Doctoral Thesis

A Study on Carrier Transport Properties of Vertically-Stacked Nanowire Transistors

A DISSERTATION SUBMITTED TO THE
DEPARTMENT OF ELECTRONICS AND APPLIED PHYSICS
INTERDISCIPLINARY GRADUATE SCHOOL OF
SCIENCE AND ENGINEERING
TOKYO INSTITUTE OF TECHNOLOGY
FOR THE DEGREE OF DOCTOR OF ENGINEERING

Kiichi Tachi

March 2011

Supervisor: Professor Hiroshi Iwai

ABSTRACT

Vertically-stacked silicon nanowire MOSFETs (SNWTs) were experimentally investigated as one of the possible solutions to achieve both high speed, low power consumption in combination with high integration capabilities for future LSI applications. To evaluate the potentials, analyze and improve the performance of these devices, source/drain series resistance for thick source/drain region were studied. Carrier transport mechanisms and the controllability of threshold voltage for vertically-stacked SNWTs with separated gates were also investigated.

The influence of *in situ* doped SEG source/drain was examined for vertically-stacked channel MOSFETs. A large enhancement, by a factor of 2 in the drive current, was obtained when *in situ* doped SEG process was adopted. Detailed parameter extraction from the electrical measurements showed the R_{SD} values can be reduced by 90 and 75% for *n*- and *p*-FETs, respectively, when in situ doped SEG is reinforced by adding ion implantation. On the other hand, by combining the ion implantation to SEG process, V_T roll-off characteristics and the effective mobility behavior are slightly degraded. Mobility analysis revealed an increase in the Coulomb scattering with L_G scaling, indicating the diffusion of dopant atoms from S/D regions. Further improvements in the performance can be sought by optimizing the S/D activation annealing step.

In order to enhance the performance of the vertically-stacked nanowire MOSFETs, the carrier transport limiting components caused by short channel effects were assessed. The optimization of drive currents will have to take into account specific effects to vertically-stacked SNWTs. In particular, the use of SiGe sacrificial layer to make vertically-stacked channels cause large mobility degradation due to the surface roughness, resulted from the damage of plasma etching. This leads to the poor ballisticity in the short channel SNWTs. Hydrogen annealing was shown to be advantageous for improving the surface-roughness limited mobility. Charge pumping measurements, however, revealed that circular-shaped SNWTs, which are formed by annealing, have a higher interface trap density (D_{it}) than rectangular ones, leading to low-field mobility degradation. This high D_{it} could be caused by the

continuously-varying surface orientation. The resulting additional coulomb scattering could partly explain the quite low mobility in 5 nm diameter SNWTs together with the already known transport limitations in NWs. Vertically-stacked SNWTs with independent gates by internal spacers between the nanowires to control threshold voltage (named Φ-FETs), were evaluated. Φ-FETs demonstrated excellent V_T controllability due to inter-gate coupling effects. Numerical simulations to optimize Φ-FETs structures show that when the spacer width is reduced, the DIBL value can be lowered by a factor of 2 compared to independent-gate FinFETs with the same silicon width. The superior scaling of Φ-FETs with narrow spacer results from a better electrostatic control which also attenuates the inter-gate coupling.

Overall it was shown that using vertical stack structure can increase the drive current density while allowing for better threshold voltage controllability. As for the performance benchmark, nanowires with a diameter of 10 nm, showed the most acceptable balance between mobility, short channel effect. However, to further improve the device performance, process induced surface damage of nanowires must be mitigated.

ACKNOWLEDGEMENTS

I would like to thank, first and foremost, my supervisor, Professor Hiroshi Iwai of Tokyo Institute of Technology for his guidance, encouragement and continuous support throughout my research.

I am also grateful to Professor Kenji Natori, Professor Kazuo Tsutsui, Professor Eisuke Tokumitsu, Professor Masahiro Watanabe and Professor Shun-ichiro Ohmi of Tokyo Institute of Technology for reviewing the thesis and for valuable advice.

All studies referred to in this thesis were performed at CEA-Leti in France during the last three years. I am indebted to everyone with whom I have worked at CEA-Leti over the years.

I am also grateful to Dr. Simon Deleonibus, Dr. Olivier Faynot and Dr. Thomas Ernst of CEA-Leti's Electronics Nanodevices Laboratory, and Professor Sorin Cristoloveanu of Grenoble Institute of Technology for their encouragement, useful advices and great help whenever I met difficult and subtle problem.

I would like to thank Professor Kuniyuki Kakushima, Professor Parhat Ahmet, Professor Nobyuki Sugii, and Professor Akira Nishiyama of Tokyo Institute of Technology for their encouragement support of my studies.

I am also grateful to research colleagues of Iwai's Laboratory for the kind friendship.

I would like to express sincere gratitude to laboratory secretaries, Ms. M. Karakawa, Ms. A. Matsumoto, and Ms. M. Nishizawa for their invaluable assistance regarding administrative issues.

This work was partially performed as part of the IBM-STMicroelectronics-CEA-LETI Development Alliance.

To my family for their kind
support and encouragement throughout my life.

Merci beaucoup!

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 5 Carrier Transport Properties of Vertically-Stacked Nanowire MOSFETs

Chapter 7 Conclusions

LIST OF SYMBOLS AND ABBREVIATIONS

A

AFM	Atomic Force Microscopy
ALD	Atomic Layer Deposition
α_μ	Mobility degradation factor

B

BEOL	Back-End-Of-the-Line
BOX	Buried Oxide
β	Gain factor of the transistor

C

CMOS	Complementary MOS
CMP	Chemical Mechanical Polishing
C	Capacitance [F]
C_{dm}	Maximum depletion-layer capacitance per area [F/cm^2]
C_{ox}	Gate oxide capacitance per area [F/cm^2]

C_{in}	Input capacitance of the next stage or stages [F]
C_{out}	Output capacitance of the switching inverter [F]
C_{wire}	Wiring capacitance [F]
C_G	Gate capacitance [F]
C_{GC}	Gate-to-channel capacitance [F]
C_L	Load capacitance [F]
$C_L(W)$	Load capacitance depending on gate width [F]

D

DG	Double gate
DIBL	Drain–induced barrier lowering
D_{it}	Interface trap density

E, ε

EI	Electrostatic integrity
EOT	Equivalent oxide thickness
E_{eff}	Transverse effective electric field [V/cm]
E_g	Band gap [eV]
$E_{lateral}$	Lateral electric field [V/cm]
E_F	Fermi energy [eV]
E_s	Electric field near source edge [V/cm]
ε_0	Vacuum permittivity [8.85 x 10^{-14} F/cm]
ε_s	Semiconductor permittivity [Si: 1.04 x 10^{-12} F/cm]
ε_{ox}	Silicon-dioxide permittivity [3.45 x 10^{-13} F/cm]
ε_r	Relative permittivity

F

FDSOI	Fully-depleted SOI
f	Frequency [Hz]

f_{clk}	Clock frequency
Φ_B	Potential barrier [eV]
Φ_{MS}	Work-function difference between metal and silicon [eV]

G

GIDL	Gate-Induced Drain Leakage
GAA	Gate-All-Around
g_m	Transconductance [S]

H

HM	Hard Mask
HP	High Performance logic
HTO	High-Temperature Oxide
H_{fin}	Fin height [cm]
H_{NW}	Nanowire hight [cm]
η_F	Reduced Fermi energy [eV]

I

IG	Independent-Gate
I/I	Ion Implantation
I_{Dlin}	Drain current at low drain voltage [A]
I_{Dsat}	Saturation current [A]
I_{DS}	Drain-to-source current [A]
I_{ON}	On-state current [A]
I_{OFF}	Off-state current [A]
I_N	nMOSFET drain current in a CMOS inverter [A]
I_P	pMOSFET drain current in a CMOS inverter [A]
I_{GIDL}	Gate-induced drain leakage current [A]
I_{cp}	Charge pumping current [A]

I_{leak}	Total leakage current including gate and junction leakages [A]
I_{gate}	Gate leakage current [A]
I_{subth}	Subthreshold current [A]
I_{Tr}	Transistor current corresponding to the threshold voltage [A]

J

J_G	Gate leakage current density [A/cm^2]

K, κ

k	Boltzmann constant [=8.617 x 10^{-5} eV/K]
κ	Scaling factor, Relative dielectric constant

L

LOP	Low Operating Power logic
LSTP	Low Standby Power logic
LSI	Large-Scale Integrated circuit
L_G	Physical gate length [cm]
L_{eff}	Electrical channel length [cm]
L_{ov}	Gate overlap length [cm]
L_m	Mask gate length [cm]
λ	Backscattering mean free path of carriers [cm]
λ	Natural length [cm]
l	Critical length of scattering [cm]

M, μ

MBCFET	Multi-Bridge Channel MOSFET
MCFET	Multi-Channel MOSFET
MOSFET	Metal-Oxide-Silicon Field-Effect-Transistor
m	Body-effect coefficient

m^*	Carrier effective conduction mass
μ_{eff}	Effective mobility [$cm^2/V \cdot s$]
μ_s	Mobility near source edge [$cm^2/V \cdot s$]
μ_{ph}	Phonon limited mobility [$cm^2/V \cdot s$]
μ_{sr}	Surface-roughness limited mobility [$cm^2/V \cdot s$]
μ_{cb}	Coulomb limited mobility [$cm^2/V \cdot s$]
μ_0	Low-field mobility [$cm^2/V \cdot s$]

N

N_{INV}	Inversion charge density [cm^{-2}]
N_{INV}^{source}	Inversion charge density near source edge [cm^{-2}]
N_a	Accepter impurity concentration [cm^{-3}]
N_d	Donor impurity concentration [cm^{-3}]
N_t	Oxide trap density

O

P

Q

q	Unit electronic charge [C]
Q_i	Inversion charge per unit gate area [C/cm^2]
Q_B	Total gate depletion charge [C]
Q_{inv}	Inversion charge [C]
θ_{eff}	Mobility reduction factor

θ_1	First order mobility reduction coefficient
θ_2	Second mobility reduction coefficient

R

RMS	Root Mean Square
RP-CVD	Reduced-Pressure Chemical Vapor Deposition
R_{SD}	Source/drain series resistance
R_S	Source resistance
R_D	Drain resistance
r	Backscattering rate near-source region

S

SEG	Selective Epitaxial Growth
SEM	Scanning Electron Microscopy
SNM	Static Noise Margin
SNWT	Silicon NanoWire field-effect Transistor
SOI	Silicon-On-Insulator
SON	Silicon-On-Nothing
SRAM	Static Random Access Memory
STI	Shallow Trench Isolation
S_Y	Y-function slope
SS	Subthreshold swing [V/decade]

T, τ

TEM	Transmission electron microscopy
T	Temperature [K]
T_{dep}	Thickness of depletion layer
t_{ox}	Gate oxide thickness [cm]
t_{Si}	Silicon thickness [cm]

t_r	Rising time [s]
t_f	Falling time [s]
τ	Average time between two collisions [s]
τ_n	nMOSFET pull-down delay [s]
τ_p	pMOSFET pull-up delay [s]

U

V

VDT	Voltage-Doing Transformation
V_T	Threshold voltage [V]
V_{DD}	Power-supply voltage [V]
V_G	Gate voltage [V]
V_S	Source voltage [V]
V_D	Drain voltage [V]
V_{DS}	Source–drain voltage [V]
V_{Dsat}	Drain saturation voltage [V]
V_{in}	Input voltage [V]
V_{out}	Output voltage [V]
V_{fb}	Flat-band voltage [V]
V_{base}	Gate pulse base level [V]
V_{Tlong}	Threshold voltage in a long-channel device [V]
V_{bi}	Built-in voltage [V]
$V_{thermal}$	Thermal voltage (= kT/q) [V]
v_s	Average carrier velocity near the source edge [cm/s]
v_{sat}	Saturation velocity [cm/s]
v_0	Ballistic velocity of carriers [cm/s]
v_{inj}	Injection carrier velocity at the top of the barrier near the source edge [cm/s]

W

W	Gate width [cm]
W_d	Depletion-layer depth [cm]
W_{dm}	Maximum depletion-layer depth [cm]
W_{space}	Lateral space between nanowires for multi-finger [cm]
W_{pitch}	Nanowire Pitch for multi-finger [cm]
W_{NW}	Nanowire width [cm]
W_{fin}	Fin width [cm]
W_m	Mask gate width [cm]
W_{TOT}	Total gate width [cm]
W_{Top}	Top-view width [cm]
W_{sp}	Spacer width [cm]

X

x_j	Source/drain junction depth [cm]

Y, ψ

ψ_B	Difference between Fermi level and intrinsic level
ψ_S	Surface potential [eV]

Z

$\mathfrak{I}_n$	Fermi integral of the *n*th order

CHAPTER 1

INTRODUCTION – MOSFET SCALING –

Silicon-based large-scale integrated circuits (LSIs) have been rapidly developed in the past 40 years with an unprecedented growth of the semiconductor industry, bringing an enormous impact on the way people work and live. This evolution is owed to the continued downsizing of metal-oxide-silicon field-effect-transistors (MOSFETs). Recently, however, the conventional miniaturization has caused various problems such as threshold voltage roll-off, subthreshold leakage, gate leakage, etc.

In this chapter, firstly, the conventional scaling method and the basic operation of MOSFET are described. Then the specific features of short-channel MOSFETs are considered. The latter half of this chapter covers the proposed solutions and challenges to continue the scaling toward purpose of this thesis.

CHAPTER 1 CONTENTS

1.1 MOSFET DOWNSIZING

Since the invention of the CMOS (complementary MOS) in 1963, which both *n*-channel and *p*-channel MOSFETs are constructed simultaneously on the same substrate, the number of transistors on a chip has increased by MOSFET downsizing in accordance with Moore's Law proposed by Gordon E. Moore in 1965 [1.1]. His prediction states that the number will double about every two years. That has happened fairly regularly up to the current time as shown in Figure 1.1 (a) [1.2]. In addition, the average transistor price has decreased markedly over the past four decades (Figure 1.1 (b)) [1.3]. Increasing transistor budgets and decreasing average price per transistor opens up the possibility for high-speed designs that were not technologically or economically feasible in the past (Figure 1.1 (c)) [1.2]. The speed improvements have been achieved by MOSFET scaling. In this section, the basic MOSFET operation principle and the scaling rule are described.

