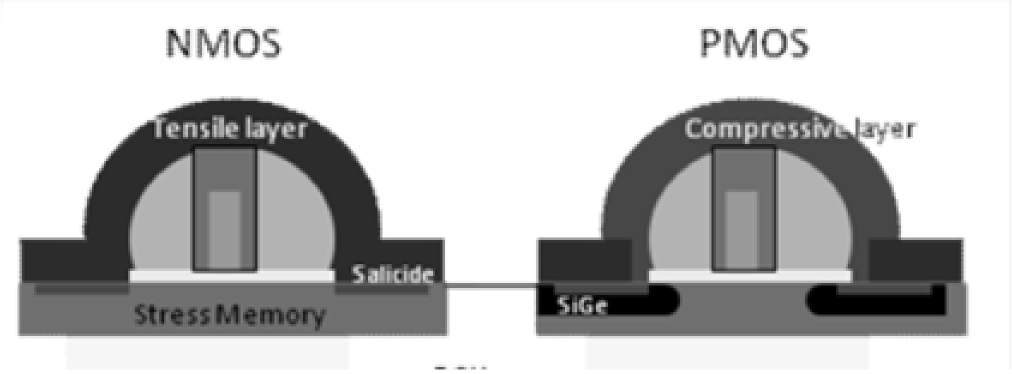


Fig. 2-5: Schematic representation of process induced strain in CMOS. Tensile strain in NMOS is obtained by tensile CESL SiN and stress memorization technique (SMT). Compressive strain in PMOS is induced by compressive CESL SiN and embedded SiGe.

Chapter 3

Review and Mechanism Proposed on Stress Memorization Technique (SMT)

3.1 SMT Literature Review and Discussion

Complementary metal oxide semiconductor (CMOS) transistors are composed of nFET transistors and pFET transistors and it is desirable that both types of transistors have high current driving abilities. Newly developed strain silicon technologies improve MOSFETs drive currents by fundamentally altering the band structure of the channel and can therefore enhance performance even at aggressively scaled channel lengths as previously reported in Chapter-2. For nFET devices, the electron mobility is enhanced in strained-Si compared to bulk-Si due to tensile strain splitting the six-fold degenerate conduction band valleys and causing the lower energy two-fold band with reduced in-plane effective mass to be preferentially filled [40, 41]. A four-fold band with increased energy is also created, which consequently contributes to the higher electron mobility through a reduction in intervalley scattering. Tensile strained-Si channels are, thus, useful for electron mobility enhancement for nFET devices. As discussed in the next sections, one can realize this tensile strained channel by using stress memorization technique (SMT).

3.1.1 SMT Enhances nFET Device Performance

The stress memorization technique (SMT) concept was firstly proposed by Ota *et al.*, using a tensile oxide cap layer deposited on the gate electrode after ion implantation was performed to form extension region and source/drain (S/D) regions as shown in Fig. 3-1 [42]. Another novel approach by utilizing a tensile strained silicon nitride cap layer was also firstly proposed by Chen *et al.* [19], opens up further the opportunities for cost-effective and process-friendly methods to achieve higher level of strain in advanced CMOS fabrication technology. Chen's SMT technique improves the nFET performance by depositing a tensile nitride stressor on top of poly gate electrode after S/D implantation. This cap strained layer is present during the subsequent S/D dopant activation annealing to re-crystallize the gate and the S/D regions of the nFET into a strained state under nitride's mechanical strain, and causing stress from the cap nitride film to be memorized in a channel region. Accordingly, in an nFET, mobility of electrons increases so that the current driving ability is enhanced. Upon subsequent nitride stripped off, a significant portion of this strain is retained (`memorized') and results into higher nFET drive current by means of the process flow shown in Fig. 3-2. The nFET device improvement can be boosted up to 10-15% by Chen's SMT technique [19] post optimized amorphized implantation. Generally, a bi-layer SMT film stack in which an thin oxide underlying film to protect silicon active region upon subsequent wet chemical etch and a thicker nitride stressor cap are deposited together to meet process integration requirements [37].

3.1.2 SMT Induced pFET Loss

However, the use of an SMT causes the current driving ability to deteriorate for pFET [37]. The reason for pFET performance deterioration is the p-type dopant boron de-activation caused by pFET covered with an SMT nitride upon S/D annealing [37, 39]. It has been proved that the device performance penalty for pFET transistors is largely related to boron dose loss around extension, and S/D junction [38, 39]. Such boron depletion can increase channel conducting resistance and thus pFET device performance can be weakened. This is because when the transistor is covered with a rigid, high density silicon nitride film serving as a cap film, hydrogen originated from SMT oxide serving as an underlying film can not diffuse outward during S/D annealing as illustrated in Fig. 3-3 [38, 39, 43, 44]. Hence, the excess hydrogen atoms thus diffuse downward to the S/D region underneath and results in a decrease in activation rate of boron contained in the pFET S/D and gate regions.

One way of avoiding this pFET degradation is by removing the nitride layer on top of pFET transistors before anneal as shown in Fig. 3-2. The “selective SMT” processes with extra litho and etch step process to remove the nitride film on the pFET transistor before activation annealing is effective to minimize pFET loss (only the nFET transistor is covered with the SMT film) [19, 37].

Another way to avoid or to lessen pFET degradation by SMT is to adopt an optimized SiN layer to minimize boron de-activation [37]. Ortolland *et al.* had proposed their novel SMT nitride cap which can lessen pFET loss and can still keep nFET device improved. They demonstrated that the understanding of coupling between nitride properties, dopant activation and poly-silicon gate mechanical stress allows enhancing nFET performance by 7% without significant pFET degradation. This approach enables a more economical implementation of the SMT process as no extra selective nitride removal steps are required [37].

Moreover, in situ N_2 plasma nitridation treatment before the SMT layer deposition has also been reported to recover pFET performance by reducing the boron-doping loss from the gate and the source/drain region [45].

3.2 Mechanism of SMT Stress Generation

The enhanced nFET transistors by SMT process is based on the basis of stress memorization. Owing to this memorized stain in the channel, the SMT process boosts electron mobility as discussed previously. The linear transconductance, G_{m_lin}, using SMT process is compared with that of non-SMT process with an around 10% superior G_{m_lin} enhancement has been reported [36]. These results suggest that the SMT process induces a permanent stress in the channel region to enhance nFET carrier mobility and results in the improvement of device performance, which is in agreement with Chen *et al.* [19]. In addition, the strain can be retained and is attributed to the stress "memorization" effect even with the removal of the nitride stressor film. In other words, SMT strain results from the residual stress memorized in the channel after nitride removal, rather than the initial mechanical stress created after nitride deposition as in the conventional CESL technique [32, 33].

Stress memorization techniques works effectively for electron mobility enhancement in essence by improving the conditions during the re-growth of heavily implantation damaged gate, drain and source areas. Hence many studies with SMT process [19, 36, 37, 46, 47] suggest that the process involves the following fundamental elements: S/D n-type dopant implantation, amophorized silicon implanted region, capping nitride deposition, S/D thermal annealing, and amorphous silicon transformation into re-crystallized silicon with plastic silicon deformation. While a detailed understanding of the fundamental mechanism of the SMT process remains the subject of active discussion, recent studies have proposed some models to deliver further insights into the mechanism behind.

3.2.1 Stress Memorized in Poly Gate Model

Adam *et al.* [47] have proposed a phenomenological model to understand the stress memorization effect. A plastic deformation model is used to simulate such irreversible shape change on the polysilicon gate. Their simulation results show that the polysilicon gate undergoes a structure deformation via a transition from the elastic to the plastic states during the S/D anneal. Consequently, this plastic deformation generates a tensile strain in the channel lateral direction and a compressive strain in the channel vertical direction and can boost electron mobility according to the silicon piezoresistance model [16, 17].

Moreover, the capping nitride film also affects the poly-Si stack deformation by the SMT process as depicted in Ref. 15, by suppressing outward expansion of the poly-Si gate. This scenario, schematically drawn in Fig. 3-4, is also supported by Miyashita *et al.* [46] in which the correlation between nitride film porosity and the resulting channel stress are analyzed. In Miyashita's work, the gate polysilicon is amorphized by S/D dopant implantation, then capped with the SMT nitride layer. During the S/D annealing, the amorphized silicon expands as it re-crystallizes; but since it is constrained by the capping nitride and sidewall spacers, compressive vertical stress is thus generated and applied to the channel underneath.

Miyashita *et al.* has investigated the stress memorization technique (SMT) using poly-gates through both physical analysis and electrical characterization [46]. The gate polysilicon is amorphized by implantation, then capped with nitride. During the source/drain anneal the amorphized silicon expands as it re-crystallizes, but since it is constrained by the cap and sidewall spacers, compressive vertical stress is thus applied to the channel underneath [46]. According to the piezoresistance model, the nFET electron mobility can be enhanced by tensile stress in the length direction, and by the compressive stress in the gate vertical direction [16, 17].

It has been clarified that channel compressive strain in the vertical direction originates from poly-gate volume expansion, which is associated with both grain growth and highly concentrated impurities implanted into gates. In addition, the gate polysilicon's volume can be further expanded with high atomic mass n-type species such as arsenic or phosphorus implantation, and different level of silicon pre-amorphized induced by different implantation species can also give rise to different level of the memorized strain [46]. In this respect, both arsenic and phosphorus S/D implants were studied. It has been found that arsenic with the greater atomic mass increases the poly gate expansion tendency much more that that of the relatively lighter phosphorus [46]. By optimizing key factors in the SMT process with arsenic S/D implantation, a more competitive nFET drive current can be obtained as compared to that with phosphorus S/D implantation. For aggressive scaling beyond 45nm node, suitable arsenic S/D implantation to achieve highly amorphized poly gate combined with optimized SMT processes is suggested [46].

3.2.2 Stress Memorized in Source-Drain Model

During the dopant activation annealing process for re-crystallizing the lattice, the growth of the crystal will occur under stress conditions created by the SMT cap nitride and result in a strained crystal. With the same S/D implantation, the substantially amorphized S/D region adjacent to the gate electrode also experienced similar re-crystallization in the presence of a rigid nitride layer deposited above the transistor structure as shown in Fig. 3-5.

After the S/D re-crystallization upon annealing, the SMT nitrider may completely removed, and a certain amount of strain may be "conserved" in the re-grown lattice portion of S/D regions as illustrated in Fig. 3-5E. This effect can also be regarded as stress memorization.

Upon re-crystallization of the substantially amorphized material, the increased volume of the amorphous material compared to the crystalline material may be substantially maintained due to the presence of the rigid SMT nitride cap layer that reduces or prevents the natural volume reduction which would usually occur during the re-crystallization, thereby causing the amorphous portion to “connect” to the surrounding silicon material in a strained state due to the rigidity of the overlying layer. Hence, the strained re-grown crystalline S/D region may also induce a corresponding tensile strain to the adjacent transistor channel region to boost electron mobility, even after the rigid SMT nitride layer is removed.

After LDD implantation, plus the thermal annealing of the subsequent offset and main spacer deposition, SMT phenomena can happen due to solid phase epitaxial regrowth (SPER) [48]. Because SMT by SPER method only requires an amorphized silicon region, spacer cap layer deposition, together with the thermal annealing during spacer deposition process. A new stress memorization technique (SMT) is used to induce tensile stress into the channel using a low temperature SPER process after each spacer deposition [49]. The temperature for this stress memorization phenomenon is usually below 700°C with a limited amount of dopant diffusion. Thus, this multiple SPER induced stress memorization technique provide an alternative method to further enhance nFET performance [49] .

3.2.3 Stress Memorized Effect with Respective to Nitride Integrity and Capping Strength

A higher density or a more rigid SMT capping nitride film plays a major role for the SMT strain generation [37, 38]. It has been suggested by Ortolland, *et al*. that harder SMT nitride, with lower porosity and higher capping strength, is more effective to conduct the memorized stress to the channel [37, 38]. Because the rigid capping tensile-strained nitride film acts as a boundary layer to suppress poly-Si volume outward expansion during the recrystallization of amorphous poly-Si upon RTA annealing [19, 47], a tensile strain along the channel direction is generated in the device itself. This is also consistent with results reporting that the harder spacer nitride film as the spacer dielectric may act as a boundary layer to inhibit poly outward expansion, and thus it is more effective to apply stress to the channel as suggested by Miyashita's investigation [46].

3.3 Summary

This chapter introduces the SMT prevails in today's semiconductor industry. The conventional stress memorization approaches for nFET transistors may typically be implemented in the overall CMOS process by taking advantage of the effect that the introduction of the n-type dopant for defining the deep S/D regions and intermediate areas may be associated with a high degree of lattice amorphization. This S/D implantation results in a substantially amorphous state of a substantial portion of the S/D regions due to the high dose and energy used. Thus, prior to performing a respective anneal process for re-crystallizing the implantation-induced damage and for activating the dopants, a stiff SMT cap nitride can be deposited over the transistor and act as a stressor during the anneal process, thereby obtaining the desired strain state of the re-grown deep S/D regions, which may therefore induce the desired tensile strain in the channel region.

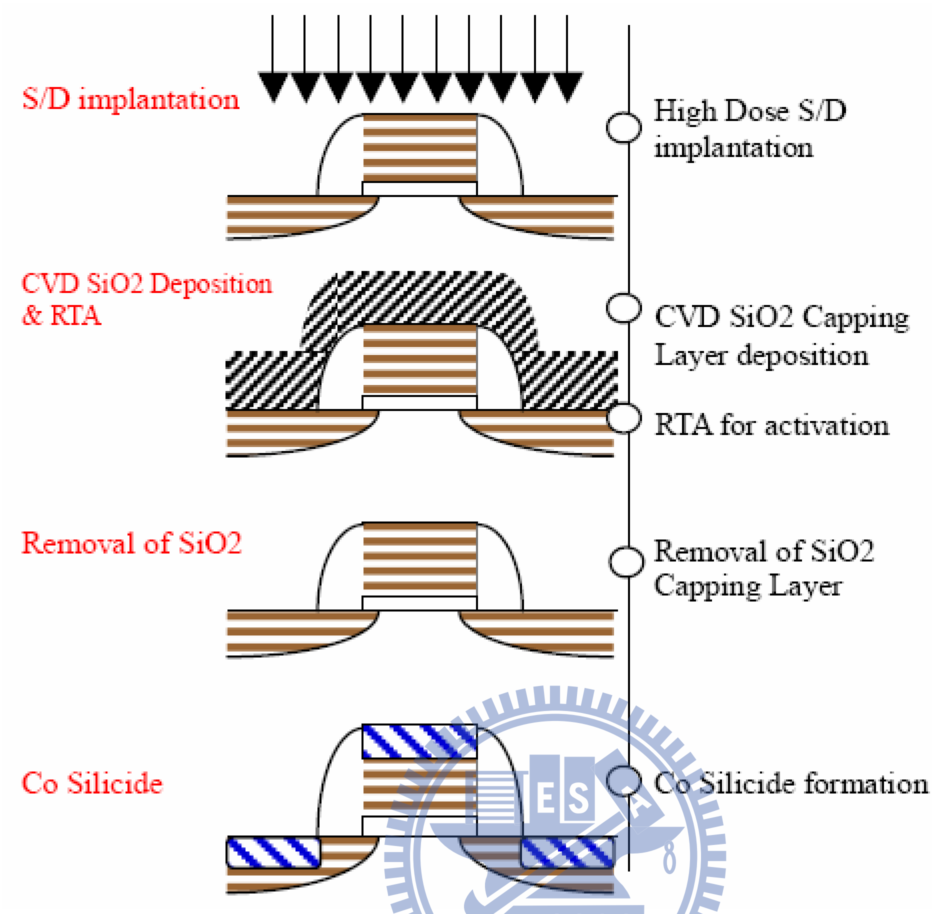


Fig. 3-1 Fabrication process for locally strained channel (LSC) transistor [42]. A cap oxide is used to generate SMT strain.

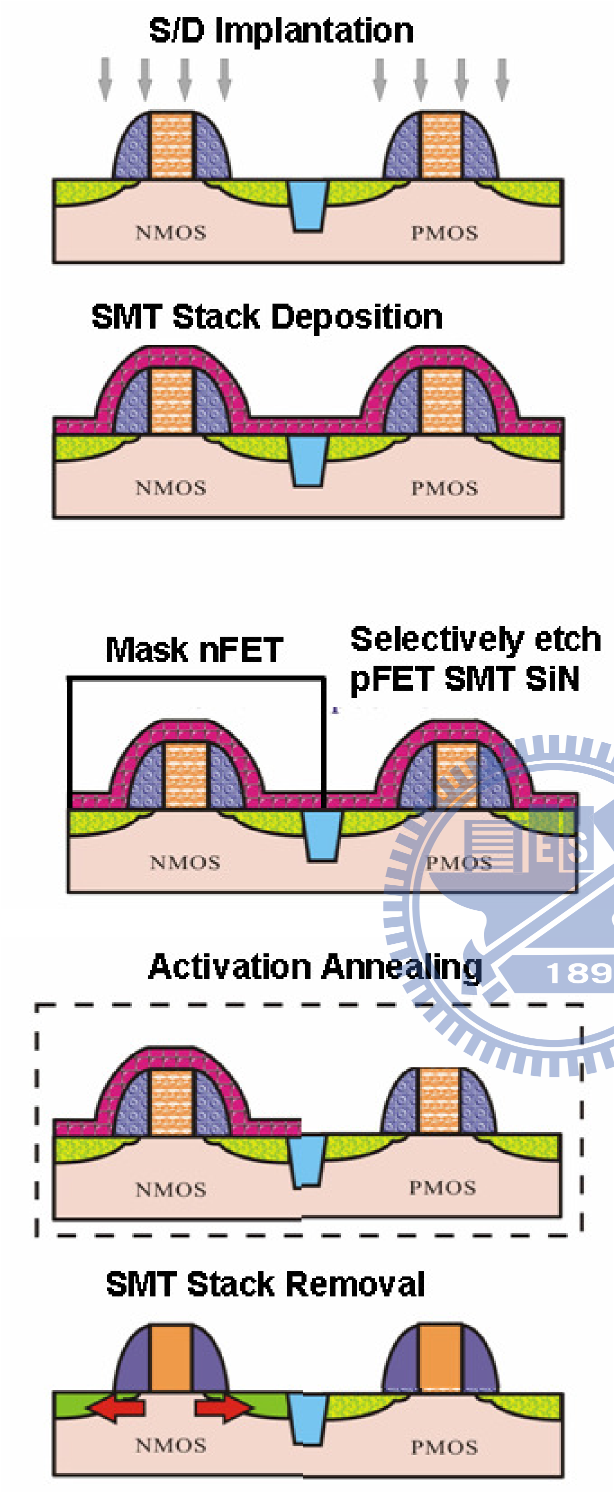


Fig. 3-2 The SMT process flow sequences with a nitride cap [19]. A selective SMT process to remove nitride on pFET before activation annealing is required to reduce pFET loss.

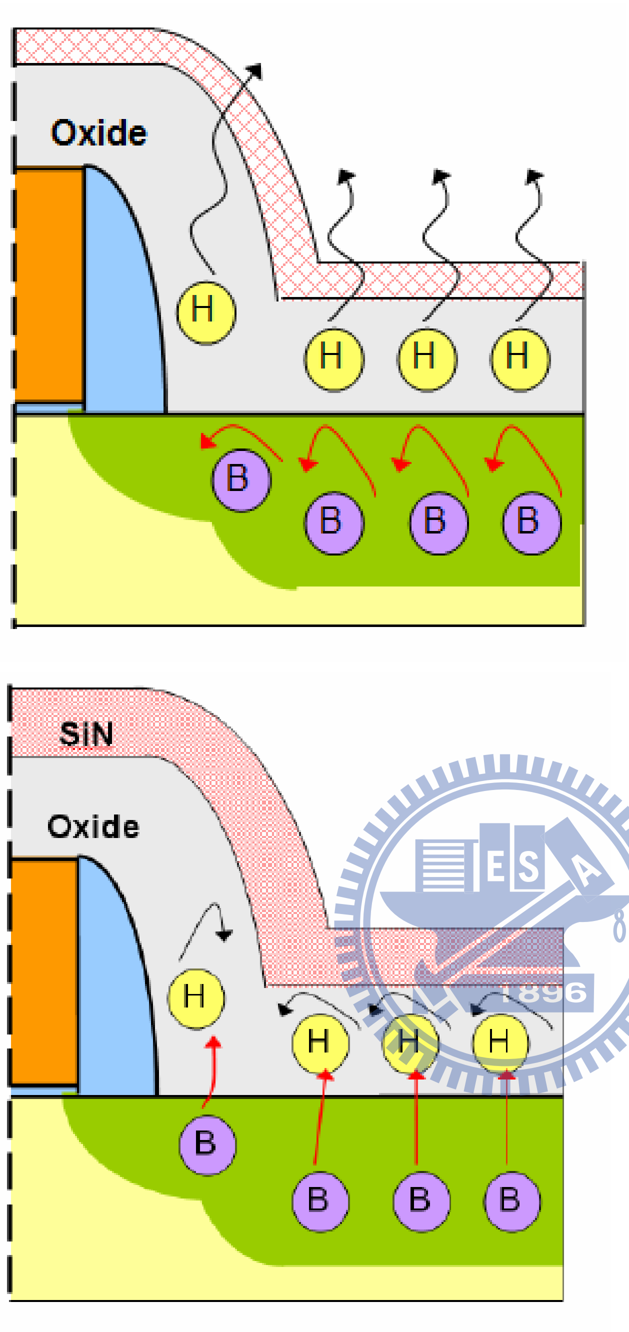


Fig. 3-3A (Top) Without nitride cap, the outgas hydrogen does not enhance boron depletion. Fig. 3-3B (Bottom) The hydrogen induced Boron depletion model for pFET with SMT cap nitride process [39]. Hydrogen is not able to diffuse outwardly via nitride layers and thus enhance boron deactivation and its depletion [38, 43, 44] .

Fig. 3-4A

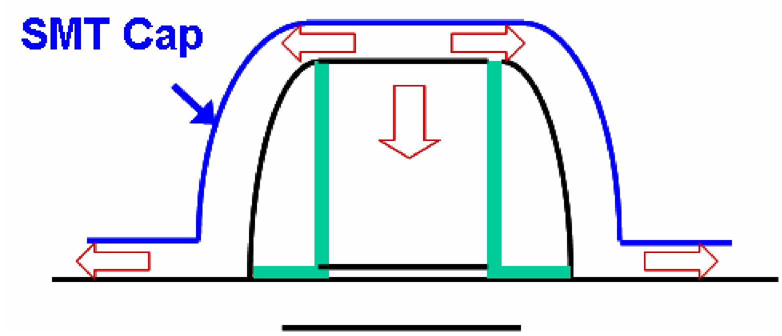


Fig. 3-4B

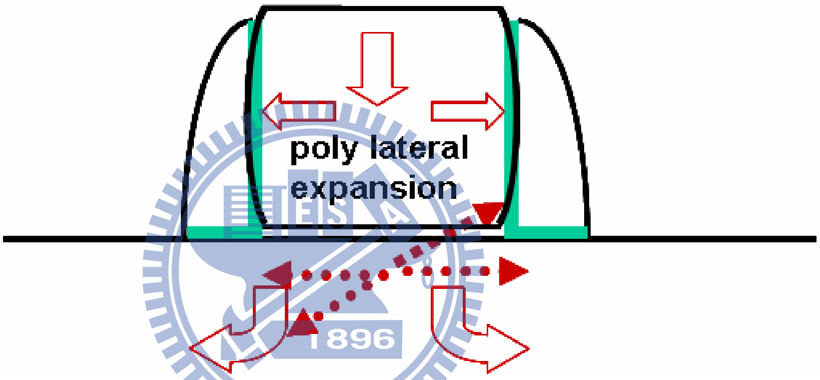


Fig. 3-4 Schematic diagram of stress memorized in gate. SMT tensile nitride tends to flatten poly after its deposition in Fig3-4A. Upon thermal annealing, the poly-gate expands along channel direction and width direction under the tensile SMT cap nitride stress. After thermal annealing, poly gate deforms and induce tensile strain retained in the channel region, (both length and width direction) in Fig. 3-4B.

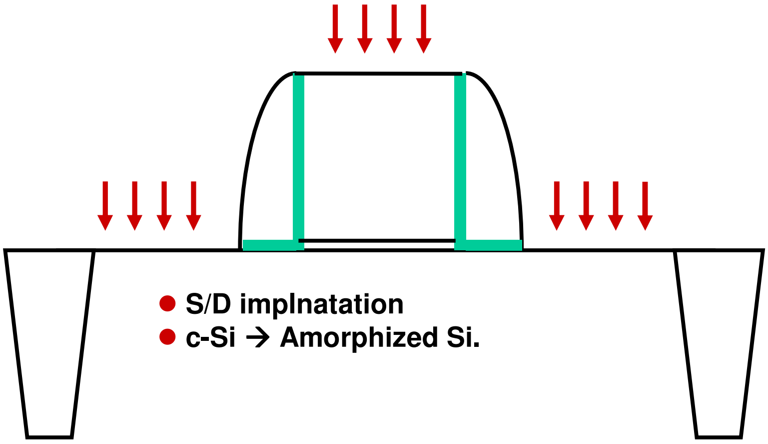


Fig. 3-5A

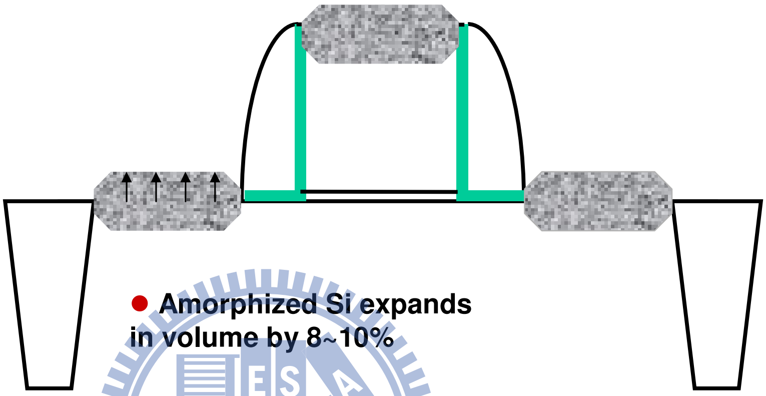


Fig. 3-5B

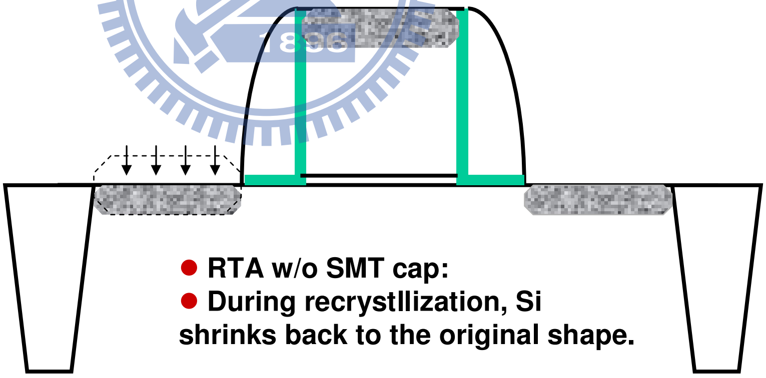


Fig. 3-5C

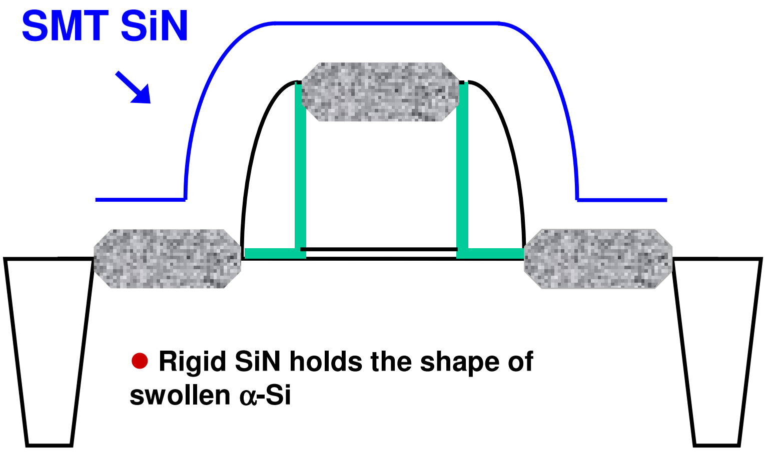


Fig. 3-5D

Fig. 3-5E

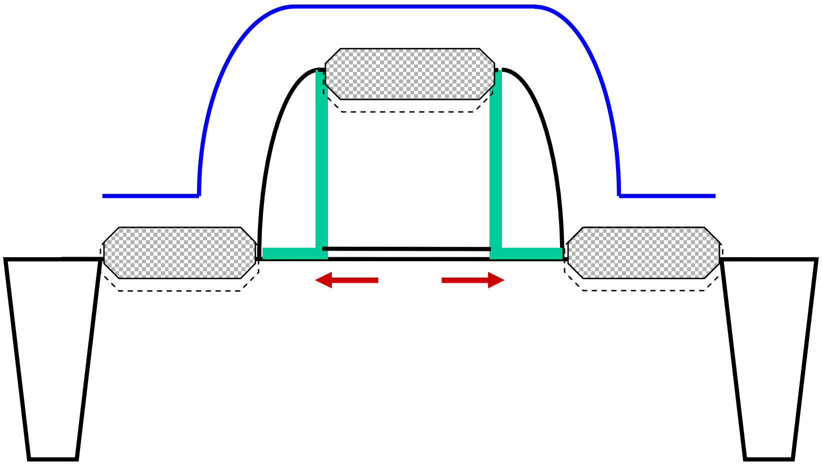


Fig. 3-5 Schematic diagram of stress memorized in source and drain active region.