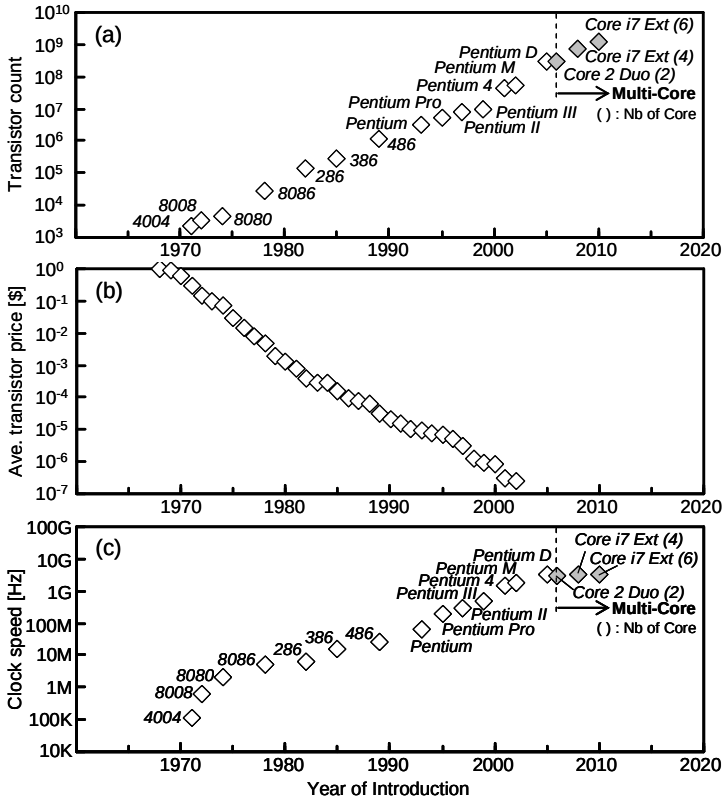


Figure 1.1 (a) Number of transistors in Intel's microprocessor chips, (b) Average transistor price by year, (c) Clock speed of Intel's microprocessor [1.2, 1.3].

1.1.1 Basic CMOS Operation

A schematic three-dimensional illustration of conventional CMOS transistors of the year early 2000 or earlier, consisting of an n-channel MOSFET (nMOSFET) and a p-channel MOSFET (pMOSFET) integrated on the same chip, is shown in Figure 1.2. The MOSFET is a four-terminal device with the terminal s designed as *gate*, *source*, *drain* and *substrate* or *body*. The nMOSFET consists of an p-type silicon (Si) substrate into which n+ regions, the source and the drain, are formed (e.g., by ion implantation). The gate electrode is usually made of heavily doped polysilicon (poly-Si) and is insulated from the substrate by a thin silicon dioxide (SiO_2) films, the gate oxide. The SiO_2 film is usually formed by thermal oxidation of silicon substrate. The surface region under the gate oxide between the source and the drain is called the *channel* region and is critical for current conduction in a MOSFET. One of the main reasons for successfully developing MOSFET is the presence of the SiO_2 film which is enable to form thermally stable and high quality interface between the gate oxide and the channel. To obtain low resistive contact, metal silicide is formed on the polysilicon gate as well as on the source and drain diffusion regions. A MOSFET is surrounded by a thick oxide called the field oxide to isolate it from the adjacent devices. The key physical parameters are gate length (L_G), gate width (W), source/drain junction depth (x_j), gate oxide thickness (t_{ox}), and channel dopant concentration (N_a, N_d).

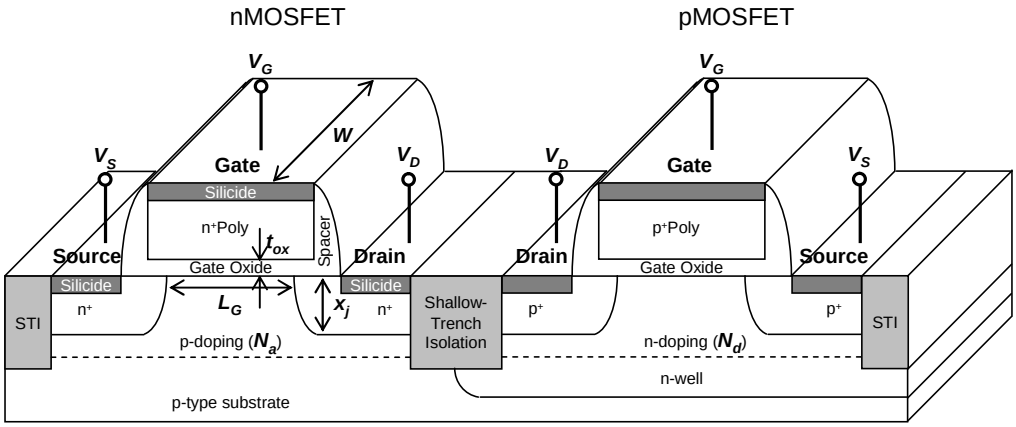


Figure 1.2 Three-dimensional view of basic CMOS structure. V_G is gate voltage, V_S is source voltage, V_D is drain voltage, L_G is gate length, t_{ox} is gate oxide thickness, N_a is accepter impurity density, N_d is donor impurity density, x_j is junction depth

■ Drain–Source Current Model

In the MOSFET, an inversion charge layer at the silicon–gate oxide interface acts as a conducting channel. For example, in an nMOSFET, the substrate is *p*-type silicon and the inversion charge consists of electrons that form a conducting channel between the n^+ ohmic source and the drain contacts. The onset of strong inversion is defined in terms of a threshold voltage (V_T) being applied to the gate electrode relative to the other terminals. Depending on the gate voltage (V_G) and source–drain voltage (V_{DS}), a MOSFET can be biased in one of the three regions: linear region, saturation region, and subthreshold region, as shown in Figure 1.3.

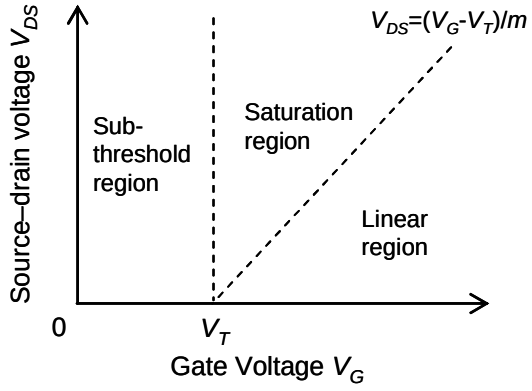


Figure 1.3 Three regions of a MOSFET operation in the V_{DS}–V_G plane [1.4].

Here the drain–source current (I_{DS}) model for long-channel MOSFETs is described. One of the key assumption is the *gradual–channel approximation*, which assumes that the transverse field in the channel in much larger than the longitudinal field, that is one-dimensional. This allows us to express I_{DS} as [1.4]

$$I_{DS} = \mu_{eff} \frac{W}{L_G} \int_0^{V_{DS}} (-Q_i(V))dV \ , \tag{1.1}$$

where μ_{eff} is the effective mobility of carriers and Q_i is the inversion charge per unit gate area. The drain–source current based on the gradual–channel approximation is valid for most of the V_{DS} regions except beyond saturation voltage (V_{Dsat}). Beyond V_{Dsat}, I_{DS} stays constant at a saturation value (I_{Dsat}), independent of V_{DS}. In addition, the use of the *charge-sheet model*, which assumes that the inversion layer is a charge sheet as zero

thickness, allows us to simply express the I_{DS} in each V_{DS} region as the following equations:

$$I_{DS} = \mu_{eff} C_{ox} \frac{W}{L_G}(V_G - V_T)V_{DS} \qquad \text{for } V_{DS} \le V_G - V_T \text{ (linear region)} \tag{1.2}$$

$$I_{Dsat} = \mu_{eff} C_{ox} \frac{W}{L_G}\frac{(V_G - V_T)^2}{2m} \qquad \text{for } V_{DS} > V_{Dsat} \text{ (saturation region)} \tag{1.3}$$

where C_{ox} (= ε_{ox}/t_{ox}) is the gate oxide capacitance per area, ε_{ox} is the oxide permittivity of the gate oxide, m is the body-effect coefficient related to doping concentration and oxide thickness.

When the gate bias is below the threshold voltage and the semiconductor is in weak inversion or depletion, the corresponding drain current is called the subthreshold current. The subthreshold region tells how sharply the current drops with gate bias. Figure 1.4 shows schematic I_{DS} –V_G curves of nMOSFETs. The I_{DS} at V_G = 0 and V_{DS} = V_{DD} is equivalent to the off-state current (I_{OFF}), while the I_{DS} at V_G = V_{DS} = V_{DD} is equivalent the on-state currents (I_{ON}). Here, the V_{DD} is the power-supply voltage. The transition from one state to another defines the V_T of the MOSFET. In Figure 1.4 (a), I_{DS} on a linear scale appears to approach zero immediately below the threshold voltage. On a logarithmic scale, however, it is seen that the descending I_{DS} remains at nonnegligible levels for several tenths of a volt below V_T (Figure 1.4 (b)).

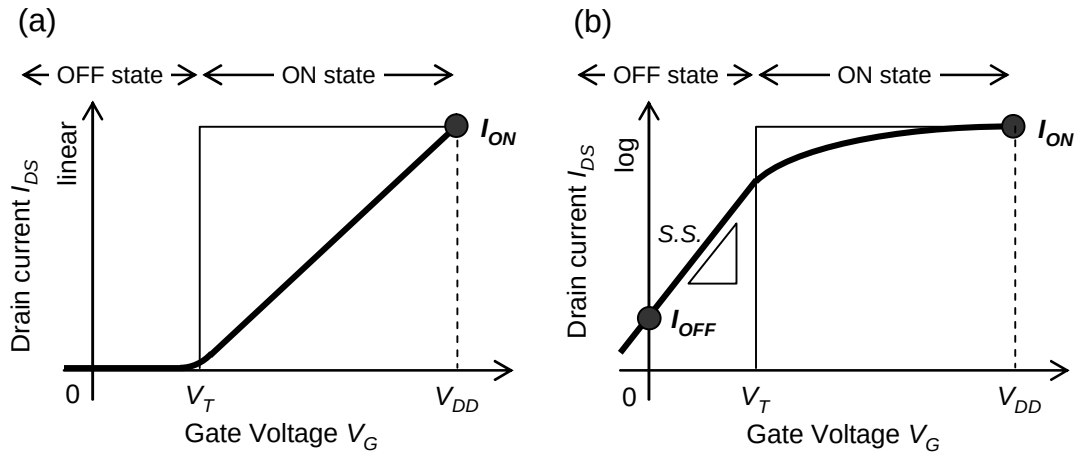


Figure 1.4 Typical I_{DS}–V_G characteristics of an nMOSFET at high drain voltages. The same current is plotted on both linear scale (a) and logarithmic scale (b).

In general, the current density in the semiconductor can be expressed by the sum of the drift current density and the diffusion current. In week inversion and depletion, however, electron charge is small and thus, the drift current is low. The drain current is dominated by diffusion. Therefore, the inversion charge density (N_{INV}) does not drop to zero abruptly. One can write the subthreshold current (I_{subth}) a function of V_G as

$$I_{subth} = \mu_{eff} C_{ox} \frac{W}{L_G} (m-1) \left(\frac{kT}{q} \right)^2 e^{q(V_G - V_T)/mkT} (1 - e^{-qV_{DS}/kT}) \quad \text{(subthreshold region)}, \quad (1.4)$$

where q is electronic charge, k is the Boltzmann constant, and T is temperature. The subthreshold current is independent of the drain voltage once V_{DS} is larger than a few kT/q, as would be expected for diffusion-dominated current transport. The dependence on gate voltage, on the other hand, is exponential. The subthreshold behaviors are primarily determined by fabrication technology considerations. The parameter to quantify how sharply the MOSFET is turned off by the gate voltage is called the subthreshold swing (SS), defined as the gate-voltage change needed to induce a drain-current change of one order of magnitude. The SS thus can be written as

$$SS = \left(\frac{d \log_{10} I_{DS}}{dV_G} \right) = \ln(10) \frac{kT}{q} \left(1 + \frac{C_{dm}}{C_{ox}} \right), \quad (1.5)$$

where C_{dm} is the maximum depletion-layer capacitance per area.

■ CMOS Inverter

The most basic component of digital static CMOS circuits is a CMOS inverter, which is composed by an nMOSFET and a pMOSFET as shown in Figure 1.5 (a). The source terminal of the nMOSFET is grounded, while that of the pMOSFET is connected to the V_{DD}. The gates of the two MOSFETs are connected as the input node, and the drains are connected as the output node. Here, C_L in Figure 1.5 is a lumped load capacitance of the output node (including the output capacitance C_{out} of the switching inverter, the input capacitance C_{in} of the next stage or stages it drives, and the wiring capacitance C_{wire}). The current through the pMOSFET ($I_P > 0$) flows from V_{DD} into the output node and tends to charge up the node voltage toward V_{DD} (i.e., pull-up), while the current through the nMOSFET ($I_N > 0$) flows out of the output node into the ground and

tends to discharge the node voltage to zero (i.e., pull-down) as shown in Figure 1.5 (b) and (c). In such a system, the complementary nature of n- and pMOSFETs allows one and only one MOSFET to conducting in one of the two stable states. Since only one of the MOSFETs is "on" in the steady state, there is little static current or static power dissipation. In principle, the power dissipation occurs only during switching transients when a charging or discharging current is flowing through the circuit [1.4].

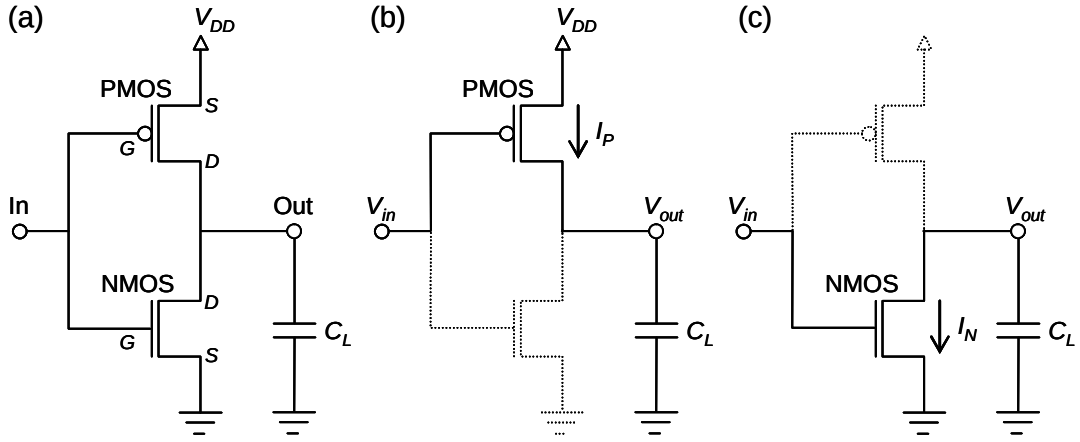


Figure 1.5 (a) Circuit diagram of CMOS inverter. (b) Charge and (c) discharge equivalent circuits.

1.1.2 MOSFET Scaling

The CMOS performance is basically evaluated by three indexes: its *switching speed*, *power consumption*, and *integration density*. These indexes have been successfully developed in the past 40 years. This progress has been achieved on the basis of MOSFET scaling rule. As the channel length decreases, the depletion width (W_d) of the source and drain becomes comparable to the channel length. This will cause short-channel effects (as described in detail in Section 1.2). Because the short-channel effects complicate device operation and degrade device performance, these effects should be eliminated or minimized so that a physical short-channel device can preserve the electrical long-channel behavior. The most-ideal scaling rule to avoid the short-channel effects is simply to reduce all device dimensions (both horizontal and vertical) and voltages and increase channel doping level by the same factor ($\kappa > 1$) as

shown in Figure 1.6, so that the electric field remains unchanged and the depletion width of the source and drain is reduced. This scaling is called "constant-field scaling" proposed by R. Dennard in 1974 [1.5]. The scaling rule is summarized in Table 1.1. In the scaled MOSFETs, the current equations (1.2, 1.3) can be rewritten as

$$I_{DS}^{'} = \mu_{eff} \frac{\varepsilon_0 \varepsilon_{ox}}{(t_{ox}/\kappa)} \frac{W/\kappa}{L_G/\kappa} \left(\frac{V_G - V_T - mV_{DS}/2}{\kappa} \right) (V_{DS}/\kappa) = \frac{I_{DS}}{\kappa} \quad \text{(linear region)} \quad (1.6)$$

$$I_{DS}^{'} = \mu_{eff} \frac{\varepsilon_0 \varepsilon_{ox}}{(t_{ox}/\kappa)} \frac{W/\kappa}{L_G/\kappa} \left(\frac{V_G - V_T}{\kappa} \right)^2 / 2m = \frac{I_{DS}}{\kappa}. \quad \text{(saturation region)} \quad (1.7)$$

Since the current is reduced by the factor of κ, the channel current per unit of channel width is unchanged by scaling. This is consistent with the same sheet density of carriers moving at the same velocity. As a result, the switching speed, the power consumption, and the integration density are improved by factor of κ, $1/\kappa^2$, and κ^2, respectively.

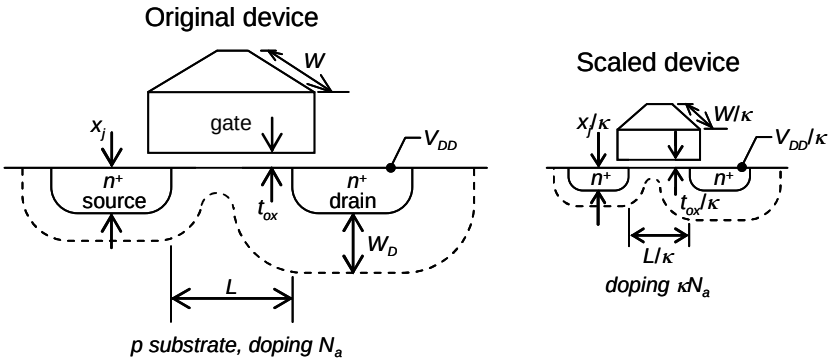


Figure 1.6 Principles of MOSFET constant-electric-field scaling.

In reality, because of the unwillingness to depart from the standardized voltage levels of the previous generation, the power-supply voltage was seldom scaled in proportion to channel length. Even though constant-field scaling provides a basic guideline to the design of scaled MOSFETs, the requirement of reducing the voltage by the same factor as the device physical dimension is too hard. Moreover, there are several factors that scale neither with the physical dimensions nor with the operating voltage. The primary reason for the nonscaling effects that neither the thermal voltage ($V_{thermal} = kT/q$) nor the silicon band gap (E_g) changes with scaling. Because of the exponential dependence of subthreshold current (see the equation (1.4)), the threshold

voltage cannot be scaled down significantly without causing a substantial increase in the off-current. In fact, even if the threshold voltage is held unchanged, the off-current per device still increases by a factor κ (from the C_{ox} factor) when the physical dimensions are scaled down by κ. Note that the *SS* factor remains essentially the same since *SS* is proportional to $1+C_{dm}/C_{ox}$ (see the equation (1.5)) and both capacitances are scaled up by the same factor κ. In addition, the process difficulties for aggressively scaled MOSFETs also limit the scaling. With the practical limitations, other scaling rules have been proposed, including constant-voltage scaling and generalized scaling [1.4]. This allows the various device parameters to be adjusted independently as long as the overall behavior is preserved. These nonideal factors, which hinder constant-field scaling, result in some form of a penalty as shown in Table 1.1, especially power consumption. Therefore, it is important to understand how factors affect the performance of a CMOS LSI chip in order to obtain an optimized device structure.

Table 1.1 Constant-field scaling and generalized scaling of MOSFET device and circuit parameters.

	Device and Circuit Parameters	Multiplicative Factor (κ>1)	
		Constant-field Scaling	Generalized Scaling
Scaling assumptions	Device dimensions (t_{ox}, L, W, x_j)	$1/\kappa$	$1/\kappa$
	Doping concentration (N_a, N_d)	κ	$\alpha\kappa$
	Voltage (V)	$1/\kappa$	α/κ
			Long Ch. Vel. Sat.
Derived scaling behavior of device parameters	Electric field (E)	1	α
	Carrier velocity (v)	1	α 1
	Depletion-layer width (W_d)	$1/\kappa$	$1/\kappa$
	Capacitance ($C=\varepsilon A/t$)	$1/\kappa$	$1/\kappa$
	Inversion/layer charge density (Q_t)	1	α
	Current, drift (I)	$1/\kappa$	α^2/κ α/κ
Derived scaling behavior of circuit parameters	Circuit delay time (τ~CV/I)	$1/\kappa$	$1/\alpha\kappa$ $1/\kappa$
	Power dissipation per circuit (P~VI)	$1/\kappa^2$	α^3/κ^2 α^2/κ^2
	Power/delay product per circuit ($P\tau$)	$1/\kappa^3$	α^2/κ^3
	Circuit density ($\propto I/A$)	κ^2	κ^2
	Power density (P/A)	1	α^3

1.1.3 CMOS Performance Indexes

The need for devices that consume a minimum amount of power was a major driving force behind the development of CMOS technologies. As a result, CMOS devices are best known for low power consumption due to the unique characteristic of zero standby power. This enables higher integration levels and makes them the technology of choice for most LSI applications. However, continuing aggressive scaling of MOSFETs and the resulting circuit density growth lead inevitably to a further increase of power consumption of CMOS integrated circuits recently. Therefore, for optimizing the performance, simply knowing that CMOS devices may use less power than equivalent devices from other technologies does not help much. It is necessary to understand the relationship between switching characteristics, integration density, and power dissipation in terms of MOSFET dimensions.

■ **Switching Characteristics**

Realistic benchmarking of the switching delay for CMOS circuits is essential to quantify technology requirements in order to continue its historical scaling trend. In the research stage, it is important to establish the connection of the device-level targets with circuit-level performance. In 1995, M. Bohr has proposed a MOSFET speed metric as an approximated gate delay [1.6]. This metric assumes a transistor circuit as shown in Figure 1.7. This circuit can be easily imagined from Figure 1.5 (b) or (c). Transistor A and B are identical. The gate delay is defined as the time it takes node C to make a voltage swing equal to V_{DD}. The only capacitance on node C is the gate capacitance C_G of transistor B, which is defined as gate area times gate oxide capacitance per unit area (C_{ox}). It is important to note that physical gate length L_G as opposed to electrical channel length L_{eff} is used to calculate C_G because L_G will include both channel and overlap capacitance. The on-state drive current I_{ON} of transistor A charges or discharges this capacitor. The gate delay can thus be approximated by the equation:

$$\tau \approx \frac{C_G \times V_{DD}}{I_{ON}}. \tag{1.8}$$

Historically, this expression (1.8) has been widely used as a benchmark of the

MOSFET speed [1.7]. This is because one can easily quantify the relative performance without fabricating a circuit. The only information needed from a MOSFET to estimate its *CV*/*I* gate delay is V_{DD}, L_G, I_{ON}, and C_{ox}. The equation (1.8) can be rewritten by using the equation (1.3):

$$\tau \approx \frac{(C_{ox}WL_G)\times V_{DD}}{\mu_{eff}C_{ox}\frac{W}{L_{eff}}\frac{(V_{DD}-V_T)^{\alpha}}{2m}} = (L_{eff}+2L_{ov})L_{eff}\frac{2m}{\mu_{eff}}\frac{V_{DD}}{(V_{DD}-V_T)^{\alpha}}, \qquad (1.9)$$

where L_{ov} is the gate overlap length and the power of α ($1 < \alpha < 2$) indicates the degree of velocity saturation in short-channel MOSFETs. It is clear that the gate delay is strongly depends on the gate length. This is a reason that I_{ON} enhancement as a function of the gate length is often used as a metric of a MOSFET speed. Note that this gate delay is independent on the gate width because both C_G and I_{ON} are linearly proportional to the gate width.

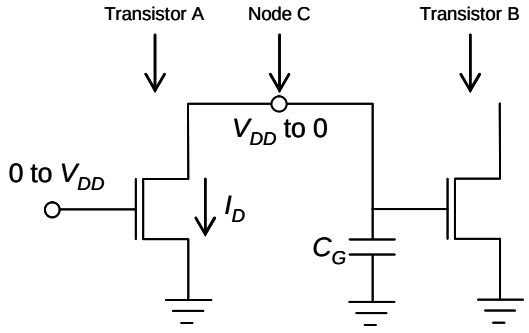


Figure 1.7 CV/I performance metric [1.6].

However, in actual CMOS circuits, the load capacitance is not so simple. The difference of geometrical dependence among each capacitance sources leads to the gate delay estimation error. Here we discuss an effect of the gate width independent components. When the gate capacitance in the equation (1.8) is replaced by a lumped load capacitance C_L, the equivalent circuit can be modified as shown in Figure 1.8. The load capacitance can separate three major components.

1. The output capacitance C_{out} that results from the drain diffusion of the driving transistor A.
2. The input capacitance C_{in} that is the gate capacitance of the transistor B being

driven by the transistor A, including the intrinsic and the overlap components.

3. The wiring capacitance C_{wire} that results from interconnects to the gates being driven.

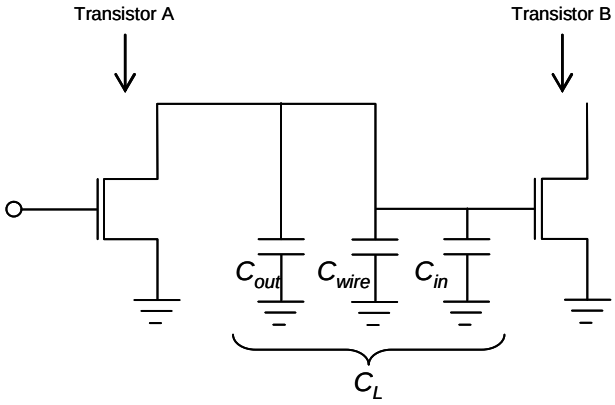


Figure 1.8 Equivalent circuit with wiring capacitance.

In general, the output of a transistor may drive more than one stage. In that case, the fan-out *FO* is 2, 3, …, which means that each inverter in the chain is driving 2, 3, … stages in parallel. The lumped load capacitance C_L thus can be written by

$$C_L = C_{out} + C_{wire} + C_{in} \times FO \,. \tag{1.9}$$

Here C_{out} and C_{in} are linearly proportional to the gate width, while C_{wire} is independent on that. The gate delay of the equation (1.8) then can be rewritten by the equation (1.9):

$$\tau \approx \frac{C_L \times V_{DD}}{I_{ON}} = \frac{(C_{out} + C_{wire} + C_{in} \times FO) \times V_{DD}}{I_{ON}}$$

$$= \frac{(C_{out} + C_{in} \times FO) \times V_{DD}}{I_{ON}} + \frac{C_{wire} \times V_{DD}}{I_{ON}}$$

$$= \frac{C_L(W) \times V_{DD}}{I_{ON}} + \frac{C_{wire} \times V_{DD}}{I_{ON}}, \tag{1.10}$$

where $C_L(W)$ is the gate width dependent capacitance, which consists of the p-n junction capacitance and the gate capacitance including the intrinsic and overlap components. Since both $C_L(W)$ and I_{ON} are linearly proportional to the gate width, the gate width effect is canceled out in the first term. On the other hand, the second term has the gate width dependence. As the gate width increases, the delay decreases. However, the

reduction in the delay must be traded off against the increased area (and power) when the width is increased.

Moreover, accuracy of the *CV*/*I* metric in approximating delay time is degraded as device scaling progresses [1.8–1.11]. The problem of the *CV*/*I* metric is that the MOSFETs in a real CMOS logic gate chain usually do not operate at the bias point (V_{GS} = V_{DS} = V_{DD}) which gives I_{ON} because an inverter is driven by output from a previous stage whose waveform has a finite rise or fall time associated with it as shown in Figure 1.9. The peak current is typically 80–90% of maximum on-current I_{ON} at V_G = V_{DS} = V_{DD}. The exact percentage depends on the detailed device parameters such as mobility, velocity saturation, short-channel effects and series resistance. In particular, the contribution of drain–induced barrier lowering (DIBL) due to the short-channel effects has been pointed out as shown in Figure 1.9 (b) [1.8–1.11]. The difference of I_{DS}–V_{DS} curves between with and without a degraded DIBL results in different bias-point trajectories during switching. Though I_{ON} is kept at the same level, the inverter chain built by devices with the degraded DIBL switches at a lower speed. To evaluate of the delay of a short-channel MOSFET, it is important to consider both the current level and the DIBL.

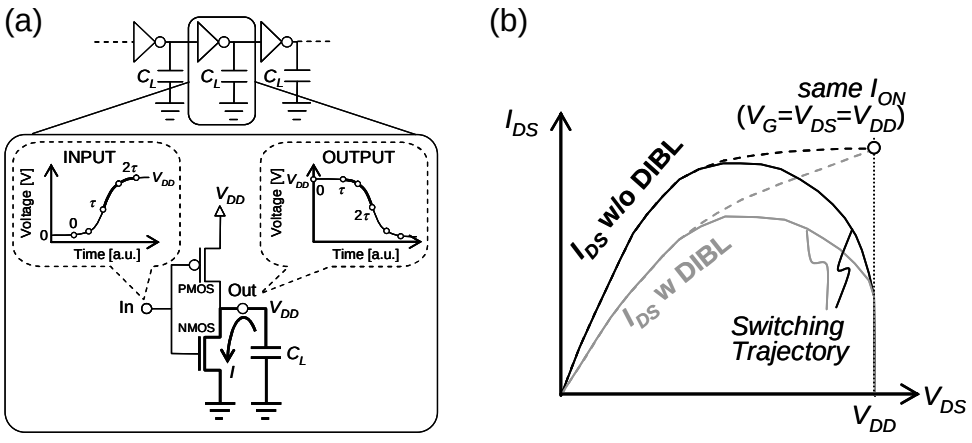


Figure 1.9 (a) Operation of a CMOS inverter in inverter chain. (b) I_{DS}-V_{DS} curve and trajectories with and without a degraded DIBL for nMOSFET [1.11].