The implanted silicon active region in Fig 3-5A becomes amorphized as shown in Fig 3-5B. Silicon recrystallizes and shrinks back to its original position (volume) without the presence of a cap layer upon thermal annealing in Fig 3-5C. However, with the presence of a SMT cap layer during activation annealing to hold the upper active silicon in Fig. 3-5D, the silicon shrinks its volume only by pulling the lower portion of silicon upwardly as drawn in Fig. 3-5E. Thus, a tensile strain builds up in crystallized silicon and is frozen in the channel to boost electron carrier mobility due to silicon plastic deformation.

Chapter 4

New Findings on SMT

4.1 Experimental and Device Fabrication

The devices reported in this chapter were fabricated in the National Nano Device Laboratories (NDL) on 6” (150 mm) silicon wafers using a conventional MOSFET process flow that including LOCOS isolation, gate oxide, spacer, source-drain (S/D) implantation, thermal annealing, and Al metallization processing as listed in Fig. 4-1. All the nFET devices characterized in this study had 3nm thick gate oxide thermally grown in a vertical furnace, and a 150nm thick poly-Si layer as the gate electrode. After self-aligned spacer formation, a 100Å plasma enhanced (PE) TEOS oxide was deposited prior to the S/D implantation to reduce the N^+ region junction depth. Prior to the S/D activation annealing, a similar SMT process [19] was employed for SMT strained nFET, while skipping this SMT nitride capping process for the non-strained nFET control devices. For the SMT strained nFETs, a 800Å tensile PE nitride film was deposited at a relatively low temperature of 300°C over the entire device structure including poly gate and S/D region. The main reaction gasses utilized for SMT nitride formation were NH_3, N_2, and SiH_4. To facilitate metallization, the SMT nitride film was then wet removed using chemical etchants H_3PO_4 and diluted 100:1 HF. Next, passivation was performed by 300nm thick TEOS oxide, followed by patterning and etching of and Al metallization..

After completion of the metallization processes, the wafers were annealed at 400°C in a forming gas ambient and the fabricated device wafers were characterized electrically. The nominal device dimension evaluated was a gate length of 0.4 μm and width of 10 μm. The device characteristics were measured by a semiconductor parameter analyzer (HP 4156A, plus Agilent ICS Software). The threshold voltage Vt was determined by the constant drain current method when drain current set to 10^{-7} • W/L (Amp) [50]. The linear and saturation mode threshold voltages, V_{t_lin} and V_{t_sat} , were extracted at drain voltages of 50mV and 2V, respectively.

The identical S/D implantation condition was chosen for the process on both blanket and device wafers. The diffusion profiles of arsenic and phosphorus on blanket wafers were studied by Secondary Ion Mass Spectrometry (SIMS) analysis and used to quantify device junction depth, and to study the arsenic and phosphorus diffusion. The dopant diffusion profiles from these SIMS results were in turn correlated to the extracted nFET V_t roll-off behavior, so as to investigate the SMT strained nFETs' SCE characteristics.

4.2 Results and Discussion

SMT strain engineering has become a standard technique to improve nFET transistor performance since the 65nm node [19]. Even in the advent of advanced gate stacks such as nitride gate oxide or high dielectric constant material, it is still necessary to boost device performance to its limits in order to gain competitive advantage. In this section, several other dopant diffusion phenomena and accompanied benefits raised by SMT engineering have been further investigated.

4.2.1 nFET Mobility Enhancement by SMT

The linear transconductance, G_{m_lin}, using SMT process is compared with that of non-SMT process. The superior G_{m_lin} with around 10% enhancement is shown in Fig. 4-2. These results suggest that the SMT process induces a permanent stress in the channel region to enhance nFET carrier mobility and results in the improvement of device performance, which is in agreement with Chen *et al.* [19]. In addition, the strain can be retained and is attributed to the stress "memorization" effect even with the removal of the nitride stressor film. In other words, SMT strain results from the residual stress memorized in the channel after nitride removal, rather than the initial mechanical stress created after nitride deposition as in the conventional CESL technique [32, 33]. Many studies with SMT process [19, 37, 46, 47] suggest that the process involves the following fundamental elements: S/D n-type dopant implantation, amophorized silicon implanted region, capping nitride deposition, S/D thermal annealing, and amorphous silicon transformation into recrystallized silicon with plastic silicon deformation. While the precise physics underlying the stress-memorization is still unclear, recent studies have given some insights into the problem. Adam *et al.* [47] have proposed a phenomenological model to understand the stress memorization effect. A plastic deformation model is used to simulate such irreversible shape change on the polysilicon gate. Their simulation results show that the polysilicon gate undergoes a structure deformation via a transition from the elastic to the plastic states during the S/D anneal. Consequently, this plastic deformation generates a tensile strain in the channel lateral direction and a compressive strain in the channel vertical direction and can boost electron mobility according to the silicon piezoresistance model [16, 17].

Moreover, the capping nitride film also affects the poly-Si stack deformation by the SMT process as depicted in [47], by suppressing outward expansion of the poly-Si gate. This scenario is also supported by Miyashita *et al.* [46] in which the correlation between nitride film porosity and the resulting channel stress are analyzed. In Miyashita’s work, the gate polysilicon is amorphized by S/D dopant implantation, then capped with the SMT nitride layer. During the S/D annealing, the amorphized silicon expands as it re-crystallizes; but since it is constrained by the capping nitride and sidewall spacers, compressive vertical stress is thus generated and applied to the channel underneath. It has also been suggested by Ortolland *et al.* [37] that harder nitride, with lower porosity and higher capping strength, is more effective to conduct the memorized stress to the channel. In addition, the gate polysilicon's volume can further expand after implantation with high atomic mass n-type species such as arsenic or phosphorus. Because the rigid capping tensile-strained nitride film acts as a boundary layer to suppress poly-Si volume outward expansion during the recrystallization of amorphous poly-Si upon RTA annealing [19, 47], a tensile strain along the channel direction is generated in the device itself. This is also consistent with results reporting that a rigid capping nitride film plays a major role in the SMT strain generation process [37].

4.2.2 New Findings on SMT: Improved nFET Short Channel Effect with the SMT Process

In addition to the enhancement of nFET transconductance and mobility by permanent strain memorization as shown in Fig. 4-2, the variations of threshold voltages with and without SMT process were further investigated. As drain voltage is increased to switch the nFET operation from the linear region to the saturation region, the depletion region of the p-n junction between the drain and body increases in size and extends under the gate. Thus, the drain assumes a greater portion of the burden of balancing depletion region charge, leaving a smaller burden for the gate. Therefore, device threshold voltage decreases with shorter channel length conditions, a phenomenon commonly referred to as "V_t roll-off" or "Short Channel Effect" (SCE). Another index commonly used to represent the SCE is drain-induced-barrier-lowering

(DIBL) as predicted based on Equation 1 by $\Delta V_t/\Delta V_{ds}$. Because of the smaller V_{t_sat} for shorter length nFET devices, DIBL increases as channel length is reduced. This suggests that the channel built-in potential or the barrier height for electrons would be decreased when channel length is highly scaled, thus sharply increasing off-state leakage. Hence, to suppress SCE is an important task for manufacturing highly scaled device geometries.

$$DIBL = \frac{\Delta V_t}{\Delta V_{ds}} = \frac{V_{t\_lin} - V_{t\_sat}}{2 - 0.05} \cdot 1000 \ \ (\frac{mV}{V}) \qquad \text{Eq. 4-1}$$

The nFET V_{t_lin} and V_{t_sat} (extracted from electrically measured I_d-V_g transfer characteristics) for both SMT strained and non-strained control devices, is plotted as a function of transistor channel length in Fig. 4-3, and the calculated DIBL value is plotted in Fig. 4-4. Evidently there is a less roll-off of the threshold voltage for the split with SMT compared to the non-SMT strained split, and thus the DIBL is improved with the adoption of the SMT process. In other words, SMT provides nFET a higher immunity to short channel effect degradation because of the suppression of the threshold voltage roll-off and the improved DIBL performance. This is another major advantage of SMT process and rarely reported in previous literatures: the improved SCE immunity enables devices fabricated with reduced pocket implantation dose (reducing mobility degradation by coulomb scattering in the channel and silicon lattice damage) and/or increased N^+ S/D dosage (reducing conducting resistance). Both of these techniques, when properly employed, can further enhance the SMT strained nFET drive current under comparable SCE conditions for scaled devices fabricated nowadays.

4.2.3 New Findings on SMT: Retarded S/D Dopant Diffusion

One of the major challenges for MOSFET scaling in the sub-130nm regime is the formation of ultra-shallow junctions (USJs). The International Technology Roadmap for Semiconductors (ITRS) indicates that future generations of CMOS technology should have source/drain-extension junctions that are <15nm deep, with sheet resistance Rs~1000 Ω/ , in order to keep pace with historical improvements in high-performance logic devices [51]. This USJ requirement stems from the need to suppress short channel effects (SCE), and dictates a very limited thermal annealing budget to suppress dopant diffusion. On the other hand, to form the S/D regions of the MOSFET transistor with low sheet resistance, it is necessary to ensure low parasitic resistance by adopting a high annealing temperature to maximize S/D dopant activation. The S/D annealing in present day advanced CMOS processes is typically performed at a temperature of between 900°C and 1100°C to activate the impurity implanted regions. It has been found, however, that during annealing to activate the implanted dopants under such high temperature, the n-type dopants can diffuse into the crystalline semiconductor substrate such that the junction depth X_j increases [52, 53].

In order to understand the root cause of the improved SCE performance observed in Fig. 4-3 and Fig. 4-4, the same S/D dopant implantation conditions and activation annealing process steps were applied to the 6” Si (100) blanket wafers. N-type dopants arsenic and phosphorus concentration profiles inside silicon substrate were characterized by Secondary Ion Mass Spectrometry (SIMS) analysis. In the following study, the junction depth X_j (depth in nm) is taken to be the depth at which the n-type dopant concentration falls to 5E18 atoms/cm3 as measured from the SIMS profiles.

The SIMS analyses for arsenic and phosphorus implanted wafers, before and after activation anneal, are shown in Fig. 4-5 and Fig. 4-6, respectively. Clearly, S/D thermal activation annealing results in junctions deeper than the as-implanted profiles. The high (~5E15 atom/cm^2) S/D implant doses applied in modern MOSFET devices nowadays are sufficient to amorphize the silicon surface region [52]. Consequently, the subsequent S/D dopant activation annealing, adopted either by high-temperature furnace or RTA annealing, will likely result in substantial transient enhanced diffusion (TED) for both arsenic and phosphorus, culminating in junction depths considerably deeper than the as-implanted profiles, as observed in Fig. 4-5 and Fig. 4-6. The origin of this enhanced dopant diffusion could be the annealing of ion-implantation damage in the silicon substrate and/or the formation of an amorphous silicon surface layer by high dose Arsenic (> $2x10^{14}cm^{-2}$) implantation as suggested by R. B. Fair *et al* [52]. Nevertheless, it is also noted that reduced diffusion is seen for the arsenic and phosphorus profiles after activation annealing with the adoption of the SMT process. More specifically, there is a 10% and 13% X_j reduction for arsenic and phosphorus by the addition of the SMT process under the same implantation and activation annealing conditions. The SIMS data shown in Fig. 4-5 and Fig. 4-6 imply that similar retarded dopant diffusion can also be obtained for the lateral or longitudinal direction underneath the spacer and gate on a real nFET structure. Therefore, the lateral encroachment of n-type S/D implants under the gate can be reduced to form a shallower n-type S/D junction. Moreover, the undesired device SCE can also be improved by reducing the junction depth to lessen the drain voltage impact onto the channel potential, enabling the channel potential to be primarily controlled by the gate voltage V_g [11].

In brief, the SIMS results in Fig. 4-5 and Fig. 4-6 suggest that SMT process retards n-type dopant diffusion and lowered junction depth X_j can be obtained when compared with wafers processed without SMT. This is consistent with previous findings that SMT technique enables the fabrication of nFETs with improved SCE as demonstrated in Fig. 4-3 and Fig. 4-4.

4.2.4 nFET Device Strain Superposition by CESL and SMT

Strained channel effect from CESL techniques can be enhanced by increasing the CESL nitride thickness [32, 33], however, the Pre Metal Dielectric (PMD) gap-fill capability limit this approach at highly scaled geometries. Therefore, combining more than one stressor to further boost device current is another promising field to be explored. A further experiment in our work addressed this by fabricating nFET devices with a process integrating both SMT and CESL techniques. Performance gains from SMT and uniaxial 800Å tensile CESL nitride onto transistor channels are demonstrated to be additive as indicated by the nFET output characteristics shown in Fig. 4-7.

4.3 Point Defect Engineering by SMT

Point defects are defects which involve a few extra or missing atoms without an ordered structure as schematically drawn in Fig. 4-8 [54]. Vacancies are sites which are usually occupied by an atom but which are unoccupied. Interstitial defect or interstitial atoms are atoms fill an empty space of full crystal structure, and such atoms stay outside regular crystal lattice sites. A nearby pair of a vacancy and an interstitial is often called a Frenkel defect or Frenkel pair, which is formed when an ion moves into an interstitial site and creating a vacancy. A substitutional atom (defect) stays neither at a vacant site nor at an interstitial site, instead, it is incorporated at a regular atomic site in the crystal structure. [54].

Dopant diffusion can take place either by simple atomic exchange or atomic migration assisted by point defects like interstitials and vacancies. For example, if a neighboring atom moves to occupy the vacant site, the vacancy can move in the opposite direction to the site which used to be occupied by the moving atom. Therefore, point-defects (vacancies and interstitials) diffusivities and equilibrium concentrations can influence implanted dopant atom diffusion behavior [55]. Furthermore, different dopants have different diffusion mechanisms as reported by Plummer *et al*. For example, boron and phosphorus diffuse primarily by interstitials (f_I~1) [56-57], while arsenic diffuses via both vacancies and interstitials (0.2~ f_I~0.55) [56-57]. The concentration and specific properties of point defects (impurity atoms), their migration parameters, equilibrium or non-equilibrium conditions, and the atomic diffusion mechanisms determine the final diffusion profile in a solid state experiment.

The presence of additional strain in the silicon lattice can further change the point defect concentrations and hence interact with dopants diffusion mechanisms. In terms of forming a lowest energy stable configuration, a silicon crystal with a vacancy prefers the lattice constant to be reduced [58-59]; while a silicon crystal with an interstitial prefers an increased lattice constant [18, 60-62]. This indicates that, in the presence of compressive strain, the vacancy equilibrium concentration will be increased and the interstitial equilibrium concentration will be decreased. On the contrary, tensile strain will increase the equilibrium interstitials and reduce the vacancy population [18].

Moreover, change in silicon point-defect diffusivity with biaxial strain is shown in Fig. 4-9. Vacancy diffusion is increased in compressive strain and is retarded with tensile strain. For interstitial diffusivity, it is increased in tensile strain and is retarded with compressive strain [18, 62]. Therefore, one can expect the typical dopant diffusivities relationship with respect to tensile and compressive strain as shown in

Fig. 4-10. A large compressive strain results in a decrease in boron and phosphorus diffusivities, whereas tensile strain results in an increase in their diffusivities. Arsenic dopant diffuses is mediated by both vacancy and interstitial mechanisms as $0.2 \sim f_I \sim 0.55$. Hence, arsenic diffusion is enhanced under both compressive and tensile strains [61], and the increase in vacancy and/or interstitial equilibrium concentrations results in an enhancement of arsenic diffusion.

The typical memorized strain behaviors and their polarities in nFET are illustrated in Fig. 4-11. After typical SMT nitride stress changing towards a more tensile state upon activation annealing [63-65], the S/D portion is left with the residual compressive. By qualitatively applying aforementioned point defect diffusion concept, one can expect that the phosphorus diffusion is retarded in the presence of such compressive strain as it is primarily an interstitial diffuser $f_I \sim 1$. Thus the retarded phosphorus S/D diffusion in Fig. 4-6 after SMT process can be explained because both the memorized compressive strain reduces interstitial point defect concentration and its diffusivity as shown in Fig. 4-9 and Fig. 4-10.

Nevertheless, the excess interstitials or vacancy generated after the conventional S/D implant lead to the enhancements in arsenic diffusivities, and is against the SIMS experimental results in Fig. 4-5. Since arsenic can diffuse via both interstitials and vacancies ($f_I \sim f_V \sim 0.5$), therefore, some other mechanisms should be used to explain this discrepancy due to the complex arsenic diffusion kinetics involved. If the arsenic interstitial diffusion mechanism dominant the diffusion process as proposed by $As + I \rightarrow As_i$, the diffusion profile can also be retarded under compressive strain state as is the case of phosphorus.

In other respects, thermal annealing of wafers implanted by high dose Arsenic ($>2\text{x}10^{14}\text{cm}^{-2}$) is believed to produce large transient concentrations of point defects that may continue to exist for the duration of the annealing [52]. Such point defects are generally highly mobile at high temperatures. They also increase the mobility of the atoms of the host lattice, and thus cause of solid state diffusion. At the same time, point defects may effectively lengthen the duration of the diffusivity transient region and give rise to lowered dopant diffusivity upon thermal activation, generating the aforementioned retarded n-type dopants diffusion behavior in the silicon substrate itself [66]. In addition, it has also been proposed and experimentally confirmed that point defects do retard n-type dopant diffusion upon thermal annealing [67]. Such point defect generated via either shallow end-of-range damage or small clusters of point defects during the annealing can dominate the whole TED process. For example, Ning Kong *et al.* [67] observed retarded arsenic diffusion by specific implant conditions which created a vacancy-rich region overlapped with the principle dopant region. Two possible mechanisms, interstitial-vacancy recombination and dopant clustering, have been proposed and analyzed based on their diffusion retardation [67] as illustrated in Fig. 4-12. The first model claims that once the more energetically favorable interstitial-vacancy are formed, it will come to IV recombination and point defect annihilation. Consequently, the arsenic diffusion can be retarded as less point defects contribute to arsenic diffusion. In the second model proposed, Ning Kong *et al.* assert that the heavy arsenic vacancy cluster atoms diffusion can also be retarded during the annealing process as shown in Fig. 4-12 due to their heavy mass. Because arsenic atoms can be trapped together in the form of As_nV_m, such clustered atoms tend to have a lower diffusivity.

With similar concept, the retardation of n-type dopant diffusion by the SMT process during high-temperature thermal annealing demonstrated in our work is likely caused by the generation of point defects in the S/D regions of the transistor after implantation and capping nitride deposition. Such excess interstitials or point defects form a retarding gradient during activation annealing, and would opposes n-type dopant diffusion. Hence, retarded and shallower n-type dopant diffusion profile is formed after SMT processing, thus improves nFET SCE characteristics. Finally in this paper, the fabrication of ultra shallow junctions on nFET devices is demonstrated with proposed SMT engineering and associated point defect manipulation approaches.

4.4 Hot Carrier Stress of SMT Strained nFET

The degradation of nFETs due to hot-carrier injection has been studied for over twenty years [68-71]. As the device feature size is highly scaled down, the lateral electric field increases when the applied operation voltage is kept the same. The lateral electric field neighboring the drain accelerates the carrier and generates hot carrier with high kinetic energy when device is operated in the saturation region.

Moreover, this hot carrier problem is especially a concern for nFET transistor since electron mobility is higher than that of hole. Because, these hot electrons will generate new electron and hole pairs as they collide with the silicon lattice. Some of these hot electrons will be injected into the gate oxide and being trapped in the oxide layer, making increased oxide fix charge and threshold voltage permanent shift. Some of these hot electrons will also travel through the oxide layer, resulting in increased gate leakage current or deteriorate SiO_2/Si interface by increasing the number of interface state defects. On the other hand, the generated holes can flow into the substrate in the form of substrate current I_{sub}. If this substrate current is large enough to make the source and substrate junction being forward biased, then the parasitic bipolar transistor in a typical MOS transistor will be triggered and leads to MOSFET malfunction. Thus, the nFET device's hot carrier reliability is normally stressed at maximum substrate current condition. Consequently, it becomes an important task to suppress such hot carrier induced device degradation for highly scaled devices fabricated with strained silicon technology.

Uniaxial strained nFET with worse hot carrier stress results have been reported by Lu *et al* [72]. Their CESL nitride layer used to induce channel strain for mobility enhancement was deposited by low-pressure chemical vapor deposition (LPCVD). They found that the nFET hot-electron degradation is worse with the addition of the SiN capping. Furthermore, they proposed that both the bandgap narrowing caused by the channel strain and the abundant hydrogen species from the precursors of CELS nitride deposition can contribute to the aggravated hot-electron effect.

The nominal nFET devices used here for our hot carrier stress characterization were width 10 μm and length 0.5μm.The nFET devices were stressed at V_d=4.5V with V_g set at a maximum substrate current condition. The shift of threshold voltage (ΔV_{th}), and the degraded peak value of transconductance (ΔG_m) as a function of stress time, are shown in Fig. 4-13 and Fig. 4-14, respectively. The SMT strained nFET with plasma enhanced CVD (PECVD) nitride has a relatively worse hot carrier stress results as compared to the non-SMT control device. These results clearly indicate that the memorized SMT strain in the channel affects the generation of channel hot electrons and their associated impact ionization process with the following possible mechanism behind.

The increased electron mobility by SMT process can also directly increase electron impact ionization rate [73] and therefore aggravate hot carrier stress effect. Besides, this hot carrier stress results can also be correlated to the gate edge damage induced by the SMT strain [74]. Such damaged gate is subsequently prone to cause an anomalously high gate tunneling current during transistor operation [74]. Moreover, the enhanced nFET's hot carrier stress reliability degradation may also be correlated to the generation of the interfacial state in the channel or around the gate edge, which is caused by enhanced wafer bowing exerted by PECVD SMT nitride's stress change upon thermal annealing [65].

4.5 Summary

We have investigated the effects of utilizing tensile strained nitride deposited as a SMT stressor to enhance nFET device performance. The SMT technique improves nFET performance and transconductance as expected owing to owing to the enhancement of electron mobility by the memorized strain for scaled nFET devices.

In addition to the aforementioned nFET mobility improvement, the SMT nitride cap layer modifies the formation of the defects created by ion implantation, and hence alters the diffusion kinetics of the implanted n type dopants. Therefore, the SMT strain technique enables better short-channel-effect control, suppresses V_t roll-off, and improves DIBL owing to a reduced junction depth X_j as demonstrated by the SIMS profiles.

Finally, when combined with the strained CESL technique, SMT provides supplementary strain to the n-channel, elucidating the prospects on strain silicon processes integration for aggressively scaled CMOS fabrication.

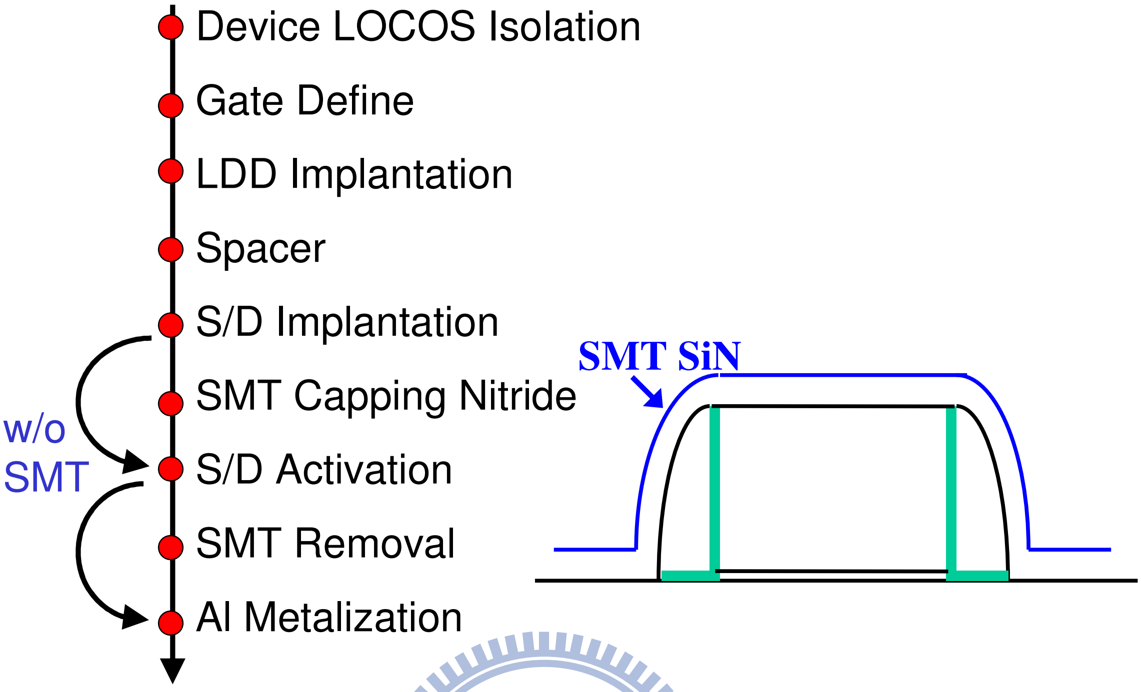


Fig. 4-1. SMT Process Flow. The SMT process deposits a nitride layer over the gate before annealing. The presence of this nitride layer during the subsequent S/D annealing and recrystallization of amorphized regions, acts to increase tensile strain in the NMOS channel. Then, this nitride layer is removed, but the strain by the nitride cap introduced during the annealing is retained or “memorized”.