■ **Integration density**

Integration density means that what number of MOSFETs is fabricated on a chip. As shown in Figure 1.1 (a), the number has been dramatically increased in the last forty years owing to MOSFET area scaling. Figure 1.10 shows a sample layout of a CMOS inverter based on lambda-based design rules [1. 15]. The lambda unit is fixed to half of the minimum available lithography of the technology, typically the minimum gate length (L_G=2λ). The layout clearly shows that the source/drain active area occupies a large area in a MOSFET, which is determined by the source/drain length and the gate width. The former is limited by the feature size of the process technology used (e.g. contact size, space between contact and gate poly). The latter depends on the carrier mobility ratio between n- and pMOSFET and the current level needed. Regarding the integration density, the key consideration is how to obtain high current level in smaller gate width as the feature size is fixed.

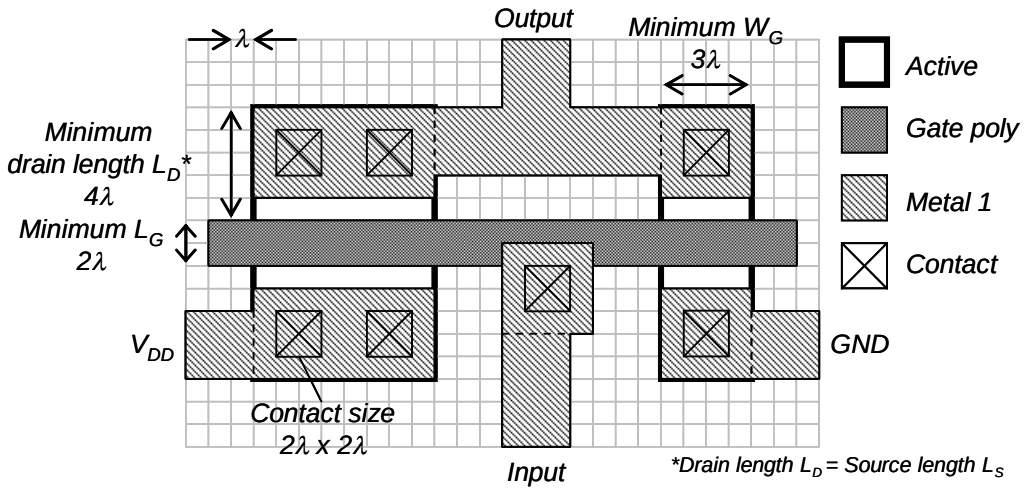


Figure 1.10 Layout of a CMOS inverter based on lambda-based design rules [1. 15]

■ **Power Consumption**

There are three major sources of power consumption in digital CMOS circuits which are summarized in the following equation [1.12]:

$$\begin{aligned} P_{consum} &= P_{switching} + P_{short-circuit} + P_{leakage} \\ &= \alpha f_{clk} C_L V_{DD}^2 + I_{sc} V_{DD} + I_{leak} V_{DD} \end{aligned} \tag{1.11}$$

The first term represents the switching component of dynamic power, where C_L is the load capacitance, f_{clk} is the clock frequency, and α is the node transition activity factor (the average number of times the node makes a power consuming transition in one clock period). The second term is due to the direct-path short circuit current, I_{sc}, which arises when both the n- and pMOSFET are simultaneously active, conducting current directly from supply to ground. Finally, the leakage component of static power is negligible in principle. This is one of the primary advantages of CMOS system. CMOS circuits do not dissipate power if they are not switching.

However, the continuing MOSFET scaling and the resulting circuit density growth has leaded to unacceptable level of static power consumption recently. Figure 1.11 shows the power consumption for CMOS logic circuits. As MOSFETs get smaller with each new process technologies, their channel lengths become shorter and the static power consumption increases over the dynamic power consumption. Nowadays, reduction of the static power consumption has become a major challenge for deep submicron CMOS. Total static power consumption can be obtained as

$$P_{static}(=P_{leakage}) = n\sum I_{leak} \times V_{DD} = n(I_{subth} + I_{gate} + I_{GIDL}) \times V_{DD}, \qquad (1.12)$$

where n is the number of transistors in the off-state, I_{gate} is the gate tunneling current, and I_{GIDL} is the gate-induced drain leakage (GIDL) current. The leakage currents can be divided in to three main groups: I_{subth}, I_{gate}, and I_{GIDL}. The increase of I_{subth} is due to the nonscaling effects and short-channel effects (see in Section 1.2). The increase of I_{gate} is due to the aggressively scaled t_{ox} which causes exponential increase of the tunneling current when the t_{ox} is scaled down to less than 2 nm as shown in Figure 1.11. This gate leakage can be reduced by replacing gate silicon dioxide (SiO_2) to high-κ material. Practically, high-κ technology has been introduced to the 45 nm process technology microprocessor since 2007 by Intel Corporation [1.13]. The I_{GIDL} is caused by band-to-band tunneling in the gate-to-drain overlap region when a large gate-to-drain voltage is biased and the band bending then is larger than the silicon band gap. Lightly doped drain (LDD) structure has been studied as a solution for the GIDL [1.14, 1.15]. Figure 1.12 shows the summary of leakage current components. In actual fact, the

off-state leakages become more and more problematic in a scaled MOSFET (as explained in detail in Section 1.2).

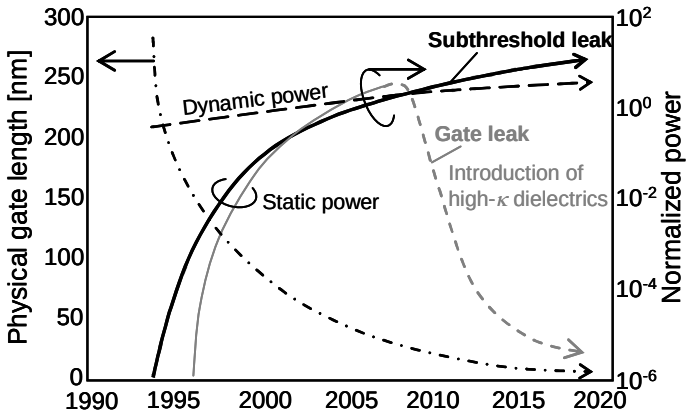


Figure 1.11 Sources of leakage current increase as the technology causes gate lengths to shrink. Data from ITRS [1.6]

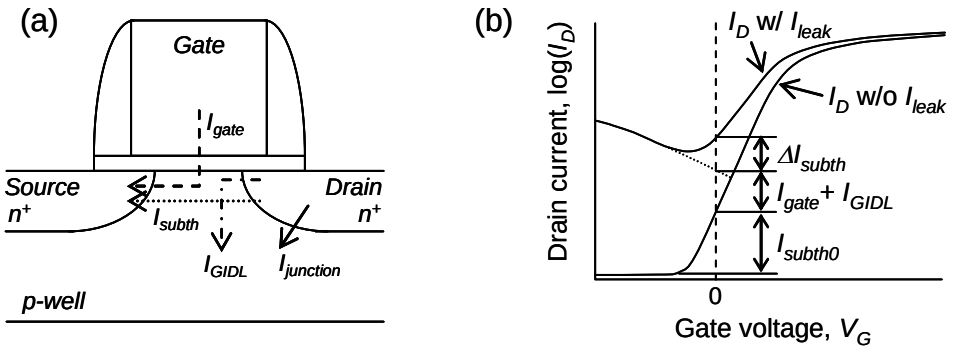


Figure 1.12 (a) Leakage current components in a nMOSFET, (b) I_D–V_G curves with and without leakage currents. I_{subth0} is the initial subthreshold current, while ΔI_{subth} is the added current due to the short-channel effects.

In this section, we discussed the basic MOSFET operation, the scaling method, and the CMOS performance indexes. CMOS LSI has been developed owing to the diminishing size of MOSFET. Especially, the impact on the gate length and width to CMOS performances was discussed in the latter half. To improve the whole performance, we have to consider the switching characteristics, the integration density,

and the power dissipation at a same time. The diminishing gate length contributes the gate delay reduction, while causing the power dissipation increase due to the short-channel effects. On the other hand, the diminishing gate width can reduce device area leading to higher integration density, while causing the gate delay increase if the gate width independent capacitances are not negligible. To make the most of the geometrical effects, the key issues will be summarized to the following two points:

1. How can we suppress the short-channel effects?
2. How can we increase the effective gate width in a given layout area?

In the next section, we will discuss the short-channel effects in detail.

1.2 SRORT-CHANNEL MOSFET

Short-channel MOSFETs differ in many important aspects from long-channel devices. This section covers the features of short-channel devices that especially are important for current MOSFETs.

1.2.1 Short-Channel Effects

The key difference between a short-channel and a long-channel MOSFET is that the field pattern in the depletion region of a short-channel MOSFET is two-dimensional. In other words, the gradual–channel approximation breaks down for short-channel devices. The two-dimensional field pattern arises from the proximity of source and drain regions. The source–drain distance is comparable to the MOS depletion width in the vertical direction, and the source–drain potential has a strong effect on the band bending over a significant portion of the device. This phenomenon will cause the device parameter changes which can be summarized as follows: (1) V_T roll-off, (2) subthreshold swing degradation, and (3) DIBL degradation.

- **Threshold voltage roll-off**

One way to describe it is to consider the net charge (ionized accepters or donors) in the depletion region of the device. The field lines terminating on these fixed charges

originate either from the gate or from the source and drain. This is referred to as the charge-sharing model as shown in Figure 1.13 [1.4]. At a low drain voltage, only the field lines terminating on the depletion charges within the trapezoidal region are assumed to originate from the gate. The rest of the field lines originate either from the source or from the drain. Total charge with in the trapezoidal region, $Q'_B \propto W_{dm} \times (L + L')/2$, is proportionally less than the total gate depletion charge, $Q_B \propto W_{dm} \times L$, in the long-channel case. As a result, it takes a lower gate voltage to reach the threshold condition of a short-channel device,

$$V_T = V_{fb} + 2\psi_B + \frac{Q'_B}{WLC_{ox}}, \tag{1.13}$$

Where W_{dm} is the maximum depletion-layer depth, V_{fb} is the flat-band voltage, and ψ_B is the difference between Fermi level and intrinsic level. Note that the horizontal depletion-layer widths y_S and y_D, are smaller than the vertical depletion-layer widths W_S, and W_D, respectively, because the transverse field strongly influences the potential distribution at the surface. Even though a simple, analytical expression for the threshold voltage can be obtained from the charge-sharing model, the division of depletion charge between the gate and the source and drain is somewhere arbitrary.

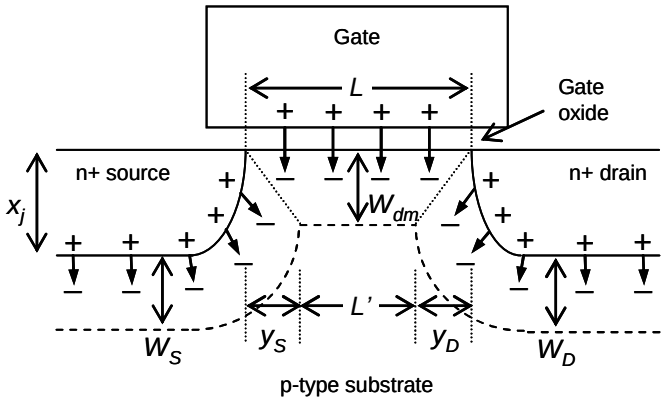


Figure 1.13 Schematic diagram of the chare-sharing model. The dashed lines indicate the boundary of the gate and source–drain depletion regions. The arrows represent electric field lines that originate from a positive charge and terminate on a negative charge. The dotted lines partition the depletion charge and form the two sides of the trapezoid.

The voltage-doing transformation (VDT) has been proposed to replace the influence of the lateral drain–source field by an equivalent reduction in channel doping concentration [1.17, 1.18]. On the basis of this model, the threshold voltage roll-off can be written as follows:

$$SCE \equiv \mid \Delta V_T \mid = 0.64 \frac{\varepsilon_s}{\varepsilon_{ox}} EI \times \Phi_D \tag{1.14}$$

where

$$EI = \left(1 + \frac{x_j^2}{L_{eff}^2} \right) \frac{t_{ox}}{L_{eff}} \frac{W_{dm}}{L_{eff}} \tag{1.15}$$

is the Electrostatic Integrity for planar bulk MOSFET, ε_s is the semiconductor permittivity, and ε_{ox} is the oxide permittivity. The VDT was proven to be successful in describing the V_T roll-off for all CMOS technologies from CMOS 1.2 μm down to CMOS 65 nm based on a comparison to numerical simulations and experimental data.

- **Subthreshold swing degradation**

The physics of the short-channel effect can be understand from a different angle by considering the potential barrier at the surface between the source and drain, as shown in Figure 1.14. Under off conditions, this potential barrier prevents electron current from flowing to the drain. The surface potential is mainly controlled by the gate voltage. When the gate voltage is below the threshold voltage, there are only a limited number of electrons injected from the source over the barrier and collected by the drain (subthreshold current). In the long-channel case, the potential barrier is flat over most part of the channel. Source and drain fields only affect the very ends of the channel. As the channel length is shortened, however, the source and drain fields penetrate deeply into the middle of the channel, which lowers the potential barrier between the source and drain as shown Figure 1.14 (b). The region of maximum potential barrier shrinks to a single point near the center of the channel. This causes a substantial increase of the subthreshold current.

Here the equation 1.5 for the subthreshold swing for long-channel devices, which comes from the direct derivation of the drain current expression, can be rewritten as a

function of the surface potential ψ_S,

$$S.S. = \frac{kT}{q}\ln(10)\cdot\left(\frac{\partial V_G}{\partial \psi_s}\right) = \frac{kT}{q}\ln(10)\cdot\left(1+\frac{1}{C_{ox}}\frac{\partial Q_B}{\partial \psi_s}\right), \tag{1.16}$$

It is clear that the potential barrier lowering in short-channel device will have an impact on subthreshold swing. On the basis of the VDT model, the subthreshold swing behavior with the gate length can be rewritten as follows [1.18]:

$$S.S. = \frac{kT}{q}\ln(10)\cdot\left[1+\frac{1}{C_{ox}}\frac{\partial Q_B}{\partial \psi_s}+\frac{\varepsilon_s}{\varepsilon_{ox}}\frac{t_{ox}}{L_{eff}}\frac{x_j}{L_{eff}}\left(1+\frac{3}{4}\frac{W_{dm}}{L_{eff}}\right)\sqrt{1+2\frac{V_{DS}}{\Phi_D}}\right]. \tag{1.17}$$

Figure 1.15 shows the subthreshold swing calculated by using MASTAR MOSFET modeling software based on the VDT model [1.19]. It is clear that the subthreshold swing increases as decreasing the gate length.

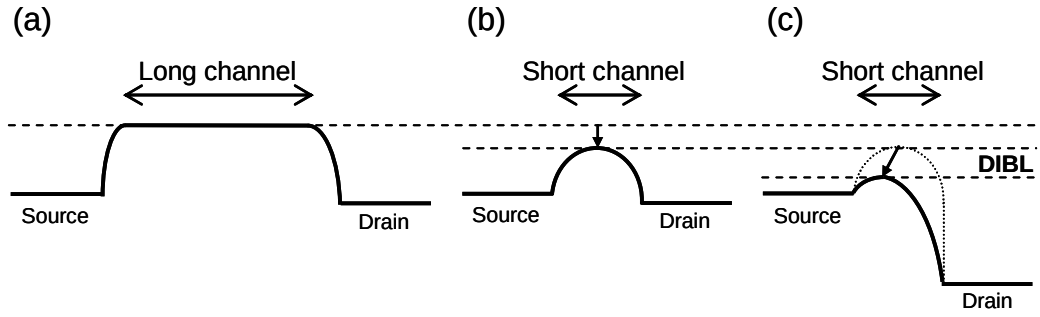


Figure 1.14 Surface potential lowering due to the short-channel effects: (a) a long-channel MOSFET, (b) a short-channel MOSFET at low drain bias, (c) a short-channel MOSFET at high drain bias.

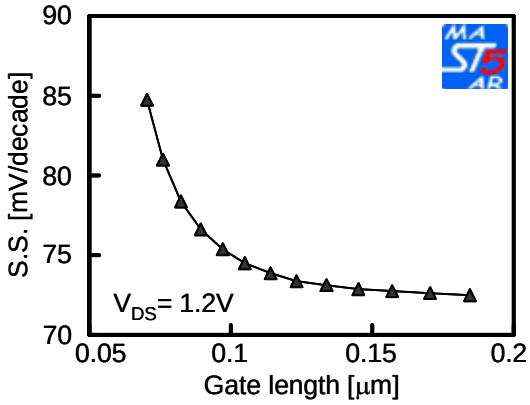


Figure 1.15 Subthreshold swing calculated by using MASTAR MOSFET modeling software [1.19]: $N_a=10^{18}cm^{-3}$, x_j=30 nm, t_{ox}= 1.3 nm, V_{DD}= 1.2 V, W=1μm.

■ **Drain-induced barrier lowering**

When a high drain voltage is applied to a short-channel device, the barrier height is lowered even more, and the point of maximum barrier also shifts toward the source end as shown in Figure 1.14 (c). The lowering of the source barrier causes an injection of extra carriers, thereby increasing the current substantially. As a result, the threshold voltage decreases. This effect is referred to as drain-induced barrier lowering (DIBL). On the basis of the VDT model, the DIBL behavior can be rewritten as follows [1.18]:

$$DIBL = 0.80 \frac{\varepsilon_s}{\varepsilon_{ox}} EI \times V_{DS} . \tag{1.18}$$

Figure 1.15 shows the DIBL calculated by using MASTAR MOSFET modeling software based on the VDT model [1.19]. The DIBL increases in short-channel MOSFET as well as the subthreshold swing.

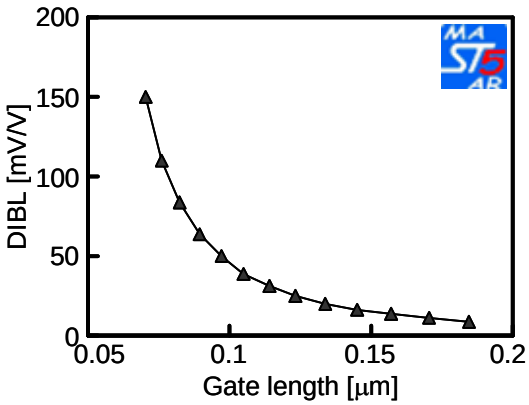


Figure 1.16 DIBL calculated by using MASTAR MOSFET modeling software [1.19]: $N_a=10^{18}cm^{-3}$, $x_j=30$ nm, $t_{ox}=1.3$ nm, $V_{DD}=1.2$ V, $W=1\mu m$.

Consequently, the threshold voltage for a short-channel device can be rewritten as follows:

$$V_T = V_{Tlong} - SCE - DIBL , \tag{1.19}$$

where V_{Tlong} is the threshold voltage in a long-channel device. Figure 1.17 shows the threshold roll-off calculated by using MASTAR MOSFET modeling software based on the VDT model [1.19].

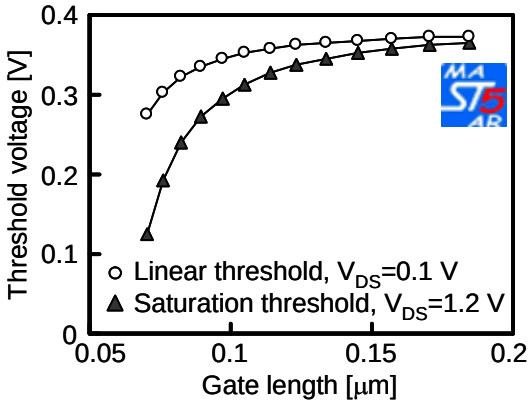


Figure 1.17 V_T roll-off calculated by using MASTAR MOSFET modeling software [1.19]: $N_a=10^{18}cm^{-3}$, $x_j=30$ nm, $t_{ox}=1.3$ nm, $V_{DD}=1.2$ V, $W=1\mu m$.

Here, we should reconsider the static power consumption in short-channel CMOS logic gates. The increase of off-state current is the sum of these short-channel effects as shown in Figure 1.18. The subthreshold current for a short-channel device can be rewritten as follows [1.18]:

$$Log(I_{subth}) = Log(I_{Tr}) - (V_{Tlong} - SCE - DIBL)/SS\,, \qquad (1.20)$$

where I_{Tr} is the transistor current corresponding to the threshold voltage. With gate length and threshold voltage scaled down the subthreshold leakage current increases exponentially. In order to suppress the increase, it is clear that we have to avoid increasing the three factors: SCE, DIBL, and *SS*. Now, we can easily understand which physical parameters are important from the equations 1.14, 1.15, 1.17, and 1.18; t_{ox}/L_{eff}, x_j/L_{eff}, and W_{dm}/L_{eff}. Thinning of gate oxide was one of the easiest solutions technologically. Until now, aggressive scaling of the gate oxide (SiO_2) thickness thus has continued. However, this resulted in rapid increase of the tunneling current when the gate oxide thickness is scaled down to less than 2 nm as shown in Figure 1.11. This gate leakage can be reduced by replacing gate silicon dioxide (SiO_2) to high-κ material. To form shallow junctions, the use of Silicon-On-Insulator (SOI) wafer is one of the most effective solutions. The source/drain junction depth in SOI thin body device is terminated by the buried oxide. Moreover, the thin body thickness can limit the maximum depletion-layer depth. These devices will be discussed in section 1.4.

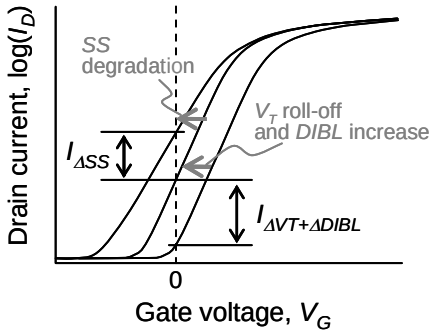


Figure 1.18 Off-current increase due to the short-channel effects.

1.2.2 Source/Drain Series Resistance in short-channel MOSFET

In the discussion of MOSFET current thus far, it was assumed that the source and drain regions were perfectly conducting. In reality, as the current flows from the channel to the terminal contact, there is a voltage drop in the source and drain regions due to the finite silicon receptivity and metal contact resistance as shown in Figure 1.19.

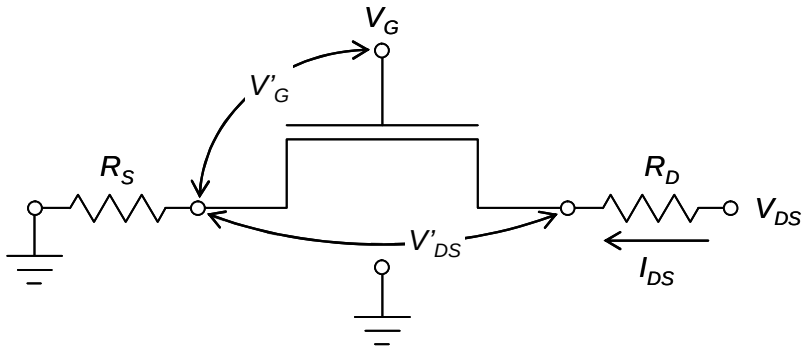


Figure 1.19 Equivalent circuit of MOSFET with source and drain resistance [1.4].

In a long-channel device, the source–drain parasitic resistance is negligible compared with the channel resistance. In a short channel device, however, the source–drain series resistance can be an appreciable fraction of the channel resistance and can therefore cause significant current degradation. The most severe current degradation by series resistance occurs in the linear region (Low V_{DS}) when the gate voltage is high because the MOSFET channel resistance is the lowest under such bias conditions. The MOSFET current in the saturation region is least affected by the resistance degradation of

source–drain voltage, I_{DS} is essentially independent of V_{DS} in saturation. The saturation current is only affected through gate–source voltage degradation by voltage drop between the source contact and the source end of the channel [1.4]:

$$V'_G = V_G - R_S I_{DS} . \tag{1.21}$$

In aggressively scaled MOSFETs, however, the saturation current degradation becomes more and more problematic. Figures 1.20 and 1.21 show the on-current degradations due to the source/drain series resistance. This degradation implies that further improvements in I_{ON} by shortening the gate length cannot be expected when the channel resistance becomes comparable to the source and drain resistance. Nowadays, the source/drain series resistance is a major concern for the MOSFET scaling. Shallow junctions in the source/drain regions are needed to minimize the short-channel effects as discussed in subsection 1.2.1. ITRS predicts that source/drain extension junction depth below sub-50 nm CMOS will be scaled further down around 10 nm to maintain acceptable short channel performance, but this may lead to a high series resistance problem [1.7, 1.20]. In other words, further down-scaling without improvement of I_{ON} is meaningless even if short-channel effects are completely suppressed by introducing shallow junction technologies (e.g. source/drain extension, SOI wafer, etc.). It is important to design MOSFET structure effective in suppression of both short-channel effects and source/drain resistance.

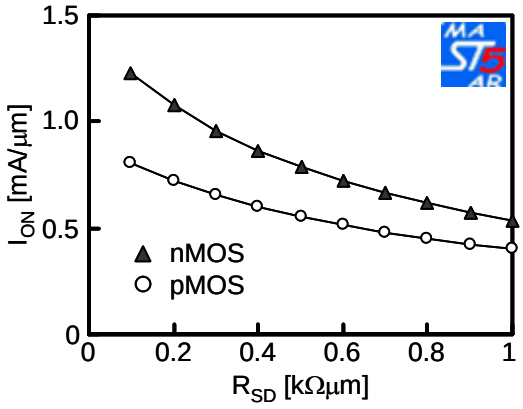


Figure 1.20 On-currents as a function of source/drain series resistance. All plots are calculated by using MASTAR MOSFET modeling software [1.19]: $N_a=10^{18}cm^{-3}$, t_{ox}= 1.3 nm, V_{DD}= 1.2 V, L_G=50 nm, W=1μm.

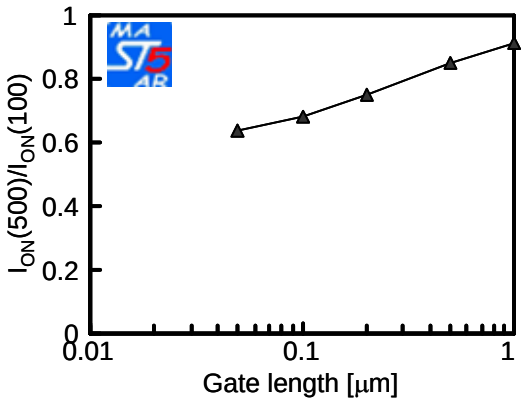


Figure 1.21 On-current lowering ratio of R_{SD}= 500 Ω.μm to 100 Ω.μm as a function of gate length. All plots are calculated by using MASTAR MOSFET modeling software [1.19]: $N_a=10^{18}cm^{-3}$, t_{ox}= 1.3 nm, V_{DD}= 1.2 V, W=1μm.

1.2.3 Carrier Transport Mechanisms in short-channel MOSFET

In short channel devices, the short-channel effects become problematic. On the other hand, carrier transport also has gate length dependence. Precise understanding of correlation between low-lateral-field mobility and high-lateral field velocity, which is more directly related to on-current, is important. In analyzing the carrier transport mechanisms depending on the gate length, we go back to the general equation of the saturation current. We start with the generalized form

$$I_{ON} \approx WqN_{INV}^{source} \cdot \upsilon_s \tag{1.22}$$

where q is the elemental charge, N_{INV}^{source} is the inversion charge density near the source edge, and υ_s is the average carrier velocity near the source edge [1.21]. While qN_{INV} is simply determined by the maximum value at the source as $C_{ox}(V_G–V_T)$, the determination mechanism of υ_s is dependent on the gate length. So the only critical parameter is υ_s. Figure 1.22 shows the schematic diagrams of carrier transport models to determine I_{ON}. In long-channel MOSFETs, the carrier mobility μ is the solely important factor in determining the velocity υ. The υ–μ relationship is given by $\upsilon = \mu * E_{lateral}$, where $E_{lateral}$ is the lateral electric field, as shown in Figure 1.23. However, the linear υ–μ relationship breaks down when the gate length becomes shorter. As the lateral

electric field is increased, the average carrier velocity and the average carrier energy increase as well. When the carrier energy increases beyond the optical phonon energy, the probability of emitting an optical phonon increases abruptly. This mechanism causes the carrier velocity to saturate with increasing electric field as shown in Figure 1.24 [1.22, 1.23]. As velocity saturation phenomenon begins to occur, μ dependence of v becomes weaker. As far as carrier scattering events in the channels sufficiently occur and the stationary transport dominate the carrier transport, the carrier transport model as shown in Figure 1.22 (a) basically holds even under the existence of velocity saturation.