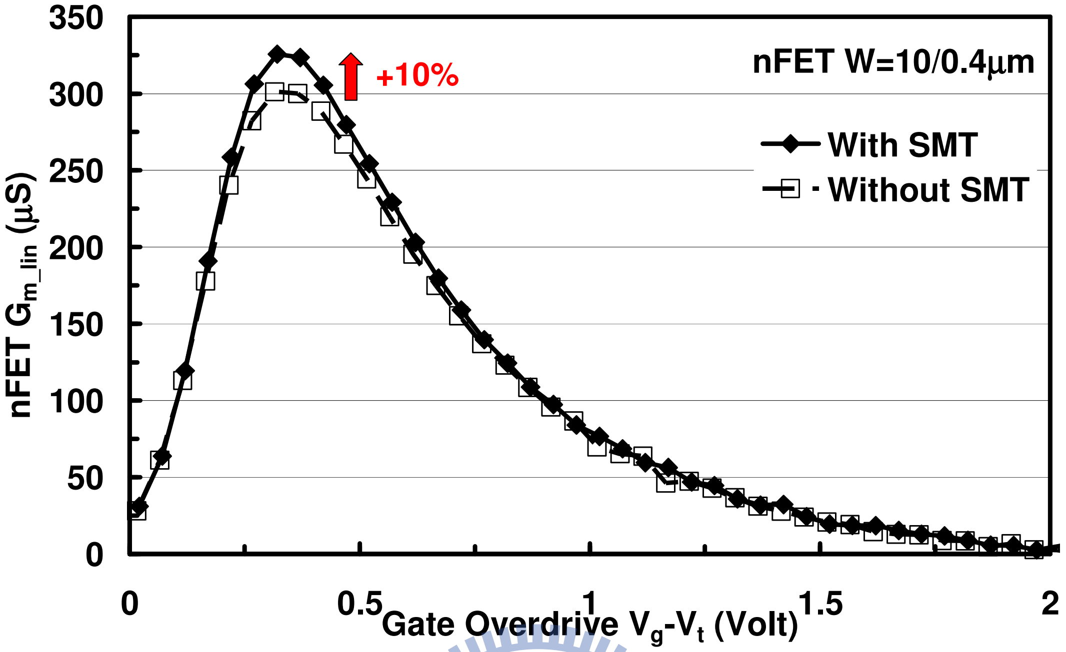


Fig. 4-2 Measured 10x0.4µm nFET G_{m_lin}. There is a ~10% improvement in the peak value of G_{m_lin} with the adoption of SMT. The nFET S/D is implanted with Arsenic.

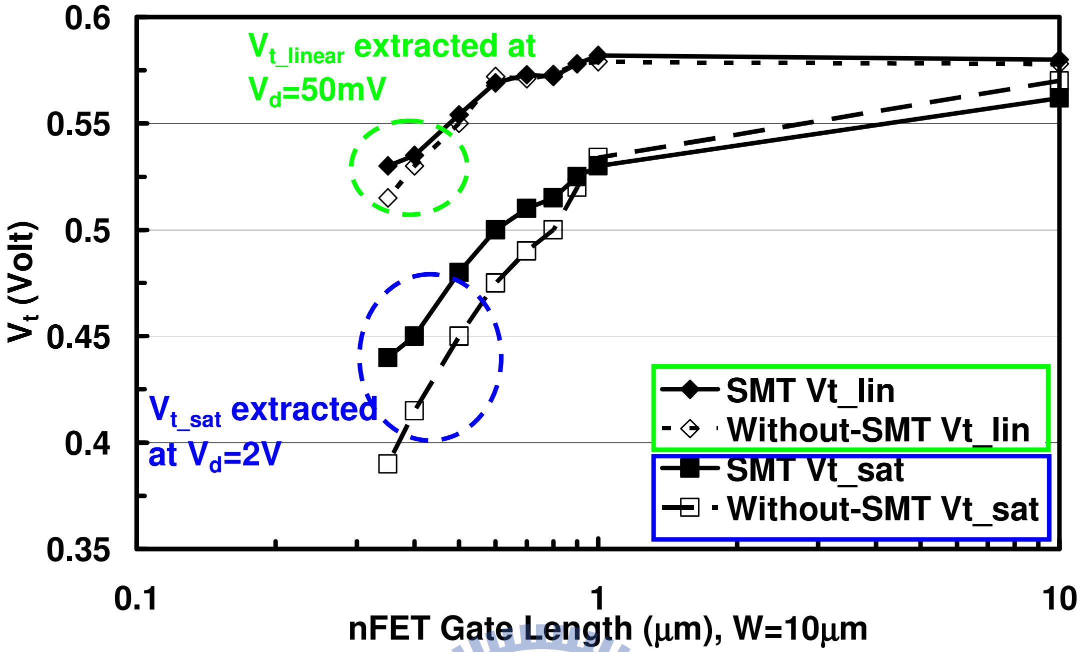


Fig. 4-3 Measured W= 10μm nFET V_t as a function of channel length; SMT improves nFET SCE, as indicated by less V_t roll-off.

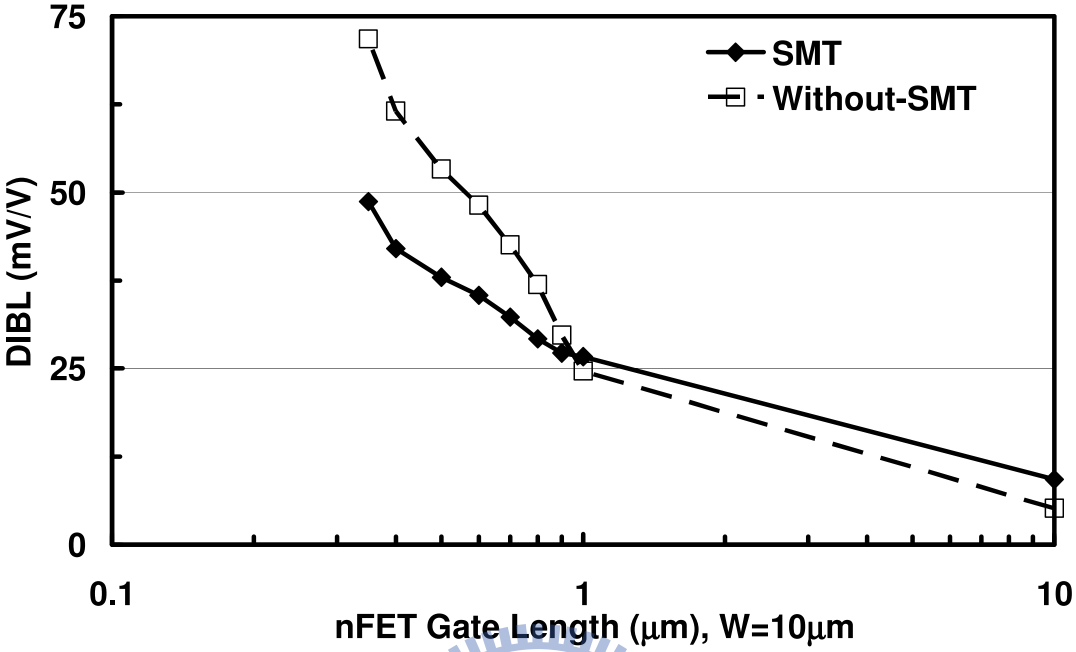


Fig. 4-4 Calculated W=10 µm nFET DIBL. Short channel effect improvement by SMT is indicated by improved short channel DIBL.

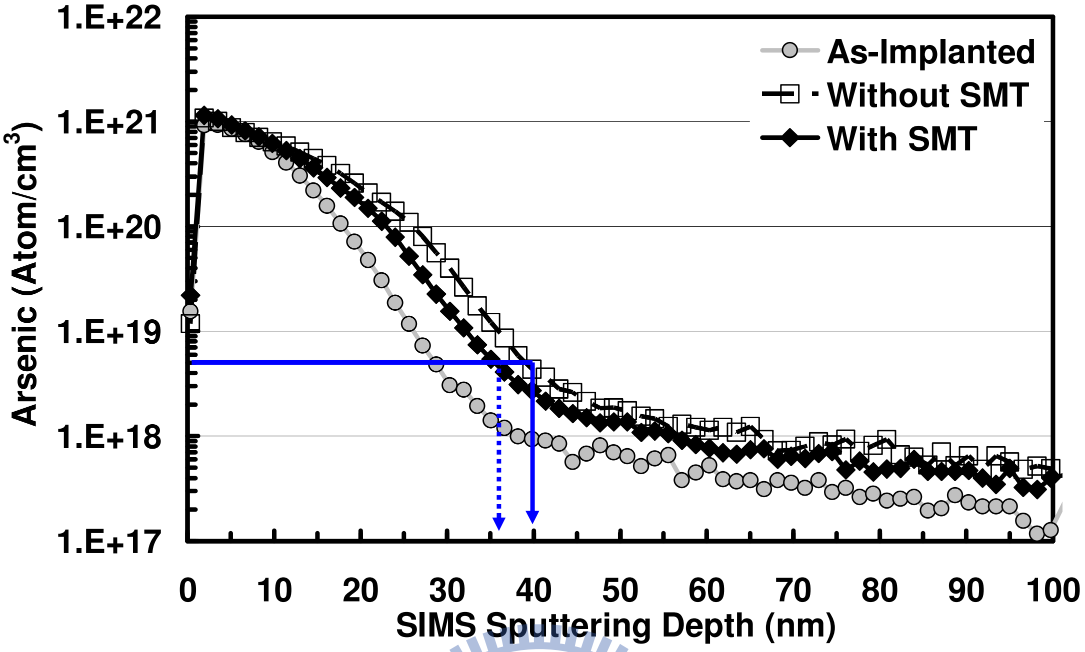


Fig. 4-5 SIMS analyses for arsenic implanted wafers. The SIMS diffusion profiles suggest that SMT retards arsenic diffusion upon thermal annealing, resulting in 10% reduction of X_j.

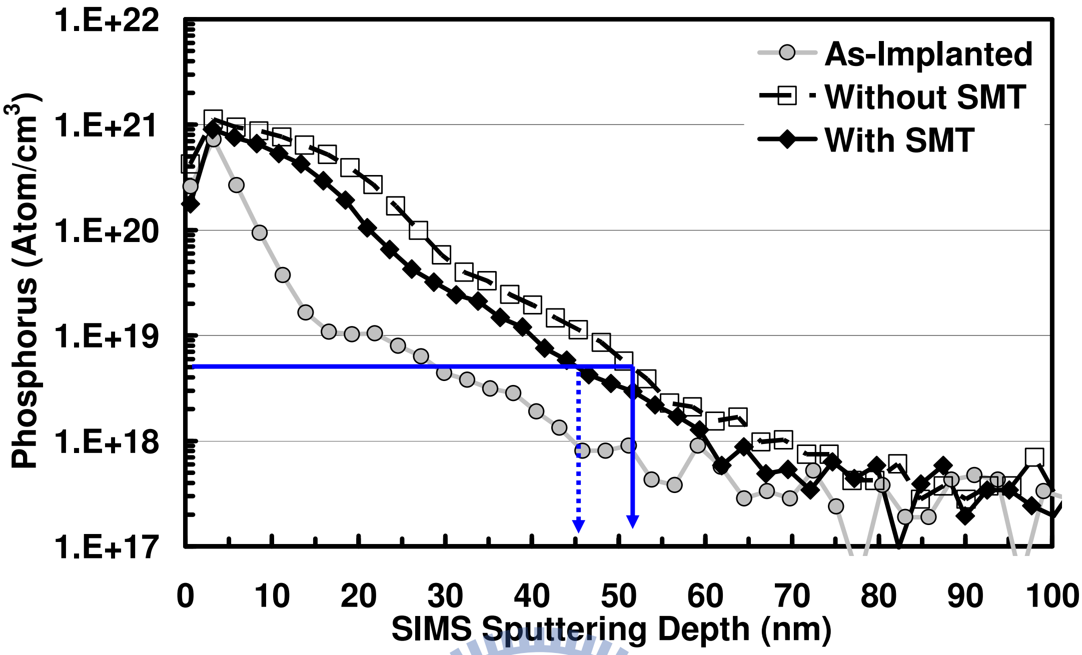


Fig. 4-6 SIMS analyses for phosphorus implanted wafers. The SIMS diffusion profiles suggest that SMT retards phosphorus diffusion upon thermal annealing, with 13% reduction of X_j.

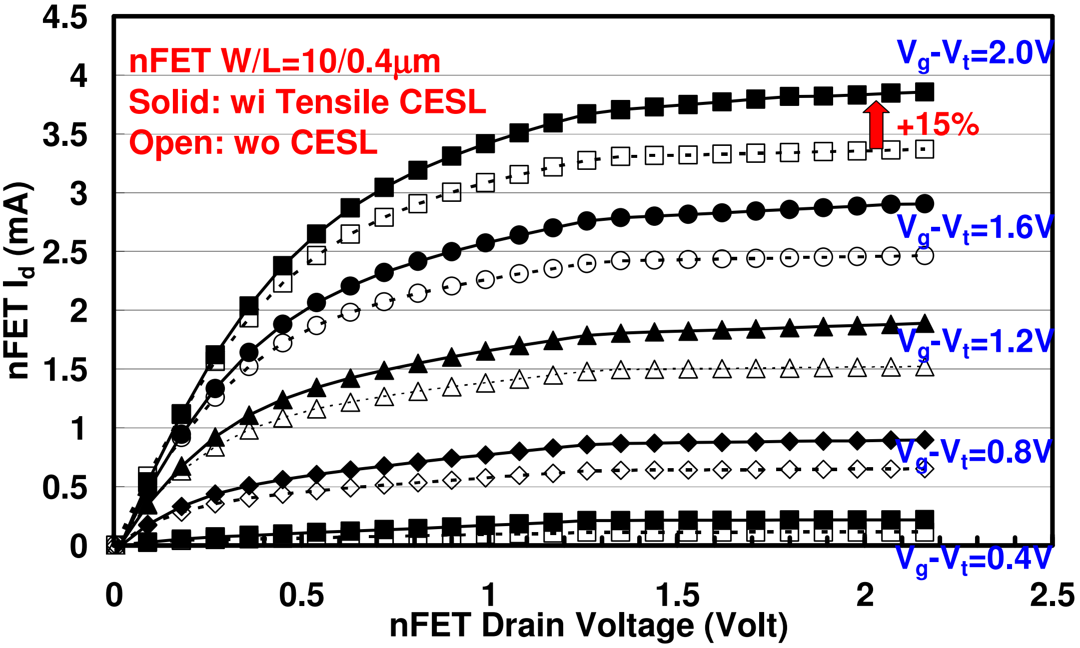


Fig. 4-7 Measured 10x0.4µm nFET I_d - V_d output characteristics. There is a further ~15% I_d improvement with the adoption of 800A tensile CESL SiN at gate overdrive = 2.0 V and V_d=2.0V.

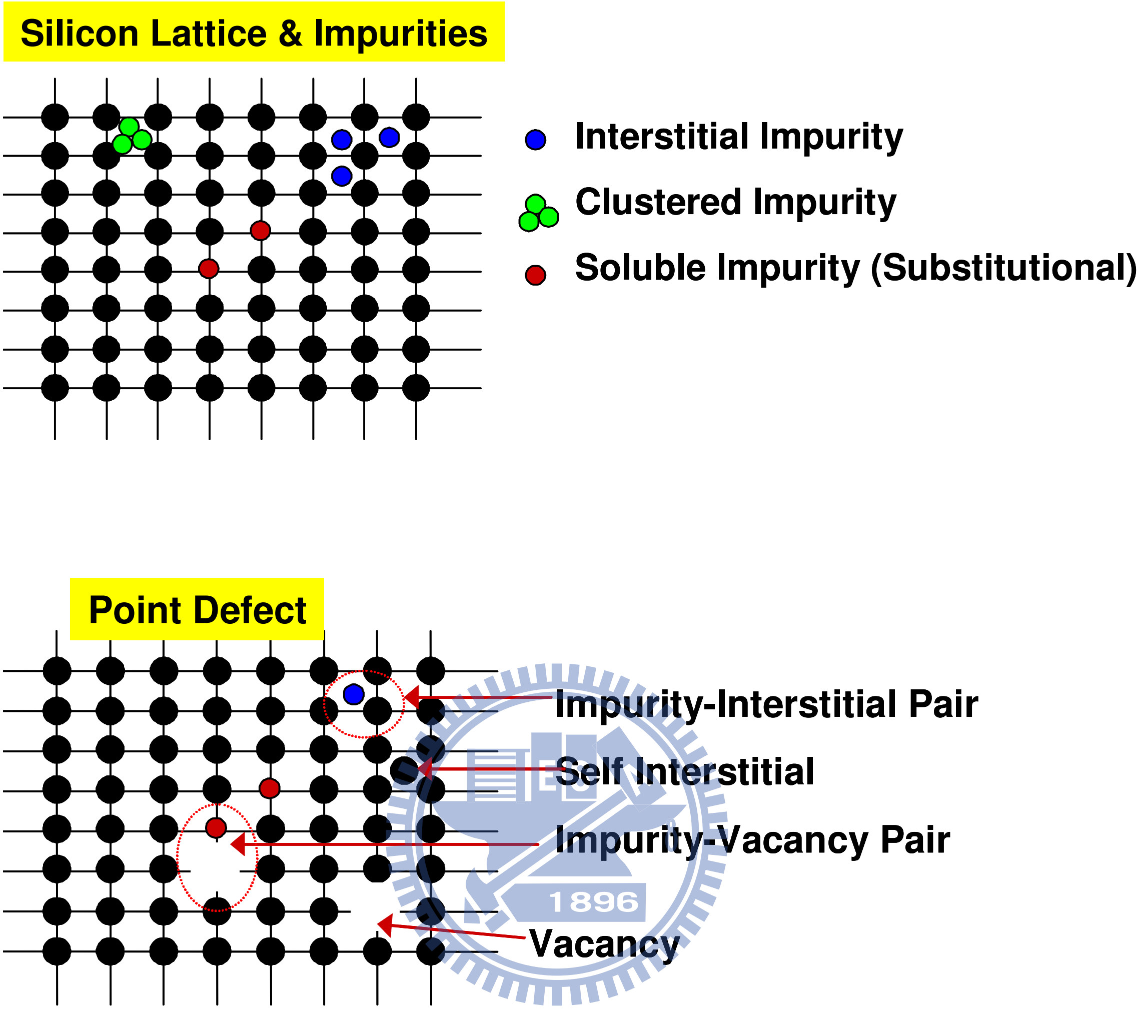


Fig. 4-8 Schematic illustration of typical simple point defect types in a monatomic solid.

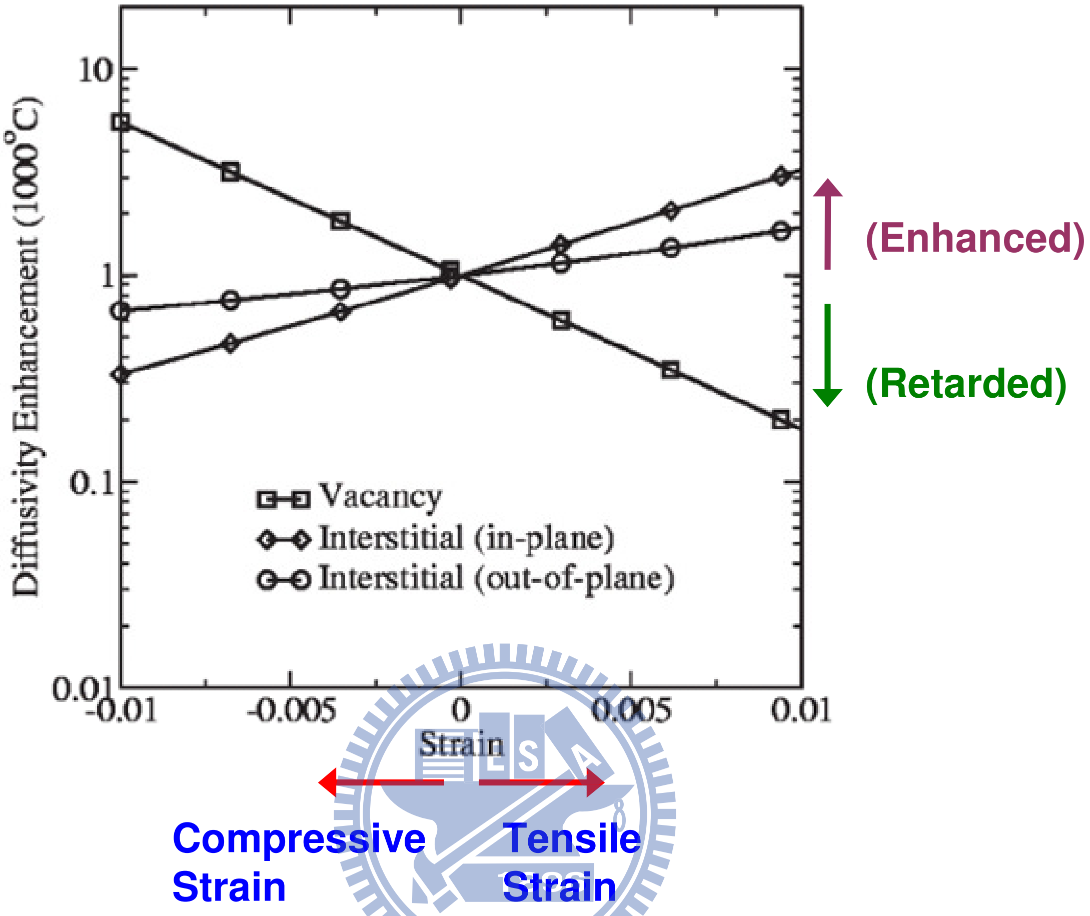


Fig. 4-9 Change of point defect diffusivity upon strain. The vacancy diffusivity is enhanced under biaxial compressive strain, while the interstitial diffusivity is enhanced under biaxial tensile strain [18, 62].

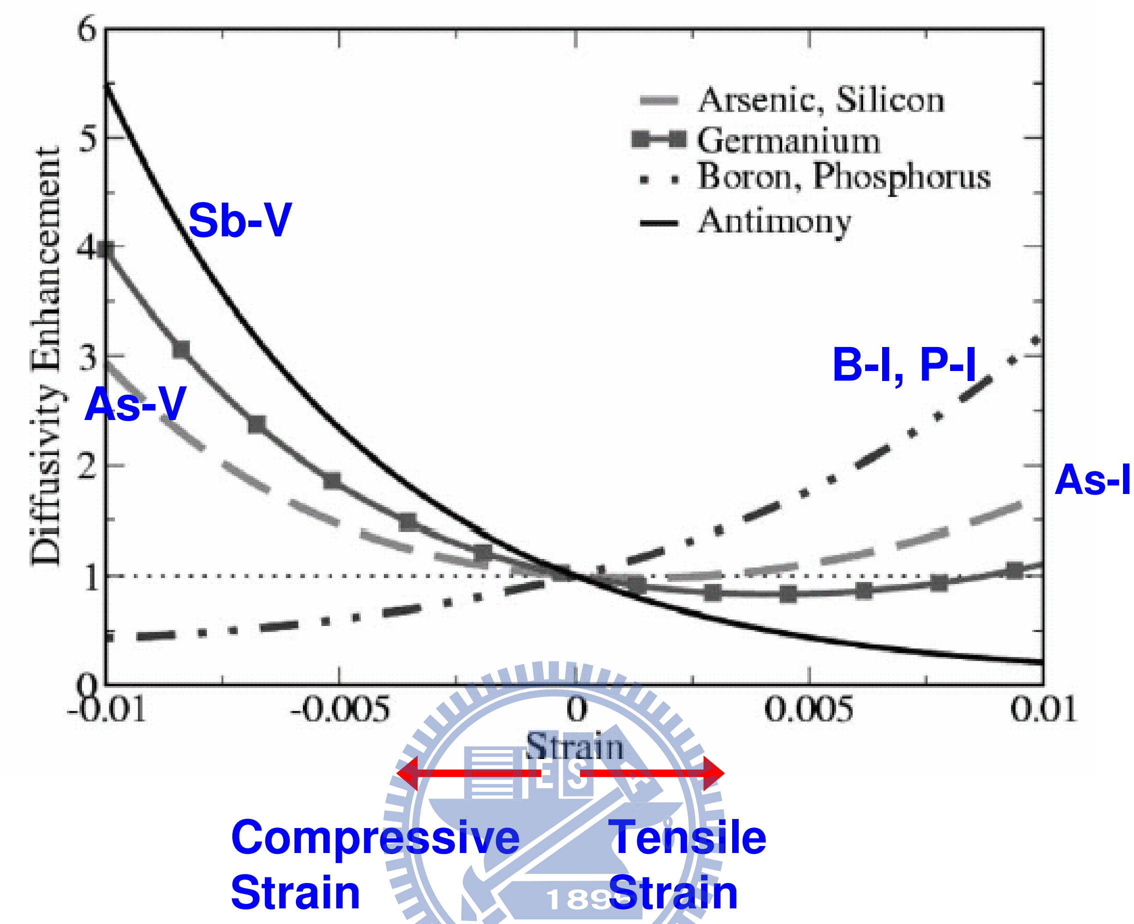


Fig. 4-10 Change of specific point defect diffusivity upon strain. Compressive strain enhances vacancy diffusers whereas interstitial diffusers are enhanced by tensile strain. A large compressive strain clearly results in a decrease in boron and phosphorus diffusivities, whereas tensile strain results in an increase in their diffusivities [56-57]. Antimony is a dominant vacancy diffuser [18, 56]. Hence, it shows enhancements under compressive strain and a retarded diffusion under tensile strain.

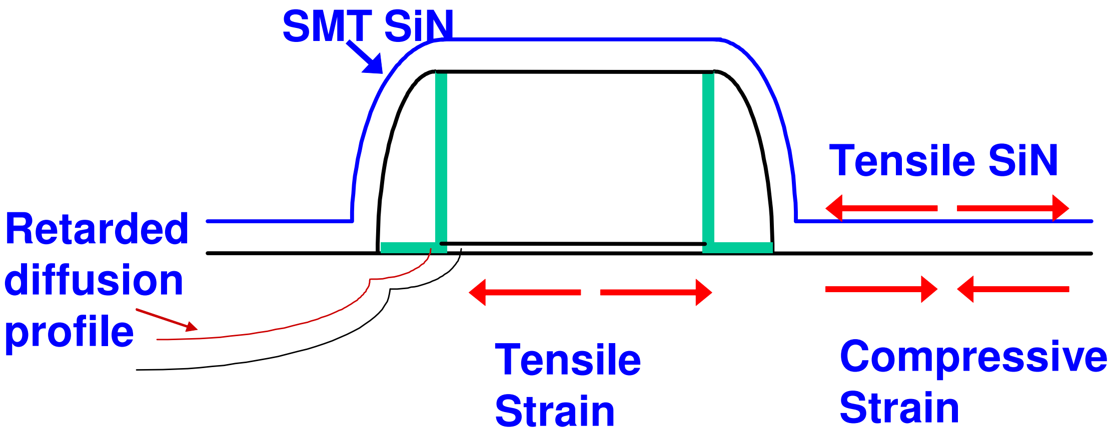


Fig. 4-11 Schematic drawing showing compressive strain is induced at source/drain after SMT nitride stress change towards a more tensile state upon activation annealing [63-65]. The retarded n-type dopant diffusion profile with reduced junction depth is confirmed by SIMS results in Fig. 4-5 and Fig. 4-6.