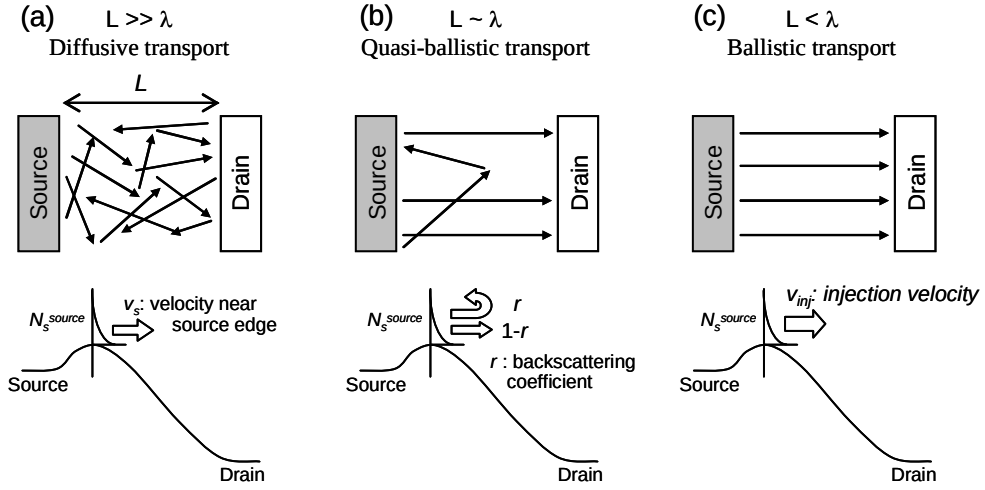


Figure 1.22 Schematic diagrams of carrier transport models to determine I_{ON}. (a) Conventional transport model. (b) Quasi-ballistic transport model. (c) Full-ballistic transport model [1.24, 1.25].

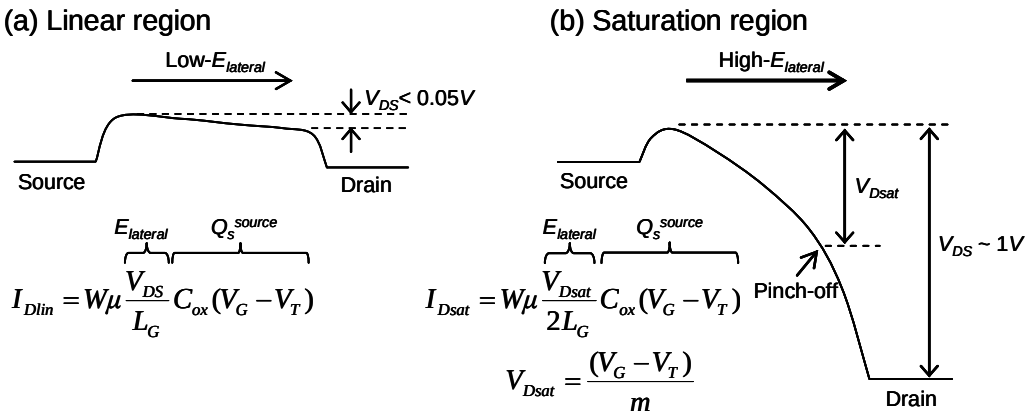


Figure 1.23 Channel potential profiles under conditions of carrier mobility μ and velocity v. (a) Linear region. (b) Saturation region.

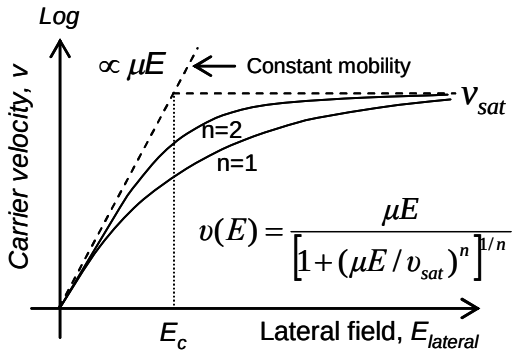


Figure 1.24 Velocity–field relationship for electrons (n=2) and holes (n=1) by the empirical form inserted [1.22]. The critical field $E_c = v_{sat}/\mu$.

As the channel length becomes shorter, nonstationary transport becomes more dominant, where sufficient numbers of scattering events do not occur inside the channels. This situation, as shown in Figure 1.22 (b), has been formulated as quasi-ballistic transport by Lundstrom *et al.* [1.21]:

$$I_{ON} = qN_{INV}^{source} \cdot \upsilon_{inj} \cdot \frac{1-r}{1+r}, \tag{1.23}$$

where υ_{inj} is the injection velocity at the top of the barrier near the source edge, and r is the backscattering rate near-source region. The fraction r of the carriers are scattered back to the source, as shown in Figure 1.22 (b). Hence the effective velocity of carriers at the barrier, called virtual source velocity, is determined by this fraction. Lundstrom's theory uses assumption that only scattering events that take place in the vicinity of the virtual source are responsible for backscattering of carriers to the source. Once a carrier passes the point where the potential has dropped by kT/q from the barrier top, the probability of return to the source is negligible. Hence, the r depends on the ration between the backscattering mean free path of carriers, λ, and the distance over which the potential drops by the thermal voltage, the so-called critical length of scattering, l:

$$r = \frac{l}{l+\lambda}. \tag{1.24}$$

The notion of mobility in short-channel devices, where the scattering mean free path is comparable to the channel length, is subject of controversy. However, a

phenomenological mobility can always be extracted at low V_{DS}. Assuming that the mobility is constant across the channel and that the carriers at the top of the barrier are in a near-equilibrium condition, Rahman related this mobility to the backscattering mean free path by matching the low V_{DS} drift-diffusion equation with the MOSFET scattering model [1.22]:

$$\lambda = \left(\frac{2\mu}{\upsilon_\theta} \frac{kT}{q} \right) \frac{\Im_0(\eta_F)}{\Im_{-1}(\eta_F)}, \tag{1.25}$$

where v_θ is the ballistic velocity of carriers which in the non-degenerate limit is equal to the thermal velocity, η_F=(E_F-ε)/kT is the reduced Fermi energy, ε is the minimum band energy, and $\Im_n$ is the Fermi integral of the nth order. This equation (1.25) means that in order for the ballistic efficiency to increase, the low-field mobility should be increased. Since r is related to μ, the enhancement of mobility can be still important in increasing I_{ON} under quasi-ballistic transport regime.

Furthermore, when channel length becomes much shorter, probably down to less than 10 nm in Si MOSFETs, and no carrier-scattering events occur inside the channel, the carrier transport is dominated by full ballistic transport, as shown in Figure 1.22(c). Here, I_{ON} in MOSFETs under this ballistic transport, which have also been formulated by Natori [1.24], is simply represented by

$$I_{ON} = qN_s^{source} \cdot \upsilon_{inj} . \tag{1.26}$$

Thus the enhancement of υ_{inj} is necessary to increase I_{ON} of ballistic MOSFETs, while the carrier mobility loses its meaning.

Recently, the v–L_G relationships have been experimentally investigated to clarify the effectiveness of μ enhancement to improve I_{ON} [1.27]. Figure 1.25 shows the experimental results of L_G dependence of velocity. It was found the velocity is not completely saturated even below L_G = 50 nm and still increases with a slope of $L_G^{-0.45}$. This can be assigned to the velocity overshoot phenomenon near the source edge due to contribution of quasi-ballistic carriers. Note that the mobility enhancement is still effective at L_G = 30 nm.

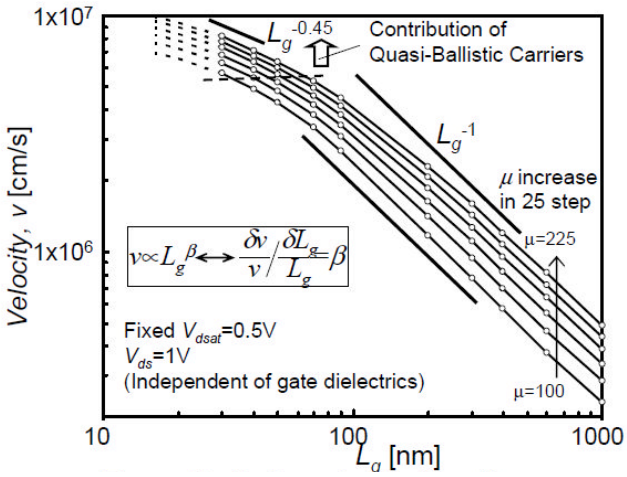


Figure 1.25 L_G dependence of velocity with μ as parameter [1.27].

In this section, we discussed several effects in short-channel MOSFETs. To suppress the off leakage current caused by the short channel effects, the minimization of t_{ox}, x_j, and W_{dm} is effective. However, the shallow junction leads to I_{ON} degradation due to the high source/drain series resistance. To avoid this trade-off, it is important to design MOSFET structure effective in suppression of both short-channel effects and source/drain resistance. Moreover, although μ dependence of v becomes weaker in short-channel MOSFETs, the μ enhancement can be still important in increasing I_{ON} under quasi-ballistic transport regime.

1.3 KEY TECHNOLOGIES TO IMPROVE MOSFET PERFORMANCE

This section introduces some solutions and challenges to continue the scaling toward purpose of this thesis. First, we consider a leading edge of CMOS technology. Figure 1.26 shows cross section of n- and pMOSFETs for 32 nm technology fabricated by Intel Corporation in 2009 [1.28]. The key device features are summarized in Table 1.2. The foundation of the 32 nm process technology is the second generation high-κ/metal gate MOSFET to suppress the short-channel effects without increasing gate leakage. The equivalent oxide thickness (EOT) of the high-κ dielectric has been

reduced from 1.0 nm on 45 nm to 0.9 nm on the 32 nm process while gate length has been reduced to 30 nm. Using a replacement metal gate flow, that is a gate-last process, enables stress enhancement techniques to be in place before removing the poly gate from the transistor. This 32 nm technology also uses 4th generation SiGe strained silicon for pMOSFET resulting in linear drive current exceeding nMOSFET. Moreover, the raised S/D regions and 2nd generation trench contacts technologies enables reduced S/D access resistance.

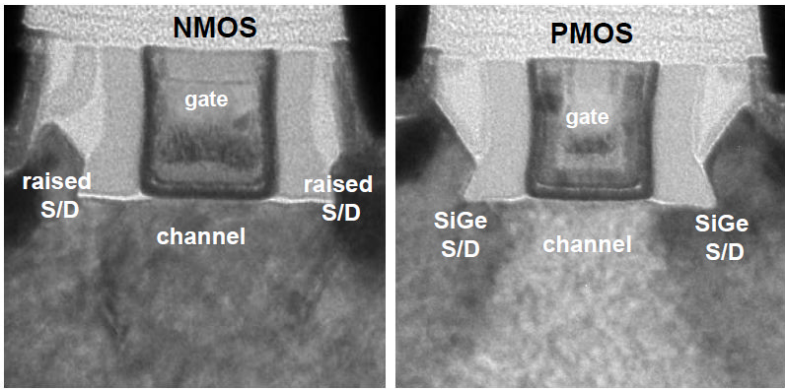


Figure 1.26 Cross section of Intel's NMOS and PMOS with 4th generation strained silicon, 2nd generation high-κ/metal gate, and raised S/D regions for 32 nm technology [1.28].

Table 1.2 Key device features of Intel 32 nm logic technology [1.28].

	NMOS	PMOS
L_G [nm]	30	30
EOT [nm]	0.9	0.9
I_{Dsat} [mA/mm] $V_{GS}=V_{DS}=1.0V$	1.62	1.37
I_{Dlin} [mA/mm] $V_{GS}=1.0V$ $V_{DS}=0.05V$	0.231	0.240
I_{OFF} [nA/mm]	100	100
V_{Tsat} [V]	0.115	-0.18
DIBL [mV/V]	~200	~200
SS [mV/dec]	~100	~100

- ☐ 30 nm gate length with 112.5 nm contacted gate pitch
- ☐ 2nd generation high-κ/metal gate
  - ✓ 0.9 nm EOT
  - ✓ Replacement metal gate approach
    - Enables stress enhancement techniques
  - ✓ Replacement high-k approach
    - Improved performance
- ☐ 4th generation SiGe strained silicon PMOS device
  - ✓ Increased Ge concentration
  - ✓ Closer proximity to channel for enhanced mobility
- ☐ Raised NMOS S/D region
  - ✓ Improved external resistance
- ☐ 2nd generation trench contacts
  - ✓ Reduced contact resistance
  - ✓ Used as local interconnects

These technologies are introduced along the performance enhancement strategies mentioned above. Note that most of these technologies are needed due to the short-channel effects. For a given short channel characteristic and constant I_{OFF}, the gate length scaling through increased threshold voltage degrades drive current. The improvement in drive current owing to the shorter gate length is offset by the reduction in overdrive voltage (V_G-V_T). Therefore, mobility enhancement becomes a key engineering factor with minimal impact to leakage. Moreover, suppression of the short-channel effects demands rapid development of high-κ/metal gate technology to minimize EOT and the raised S/D technology to obtain low S/D resistance with shallow junction as well as several channel doping technologies. In addition, increasing the impurity concentration in the channel region is also necessary to suppress the short-channel effects. However, this causes degradation of carrier mobility due to impurity scattering, which result in obstruction to increase the drive current [1.29]. Moreover, random dopant fluctuation in such short channel results in variation of the threshold voltage [1.30]. To continuously improve CMOS performance, comprehensive reforms of the CMOS scaling strategy will be required to achieve high immunity of short-channel effects.

1.3.1 Gate-All-Around Silicon Nanowire MOSFET

Fully-depleted (FD) silicon-on-insulator (SOI) MOSFETs are considered as possible successors for bulk MOSFETs because their thin body dimensions provide geometric electrostatic confinement for controlling short-channel effects as well as less stringent requirements of EOT scaling over conventional bulk Si. Figure 1.27 shows a typical planar FDSOI MOSFET structure. The thin layer of silicon is separated from buried oxide (BOX) film, thus electrically isolating the devices from the underlying silicon substrate. In this structure, the source and drain junction capacitance is almost entirely eliminated. Here FD structure means that the silicon body film is thin enough that the entire film is depleted before the threshold condition is reached. An important merit of SOI technology is that it provides the cornerstone for new FD device structures such as multi-gate MOSFETs, which includes more than one gate into a single device as

shown in Figure 1.28 [1.32].

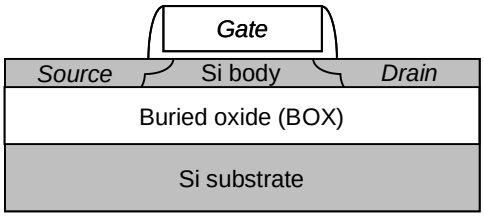


Figure 1.27 Cross-section of a planar FDSOI MOSFET.