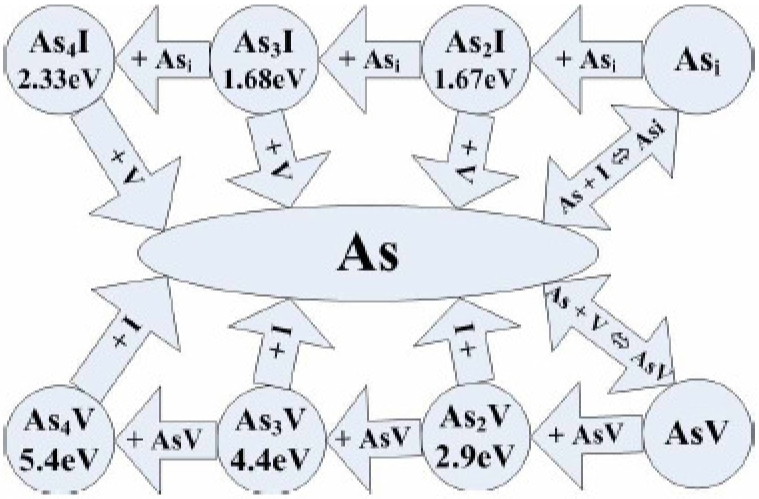


Fig. 4-12 Schematic drawing showing physical models about arsenic pairing and clustering [67].

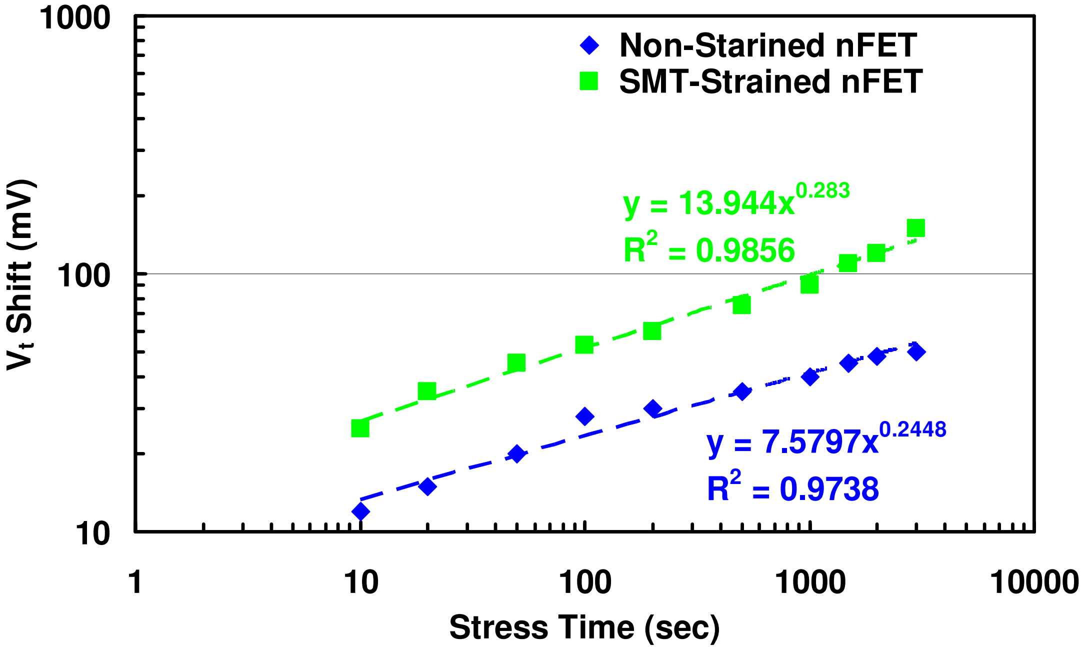


Fig. 4-13 Measured 10x0.5µm nFET hot carrier degradation Vt characteristics. The nFET was stressed at V_{ds}=4.5V with Vg biased at a maximum substrate current condition. SMT stained nFET behaves a larger threshold voltage shift.

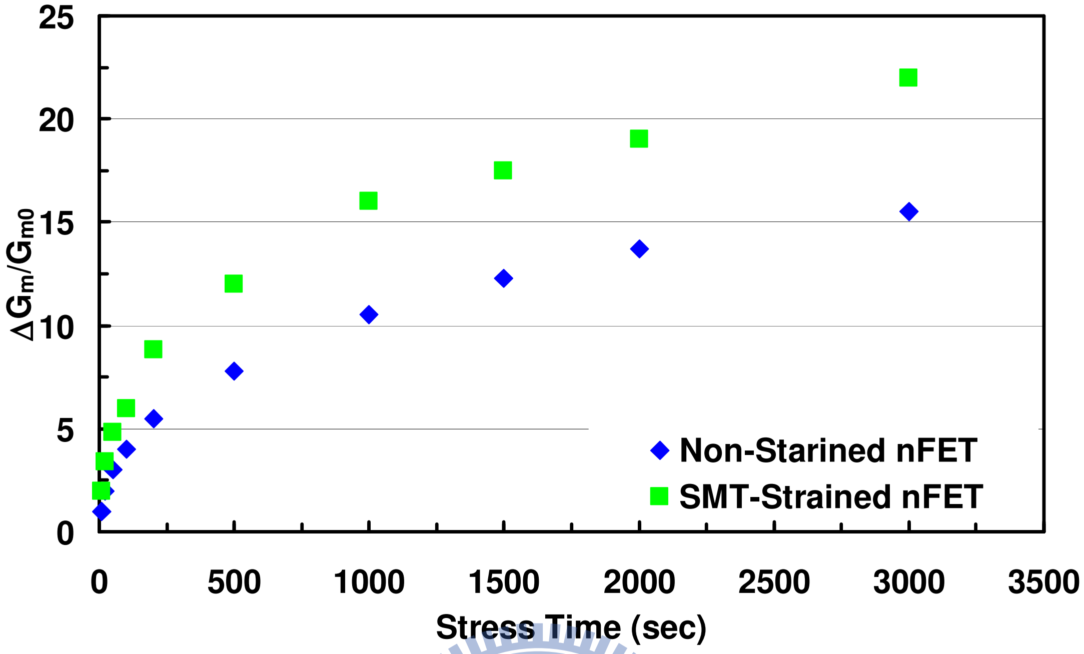


Fig. 4-14 Measured 10x0.5μm nFET hot carrier degradation G_m degradation characteristics. SMT stained nFET behaves a larger G_m degradation.

Chapter 5

Characterization of nFET Devices with Different SMT Nitrides

5.1 SMT Devices with Two Different Kinds of Nitride

SMT provides additional strain to the n-channel, creating prospects for using strained silicon technology to achieve the device performance targets of aggressively scaled CMOS design rules. Therefore, to understand and maximize the benefit obtained from the highly strained nitride on the silicon channel, SMT strained nFET device performance including threshold voltage, transconductance, electron carrier mobility and the associated interface state density D_{it}, have been characterized by electrical measurement techniques with two different kinds of SMT nitrides. Moreover, the wafer bowing measurements are presented to quantify the evolution of wafer curvature throughout the entire SMT fabrication process.

The devices reported in this chapter were fabricated in the National Nano Device Laboratories (NDL) on 6” (150 mm) Si wafers using a conventional MOSFET process flow including LOCOS isolation, gate oxide, spacer, source-drain (S/D) implantation, thermal annealing, and Al metallization processing as listed in Fig. 5-1. All the nFET devices characterized in this study had 3nm thick gate oxide thermally grown in a vertical furnace, and a 150nm thick N^+ poly-Si layer as the gate electrode. After gate dielectric and gate electrode deposition and patterning, desired ion species were implanted through a 100Å plasma enhanced (PE) TEOS screen oxide to form the N^+ S/D region. This screen oxide was employed to reduce the N^+ region junction depth.

Prior to the S/D activation annealing, a SMT process flow for strained nFET similar to that proposed by Chen *et al* [19] was employed. For the SMT strained nFET nitride capping layer, a ~110 nm tensile nitride film was deposited on the whole device structure including poly gate and S/D region. The SMT strain generation by two different Chemical Vapor Deposition (CVD) processes are comparatively studied. The first type of nitride film was deposited under a high temperature 780°C (hereafter denoted as HT-SiN) LPCVD system using a batch type furnace tool, and thus the HT-SiN nitride film coated both the front and back sides of the wafer. On the other hand, the second nitride film was deposited at a relatively low temperature (hereafter denoted as LT-SiN) 300°C process condition using a single wafer Plasma Enhanced CVD (PECVD) system with NH_3, N_2, SiH_4 as main reaction gasses, thus depositing the LT-SiN nitride film only on the front side of the wafer.

After S/D annealing, a permanent stress was created in the channel region, resulting in strain-enhanced carrier mobility. Afterwards, to facilitate metallization, the SMT nitride films were removed using a wet chemical etchant H_3PO_4. Passivation was performed by depositing a 300nm thick TEOS oxide, followed by patterning and etching of contact holes, and finally metallization. After completion of the metallization processes, the wafers were annealed at 400°C in a forming gas ambient and the electrical characteristics of the fabricated devices were measured. The nominal device dimension evaluated was a gate length of 0.4 μm and width of 10 μm. The device characteristics were measured using a semiconductor parameter analyzer (HP 4156A, plus Agilent ICS Software). The threshold voltage was determined by the constant drain current method when drain current equals $10^{-7} \cdot W/L$ (Amp) [50]. The linear and saturation mode threshold voltages, V_{t_lin} and V_{t_sat} , were extracted at drain voltages of 50 mV and 2V, respectively.

5.1.1 Study of Nitride Film Thickness Shrinkage

The thickness of LT-SiN and HT-SiN films was measured by a SOPRA ellipsometer. Generally speaking, as compared to LPCVD films, PECVD materials have much lower thermal budget and their film properties can be tuned and controlled over a much wider range by varying appropriate process parameters. In this work, the LT-SiN is formed at 300°C by PECVD with a mixed gas containing SiH4, NH3 and N2, while the HT-SiN is formed at 780°C in a LPCVD furnace with a mixed gas containing SiH4Cl2 and NH3.

Table 5.1 depicts the LT-SiN and HT-SiN nitride film thicknesses before and after S/D annealing. The typical as-deposited nitride thickness (measured on unpatterned monitor wafers) was ~ 1150Å and ~ 1100Å for for LT-SiN and HT-SiN, respectively. However, after annealing, the thicknesses of both of nitride films reduced to ~ 1050Å. The calculated thickness shrinkage rate after annealing for the LT-SiN is ~12% as compared to ~ 3% for that of the HT-SiN.

5.1.2 Study of Nitride Film Stress Change

For strain nitride for CESL application [32, 33], the magnitude of the thin nitride film stress is one of the key properties that can determine the device performance and reliability lifetime. The deposition tool and process parameters largely determine the type and level of nitride stress that develops on the wafer.

Measurement of stress in thin film generally begins by preparing a blanket wafer with a known initial wafer bow. After depositing a thin film on the blanket wafer, the wafer bowing is measured again to see if it is changed by the deposition. The stress of the film then can be calculated by Stoney's equation (shown in equation 5-1) using the difference in the amount of wafer bowing before and after film deposition [75]. Wafer bowing is measured using the reflection angle of the incident laser on the film as illustrated in Fig. 5-2. The laser is incident to the film with a known incident angle, and the reflection angle of the light is measured using a location-sensitive photo diode. The change in reflection angle with respect to laser scan distance yields the curvature of the wafer.

The sign of stress (i.e., tensile or compressive) is determined by the shape of the wafer bow or curvature. When the film is stretched to fit the substrate, the film’s tensile stress causes compressive strain on the underlying substrate. When the substrate is allowed to bend, the film makes the wafer concave. When the film is squeezed to fit the substrate, the film has a compressive strain and the substrate is strained in the tensile mode; the wafer is then convex when the substrate is released.

Because of its simplistic nature, the curvature measurement scheme is traditionally the most widely acknowledged method, and the film stress is determined by converting the measured curvatures using Stoney's equation [75] as formulated in equation 5-1. A highly stressed film will cause wafer deformation as schematically drawn in Fig. 5-2.

$$\sigma = \frac{E}{6\cdot(1-\upsilon)}\cdot\frac{t_s^{\,2}}{t_f}\cdot\left(\frac{1}{R_1}-\frac{1}{R_2}\right) \qquad \text{Eq. 5-1}$$

where

σ is the film normal stress, E is the elastic modulus of substrate,

t_s is the substrate thickness, t_f is the film thickness, υ is Poisson’s ratio of the substrate, and $(1/R_1-1/R_2)$ is the local curvature caused by intrinsic stress before and after film deposition.

SMT nitride stress and stress change behavior before and after annealing are measured and shown in Table 5.2. All stress levels are positive, indicating tensile stress. The LT-SiN shows a relatively large increase in tensile stress ~ 700 MPa, (from 360 MPa to 1050 MPa), while the HT-SiN shows a much smaller 60 MPa stress increase (from 1010 MPa to 1070 MPa).

5.1.3 Study of Nitride Film Chemical Properties

Fourier transformation infrared reflectivity (FTIR) spectra were also measured to provide information concerning the bonding structure of the SMT nitride films. The FTIR spectra were obtained by subtracting the absorbance obtained before deposition from that measured after deposition using a Bio-Rad QS-300 spectrometer. The representative SiN layer FTIR spectra after CVD deposition, and after S/D annealing are shown in Fig. 5-3. The FTIR analysis results obtained on LT-SiN and HT-SiN films are shown in Fig. 5-3b and Fig. 5-3c, respectively. The chemical bonding present in CVD deposited nitrides has been characterized by FTIR according to Lucosky *et al* [76]. The dominant spectral feature in each film is at approximately 820~850 cm^{-1} and is associated with the Si-N asymmetric bond-stretching mode. The remaining features in these spectra are due to vibrations involving Hydrogen atom motion: (i) a N-H bond stretching mode at ~ 3350 cm^{-1}; (ii) a N-H bond bending mode at ~ 1150 cm^{-1}; and (iii) a Si-H bond stretching mode at ~ 2150 cm^{-1}, respectively. The main stretching vibrations of Si-N, N-H, and Si-H bonds in nitride are illustrated in Fig. 5-3a.

The as-deposited LT-SiN spectrum of Fig. 5-3b shows substantially stronger Si-H and N-H bonds as compared with the as-deposited HT-SiN spectrum of Fig. 5-3b. This results from the higher deposition temperature of the LPCVD process.

Hydrogen atom is more mobile and active, and less likely to remain in a stable Si-H or N-H bond.

The FTIR spectra for nitrides with LT-SiN and HT-SiN after S/D annealing are also shown in Fig. 5-3b and Fig. 5-3b, respectively. We observe that the intensity of Si-H and N-H bond peaks become negligible after annealing, suggesting hydrogen removal during the S/D activation anneal.

The intensities of the Si-H and N-H FTIR absorption lines are directly related to the hydrogen contents of the nitride layer, thus providing a fast and non-destructive method of determining the hydrogen concentration. The hydrogen concentration can be correlated to the intensities of the absorption lines [77]. The areas of N-H and Si-H absorption bands can be used to estimate the N-H bonds density per centimeter square. Later, via the measured nitride thickness, the N-H and Si-H bond concentrations per centimeter cube can be calculated as listed in Table 5.3 and Table 5.4. The total hydrogen concentration is therefore the sum of the N-H and Si-H concentrations as shown in Fig. 5-4. With typical nitride density (1.05E23 atom/cm^3) referenced, one can further estimate the hydrogen atomic (%) in these nitride films as shown in Fig. 5-4 as well.

The change is much more dramatic in LT-SiN because of the large amount of hydrogen initially incorporated during deposition of the PECVD film. The magnitude of the reduction in hydrogen content is directly responsible for the irreversible increase in tensile stress of the film due to the volume reduction as reported by Hughey *et al*. [63, 64, 78]. As observed in Fig. 5-3b, the FTIR spectra show a reduction of the N-H and Si-H bond peaks which is in agreement with this scenario. This phenomenon is typical of the LT-SiN formed by PECVD, and is well correlated to the LT-SiN film's massive stress change towards a tensile equilibrium state after hydrogen removal as shown in Table 5.2. It appears that the smaller stress change observed in Table 5.2 for HT-SiN layer, is mainly due to less initial hydrogen concentration during nitride layer deposition, and less hydrogen bonding change as indicated by the FTIR spectra of Fig. 5-3c.

5.1.4 Device Wafer Bow Characterization

Wafer curvatures or bowing measurements were performed using a Tencor FLX-2320 thin-film stress metrology instrument and film stress was subsequently obtained by transforming this wafer curvature using a conversion equation such as Stoney's formula [75]. For highly scaled modern semiconductor devices, such wafer level thin-film residual stress characterization is crucial, and the strain level remaining in the device itself is a major factor in determining the final electrical characteristics, yield, and reliability. This curvature measurement technique was also applied to the patterned electrical wafers to characterize the generated SMT strain.

Wafer bowing refers to the curvature of the wafer substrate, i.e. the height difference between the center and the edge of the wafer. A graphical description of wafer bow is shown in Fig. 5-2 and formulated in equation 5-2.

$$Bow = \left(|R| - \sqrt{R^2 - r^2}\right) \cdot Sign(R) \qquad \text{Eq. 5-2}$$

where R is the radius of curvature, and r is the scan distance.

When applying strain silicon technology to increase carrier mobility, wafer bow is a key metrology item for chipmakers as they transition to production on 300 mm or 450 mm wafer sizes, since larger wafers can bow, or bend, more than twice as much as the 150 mm wafers used here on deposition of a given film. The larger global wafer bow accumulates larger stress on the wafer, and increases the device's susceptibility to film cracking, de-lamination and electrical performance degradation, which would all negatively impact transistor structural integrity and can result in serious circuit yield loss.

Wafer bow can be modulated by the presence of a stressed thin film deposited on the wafer. Therefore, wafer bow is a quantitative index for the bending (curvature) of a wafer during the device fabrication processes. For instance, wafer bow due to electroplated copper films with different copper seed layers has been studied [79]. Furthermore, previous observations by Yu *et al.* [80] have also revealed a correlation between the wafer bow height and defect injection into strained epitaxial SiGe upon activation laser and RTA annealing. This study further indicates that the final wafer curvature was determined by the initial shape of the wafer, plus the deformation caused by the subsequent fabrication processes.

In order to successfully apply SMT technology, it is essential to utilize wafer bow measurement to understand SMT strain behavior, and to minimize the SMT strain induced defect creation and/or strain relaxation after annealing when large SMT strain is present. Since any wafer will have some initial curvature and surface irregularities are caused by several front-end processing steps occurring before SMT, a curvature measurement was performed before deposition of the SMT capping nitride film. Further curvature measurements were performed after each of the subsequent SMT process steps such as nitride film deposition, RTA for S/D annealing, and nitride strip. These data were compared to the initial curvature. Fig. 5-5 plots the increased wafer bow at each SMT process step, and demonstrates that wafer curvature represents the SMT strain generated in the fabricated silicon device. The LT-SiN SMT film stress causes significantly higher levels of curvature in device wafers after SMT deposition and the subsequent S/D activation annealing. It indicates that wafers processed with SMT using LT-SiN are significantly warped compared to those processed with HT-SiN. This is likely due to the fact that HT-SiN is processed with a double sided coating furnace batch system, so the stress from the HT-SiN films on the front and back sides of the wafer cancel each other out to produce much less macroscopic wafer bending. This does not imply that there is no SMT strain generated by HT-SiN. However, with SMT processing using LT-SiN, the wafer bow is increased further after S/D activation annealing due to the large increase in film stress reported in Table 5.2.

It has been reported that recrystallization of amorphized silicon upon thermal annealing leads to strain and stress generation in the silicon material itself [81-84]. Because the transistor's poly-silicon gate and the S/D region were all amorphorized after high dosage S/D implantation, this increased tensile wafer curvature after annealing is likely to result from the SMT capping nitride film stress change and the strain exerted by the recrystallization of the amorphized S/D regions upon S/D activation annealing. Nevertheless, wafer warpage continues to exist even after the SMT nitride is subsequently removed, both for wafers processed with SMT by LT-SiN and for those with HT-SiN. This can be viewed as the generation of the SMT strain which remains present in the wafer after completing the SMT process. Moreover, as the electron mobility in the channel can be increased by tensile strain in the in-plane longitudinal direction and compressive strain in the out-of-plane vertical direction according to the piezoresistance model [17, 18, 27], this "memorized" SMT strain in the poly-Si gate and S/D region can thus enhance electron mobility and nFET drive current [19, 36, 37, 46].

5.2 nFET Device Characterization and Discussion

In addition to the enhanced wafer curvature exerted by the memorized strain as shown in Fig. 5-5, electrical characterization was preformed to study the impact of

SMT on nFET device (W/L=10/0.4 μm) transconductance and mobility.

5.2.1 Extraction of nFET Mobility and SCE

Fig. 5-6 plots the linear transconductance G_{m_lin} improvement resulting from the HT-SiN SMT process. Compared to LT-SiN process, a maximum Gm_lin enhancement of around 13% under a drain voltage of 50 mV is obtained. The nFET field-effect electron mobility, measured at a drain voltage 50 mV by the conventional split C-V method [85], is shown in Fig. 5-7. The maximum field-effect electron mobilities are 292 $cm^2/V*s$, and 264 $cm^2/V*s$ respectively in these two samples. Similarly to the findings in Fig. 5-6, the field effective mobility of the device with HT-SiN SMT is also over 10% higher than the one of the device integrating LT-SiN. From Fig. 5-6 and Fig. 5-7, it can be noted that there is an obvious transconductance and carrier mobility improvement when HT-SiN is used as an SMT layer. These results clearly suggest that the HT-SiN SMT process induces a higher permanent stress in the channel region and the electron mobility is thus enhanced after completing S/D annealing. Thus, it can also be concluded from this work that the "optimized" SMT process can further enhance nFET carrier mobility and device performance as previously reported [19, 37, 46]. The retained transconductance G_{m_lin} and mobility improvement, even after the removal of the stressor film, can be attributed to a stress "memorization" effect.

The linear and saturation threshold voltages (V_{t_lin} and V_{t_sat}) are plotted as a function of channel length in Fig. 5-8. Regarding short channel effect (SCE) control, neither HT-SiN nor LT-SiN films demonstrate significant nFET V_t roll-off from long channel (L=10μm) down to short channel (L=0.4μm), which indicates that processing with the capped SMT nitride improves nFET SCE performance by retarding n-type dopant diffusion upon subsequent activation annealing [36]. A large shift of the threshold voltage V_t can often indicate a change of the device's dopant distribution profile, or the generation of bulk traps or interface states [50] due to un-optimized stressor processing techniques, which can adversely impact device performance without strain engineering benefit. However, from the threshold voltage data in Fig. 5-8, the linear threshold voltage (V_{t_lin}) difference between wafers processed by the HT-SiN and LT-SiN SMT processes is less than 15 mV, and the saturation threshold voltage (V_{t_sat}) difference is also less than 25 mV. Such small V_{t_lin} and V_{t_sat} differences between HT-SiN and LT-SiN SMT processes suggests that the enhanced thermal budget by furnace deposition system does not significantly deteriorate S/D dopant distribution in the transistor itself.

Moreover, Drain Induced Barrier Lowering (DIBL) can be calculated from equation 5-3 as the threshold voltage difference between linear and saturation modes, divided by the difference of the applied drain voltage.

$$DIBL = \frac{\Delta V_t}{\Delta V_{ds}} = \frac{Vt_{lin} - Vt_{sat}}{2 - 0.05} \cdot 1000 \;\; (\frac{mV}{V}) \qquad \text{Eq. 5-3}$$

The calculated DIBL is plotted as a function of channel length in Fig. 5-9. DIBL reflects the influence of the drain voltage on the carrier potential in the channel [11]. The slightly increased DIBL by HT-SiN observed in Fig. 5-9 can be attributed to increased dopant diffusion encroaching from the S/D extension during the deposition of HT-SiN due to its higher deposition temperature. Therefore, to lower the DIBL and improve SCE performance, a lowered thermal budget process is always preferred to suppress dopant diffusion.

5.2.2 Mobility Degradation upon Interfacial State Deterioration

As predicted by the Matthiessen's Rule in equation 5-4 [86], there are several key factors that contribute to the carrier transport mobility variation. The mobility of a traveling carrier depends on the scattering events. Lattice vibrations scatter electrons because the vibrations distort the crystal. Imperfections such as impurity atoms, interstitials, dislocations, and grain boundaries scatter conduction electrons because in their immediate vicinity the electrostatic potential differs from that of the perfect crystal. The interface states is a major factor influencing inversion channel electron mobility, in which the maximum mobility depends on the interface-state density.

$$\frac{1}{\mu_{Effective}} = \frac{1}{\mu_{Lattice}} + \frac{1}{\mu_{impurity}} + \frac{1}{\mu_{Dit}} + \frac{1}{\mu_{Surface\_Roughness}} \qquad \text{Eq. 5-4}$$

Therefore, the high-low Capacitance-Voltage (CV) technique [87] is utilized to characterize the density of interface traps D_{it}. The measured D_{it} results in Fig. 5-10 indicate that the interface quality deteriorated with LT-SiN when trap energy range (E_C – E) is greater than 0.25eV, where E_C is the energy of the conduction band edge, and E is the interface state energy. At the same time, the LT-SiN and HT-SiN SMT processes have comparable T_{ox} values, extracted from CV analyses, and similar threshold voltage values as previously shown in Fig. 5-8. As the MOSFET parameters are related through interface states and oxide traps near interface, the increased interface trap D_{it} generated at the oxide/silicon substrate interface under inversion mode, as shown in Fig. 5-10, can lead to the reduced transconductance in Fig. 5-6 and lowered electron mobility in Fig. 5-7. Therefore, to further improve nFET device performance, the SMT process should be optimized with respect to these parameters.

5.3 Wafer Warpage Induced Gate Edge Damage and Hydrogen Induced Interface Deterioration

As shown in Fig. 5-10, when the double sided HT-SiN was utilized as the SMT capping layer, it reduced the density of interface states. Hence both the nFET transconductance in Fig. 5-6 and electron mobility in Fig. 5-7 were improved. Thus, interfacial quality degradation must be considered with the adoption of advanced nitride film in pursuing performance gain for SMT technology.

Intrinsic stress results from the microstructure created in thin films when atoms of these thin films are deposited on the substrate via various kinds of deposition tool and processes. Thus, the thin film's formation process environment, deposition rate, and its impurities therein, determine the sign and magnitude of the residual stress in the silicon substrate, and more specifically, in the silicon devices fabricated. As the LT-SiN is formed by a PE-CVD technique at low temperature, it incorporates a higher hydrogen concentration as compared to HT-SiN [63, 64]. The resulting LT-SiN nitride layer has substantially higher Si-H and N-H bonding concentration as compared with a HT-SiN nitride layer formed at high temperature, resulting in a larger stress increase after annealing [64]. Moreover, the higher temperature condition in a furnace system used to fabricate HT-SiN can help to passivate the dangling bonds at the interface. Thus a lower D_{it} level can be achieved by HT-SiN as compared to that formed by LT-SiN.

From the physical property point of view, the HT-SiN nitride is thermally stable because the layer is formed with less hydrogen bonding as evidenced in Fig. 5-3c, exhibited little thermal stress variation as in Table 5.2, and showed lower thickness shrinkage upon annealing as in Table 5.1. At the same time, HT-SiN is deposited on both the front and backside of the wafer, and thus the stress change from the two sides of the wafer cancel each other out. Therefore, wafer bowing can also be suppressed, with less curvature induced after S/D RTA as in Fig. 5-5. Such unique physical properties prevent the cracking phenomena and reduce the stress applied to the underlying layer, and less gate edge damage during SMT formation process can be expected as illustrated in Fig. 5-11.

Moreover, from the chemical property point of view, the formation of N-H bonds is more desirable than that of Si-H bonds as less hydrogen atom evolved from the higher temperature nitride upon thermal annealing. This is because the bonding strength of N-H (4.7 eV) is larger than that of Si-H (3.9 eV) [88] and Si-H is more likely to release hydrogen after bond breakage than a N-H bond. In this respect, Yamamura *et al.* also confirms that their novel strained CESL SiN, with a reduced amount of unstable Si-H bond and incorporating more stable N-H bond, is able to improve gate oxide interface quality and NBTI performance [89].

It has been reported that excess released hydrogen can lead to Si/SiO_2 interface quality degradation D_{it} [90-93]. Hydrogen is well known to play a crucial role in the fabrication of high quality Si/SiO_2 interfaces in MOSFET electronic devices in that the interface-state buildup at the device’s substrate/oxide interface is highly related to the transport of atomic hydrogen during its fabrication, and suppression of hydrogen transport aids in improving the electrical integrity of devices [90, 91]. Since hydrogen is always present in the oxides of these transistors, the interface is unavoidably exposed to hydrogen depassivation process. As hydrogen atoms impinge on the $Si(111)/SiO_2$ interface to depassivate Si-H bond, the dangling Si bond and resulting interface traps can thus be formed via the chemical reaction in equation 5-5 shown below [92, 93].

$$Si - H + H \rightarrow Si \bullet + H_2 \qquad \text{Eq. 5-5}$$

For the SMT application here, the hydrogen concentration evolved from the SMT nitride can also influence interface quality. Since the HT-SiN film incorporates less hydrogen and creates fewer Si-H bonds, it is much more compositionally stable. Less hydrogen is released upon annealing as shown by the FTIR spectra in Fig. 5-3c, and less interfacial degradation is obtained (Fig. 5-10). The mechanism responsible for this improvement is illustrated schematically in Fig. 5-11a. On the contrary, LT-SiN involves large quantities of hydrogen which will be released upon annealing as confirmed by the FTIR in Fig. 5-3b. Since the gate oxide quality at the gate edge has already been weakened by the mechanical stress exerted by the wafer bow due to LT-SiN film's stress increase upon thermal annealing as illustrated in Fig. 5-11b, the interface is especially vulnerable to hydrogen attack with hydrogen diffusion paths though the poly-gate, spacer and S/D region (Fig. 5-11b), resulting in a higher D_{it} level (Fig. 5-10).

5.4 Summary

In this chapter, SMT strain silicon technique has been experimentally studied for two different nitride films, with respect to film stress, wafer curvature, nFET device performance, transconductance, and mobility. Moreover, the influence of interface states on the characteristics of nFET devices, and the associated electron mobility behavior with these two different kinds of capping nitrides have also been quantitatively analyzed.

The strain from the SMT nitride can be tailored by nitride hydrogen composition, deposition method, and film thickness and shrinkage change upon thermal annealing. Furthermore, nFET Vt and DIBL have a similar level when using HT-SiN or LT-SiN for SMT application. Nevertheless, nFET devices with higher device performance, carrier mobility and lowered interface state density can be fabricated by the HT-SiN as the SMT capping nitride. The lowered D_{it} level achieved by HT-SiN as SMT capping nitride is likely due to the fact that HT-SiN i) has less hydrogen in nitride itself, ii) is more thermally stable with low thickness shrinkage and its stress is stable upon S/D annealing, and thus iii) less wafer curvature is induced to suppress gate edge damage. Thus, it is important to optimize the nitride physical and chemical properties to maximize SMT strain memorization effect and to minimize other side effects to achieve the best performance simultaneously.

It is generally but not always true, that a higher stress film deposited onto the transistor can produce a high mobility channel, nor is it always guaranteed that a improved device performance can be obtained with such a higher film stress when adopting strain silicon technology. This is because that excess force may lead to stress relaxation, permanent wafer deformation or transistor structural damage. Furthermore, it has been found that interface state degradation must be balanced with the adoption of advanced nitride films for SMT in pursuing further nFET performance gain.

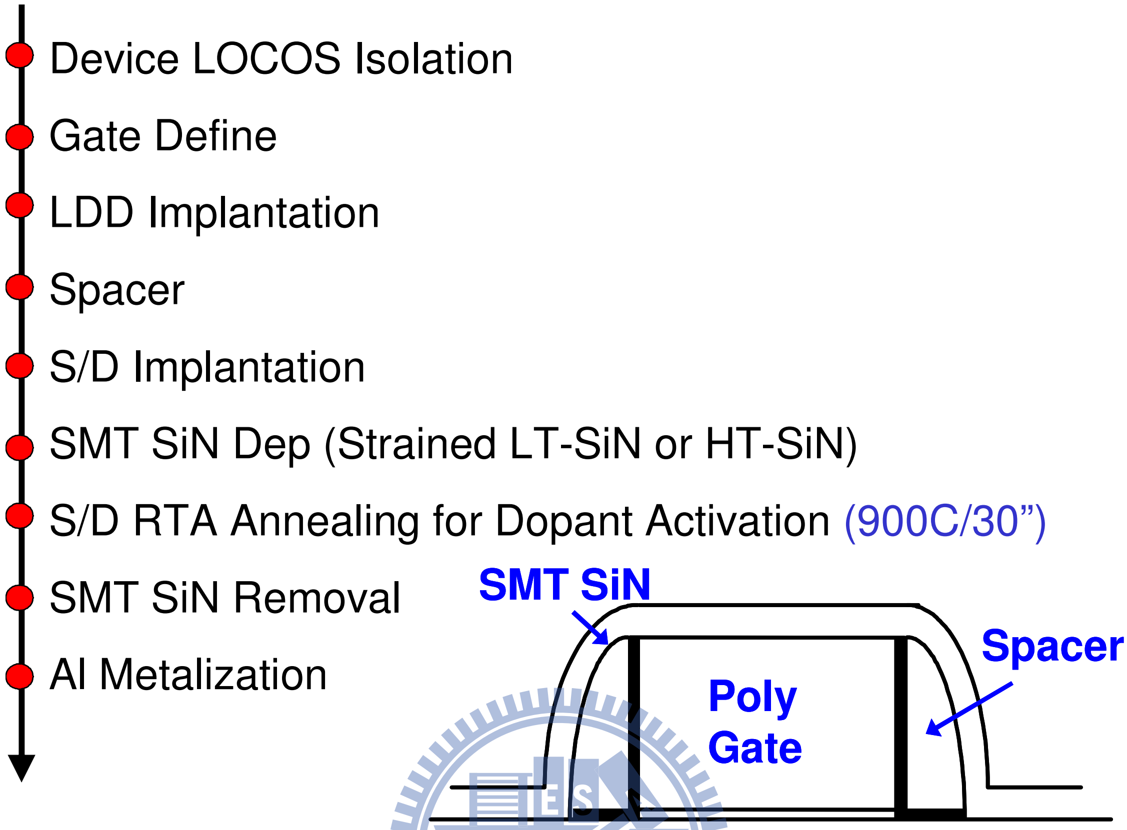


Fig. 5-1 SMT process flow. The SMT process deposits a nitride layer over the gate before annealing. The presence of this nitride layer during the subsequent S/D annealing acts to increase tensile strain in the nFET channel. Though this SMT nitride is removed after the annealing, the stress memorized in the channel region remains and mobility of electrons increases.

Thickness (A)	LT-SiN	HT-SiN
As Deposited	1163	1075
After Annealing	1028	1044
Thickness Shrinkage (%)	11.6	2.9

Table 5.1. Measured nitride thickness and associated thickness shrinkage rate (%) before and after S/D activation annealing.

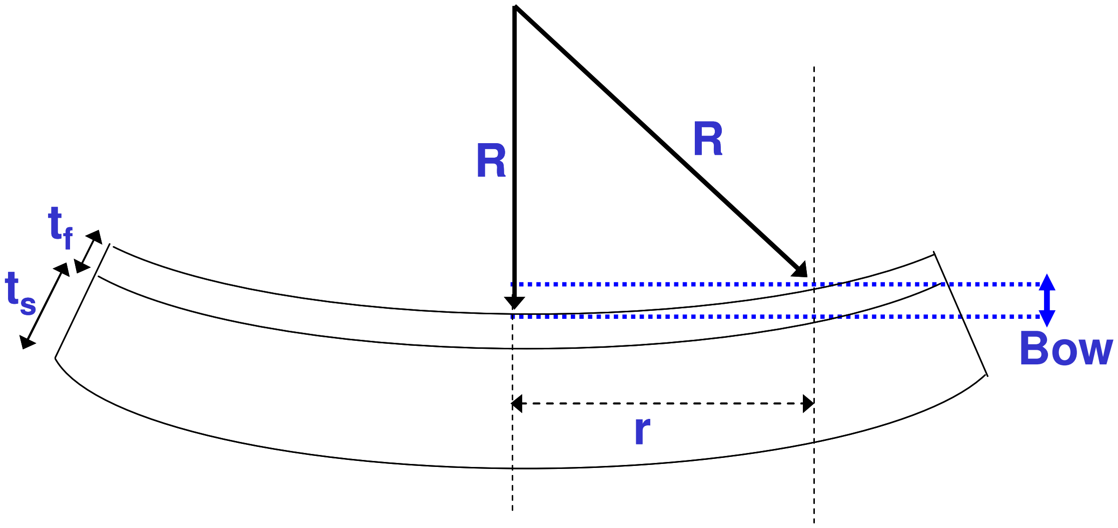


Fig. 5-2 Principle of wafer bowing measurement. Bow is characterized in terms of the curvature of the substrate, which is the height difference between the center and the edge of the wafer resulted from the wafer manufacturing process or the presence of a residually stressed thin film on one side of the wafer. In Fig. 5-2, the film is under positive (tensile) stress, and the R (radius of curvature) is positive. Therefore, the bow is also positive according to equation 5-2. The method to extract the stress of the film using wafer bowing is described in equation 5-1.

Stress (MPa)	LT-SiN	HT-SiN
As Deposited	360	1010
After Annealing	1050	1070
Delta Stress	690	60

Table 5.2. Measured nitride stress variation before and after S/D activation annealing.

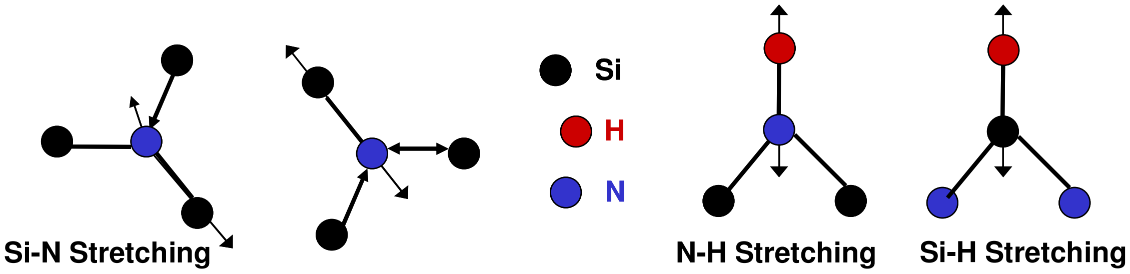


Fig. 5-3a Schematic drawing of stretching vibrations of Si-N, N-H, and Si-H bonds in nitride.

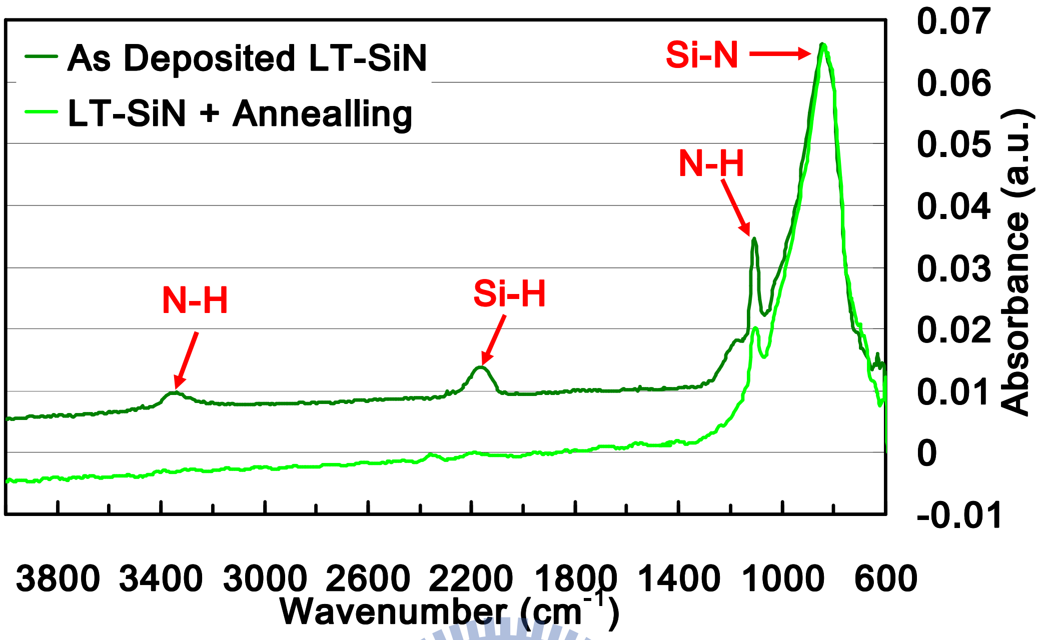


Fig. 5-3b LT-SiN film FTIR analysis results before and after S/D activation annealing. FTIR spectra indicated features associated with [76]

i) the Si-N asymmetric bond-stretching vibration at 820~850 cm^{-1};

ii) the bond-bending vibration of the N-H group at ~1150 cm^{-1},

iii) the bond stretching vibration of the Si-H group at ~2150 cm^{-1}, and

iv) the bond-stretching vibration of the N-H group at ~3350 cm^{-1}.

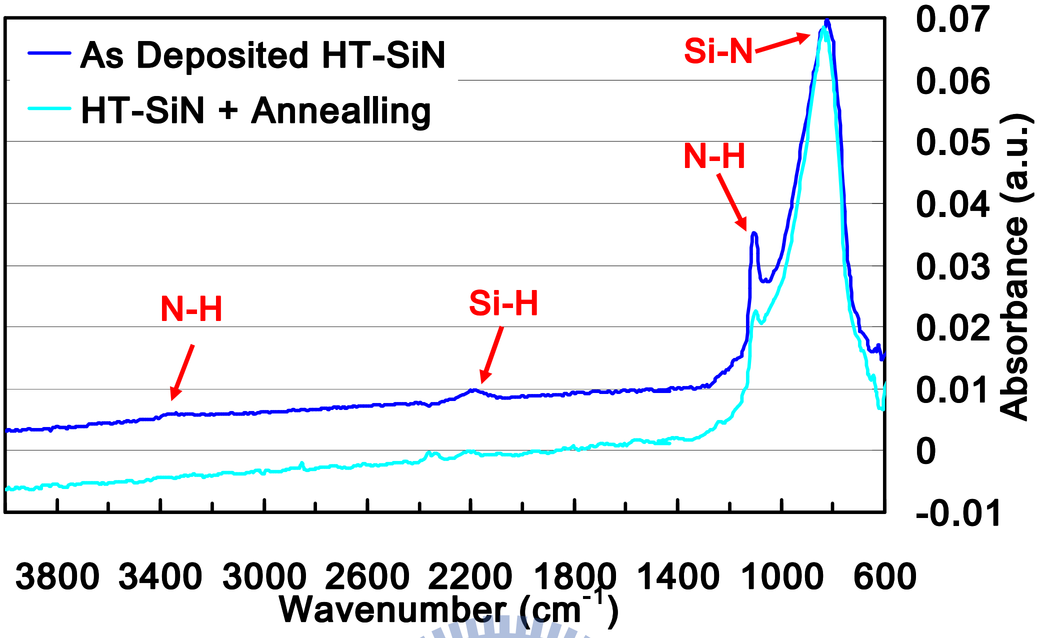


Fig. 5-3c HT-SiN film FTIR analysis results before and after S/D activation annealing. HT-SiN sample contains much less Si-H and N-H bonds as compared to those of LT-SiN sample.

[H] Concentration (atom/cm3)		
As Deposited	LT-SiN	HT-SiN
Si-H	8.67E+21	3.97E+21
N-H	8.89E+21	2.28E+21
Total [H]	1.76E+22	6.25E+21
H atomic (%)	16.7	6.0

Table 5.3. Estimated Si-H, N-H and Total Hydrogen bonds concentration in LT-SiN and HT-SiN before S/D activation annealing. HT-SiN sample contains much less Si-H and N-H bonds than those of LT-SiN sample.

[H] Concentration (atom/cm3)		
After Annealing	LT-SiN	HT-SiN
Si-H	2.27E+21	2.37E+21
N-H	2.92E+21	1.27E+21
Total [H]	5.19E+21	3.64E+21
H atomic (%)	4.9	3.5

Table 5.4. Estimated Si-H, N-H and Total Hydrogen bonds concentration in LT-SiN and HT-SiN after S/D activation annealing.

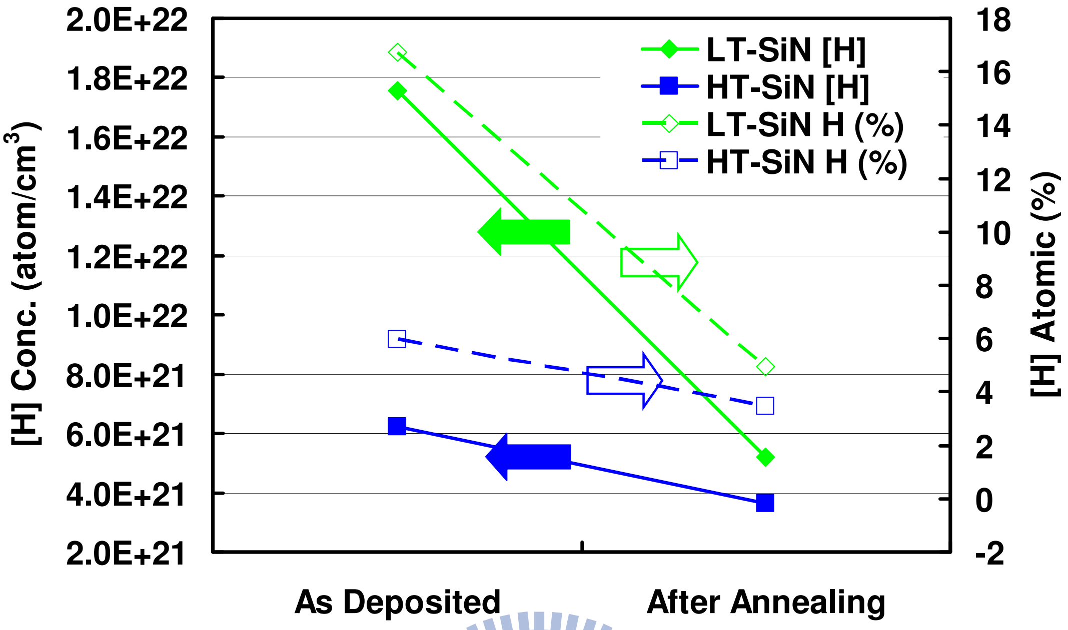


Fig. 5-4 Estimated LT-SiN and HT-SiN hydrogen concentration, and hydrogen atomic (%) based on the FTIR results before and after S/D activation annealing. HT-SiN sample contains much less hydrogen bonds than that of LT-SiN sample.

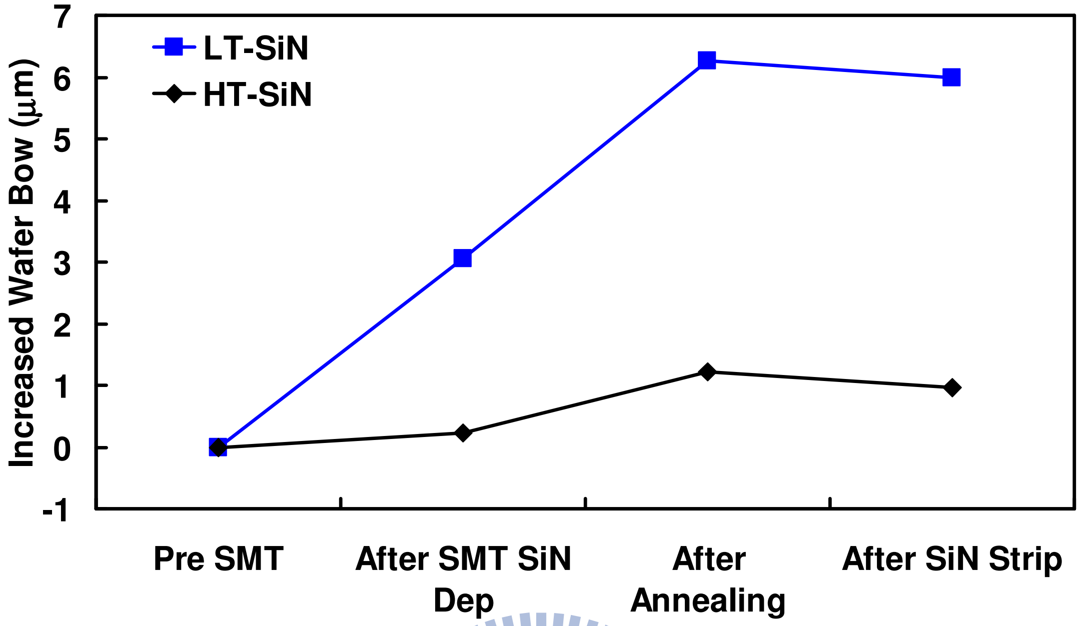


Fig. 5-5 Increased device wafer bow height measured on patterned wafers at different SMT stages to evaluate the SMT strain behavior.

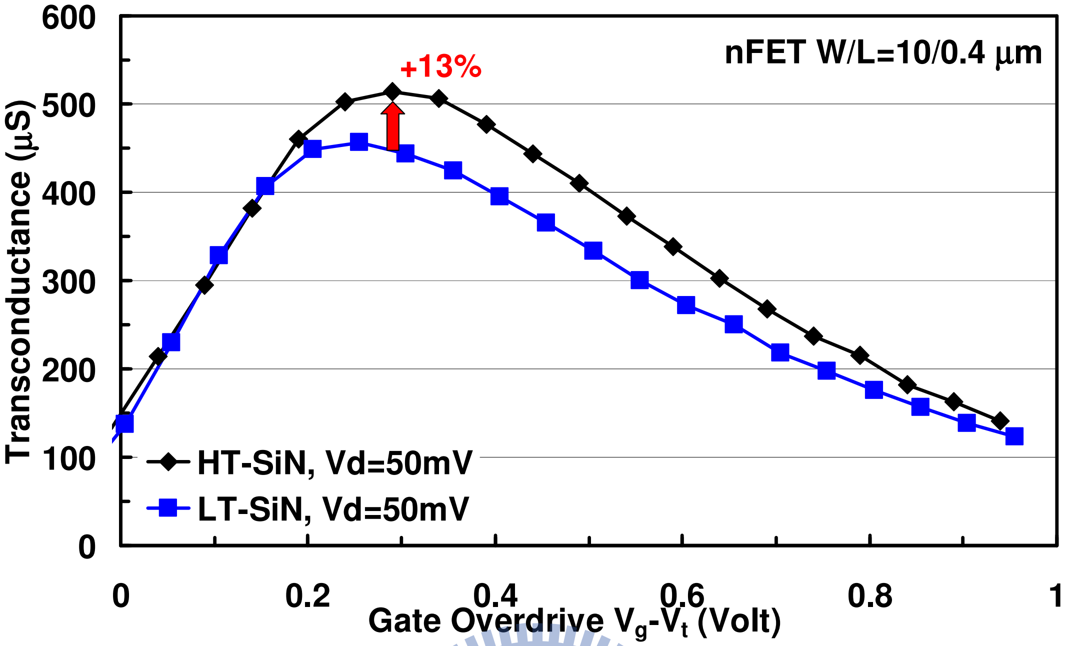


Fig. 5-6 Measured 10x0.4μm nFET G_{m_lin} at V_d=50 mV. There is a ~13% improvement in the peak value of G_{m_lin} with the adoption of HT-SiN as the SMT capping layer.

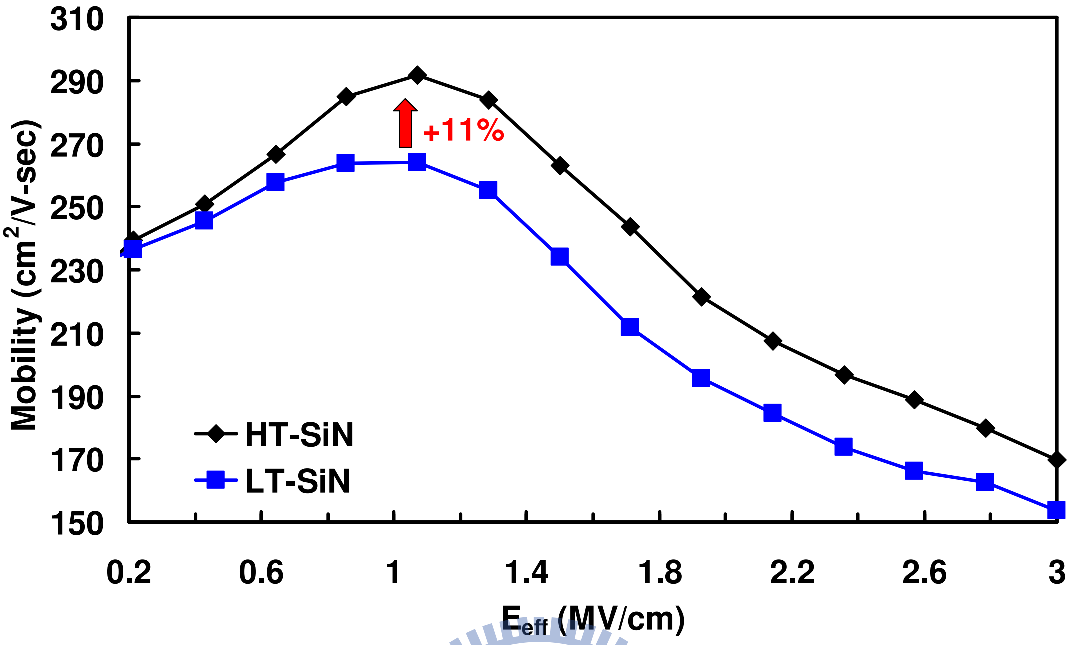


Fig. 5-7 Measured 10x0.4μm nFET mobility at V_d=50 mV. There is a ~11% improvement in the peak value of electron mobility with the adoption of HT-SiN as the SMT capping layer.