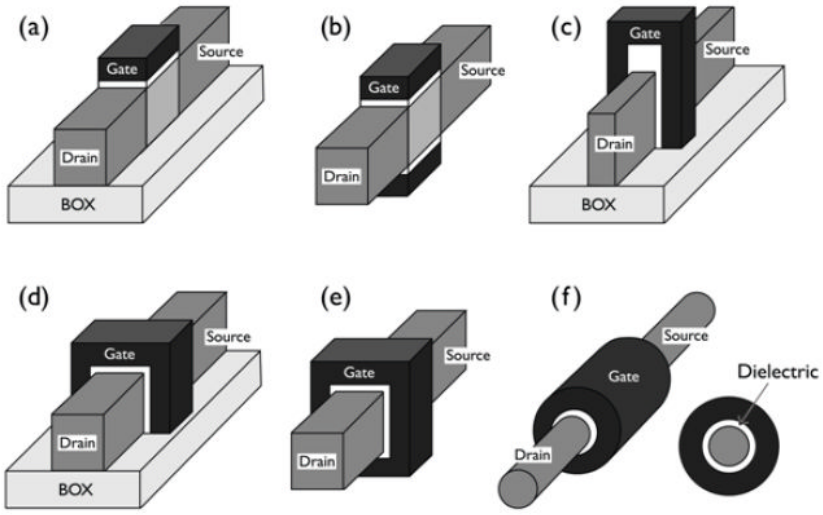


Figure 1.28 Various SOI device: (a) Single gate SOI FET, (b) double gate planar SOI FET, (c) double gate non-planar FinFET, (d) tri-gate FET, (e) quadruple-gate (or gate-all-around) FET, and gate-all-around (or surrounding gate) FET (nanowire FET).

Figure 1.29 displays the electric field liens from the drain in different MOSFET structures. More is the penetration of field lines from the drain towards the source, greater is the interference of the drain on the function of the gate. The conventional MOSFET geometry, shown in Figure 1.29 (a), can hardly be scaled down because of strong permeation of the field lines towards the source. The behavior of FDSOI structure of Figure 1.29 (b) is not encouraging because the buried oxide does not terminate the drain field lines. Only in the double-gate (DG) MOSFET structure shown in Figure 1.29 (c), the field lines are not able to reach near the source [1.31]. Hence the effect of the drain field on the channel is minimized providing far superior scalability.

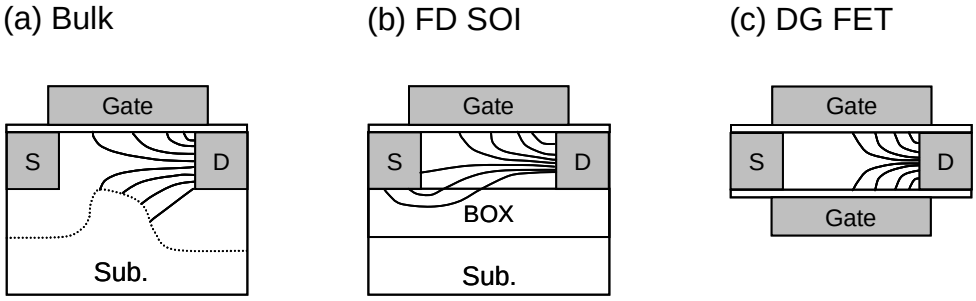


Figure 1.29 Illustration of electric field liens from drain of different device types: (a) bulk, (b) FD SOI, and (c) double gate (DG).

Yan *et al.* proposed a unique scaling theory for double-gate SOI MOSFETs as a design guideline [1.33]. According to their theory, the device should be designed maintaining

$$\alpha = \frac{L_{eff}}{2\lambda}, \tag{1.27}$$

where λ is the so-called natural length which governs the influence of lateral field on the channel potential and depends on device geometry and boundary conditions. To This natural length is an easy guide for choosing device structure and parameters, and has simple physical meaning that a small natural length corresponds to superb short channel effect immunity. Table 1.3 shows the natural length for different gate configurations [1.32]. Here a small α gives degraded short-channel effect immunity. For instance, to achieve SS < 75 mV/decade and DIBL < 50 mV/V, a needs to be larger than 2.2 for all devices when the gate oxide is 2 nm and the channel doping concentration is $1x10^{18}cm^{-3}$ [1.34]. Figure 1.30 shows the maximum allowed silicon film thickness (and device width in a four-gate device with $W = t_{si}$) to avoid short-channel effects [1.35]. As a consequence, gate-all-around (GAA) silicon nanowire MOSFETs (SNWTs) have the best short-channel effect immunity among all the FDSOI architectures for the same body dimensions. The experimental data reported by Bangsaruntip *et al.* also shows better short-channel immunity for GAA SNWTs than that of FDSOI by as shown in Figure 1.31 [1.36].

Table 1.3 Natural length in devices with different geometries [1.32].

Single gate	$\lambda_1 = \sqrt{\frac{\varepsilon_{si}}{\varepsilon_{ox}} t_{si} t_{ox}}$
Double gate	$\lambda_2 = \sqrt{\frac{\varepsilon_{si}}{2\varepsilon_{ox}} t_{si} t_{ox}}$
Quadruple gate (square section)	$\lambda_4 \cong \sqrt{\frac{\varepsilon_{si}}{4\varepsilon_{ox}} t_{si} t_{ox}}$
Surrounding gate (circular section)	$\lambda_O \cong \sqrt{\frac{2\varepsilon_{si} t_{si}^2 \ln(1+2t_{ox}/t_{si})}{16\varepsilon_{ox}} + \varepsilon_{ox} t_{si}^2}$

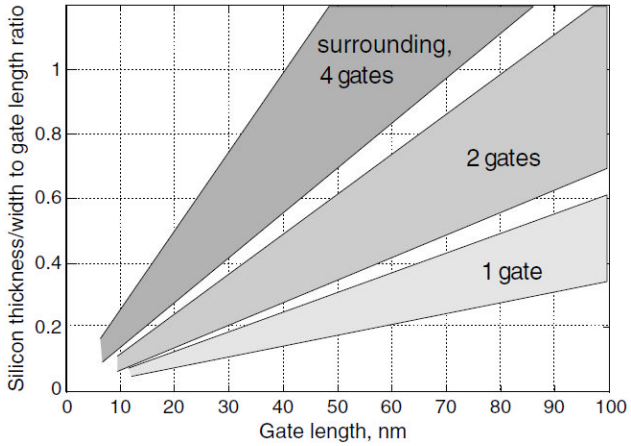


Figure 1.30 Maximum allowed Si thickness and device width vs. gate length to avoid short-channel effects in single-, double- and quadruple-gate SOI MOSFETs [1.35].

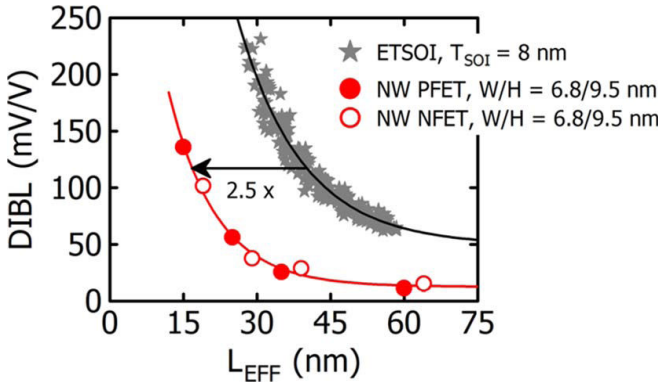


Figure 1.31 Comparison of DIBL of elliptical GAA SNWTs (width W = 6.8 nm and height H = 9.5 nm) and single-gate ETSOI FETs (SOI thickness t_{si} = 8 nm) with similar body dimensions. [1.36].

1.3.2 Vertically-Stacked Channel MOSFET

In a GAA SNWT, the current drive is essentially equal to the sum of the currents flowing along all the interfaces covered by the gate electrode [1.32]. To dive large currents, multi-finger pattern are need as shown in Figure 1.32. When an ideal current transport (without corner effect, volume inversion effect, and so on) are considered in rectangular GAA SNWT, the current level depends on the width W_{NW} and height H_{NW}, of a nanowire, the top and bottom surface mobility μ_{top}, the side surface mobility μ_{side}, and the number of the nanowires for given layout area. The on-current in GAA SNWT with multi-finger are given by

$$I_{ON}^{nanowire} = n \times (2W_{NW}\mu_{top} + 2H_{NW}\mu_{side}) \frac{C_{ox}}{L_{eff}} \frac{(V_{DD} - V_T)^{\alpha}}{2m}, \qquad (1.28)$$

where n is the number of the nanowires and the power of α ($1 < \alpha < 2$) indicates the degree of velocity saturation in short-channel MOSFETs. Considering a nanowire pitch W_{pitch}, the current per unit device width is given by

$$\frac{I_{ON}^{nanowire}}{W_{pitch}} = \frac{I_{ON}^{planar}}{W} \cdot \frac{2W_{NW}\mu_{top} + 2H_{NW}\mu_{side}}{\mu_{top}W_{pitch}}, \qquad (1.29)$$

where I_{ON}^{planar} is the current in the single-gate, planar MOSFET occupying the same area as the multi-finger device as shown in Figure 1.32.

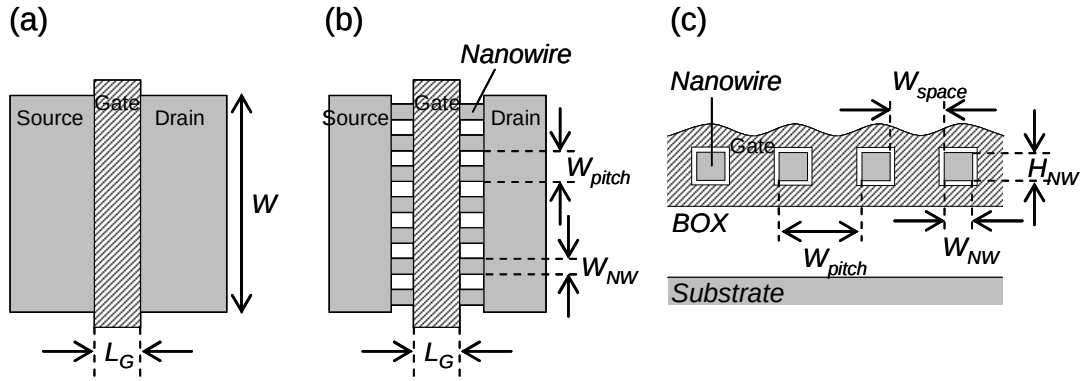


Figure 1.32 MOSFET layout; (a) planar MOSFET, (b) GAA SNWT with multi-finger, and (c) cross-section of GAA SNWT.

To realize a benefit of the multi finger layout, the current of GAA SNWT per layout surface normalized to planar MOSFET are calculated with several dimensions.

Here we assume the nanowire direction of <110>, the top surface orientation of (100), and the sidewall surface orientation of (110). The values of the mobility for (100) and (110) used thus are 300 cm^2/Vs and 150 cm^2/Vs, respectively. Figure 1.33 shows its nanowire pitch dependence. The nanowire width W_{NW} is equal to a half of the pitch width. Because of that, the currents of GAA SNWT approach to the planar one as increase the pitch. If the cross section of the nanowire is square (W_{NW} = H_{NW}), the current is independent on the pitch, and we have $I_{ON}^{nanowire}$ = 1.5x I_{ON}^{planar}. In actual case, since each nanowires are wrapped by the gate dielectrics and the space between them must be filled by gate metal, the space between nanowires W_{space} may need at least 15 nm. In that case, the diminishing cross-section to obtain the better short-channel effect immunity is traded off against the current density gain. When the space is equal to 15 nm, there is no gain of the current density for GAA SNWT with W_{NW} = 5 nm and H_{NW} = 10 nm to the planar one, for instance.

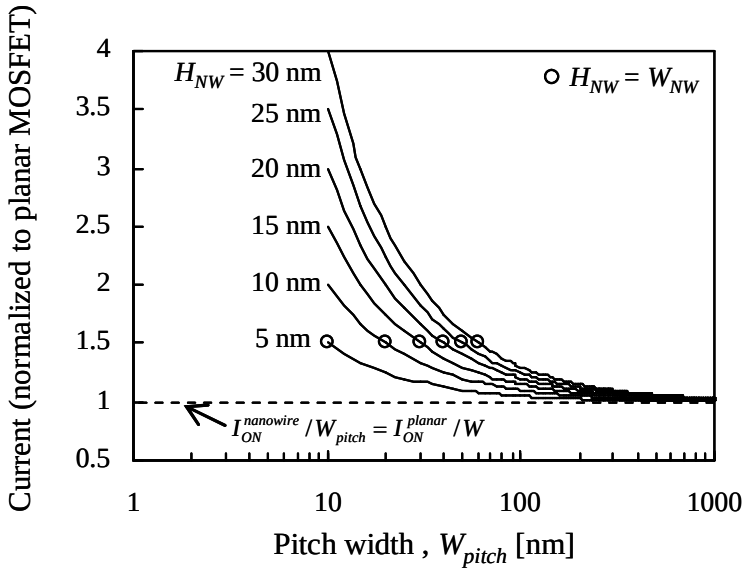


Figure 1.33 Normalized current of a rectangular GAA SNWT as a function of multi-finger pitch width. W_{NW}=W_{pitch}/2. The top interface mobility is 300 cm^2/Vs and sidewall mobility is 150 cm^2/Vs.

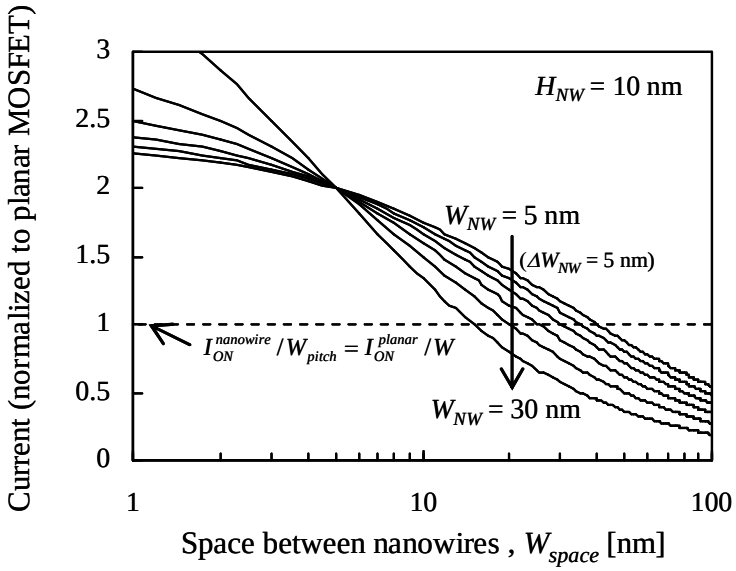


Figure 1.34 Normalized current of a rectangular GAA SNWT as a function of space between nanowires. H_{NW}=10 nm. The top interface mobility is 300 cm^2/Vs and sidewall mobility is 150 cm^2/Vs.

To obtain larger current density per a given layout with high immunity of the short-channel effects, the vertical integration of nanowires is effective as shown in Figure 1.35. In this structure, the current density can be proportionally increased to the stacking level (three levels in the case of Figure 1.35) without the layout surface area penalty as seen in Figure 1.36. This structure enables to achieve both high integration density and low power dissipation. Note that since this current increment is due to the effective surface enlargement, the intrinsic gate delay cannot be directly reduced as discussed in Subsection 1.13. The propagation delay, however, can be improved as the interconnect capacitance becomes larger. The suppressed short channel effects are also effective to reach the ideal switching trajectory.