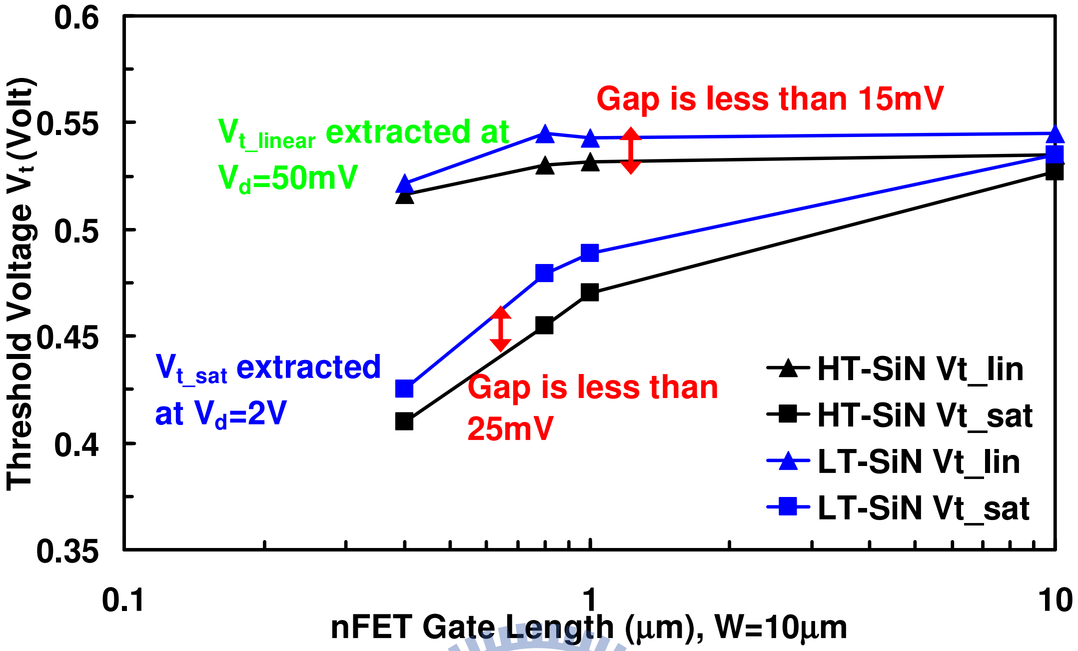


Fig. 5-8 Measured W= 10µm nFET V_t as a function of channel length; LT-SiN as SMT capping layer improves nFET SCE with a less V_t roll-off.

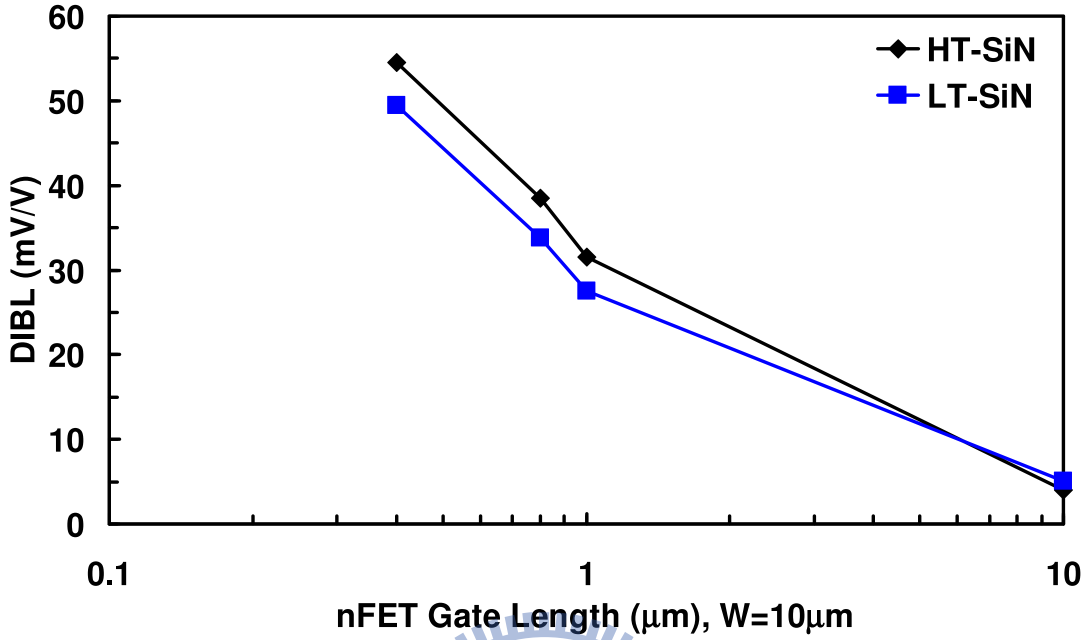


Fig. 5-9 Calculated W=10 μm nFET DIBL. Slightly improved DIBL at reduced gate length nFET devices can be observed with LT-SiN as the SMT capping layer.

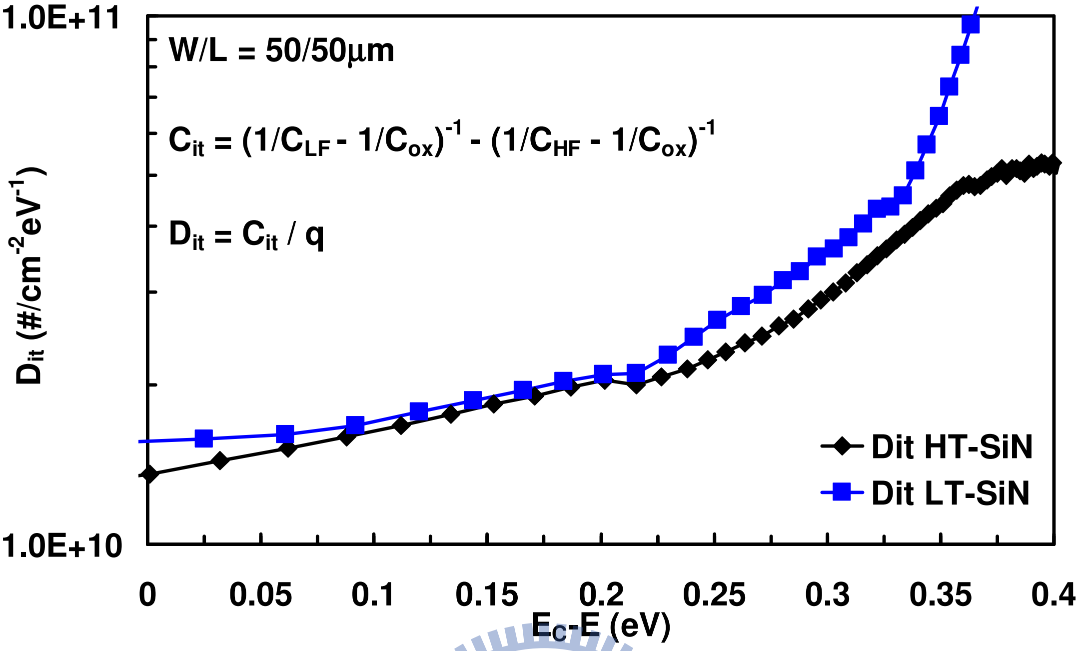


Fig. 5-10 Measured interface state density (D_{it}) characteristics as a function of trap energy near the conduction band edge using 50μm square MOS capacitors by High-Low CV technique [87]. The interfacial quality is degraded with the adoption of LT-SiN as the SMT capping layer when trap energy range (E_C – E) is greater than 0.25eV, where E_C is the energy of the conduction band edge, and E is the interface state energy.

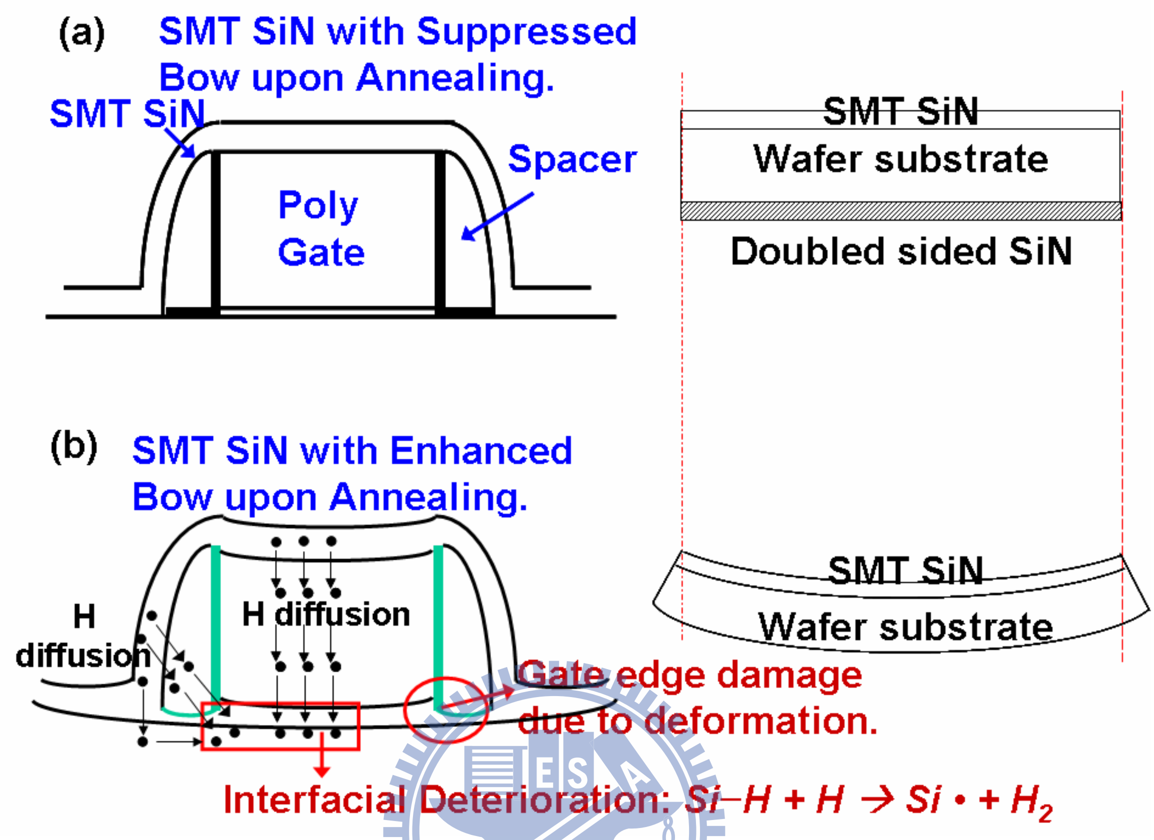


Fig. 5-11 (a) top (b) bottom.

Schematic illustration of wafer bend induced gate edge damage and gate oxide interfacial quality deterioration. This wafer bowing is due to high stress change in the SMT capping nitride in Fig. 5-11b.

Chapter 6

Ballistic Transport for Highly Scaled nFET Devices

6.1 Introduction of Ballistic Transport

The transport of a carrier in the device is closely related to the relative dimension of the device to the mean free path of the carrier, as illustrated in Fig. 6-1. When the device size L is sufficiently larger than the mean free path λ, the carrier flow is controlled by the diffusive transport and is well-characterized by the conventional mobility theory. When analyzing conventional long channel devices, the traditional approach assumes that the averaged distance between each collision, the so-called mean-free-path λ, is much shorter than the device geometry L. When these conditions hold, drift-diffusion equations can be used to describe carrier transport behavior. When gate voltage is above threshold, the drift current dominates the traditional long channel MOSFET I-V characteristic [11]. On the other hand when gate voltage is below threshold, it is usually assumed that diffusion dominates under the sub-threshold condition [11].

From device performance point of view, the success of scaling depends on achieving higher on-state drive current I_{on}, while maintaining as low as possible off-state leakage current. I_{on} can be increased in three major ways: increasing areal gate capacitance C_{ox}, decreasing gate length L_g, or increasing carrier mobility μ. These three ways can be qualitatively expected from the long-channel saturation mode MOSFET equation 6-1 below [11].

$$I_{on} = \frac{W}{2 \cdot L_g} \cdot \mu \cdot C_{ox} \cdot \left(V_{gs} - V_t\right)^2 \qquad \text{Eq. 6-1}$$

Moreover, guided by Moore's law, the performance of integrated circuits has been dramatically improved by geometric scaling of Si CMOS devices. As CMOS devices are being further and further scaled down, the device dimensions shall become comparable to or smaller than the mean-free-path λ. In this regime, the channel length is much shorter than the mean-free-path ($L \leq \lambda$), and the scattering probability in the channel is small can be ignored completely. That can be described as the ballistic transport [94-98]. Most of the carriers pass through the channel toward drain without scattering and the device current is completely controlled by the carrier injection from source into the channel as illustrated in Fig. 6-1.

The nanoscale transistors fabricated nowadays are in the range where L (<100nm) is comparable to λ, and can be characterized by the quasi-ballistic transport. Therefore, they are operated between the drift-diffusion and ballistic regimes. In this quasi-ballistic transport regime, carriers encounter a limited amount of scattering from source and drain, and the probability of a carrier encountering scattering events within the channel rapidly diminishes when the device feature are scaled down. The mobility theory no longer describes the transport on the one hand, and the presence of scattering distinguishes their behavior from ideal ballistic characteristics. Drift diffusion theory is no longer strictly valid, but it continues to provide insights into the operation of small semiconductor devices. A ballistic treatment is not strictly valid either, but it provides new useful insights when these new devices are explored.

6.2 Temperature Dependent Ballistic Transport Model

As the channel length of a MOSFET scaled into the nanoscale regime, the I_{d_sat}

of MOSFET is primarily dictated by the channel backscattering phenomena as illustrated in Fig. 6-2 [94]. The channel backscattering theory as formulated in equation 6-2 [94] relates the saturation drive current I_{d_sat} of a nanoscale MOSFET to the source injection velocity and carrier backscattering phenomena by

$$I_{d\_sat} = W \cdot Q_{inv} \cdot v_{inj} \cdot \left(\frac{1-\gamma_{sat}}{1+\gamma_{sat}} \right) \qquad \text{Eq 6-2}$$

where

W is the channel width, Q_{inv} is the inversion charge density,

v_{inj} is the injection velocity, and γ_{sat} is the carrier backscattering coefficient.

When carriers are injected from the source-end into the channel, some carriers will encounter backscattering events within a critical length l_0 from the peak in the conduction band. Thus, the backscattering coefficient γ_{sat} for non-degenerate carrier in the saturation regime can be estimated from

$$\gamma_{sat} = \frac{l_0}{l_0 + \lambda_0} = \frac{1}{1 + \lambda_0 / l_0} \qquad \text{Eq. 6-3}$$

where

λ_0 is the near equilibrium mean free path for backscattering, and l_0 is the critical length for for carrier potential to fall by K_BT below the conduction band maximum. Through substitution of equation 6-3 into equation 6-2, the I_{d_sat} can be expressed in terms of λ_0 and l_0 as

$$I_{d\_sat} = W \cdot Q_{inv} \cdot v_{inj} \cdot \left(\frac{\lambda_0 / l_0}{2 + \lambda_0 / l_0} \right) \qquad \text{Eq. 6-4}$$

where v_{inj} in Equation 6-4 represents the theoretical injection velocity at the beginning of the channel with the maximum value approaches the thermal injection velocity ($2*10^7$ cm/sec). Because the carriers at the beginning of the channel are injected from thermal equilibrium source, the injection velocity is mainly determined by the lattice temperature and their effective mass. Furthermore, the terms v_{inj}, λ_0 and l_0 have temperature dependence as given by

$$v_{inj} = \sqrt{\frac{2K_B T}{\pi m^*}} \qquad \text{Eq. 6-5}$$

$$\lambda_0 = \frac{2\mu K_B T}{q v_{inj}} \qquad \text{Eq. 6-6}$$

$$l_0 = \frac{K_B T}{qE(0)} \qquad \text{Eq. 6-7}$$

It is noted that as temperature T raises, the λ_0 / l_0 ratio is thus decreased, and leads to a increased γ_{sat} backscattering coefficient in the saturation regime according to equation 6-8.

$$\frac{\lambda_0}{l_0} = \frac{2\mu E(0)}{v_{inj}} \propto T^{-0.5} \qquad \text{Eq. 6-8}$$

Differentiation of Eq 6-4 with respect to temperature would then produce an analytical expression [49, 99] for the temperature dependence of I_{d_sat} as denoted by α,

$$\frac{\partial I_{d\_sat}}{\partial T} = I_{d\_sat}\left[\frac{1}{v_{inj0}}\frac{\partial v_{inj0}}{\partial T} + \frac{1+\gamma_{sat0}}{1-\gamma_{sat0}}\frac{\partial}{\partial T}\left(\frac{1-\gamma_{sat0}}{1+\gamma_{sat0}}\right) + \frac{1}{V_G - V_{t\_sat}}\frac{\partial(V_G - V_{t\_sat})}{\partial T}\right]$$

$\therefore$

$$\frac{\partial I_{d\_sat}}{\partial T} = I_{d\_sat}\left[\frac{1}{2T} - \left(\frac{1}{1+\gamma_{sat0}} + \frac{1}{1-\gamma_{sat0}}\right)\frac{\partial \gamma_{sat0}}{\partial T} - \frac{\eta}{V_G - V_{t\_sat}}\right] = I_{d\_sat}\alpha$$

Eq. 6-9

Therefore, one can get

$$\alpha = \frac{\Delta I_{d\_sat}}{I_{d\_sat}\Delta T} = \left(\frac{1}{2} - \frac{4}{2+\lambda_0/l_0}\right)T^{-1} - \frac{\eta}{(V_G - V_{t\_sat})} \qquad \text{Eq. 6-10}$$

where η is the temperature dependence of threshold voltage as given by

$$\eta = \frac{\Delta V_{t\_sat}}{\Delta T} \qquad \text{Eq. 6-11}$$

Both α and η can be experimentally extracted from the best fitted gradient of ΔI_{d_sat} and ΔV_{t_sat} as a function of temperature, respectively. Then, the ratio of carrier mean free path to the critical length for backscattering λ_0 / l_0 can then be deduced from equation 6-10 which yields the following expression.

$$\frac{\lambda_0}{l_0} = \frac{4}{\frac{1}{2} - \left[\alpha + \frac{\eta}{(V_G - V_{t\_sat})}\right]T} - 2 \qquad \text{Eq. 6-12}$$

By employing the λ_0 / l_0 ratio and the backscattering coefficient γ_{sat}, the ballistic efficiency B_{sat} can be estimated using

$$\nu_{eff} = \nu_{inj}\left(\frac{1-\gamma_{sat}}{1+\gamma_{sat}}\right) \qquad \text{Eq. 6-13}$$

$$B_{sat} = \frac{\nu_{eff}}{\nu_{inj}} = \left(\frac{1-\gamma_{sat}}{1+\gamma_{sat}}\right) = \frac{\lambda_0/2l_0}{1+\lambda_0/2l_0} \qquad \text{Eq. 6-14}$$

The ballistic efficiency B_{sat} approaches close to "one" as the backscattering coefficient γ_{sat} approaches close to "zero", and more backscattering flux or increased γ_{sat} results in a reduced drive current as predicted by equation 6-2 and equation 6-14. As MOSFET carrier transport becomes more ballistic, the carrier transport can be limited by carrier injection at the source side, and source injection velocity ν_{inj} plays an important role in affecting the drive current of a MOSFET in addition to carrier backscattering [97, 98]. Overall, reducing γ_{sat} and increasing source injection velocity ν_{inj} are both desirable in order to enhance I_{d_sat} performance for device operated under ballistic transport region.

6.3 Experimental Results and Discussion

For modern process induced stained silicon techniques, the mobility of short channel device generally increases more than that of long channel devices. Therefore, the short channel nFET fabricated on industry standard 12inch wafers of drawn length from 10 to 0.04μm are measured and discussed in this section. Two kinds of nFET devices with different CESL films, one with strained nitride and the other with neutral SiON, are fabricated as shown in Fig. 6-3. The transport properties of electrons in our nFET devices are studied in the linear region (V_d=50mV) and the saturation region (V_d=1V). The threshold voltage V_t is determined by the constant current method [49]. Suss PAC 300 cryogenic probe station was used to characterize low temperature nFET performance with liquid nitrogen as the coolant, and the temperature range used in the measurements varies from 300K to 150K.

6.3.1 Ballistic Measurement on Strained and Non-Strain nFET

In order to study the impact of channel strain engineering on the carrier transport characteristics of sub-100 nm nFETs, a similar approach based on the temperature dependent backscattering model [99-101] as described in section 6-2 was employed. The impact of strain on carrier mobility and drive current enhancement are investigated in terms of mean free path, backscattering coefficient, ballistic efficiency, and carrier injection velocity.

In order to extract these backscattering related parameters, α and γ were first experimentally extracted from measurements of the change in drive current I_{d_sat} and the threshold voltage V_{t_lin}, as a function of temperature.

Typical temperature dependency results of V_{t_lin} and the change in I_{d_sat} at different temperature for our 1x0.8μm nFET are shown in Fig. 6-4 and Fig. 6-5, respectively. The threshold voltage increases with reduced temperature [11]. Moreover, carrier mobility is determined by two main scattering mechanisms, electron-phonon scattering and impurity scattering. Electron-phonon scattering is due to scattering in the silicon lattice, while impurity scattering is due to ionized impurities. As temperature decreases, mobility increases due to a decrease in the electron-phonon scattering rate [102]. With higher mobility, transistor drive current can be improved; while increased threshold voltages may degrade transistor drive current. Therefore, the variations in threshold voltage and mobility are opposing forces against each other. Nevertheless, the measured nFET I_{d_sat} ratio still increases with the lower temperature as shown in Fig. 6-5.

A typical extraction results using equation 6-12 for 1x0.8μm nFET are shown in Table 6.1. Continuing on to extract these α and γ values at different transistors' length yields the λ_0/l_0 ratio characteristics as shown in Fig. 6-6. In general, the λ_0/l_0 ratio increases with decreasing gate length and is larger for the strained devices than the unstrained control devices. By scaling the gate length, a smaller l_0 will be achieved which contribute to a reduced carrier backscattering and contribute to a higher drive current performance [103].

6.3.2 Improved Backscattering Coefficient and Ballistic Efficiency

Fig. 6-7 plots the backscattering coefficient γ_{sat} as a function of gate length for strained and control transistors. For enhanced drive current performance, a smaller γ_{sat} is desirable as it implies reduced backscattering when the electrons are injected from the source into the channel. It can be seen that strained nFETs show an improvement or reduction in γ_{sat} over the control devices, due to a reduced critical length for backscattering l_0. However, it was reported that reducing l_0 has a more profound effect than reducing the entire channel length in affecting the drive current performance as carriers injected into the channel and passed beyond l_0 will not be backscattered into the source [104].

Using the backscattering coefficient γ_{sat}, the ballistic efficiency B_{sat} can be calculated using equation 6-14. When B_{sat} tends to unity, it implies that carriers injected from source into the transistor channel will not encounter any backscattering events and the transistor will be considered to operate in the fully ballistic regime. As a result of reduced γ_{sat}, strained devices show substantial improvement in the ballistic efficiency B_{sat} over the control devices in Fig. 6-8. This improvement is observed to increase with decreasing gate length and approaching $B_{sat} \sim 0.3$ for a strained device with gate length down to 0.1 μm. The continual enhancement of B_{sat} with reduced gate length indicates that channel backscattering phenomena will continue to dominate the I_{d_sat} performance of a sub-0.1 μm gate length transistor.

6.3.3 Carrier Injection Velocity Enhancement

Since I_{d_sat} can be precisely measured and B_{sat} can also be estimated, ν_{inj} can thus be obtained via equation 6-2. The injection velocity ν_{inj} plays an important role in affecting the saturation drive current performance of nanoscale ballistic MOSFET in that higher ν_{inj} leads to higher I_{d_sat}. The dependency of source injection velocity ν_{inj} as a function of gate length is further investigated next.

Fig. 6-9 shows that strained devices have a higher calculated source injection velocity ν_{inj} than that of the unstrained control devices. It is noted that ν_{inj} of a tensile CESL nitride strained transistor increases with decreasing gate length and approaches the thermal injection velocity ($2*10^7$ cm/s) at sub 100 nm gate lengths. This significant ν_{inj} enhancement could be highly related to the strain-induced modification of the electronic structure. Fig. 6-10 illustrates that the conduction band valleys are degenerate for an unstrained silicon. The presence of uniaxial tensile strain effect lifts the six-fold degeneracy at the conduction band valleys, leading to the lowering of the two-fold perpendicular valleys (Δ_2) with respect to the four-fold in-plane valleys (Δ_4). This leads to preferential electron population at the lower energy Δ_2 valleys where the in-plane transport mass is lower ($\sim 0.19\ m_0$) [40, 41].

The carrier injection velocity ν_{inj} is mainly determined by the lattice temperature and the effective mass as formulated in equation 6-5. Therefore, the electron injection velocity is actually inversely proportional to the square root of the electron effective mass, and thus the change of effective mass still plays the main role for devices operated in ballistic transport. Hence, the strain induced reduction of the effective mass will still enhance the I_{d_sat} when device scaling enters into the quasi-ballistic or ballistic transport region [103].

6.4 Summary

Since the 90nm node, strain-enhanced CMOS scaling has been very important in continuing Moore's Law. Mechanical strain modifies the crystallographic geometry in the channel, modulating carrier properties by modifying energy band structure.

Strain in the channel region affects the effective mass of carriers and therefore their mobility. Furthermore, strain also changes the carrier backscattering rate γ_{sat} at the carrier injection point from the source side [100, 101]. Strain in the channel has also been shown to reduce effective mass to increase the injection velocity ν_{inj} and ballistic efficiency B_{sat} of the carrier transport by reducing the γ_{sat} [103]. As expressed in Equation 6-2, both ν_{inj} and B_{sat} improvement are observed to account for the I_{d_sat} enhancement achieved in the strained devices.

As MOSFET carrier transport becomes more ballistic, the carrier transport can be limited by carrier injection at the source as the scattering event happening in the channel is greatly reduced [94-97]. Since ν_{inj} is inversely proportional to the square root of the effective mass as predicted in equation 6-5, therefore, the strain induced reductions in the effective mass will still enhance the drive current in ballistic MOSFETs. The reduction of carrier conductivity mass by novel strain silicon technology can therefore increase the source injection velocity ν_{inj}, and can still contributes to I_{d_sat} enhancement for highly scaled strained nFET devices.

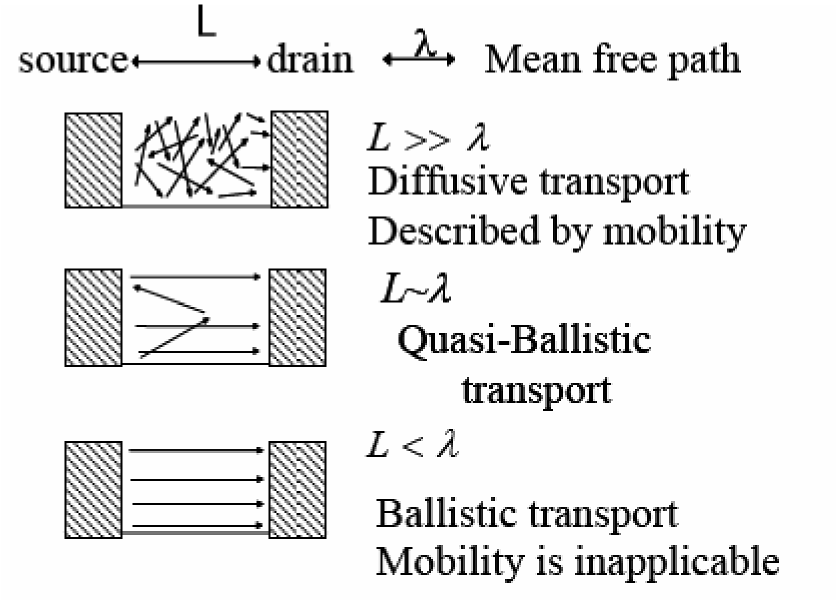


Fig. 6-1 Carrier transport in MOSFETs depends on relative dimension of the device size and the mean free path. In general, carriers will exhibit ballistic conduction when L<=λ, where L is the length of the active part of the device (i.e., a channel in a MOSFET). λ is the mean scattering length for the carrier which can be given by Matthiessen's Rule.

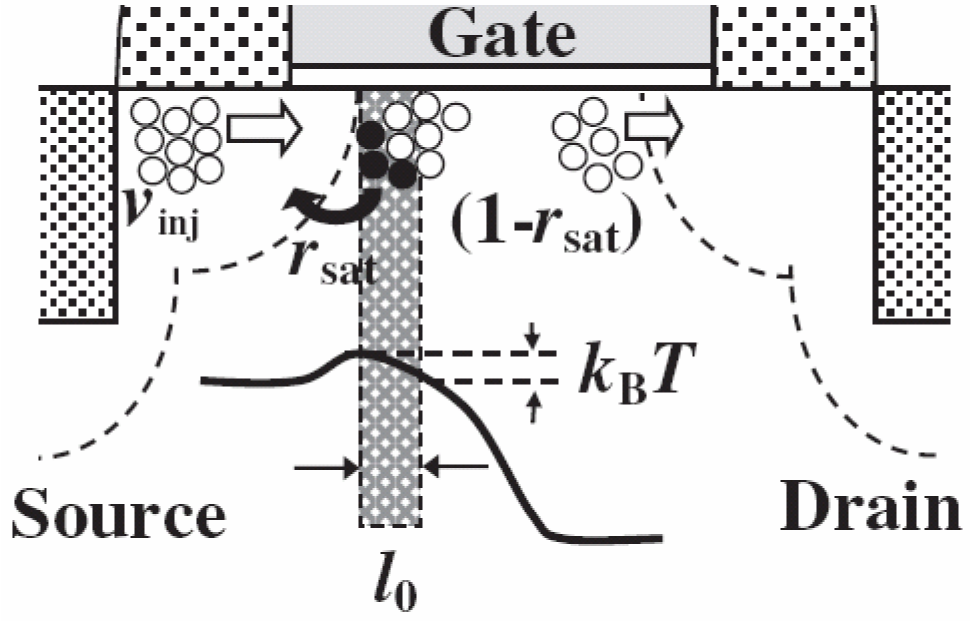


Fig. 6-2 Illustration of channel backscattering phenomenon. When carriers are injected from source into channel, they may encounter scattering within a thickness l_0 with a potential drop of k_BT/q.

Some carriers γ_{sat} may be scattered back to the source, and the remaining carriers (1- γ_{sat}) continue their march toward the drain.

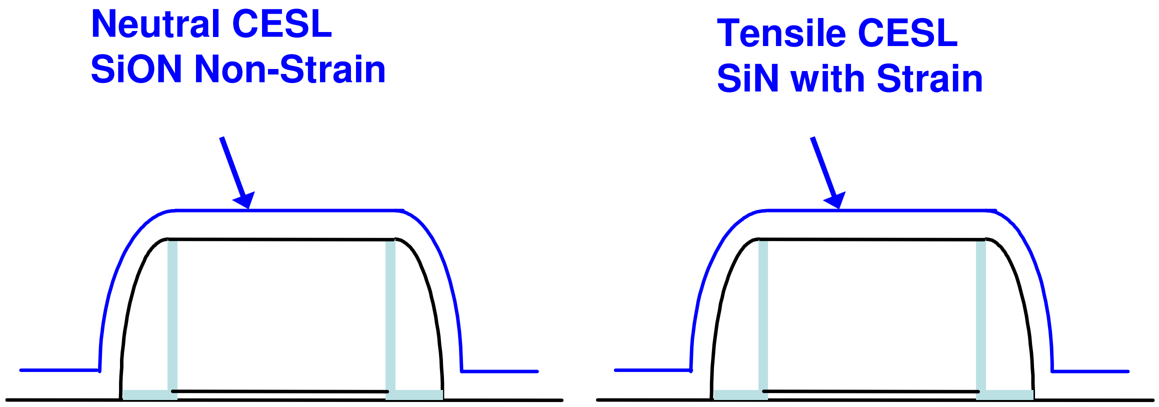


Fig. 6-3 nFET devices with CESL by strained nitride and neutral SiON are fabricated for ballistic conduction study.

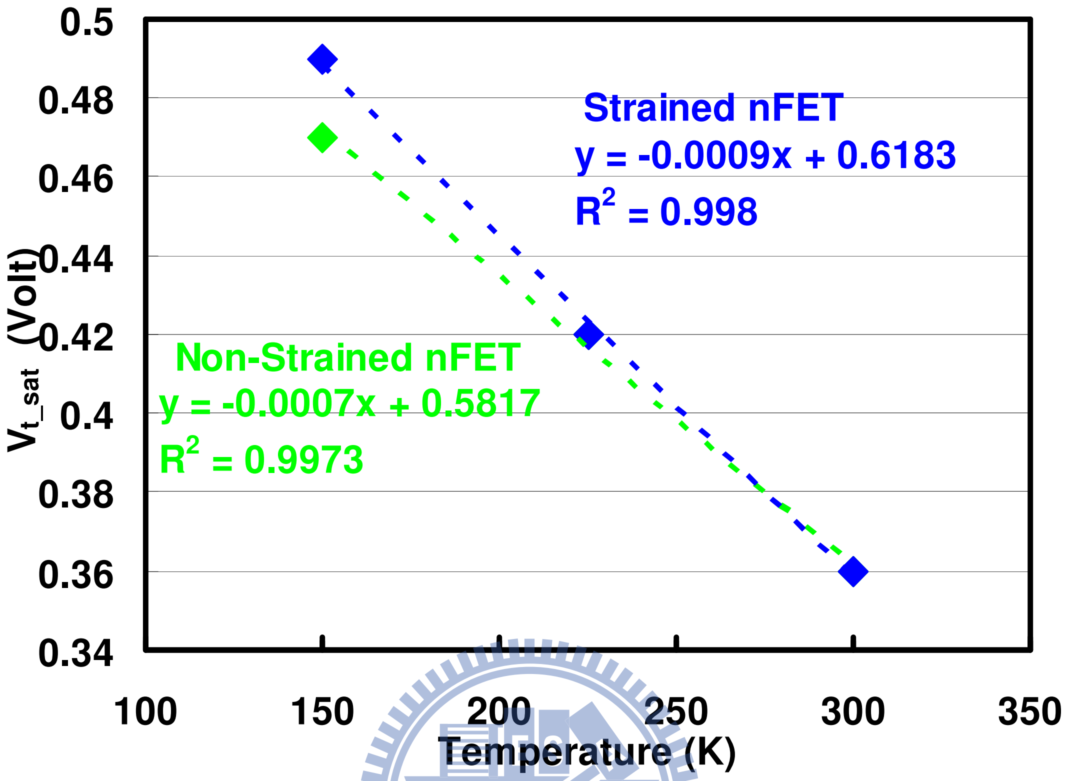


Fig. 6-4 Measured 1x0.8μm nFET devices threshold voltage as a function of temperature for ballistic conduction study.

η is related the temperature dependence of V_{t_sat}, and can be obtained from the slope of the best-fit lines.

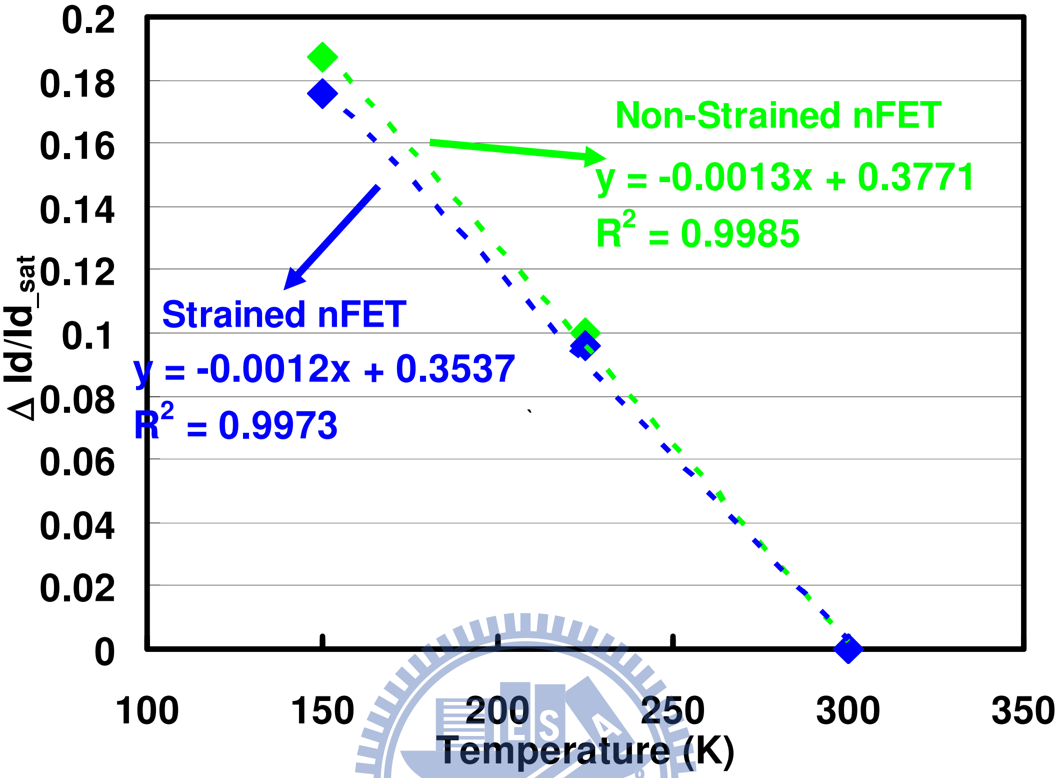


Fig. 6-5 Measured 1x0.8μm nFET devices drive current as a function of temperature for ballistic conduction study.

α is related the temperature dependence of I_{d_sat}, and can be obtained from the slope of the best-fit lines.

	Non-Strained	Strained
α	-7.00E-04	-9.00E-04
η	-1.30E-03	-1.20E-03
λ_0 / l_0	1.03	1.12
γ_sat	0.49	0.47
B_sat	0.34	0.36

Table 6.1. Extracted 1x0.8μm nFET devices η, α and ballistic transport parameters by the proposed temperature dependent ballistic transport model [99-101].

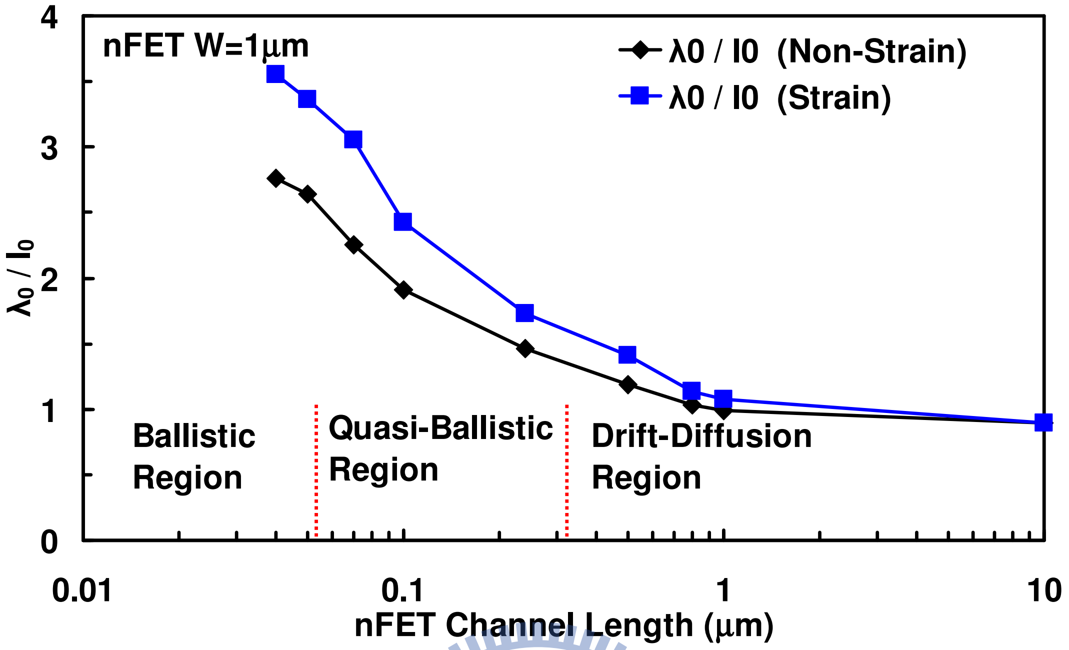


Fig. 6-6 Extracted λ_0/l_0 ratio at different gate length. Strained nFETs show higher ratio of mean free path λ_0 to the critical length for backscattering l_0 than the unstrained control devices. Improvement in λ_0/l_0 ratio is observed to increase with decreasing gate length.

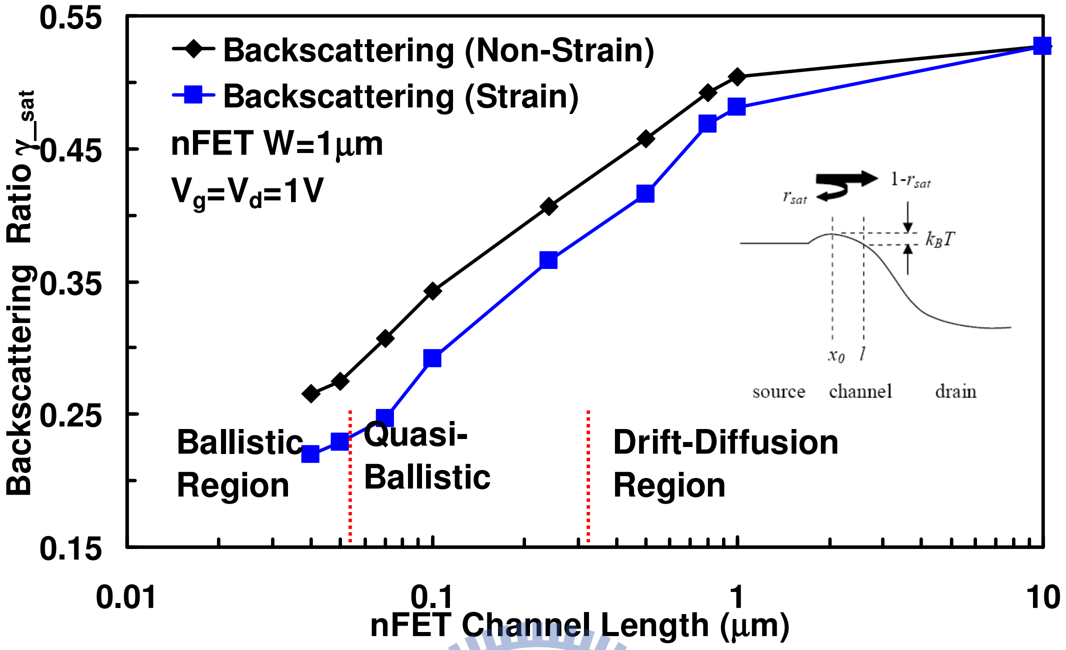


Fig. 6-7 Extracted γ_{sat} at different gate length. Uniaxial tensile strained nFETs show significant improvement in γ_{sat} over control devices due to the reduced critical length for backscattering l_0.

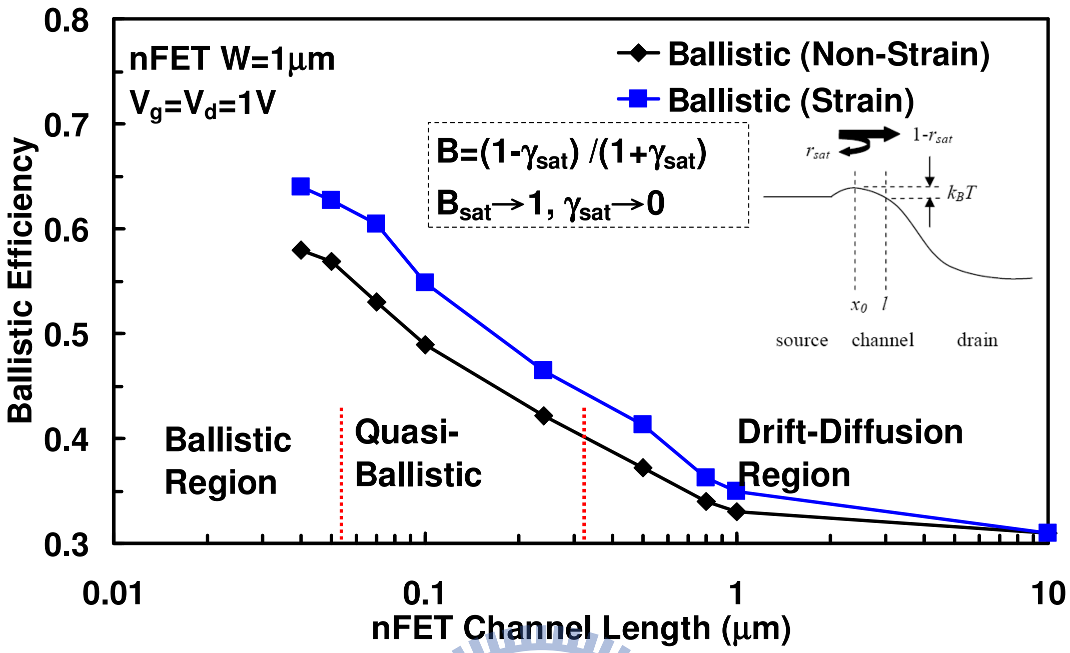


Fig. 6-8 Extracted B_{sat} at different gate length. The suppression of carrier backscattering into the source contributes favorably to the improvement in ballistic efficiency B_{sat} in strained nFETs which increases with decreasing gate length.

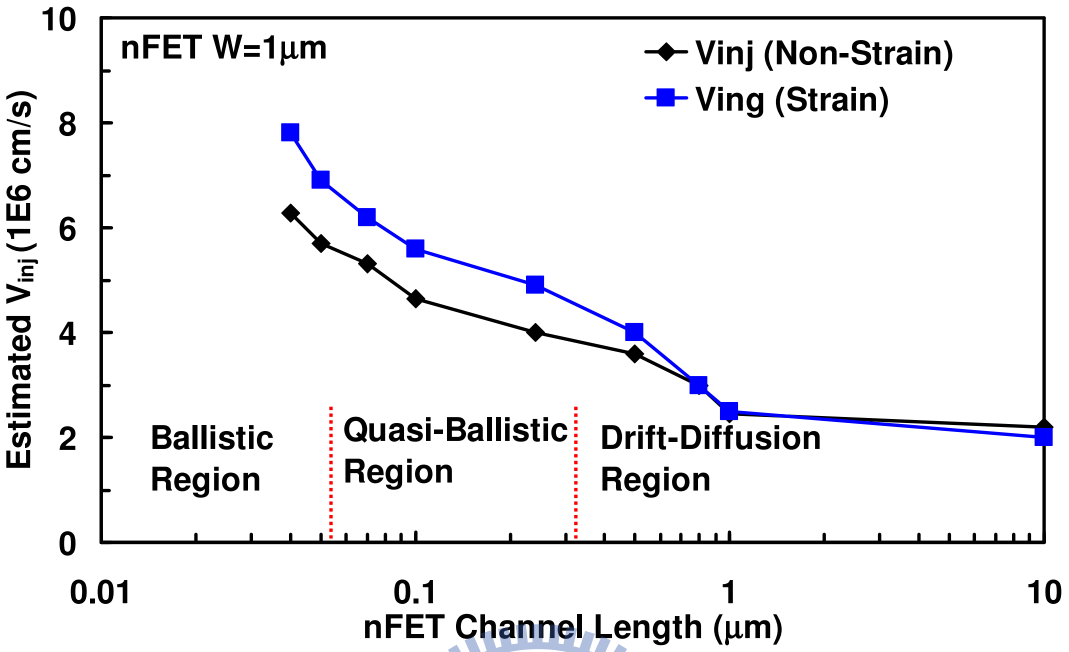


Fig. 6-9 Extracted ν_{inj} at different gate length. Strained nFETs show higher source injection velocity ν_{inj} compared to that of unstrained control devices. The improvement in ν_{inj} is attributed to the reduction of electron effective mass as carriers repopulate in the lower energy Δ_2 valleys.

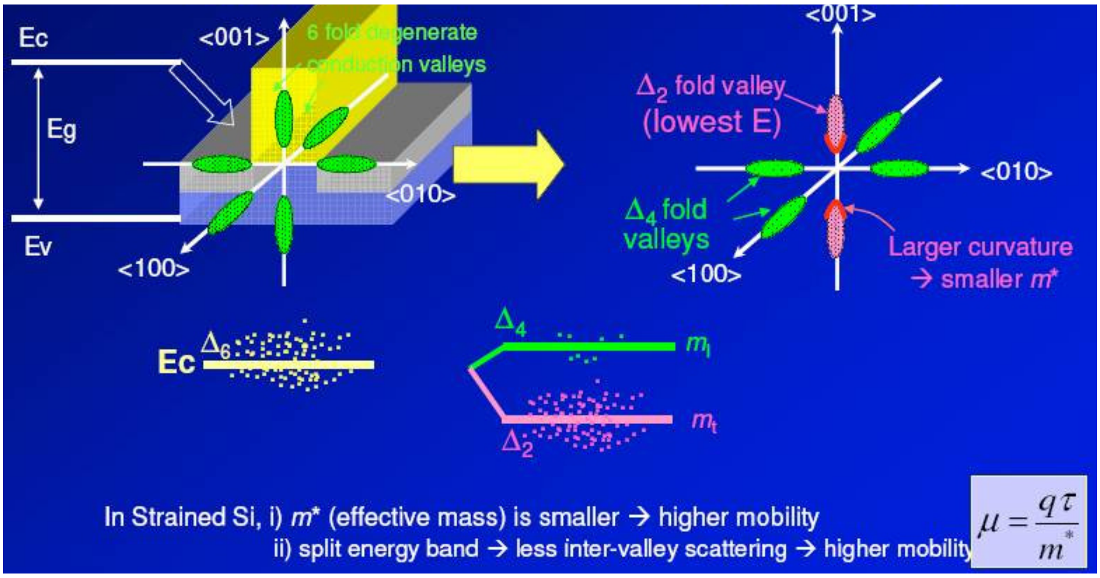


Fig. 6-10 Change of the conduction band upon uniaxial tensile strain. The presence of uniaxial tensile strain effect lifts the six-fold degeneracy at the conduction band valleys, leading to the lowering of the two-fold perpendicular valleys (Δ_2) with respect to the four-fold in-plane valleys (Δ_4). Most of the carrier repopulated to the lowered Δ_2 valley and intervalley scattering can be suppressed for strained nFETs. Furthermore, the electron residence in the Δ_2 valley has lower effective mass, and can boost electron mobility [40, 41].

The improvement in ν_{inj} shown in Fig. 6-9 can also be attributed to the reduction of electron effective mass.

Chapter 7

Conclusions and Further Recommendations

7.1 Conclusions

In the last period, various process-induced strain techniques have been developed and employed in CMOS devices. These techniques have been reported to improve carrier mobility by altering the energy band structure in the channel region through strain induced by adjacent materials.

In this thesis, we have investigated the effects of utilizing tensile strained nitride deposited as a SMT stressor to enhance nFET device performance. The SMT technique improves nFET performance and transconductance as expected owing to owing to the enhancement of electron mobility by the memorized strain for scaled nFET devices. In addition to the aforementioned nFET mobility improvement, the SMT nitride cap layer modifies the formation of the defects created by ion implantation, and hence alters the diffusion kinetics of the implanted n type dopants. Therefore, the SMT strain technique enables better short-channel-effect control, suppresses V_t roll-off, and improves DIBL owing to a reduced junction depth X_j as demonstrated by the SIMS profiles. Finally, when combined with the strained CESL technique, SMT provides supplementary strain to the n-channel, elucidating the prospects on strain silicon processes integration for aggressively scaled CMOS fabrication.

Moreover, the influence of interface states on the characteristics of nFET devices, and the associated electron mobility behavior with two different kinds of LPSiN (HT-SiN) and PESiN (LT-SiN) capping SMT nitrides have also been quantitatively analyzed. The strain from the SMT nitride can be tailored by nitride hydrogen composition, deposition method, and film thickness and shrinkage change upon thermal annealing. Furthermore, nFET Vt and DIBL have a similar level when using HT-SiN or LT-SiN for SMT application. Nevertheless, nFET devices with higher device performance, carrier mobility and lowered interface state density can be fabricated by the HT-SiN as the SMT capping nitride. Thus, it is important to optimize the nitride physical and chemical properties to maximize SMT strain memorization effect and to minimize other side effects to achieve the best performance simultaneously.

As MOSFET carrier transport becomes more ballistic, the carrier transport can be limited by carrier injection at the source as the scattering event happening in the channel is greatly reduced. Strain in the channel region affects not only the effective mass of the carrier, but also changes the backscattering rate γ_{sat}. The reduction of carrier conductivity mass therefore increases the source injection velocity V_{inj} in the highly scaled strained nFETs can still contributes to drive current enhancement.

7.2 Further Recommendations

Because of the limited timing for my PhD academic study and the heavy loading of my job, there is still many SMT related topics left that I could not finish in time. Nevertheless, they should be very interesting research work to be explored by the successors.

7.2.1 Correlation of Sidewall Spacer to the SMT Strain

For advanced semiconductor devices used in sub-90nm technology nodes, complex dopant profiles may be required for the S/D regions both in the lateral and vertical direction. This would require an implantation sequence in which the shielding effect of the gate electrode and spacer structure may be adapted in the lateral direction and subsequently performing a respective implantation process.

For example, after forming the gate electrode structure, and a thin offset spacer for protecting the gate sidewalls, dopants may be introduced to define shallower and reduced concentration lightly drain and source (LDD) regions. Then a main spacer with desired lateral distance to the gate electrode is formed. In a subsequent implantation process, an increased implantation energy and dose is used to form the drain and source regions.

It can be interesting to study the SMT effects after such spacer nitride deposition as they all involved the fundamental elements of SMT such as amorphized silicon post dopant implantation, cap layer deposition and thermal annealing. Thus, it is possible to demonstrate further nFET enhancement via multiple SMT techniques after each spacer formation steps [49].

For the conventional SMT applied after S/D implantation [19] widely adopted nowadays, it is suspected that the presence of the sidewall spacer elements during the anneal process may have a negative effect on the strained re-crystallization of the drain and source regions. That is, significant damage may also be created in the spacer elements after the high energy and dose implantation process for forming the LDD and the deep S/D regions, thereby possibly affecting their capability of suppressing a reduction in volume upon the re-crystallization process. Consequently, the S/D areas may re-grow with a less pronounced strain, and reducing the overall SMT strain transfer mechanism. The potential solution can be that the spacer elements may be considered to be removed prior to the deposition of the conventional SMT nitride material, so that the strained re-crystallized state of the S/D regions may be positioned closer to the channel region, thereby enhancing the overall strain-inducing mechanism.

7.2.2 Study Transistor Layout Effect upon SMT

The stress sources induce strain variations in the transistor channel because of different intrinsic stresses, transistor structure, and layout boundary conditions. Directional dependency occurs because the mobility sensitivity to stress (piezoresistance effects) may change drastically in sign and magnitude from one direction to another. Such layout-dependent variations significantly affect SPICE device modeling, and add to circuit design complexity. For SMT strained nFET devices, the nFET threshold voltage and carrier mobility can be also impacted by different layout [37]. As SMT involves a lot of process steps and variations in these process are magnified by the increasingly smaller features, it is suggested to study the layout effect thoroughly.

7.2.3 Integrate SMT Strain with High-K Gate and Metal Gate Technology

Incorporating a metal gate and high k into the CMOS process could allow another dimension to the novel process induced strain technique, as the metal gate could induce additional strain in the channel [105, 106], which is added to other strain-enhanced performance improvements. The strain from the metal gate can be tailored by its composition, deposition method, and film thickness and other process parameters. It should be interesting to co-optimize the conventional SMT strain and metal gate-induced strain simultaneously to achieve the best device performance at the same time.

References

[1] G. E. Moore, “Cramming more components onto integrated circuits”, Electronics, vol. 38, pp. 114, 1965.

[2] Y. Taur and T. H. Ning, Fundamentals of Modern VLSI Devices, Chapter One: Introduction, Cambridge Univ. Press, New York, 1998.

[3] W. Chang, C. C. Chen, J. C. Lu, S. J. Liou, W. J. Tsai, *et al.*, “300 mm process integration for 0.13µm generation with Cu/low-k interconnect technology”, IEDM Tech. Dig., pp. 603-606, 2001.

[4] J. C. Lu, W. Chang, S.M. Jang, C.H. Yu, and M.S. Liang, “Development of 300 mm low-k dielectric for 0.13 µm BEOL damascene process”, IEEE IITC, pp. 63-65, 2002.

[5] R. Choi, K. Onishi, C. Kang, S. Gopalan, R. Nieh, *et al.*, “Fabrication of high quality ultra-thin HfO_2 gate dielectric MOSFET’s using deuterium anneal”, IEDM Tech. Dig., pp. 613–616, 2002.

[6] Y. C. Yeo, Q. Lu, P. Ranade, H. Takeuchi, K. J. Yang, *et al.*, “Dual-metal gate CMOS technology with ultra-thin silicon nitride gate dielectric”, IEEE Electron Device Lett., vol. 22, pp. 227–229, May 2001.

[7] S. Sivakumar, “Lithography challenges for 32nm technologies and beyond”, IEDM Tech. Dig., pp. 1-4, 2006.

[8] T. Tezuka, N. Sugiyama, and S. Takagi, “Fabrication of strained Si on an ultrathin SiGe-on-insulator virtual substrate with a high-Ge fraction”, Appl. Phys. Lett., Vol. 79, pp. 1798, 2001.

[9] K. Kim, C. T. Chuang, K. Rim, and R.V. Joshi, “Performance assessment of scaled strained-Si channel-on-insulator (SSOI) CMOS”, IEEE Int. SOI Conf., pp. 17-19, 2002.

[10] B. H. Lee, A. Mocuta, S. Bedell, H. Chen, D. Sadana, *et al.*, “Performance enhancement on sub-70 nm strained silicon SOI MOSFETs on ultra-thin thermally mixed strained silicon/SiGe on insulator (TM-SGOI) substrate with raised S/D”, IEDM Tech. Dig., pp. 946–948, 2002.

[11] Y. Taur and T. H. Ning, Fundamentals of Modern VLSI Devices, Chapter Three: MOSFET Devices, Cambridge Univ. Press, New York, 1998.

[12] Y. C. Yeo, Q. Lu, W.-C. Lee, T.-J. King, and C. Hu, “Direct tunneling gate leakage current in transistors with ultrathin silicon nitride gate dielectric”, IEEE Electron Device Lett., vol. 21, pp. 540–542, 2000.

[13] Y. C. Yeo, T. J. King, and C. Hu, “Direct tunneling gate leakage and scalability of alternative gate dielectrics”, Appl. Phys. Lett., Vol. 81(11), pp. 2091-2093, 2002.

[14] J. L. Hoyt, H. M. Nayfeh, S. Eguchi, I. Aberg, G. Xia, *et al.*, "Strained silicon MOSFET technology", IEDM Tech. Dig. pp. 23-26, 2002.

[15] S. E. Thompson, M. Armstrong, C. Auth, S. Cea, R. Chau, *et al.*, "A logic nanotechnology featuring strained-silicon", IEEE Elec. Dev. Lett. 25(4), pp. 191-193, 2004.

[16] S. E. Thompson, G. Sun, Y. S. Choi, and T. Nishida, “Uniaxial-process-induced strained-Si: extending the CMOS roadmap”, IEEE Trans. Electron Devices, vol. 53, pp. 1010-1020, May 2006.

[17] S.E. Thompson, M. Armstrong, C. Auth, M. Alavi, M. Buehler, *et al.*, “A 90-nm logic technology featuring strained-silicon”, IEEE Trans. Electron Devices, vol. 51, pp. 1790-1797, Nov. 2004.

[18] P. R. Chidambaram, C. Bowen, S. Chakravarthi, C. Machala, and R. Wise, “Fundamentals of silicon material properties for successful exploitation of strain engineering in modern CMOS manufacturing”, IEEE Trans. Electron Devices, vol. 53(5), pp. 944-964, May 2006.

[19] C. H. Chen, T. L. Lee, T. H. Hou, C. L. Chen, C. C. Chen, *et al.*, “Stress memorization technique (SMT) by selectively strained-nitride capping for sub-65nm high-performance strained-Si device application”, Symposium on VLSI Technology, pp. 56-57, 2004.

[20] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, E. A. Fitzgerald, *et al.*, "Carrier mobilities and process stability of strained Si n- and p-MOSFETs on SiGe virtual substrates", J. Vac. Sci. Technol. B Volume 19, Issue 6, pp. 2268-2279, November 2001.

[21] Y. Taur and T. H. Ning, Fundamentals of Modern VLSI Devices, Chapter Two: Basic Device Physics, Cambridge Univ. Press, New York, 1998.

[22] M. V. Fischetti, F. Gámiz, and W. Hänsch, “On the enhanced electron mobility in strained-silicon inversion layers”, J. Appl. Phys.,Vol. 92, pp. 7320, 2002.

[23] M. V. Fischetti and S. E. Laux, “Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys”, J. Appl. Phys.,Vol. 80, pp. 2234-2252, 1996.

[24] M. L. Lee, E. A. Fitzgerald, M. T. Bulsara, M. T. Currie and A. Lochtefeld, “Strained Si, SiGe, and Ge channels for high mobility metal oxide semiconductor field effect transistors”, J. Appl. Phys. Vol. 97, 011101, 2005.

[25] N. Mohta and S. E. Thompson, "Mobility enhancement: the next vector to extend Moore's law", IEEE Circuits & Devices Magazine, pp. 18-23, Sep/Oct 2005.

[26] S. E. Thompson, G. Sun, K. Wu, J. Lim, and T. Nishida, "Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs", IEDM Tech. Dig., pp. 221-224, 2004.

[27] C. S. Smith, "Piezoresistance effect in germanium and silicon", Phys. Rev., vol. 94, no. 1, pp. 42-49, 1954.

[28] Y. Kanda, "Grahical representation of the piezoresistance coefficient in silicon",

IEEE Trans. Elec. Dev. 29, pp. 64, 1982.

[29] P. R. Chidambaram, B. A. Smith, L. H. Hall, H. Bu, S. Chakravarthi, *et al.*, “35% Drive Current Improvement from Recessed-SiGe Drain Extensions on 37nm Gate Length PMOS”, Symposium on VLSI Technology, pp. 48-49, 2004.

[30] K. W. Ang, K. J. Chui, H. C. Chin, Y. L. Foo, A. Du, *et al.*, "50nm silicon-on-insulator N-MOSFET featuring multiple stressors: silicon-carbon source/drain region and tensile stress silicon nitride liner", Symposium on VLSI Technology, pp. 66-67, 2006.

[31] Y. C. Yeo, “Enhancing CMOS transistor performance using lattice-mismatched materials in source/drain regions”, Semicond. Sci. Technol., 22(1), S177, 2007.

[32] A. Shimizu, K. Hachimine, N. Ohki, M. Koguchi, Y. Nonaka, *et al.*, “Local mechanical-stress control (LMC): a new technique for CMOS-performance enhancement”, IEDM Tech. Dig., pp.433-436, 2001.

[33] H. S. Yang, R. Malik, S. Narasimha, Y. Li, R. Divakaruni, *et al.*, “Dual stress liner for high performance sub-45nm gate length SOI CMOS manufacturing”, IEDM Tech. Dig., pp. 1075-1078, 2004.

[34] C. L. Cam, F. Guyader, C. D. Buttet, P. Guyader, G. Ribes, *et al.*, "Low cost drive current enhancement technique using shallow trench isolation induced stress for 45-nm node", Symposium on VLSI Technology, pp. 82-83, 2006.

[35] L. Washington, F. Nouri, S. Thirupapuliyur, G. Eneman, P. Verheyen, *et al.*, “pMOSFET with 200% mobility enhancement induced by multiple stressors”, IEEE Elec. Dev. Lett., 27(6), pp. 511-513, 2006.

[36] C. C. Lu, J. J. Huang, W. C. Luo, T. H. Hou, and T. F. Lei, “Strained silicon technology: mobility enhancement and improved short channel effect performance by stress memorization technique on nFET devices”, J. Electrochem. Soc., Volume 157, Issue 5, pp. H497-H500, 2010.

[37] C. Ortolland, P. Morin, C. Chaton, E. Mastromatteo, C. Populaire, *et al*., “Stress memorization technique (SMT) optimization for 45nm CMOS”, Symposium on VLSI Technology, pp. 78-79, 2006.

[38] C. Ortolland, Y. Okuno, P. Verheyen, C. Kerner, C. Stapelmann, *et al*., ”Stress memorization technique - fundamental understanding and low-cost integration for advanced CMOS technology using a nonselective process”, IEEE Transactions on Electron Devices, vol. 56, issue 8, pp. 1690-1697, 2009.

[39] P. Morina, F. Wacquanta, M. Juhela, C. Lavironb, and D. Lenoblea, “Influence of the spacer dielectric processes on PMOS junction properties”, Materials Science and Engineering: B Volumes 124-125, pp. 319-322, 2005.

[40] K. Rim, J. Chu, H. Chen, K.A. Jenkins, T. Kararsky, *et al*., “Characteristics and device design of sub-100-nm strained-Si N- and PMOSFETs”, Symposium on VLSI Technology, pp. 98-99, 2002.

[41] S. I. Takagi, J. K., and A. Toriumi, “Subband structure engineering for performance enhancement of Si MOSFETs”, IEDM Tech. Dig., pp. 219-222, 1997.

[42] K. Ota, K. Sugihara, H. Sayama, T. Uchida, H. Oda, *et al*., “Novel locally strained channel technique for high performance 55nm CMOS”, IEDM Tech. Dig., pp. 27-30, 2002.

[43] P. Kohli, S. Chakravarthi, A. Jain, H. Bu, C. Mehrota, *et al*., “Fundamental characterization of the effect of nitride sidewall spacer process on boron dose loss in ultra-shallow junction formation”, Mater. Sci. Eng. B 114–115, pp. 390-396, 2004.

[44] P. Kohli, A. Jain, H. Bu, S. Chakravarthi, C. Machala, *et al*., “Effect of nitride sidewall spacer process on boron dose loss in ultra shallow junction formation”, J. Vac. Sci. Tech. B, 22 (1), pp. 471, 2004.

[45] J. Yuan, V. Chan, N. Rovedo, V. Sardesai, N. Kanike, *et al*., “Blanket SMT with in situ N_2 plasma treatment on the <100> wafer for the low-cost low-power technology application”, IEEE Electron Device Letters, vol. 30, Issue 9, pp. 916-918, 2009.

[46] T. Miyashita, T. Owada, A. Hatada, Y. Hayami, K. Ookoshi, *et al*., "Physical and electrical analysis of the stress memorization technique (SMT) using poly gates and its optimization for beyond 45nm high performance applications", IEDM Tech. Dig., pp. 55-58, 2008.

[47] L. S. Adam, C. Chiu, M. Huang, X. Wang, Y. Wang, *et al*., “Phenomenological model for stress memorization effect from a capped-poly process”, IEEE International Conference on Simulation of Semiconductor Processes and Devices, Simulation of Semiconductor Processes and Devices (SISPAD) 2005, pp. 139-142, Sept. 2005.

[48] Y. G. Shin, J. Y. Lee, M. H.Park, and H. K. Kang, “The effect of stress on solid-phase epitaxial regrowth in As^+-implanted two-dimensional amorphized Si”, Jpn. J. Appl. Phys., 40, pp. 6192-6197, 2001.

[49] A. Wei, M. Wiatr, A. Mowry, A. Gehring, R. Boschke, *et al*., "Multiple stress memorization in advanced SOI CMOS technologies", Symposium on VLSI Technology, pp. 216-217, 2007.

[50] D. K. Schroder, Semiconductor Material and Device Characterization, 3rd edition, Chapter Four, John Wiley & Sons, New York, 1998.

[51] International Technology Roadmap for Semiconductors, 2006 Update.

[52] R. B. Fair, J. J. Wortman and J. Liu, “Modeling rapid thermal diffusion of arsenic and boron in silicon”, J. Electrochem. Soc., Volume 131, Issue 10, pp. 2387-2394, 1984.

[53] Y. Sasaki, K. Itoh, and S. I. Tanuma, “Enhanced diffusion of implanted arsenic in

silicon at an early stage of annealing", Jpn. J. Appl. Phys. 28, pp. 1421-1425, 1989.

[54] J. D. Plummer and M. D. Deal, Silicon VLSI Technology: Fundamentals, Practice and Modeling, Chapter-7, Prentice Hall, 2002.

[55] P. M. Fahey, P. B. Griffin, and J. D. Plummer, "Point defects and dopant diffusion in silicon", Rev. Mod. Phys., vol. 61, no. 2, pp. 289–384, Apr. 1989.

[56] A. Ural, P. B. Griffin, and J. D. Plummer, "Fractional contributions of microscopic diffusion mechanisms for common dopants and self diffusion in silicon", J. Appl. Phys., vol. 85, no. 9, pp. 6440–6446, May 1999.

[57] H. J. Gossmann, T. E. Haynes, P. A. Stolk, D. C. Jacobson, H. Gilmer, *et al.*, "The interstitial fraction of diffusivity of common dopants in Si", Appl. Phys. Lett., vol. 71, no. 26, pp. 3862–3864, Dec. 1997.

[58] A. Antonelli and J. Bernholc, "Pressure effects on self-diffusion insilicon", Phys. Rev. B (Condensed Matter), vol. 40, no. 15, pp. 10643–10646, Nov. 1989.

[59] O. Sugino and A. Oshiyama, "Microscopic mechanism of atomic diffusion in Si under pressure", Phys. Rev. B (Condensed Matter), vol. 46, no. 19, pp. 12335–12341, Nov. 1992.

[60] M. Tang, L. Colombo, J. Zhu, and T. D. de la Rubia, "Intrinsic point defects in crystalline silicon: tight-binding molecular dynamics studies of self-diffusion, interstitial-vacancy recombination, and formation volumes", Phys. Rev. B (Condensed Matter), vol. 55, no. 21, pp. 14279–14289, Jun. 1997.

[61] S. T. Dunham, M. Diebel, C. Ahn, and C. L. Shih, "Calculations of effect of anisotropic stress/strain on dopant diffusion in silicon under equilibrium and nonequilibrium conditions", J. Vac. Sci. Technol., vol. 24, no. 1, pp. 456–461, Jan./Feb. 2006.

[62] M. Diebel and S. T. Dunham, "Ab-initio calculations to predict stress effects on

defects and diffusion in silicon”, in Proc. SISPAD, pp. 147, 2003.

[63] M. P. Hughey, and R. F. Cook, “Massive stress changes in plasma-enhanced chemical vapor deposited silicon nitride films on thermal cycling”, Thin Solid Films, Vol. 460, Issue 1-2, pp. 7-16, 2004,

[64] M. P. Hughey, and R. F. Cook, “Stress development kinetics in plasma-enhanced chemical-vapor-deposited silicon nitride films”, J. Appl. Phys., Vol. 97, 114914, 2005.

[65] C. C. Lu, J. J. Huang, W. C. Luo, T. H. Hou, and T. F. Lei, “Characterization of highly strained nFET device performance and channel mobility with SMT”, J. Electrochem. Soc., Volume 157, Issue 7, pp. H705-H710, 2010.

[66] Qi, Xiang, ”Strained siliocn MOSFETs having reduced diffusion of n-type dopants”, US patent 0054164 A1., 2005.

[67] N. Kong, S. K. Banerjee, T. A. Kirichenko, S. G. H. Anderson, and M. C. Foisy, “Enhanced and retarded diffusion of arsenic in silicon by point defect engineering”, Appl. Phys. Lett. 90, 062107, 2007.

[68] T. H. Ning, P. W. Cook, R. H. Dennard, C. M. Osburn, S. E. Schuster, *et al.*, “1µm MOSFET VLSI technology: Part IV - hot-electron design constraints", IEEE Transactions on Electron Devices, vol. ED-26, no. 4, pp. 346-352, Apr. 1979.

[69] E. Takeda, H. Kume, T. Toyabe, and S. Asai, “Submicrometer MOSFET structure for minimizing hot-carrier generation", IEEE Transactions on Electron Devices, vol. ED-29, no. 4, pp. 611-618, Apr. 1982.

[70] K. N. Quader, E. R. Minami, W. J. Huang, P. K. Ko, and C. Hu, “Hot-carrier-reliability design guidelines for CMOS logic circuits", IEEE Journal of Solid State Circuits, vol. 29, no. 3, pp. 253-261, Mar. 1994.

[71] E. Takeda, “A cross section of VLSI reliability-hot carriers, dielectrics and

metallization", Semicond. Sci. Technol., vol. 9, pp. 971-987, 1994.

[72] C. Y. Lu, H. C. Lin, Y. J. Lee, and C. C. Chao, "Impacts of SiN-capping layer on the device characteristics and hot-carrier degradation of nMOSFETs", IEEE Transactions on Device and Materials Reliability, Vol. 7, Issue 1, pp. 175-180, Mar 2007.

[73] T. Y. Lu, C. M. Wang, and T. S. Chao, "Enhancement of stress-memorization technique on nMOSFETs by multiple strain-gate engineering", Electrochem. Solid-State Lett., Volume 12, Issue 1, pp. H4-H6, 2009.

[74] P. T. Liu, C. S. Huang, P. S. Lin, D. Y. Lee, S. W. Tsao, *et al.*, "Anomalous gate-edge leakage induced by high tensile stress in NMOSFET", IEEE Electron Device Letters, vol. 29(11), pp. 1249-1251, Oct 2008.

[75] G. G. Stoney, "The tension of metallic films deposited by electrolysis", Proc. R. Soc. London Ser., Vol. A82, pp. 172-175, 1909.

[76] Z. Lu, P. S. Filho, G. Stevens, M. J. Williams, and G. Lucovsky, "Fourier transform infrared study of rapid thermal annealing of a-Si:N:H(D) films prepared by remote plasma-enhanced chemical vapor deposition", J. Vac. Sci. Technol. A, Vol. 13, No. 3, pp. 607-613, May/Jun 1995.

[77] W. A. Lanford and M. J. Rand, "The hydrogen content of plasma-deposited silicon nitride", J. Appl. Phys., Vol. 49 (4), pp. 2473, 1978.

[78] M. P. Hughey, and R. F. Cook, "Hydrogen diffusion as the rate-limiting mechanism of stress development in dielectric films", Appl. Phys. Lett., Vol. 85, pp. 404, 2004.

[79] K. Ueno, T. Ritzdorf and S. Grace, "Seed layer dependence of room-temperature recrystallization in electroplated copper films", J. Appl. Phys., Vol. 86, pp. 4930, 1999.

[80] M. H. Yu, J. H. Li, H. H. Lin, C. H. Chen, K. C. Ku, *et al.*, "Relaxation-free

strained SiGe with super anneal for 32nm high performance PMOS and beyond", IEDM Tech. Dig., pp. 867-870, 2006.

[81] H. Miura, H. Ohta, N. Okamoto, and T. Kaga, "Crystallization-induced stress in silicon thin films", Appl. Phys. Lett., Vol. 60, pp. 2746-2748, 1992.

[82] H. Miura, and N. Okamoto, "Crystallization-induced stress in phosphorus-doped amorphous silicon thin films", J. Appl. Phys., Vol. 75, pp. 4747-4749, 1994.

[83] Y. Kimura, M. Kishi and T. Katoda, "Effects of elastic stress introduced by a silicon nitride cap on solid-phase crystallization of amorphous silicon", J. Appl. Phys., Vol. 86, pp. 2278-2280, 1999.

[84] Y. Kimura, M. Kishi, and T. Katoda, "The model of solid phase crystallization of amorphous silicon under elastic stress", J. Appl. Phys., Vol. 87, pp. 4017-4020, 2000.

[85] D. K. Schroder, Semiconductor Material and Device Characterization, 3rd edition, Chapter Eight, pp. 225, John Wiley & Sons, New York, 1998.

[86] K. Lee, J. S. Choi, S. P. Sim, and C. K. Kim, "Physical understanding of low-field carrier mobility in silicon MOSFET inversion layer", IEEE Trans. Electron Devices, Vol. 38, Issue 8, pp. 1905-1912, Aug 1991.

[87] S. M. Sze, and K. K. Ng, Physics of Semiconductor Devices, 3rd edition, Chapter Four, John Wiley & Sons, Hoboken, New Jersey, 2007.

[88] R. O. Weast, CRC Handbook of Chemistry and Physics, 68th Ed., CRC, Boca Raton, FL, 1987–1988.

[89] M. Yamamura, T. Matsuki, T. Robata, T. Watanabe, S. Inumiya, *et al.*, "Improvement in NBTI by catalytic-CVD silicon nitride for HP-65nm technology", Symposium on VLSI Technology, pp. 88-89, 2005.

[90] E. Cartier, D. A. Buchanan, and G. J. Dunn, "Atomic hydrogen-induced interface degradation of reoxidized-nitrided silicon dioxide on silicon", Appl. Phys. Lett.,

Vol. 64, pp. 901-903, 1994.

[91] R. E. Stahlbush, and E. Cartier, “Interface defect formation in MOSFETs by atomic hydrogen exposure”, IEEE Trans. Nuclear Science, Vol. 41, Issue 6, Part 1, pp. 1844-1853, Dec 1994.

[92] E. Cartier, J. H. Stathis, and D. A. Buchanan, “Passivation and depassivation of silicon dangling bonds at the Si/SiO_2 interface by atomic hydrogen”, Appl. Phys. Lett., Vol. 63, pp. 1510-1512, 1993.

[93] D. L. Griscom, “Diffusion of radiolytic molecular hydrogen as a mechanism for the post-irradiation buildup of interface states in SiO_2-on-Si structures”, J. Appl. Phys., Vol. 58, pp. 2524, 1985.

[94] M. S. Lundstrom, “Elementary scattering theory of the Si MOSFET”, IEEE Electron Device Lett., vol. 18, pp. 361–363, July 1997.

[95] F. Assad, Z. Ren, S. Datta, and M. S. Lundstrom, “Performance limits of silicon MOSFET’s”, IEDM Tech. Dig., Dec. 1999, pp. 547–550.

[96] F. Assad, Z. Ren, D. Vasileska, S. Datta, and M. S. Lundstrom, “On the performance limits for Si MOSFET’s: A theoretical study”, IEEE Trans. Electron Devices, vol. 47, pp. 232–240, Jan. 2000.

[97] M. S. Lundstrom and Z. Ren, “Essential physics of carrier transport in nanoscale MOSFET’s”, IEEE Trans. Electron Devices, vol. 49, pp. 133–141, Jan. 2002.

[98] M. S. Lundstrom, Fundamentals of Carrier Transport, 2nd ed., Cambridge Univ. Press, Cambridge, U.K, 2000.

[99] M. J. Chen, H. T. Huang, K. C. Huang, P. N. Chen, C. S. Chang, *et al.*, “Temperature dependent channel backscattering coefficients in nanoscale MOSFETs”, IEDM Tech. Dig., pp. 39–42, 2002.

[100] H. N. Lin, H. W. Chen, C. H. Ko, C. H. Ge, H. C. Lin, *et al.*, "Characterizing the channel backscattering behavior in nanoscale strained complementary metal

oxide semiconductor field-effect transistors", Japanese Journal of Applied Physics. 45 (11), pp. 8611-8617, 2006.

[101] H. N. Lin, H. W. Chen, C. H. Ko, C. H. Ge, H. C. Lin, *et al.*, “The Impact of Uniaxial Strain Engineering on Channel Backscattering in Nanoscale MOSFETs”, Symposium on VLSI Technology, pp. 174–175, 2005.

[102] F. H. Gaensslen, V. L. Rideout, E. J. Walker, J. J. Walker, “Very Small MOSFET’s for low-temperature operation”, IEEE Trans. on Elec. Dev. 24(3), 218, 1977.

[103] K. W. Anga, H. C. China, K. J. Chuia, M. F. Lia, G. S. Samudraa, *et al.*, “Carrier backscattering characteristics of strained silicon-on-insulator n-MOSFETs featuring silicon–carbon source/drain regions”, Solid-State Electronics, Volume 51, Issues 11-12, pp. 1444-1449, 2007.

[104] F. Assad, Z. Ren, S. Datta, and M. Lundstrom, ”Performance limits of silicon MOSFETs”, IEDM Tech. Dig., pp. 547–550, 1999.

[105] C. Y. Kang, R. Choi, M. M. Hussain, J. Wang, Y. J. Suh, *et al.*, “Effects of metal gate-induced strain on the performance of metal-oxide-semiconductor field effect transistors with titanium nitride gate electrode and hafnium oxide dielectric”, Appl. Phys. Lett., Vol. 91 (3), 033511, 2007,

[106] C. Y. Kang, R. Choi, S. C. Song, K. Choi, B. S. Ju, *et al.*, “A novel electrode-Induced strain engineering for high performance SOI FinFET utilizing Si (110) channel for both N and PMOSFETs”, IEDM Tech. Dig., pp. 885–888, 2006.

Curriculum Vitae

姓名: **陸志誠**

性別: 男

出生: 民國 61 年 08 月 27 日

籍貫: 台灣省台北縣

住址: 300, 新竹市光復路一段 476 巷 56 號 8 樓

學歷:

1) **國立成功大學電機系學士班** [民國 79 年 9 月-民國 83 年 6 月]

2) **University of Wisconsin-Madison, USA**
Electrical and Computer Engineering 電機研究所碩士班

[民國 86 年 9 月- 民國 88 年 6 月]

3) **國立交通大學電子研究所博士班** [民國 92 年 9 月-民國 99 年 7 月]

博士論文題目:

應力記憶技術應用於奈米尺寸 N 型金氧半場效電晶體特性之研究

Study of Stress Memorization Technique on the Characteristics of Nanoscale nMOS Transistors

Publication Lists

1. International Journal:

[1] **Chih-Cheng Lu**, Tan-Fu Lei, Tuo-Hung (Alex) Hou, Chao-Hsin Chien, M. H. Liao, T. L. Lee and S. M. Jang, “An Investigation of Strain Silicon Technology on Highly Strained, Highly Scaled nFET Devices”, *International Journal of Electrical Engineering*, Vol.16, No.4, 2009.

[2] **Chih-Cheng Lu**, Jiun-Jia Huang, Wun-Cheng Luo, Tuo-Hung Hou, Tan-Fu Lei, “Strained Silicon Technology: Mobility Enhancement and Improved Short Channel Effect Performance by Stress-Memorization-Technique on nFET Devices”, *Journal of Electrochemical Society*, Vol. 157, Issue 5, pp. H497-H500, 2010.

[3] **Chih-Cheng Lu**, Jiun-Jia Huang, Wun-Cheng Luo, Tuo-Hung Hou, Tan-Fu Lei, “Characterization of Highly Strained nFET Device Performance and Channel Mobility with Stress-Memorization-Technique”, *Journal of Electrochemical Society*, Vol. 157, Issue 7, pp. H705-H710, 2010.

2. International Letter: N/A.

3. International Conference:

[1] **Chih-Cheng Lu**, M.H. Liao, Lingyen Yeh, Tan-Fu Lei, Chao-Hsin Chien, T.L. Lee, S.M. Jang, “Investigation on the Limitation of Strain Silicon Technology,” *2008 International Electron Devices and Materials Symposia,* CP-428, 2008.

4. Local Conference: N/A.