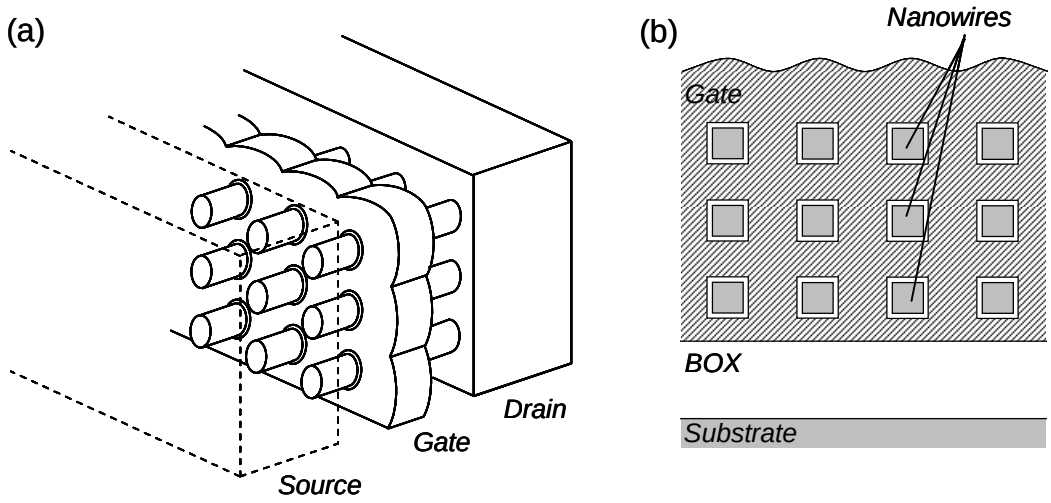


Figure 1.35 Structure of vertically-stacked GAA SNWT(a) and its cross-section (b)

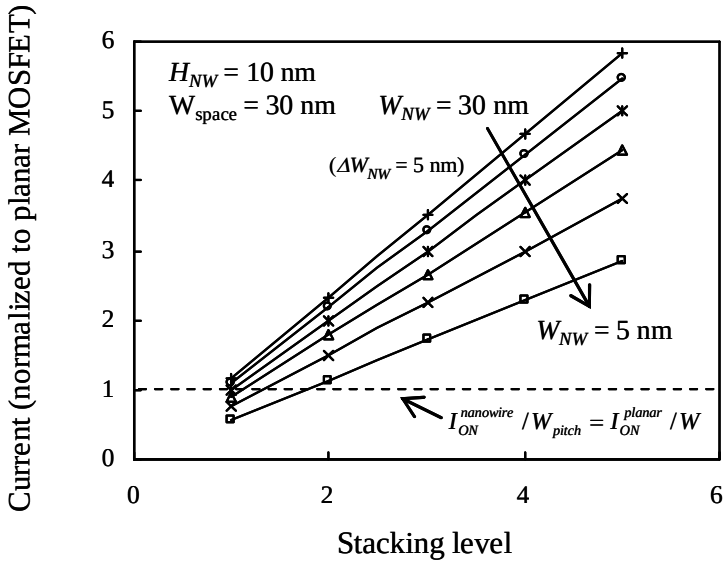


Figure 1.36 Normalized current of a rectangular GAA SNWT as a function of staking level of nanowires. H_{NW}=10 nm. W_{space}=30 nm. The top interface mobility is 300 cm^2/Vs and sidewall mobility is 150 cm^2/Vs.

1.4 PURPOSE AND CONTENTS OF THIS STUDY

The operation speed and integration density of a microprocessor chip have doubled every two or three years in line with Moore's law. For each generation, in order

to achieve the target performance, prospective problems were analyzed in detail and solved by introducing new technologies. Looking back at the history, it seems that many problems result from the short-channel effects. Now we are faced with the critical issue of power dissipation due to subthreshold leakage. This power dissipation, which surpasses dynamic power consumption, can no longer be ignored. Therefore comprehensive reform of the CMOS scaling strategies is required to continuously improve speed, integration density, and power dissipation of LSI all at the same time.

The purpose of this thesis thus is to characterize vertically-stacked GAA SNWT as one of the most promising MOSFETs for future CMOS circuits. The superior immunity to short-channel effects and the high integration density can be straightforwardly expected from this structure. However, the body dimensions must be considered with drive current enhancement. In addition, it is possible that the thick source/drain regions of the vertically-stacked channel MOSFET need special treatments in the fabrication process to obtain low access resistance. As other critical issue in this type of three dimensional devices, body bias techniques are not available for power management of digital circuits. Thus the studies are classified into three issues:

- How do we achieve low access resistance with uniform doping profile in the thick source/drain regions?
- How are the carrier transport properties of vertically-stacked nanowire channels?
- How can we realize the controllability and flexibility of the threshold voltage?

Figure 1.37 shows the outline of this thesis. First, the basic device fabrication process of vertically-stacked channel structure is presented and specific technological issues are evidenced in Chapter 2. Next, the electrical characterization methods are described in detail in Chapter 3. Chapter 4 covers the studies source/drain resistance reduction technique for thick source/drain regions in vertically-stacked channel devices. Chapter 5 covers the carrier transport properties of vertically-stacked GAA SNWTs in detail, especially the dimension effects and the impacts on the fabrication process. Chapter 6 covers the studies of the threshold voltage tuning technique by the independently separated gates. The studies referred to in each chapter were done in

order to achieve CMOSFET with high performance, high integration density, and low power dissipation for future generation. Details are as follows.

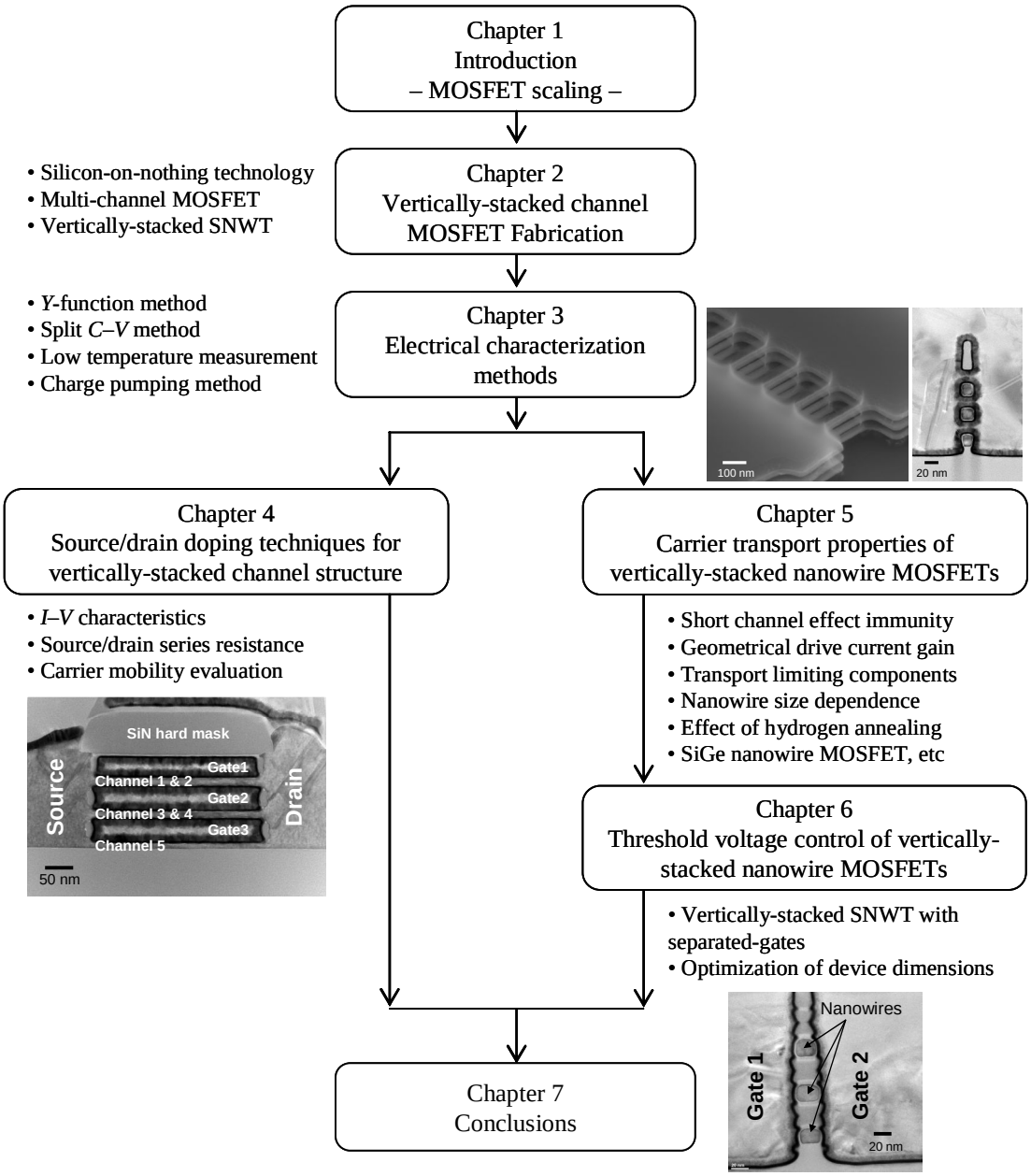


Figure 1.37 Outline of each chapter in this thesis.

a) Source/drain doping techniques for vertically-stacked channel structure (Chapter 4)

The reduction of parasitic access resistances serially connected at source and

drain region of MOSFETS are one of the challenging technology to meet the performance required in the roadmap as discussed in Chapter 1. The voltage drops at the resistances reduces the applied voltages at the drain and gate electrodes in the transistors, resulting in the decrease in the overdrive voltage to lower the on-current. These resistances also cause degradation in the time constant, commonly refereed as RC delay, to lower the switching speed. Therefore, as the channel resistance reduces with the scaling in the gate length, the parasitic series resistance should be further reduced not to increase its proportion in the total resistance at on-state. In addition, the vertically-stacked channel structure suffers from the uniformity of the doping profile due to the thick source/drain regions.

In Chapter 4, a novel process to decrease the resistivity of the source and drain regions is presented using *in situ* doped selective epitaxial growth in combination with conventional ion-implantation as a novel process for source and drain formation. The effect of source and drain formation process on the series resistance and carrier mobility will be discussed through electrical characteristics.

b) Carrier transport properties of vertically-stacked nanowire MOSFETs (Chapter 5)

Recently, short channel GAA SNWTs have been successfully fabricated with diameter of less than 10 nm using several top-down CMOS compatible processes [1.37–1.39]; they showed excellent short-channel effects immunity. On the other hand, transport property degradation in SNWTs was also reported by several groups [1.40–1.42]. However, the mobility behavior when the width is reduced has been remained unclear. Carrier transport in nanowire is commonly discussed in terms of two main mechanisms; one is one-dimension transport model, and the other is a facet-dominated transport model. The former can be adapted to sub-10 nm diameter, and the latter to more than 20 nm one. In the range of between them, the mixed transport properties are expected. The range of between them is production-friendly, provided that the short-channel effects under aggressively scaled gate length are suppressed, while carrier transport model becomes complicate due to the mixed properties.

In Chapter 5, carrier transport limiting components for vertically-stacked GAA SNWTs will be discussed in detail to obtain better performance with suppressing short channel effects.

c) Threshold voltage control of vertically-stacked nanowire MOSFETs (Chapter 6)

Another issue for the vertically-stacked SNWTs is how to control the threshold voltage. For various CMOS applications and power management, it is important to achieve controllable and flexible threshold voltage in a transistor. As one of the possible techniques for three-dimensional devices, separated-gate structures have been proposed [1.43, 1.44]. In Chapter 5, the possibility of the flexible threshold voltage for vertically-stacked GAA SNWTs with separated gate will be discussed.

Finally, in Chapter 7, the results obtained in these studies are summarized and conclusions are presented. The perspective of vertically-stacked GAA SNWTs for future LSI applications is described.

1.5 REFERENCE

[1.1] G. E. Moore, “Cramming more components onto integrated circuits,” Electronics, Vol. 38, No. 8, 1965.

[1.2] Data source: http://www.intel.com/technology/timeline.pdf

[1.3] G. E. Moore, “No Exponential is Forever: But “Forever” Can be Delayed!,” in *ISSCC Tech. Dig.*, Feb. 2003, vol.1, pp.20–23.

[1.4] Y. Taur and T. H. Ning, *Fundamental of Modern VLSI Devices*, 2nd edition, Cambridge University Press, 2009.

[1.5] R. H. Dennard, F. H. Gaensslen, H-N, Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFET’s with very small physical dimensions," *IEEE J. Solid-State Circuits*, vol. SC-9, no. 5, pp. 256–268, 1974.

[1.6] M. Bohr, “MOS Transistor: Scaling and Performance Trends,” Semiconductor International, vol. 18, no. 6, pp.75–80.

[1.7] International Technology Roadmap for Semiconductors (ITRS): http://www.itrs.net/reports.html

[1.8] M. H. Na, E. J. Nowak, W. Haensch, and J. Cai, "The effective drive current in CMOS inverters," in *IEDM Tech. Dig.*, 2002, pp. 121–124.

[1.9] E. Yoshida, Y. Momiyama, M. Miyamoto, T. Saiki, M. Kojima, S. Satoh, and T. Sugii," Performance Boost using a New Device Design Methodology Based on Characteristic Current for Low-Power CMOS," in *IEDM Tech. Dig.*, 2006, pp. 195–198.

[1.10] L. Wei, F. Boeuf, D. Antoniadis, T. Skotnicki, and H.-S. P. Wong, "Exploration of Device Design Space to Meet Circuit Speed Targeting 22nm and Beyond," in extended abstracts of SSDM, 2009, pp.808–809.

[1.11] T. Skotnicki and F. Boeuf, "How Can High Mobility Channel Materials Boost or Degrade Performance in Advanced CMOS," in *VLSI Tech. Dig.*, 2010, pp. 153–154.

[1.12] T. Sakurai, "Perspectives of power-aware electronics (Plenary)," in *Proc. ISSCC Dig. Tech. Papers*, 2003, pp. 26–29.

[1.13] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neirynck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, C. Thomas, T. T roeger, P. Vandervoorn, S. Williams, and K. Zawadzki, "A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," in *IEDM Tech. Dig.*, 2007, pp. 247–250.

[1.14] S. A. Parke, J.E. Moon, H. C. Wann, P. K. Ko, and C. Hu, "Design for suppression of gate-induced drain leakage in LDD MOSFET's using a quasi-two-dimensional analytical model," *IEEE Trans. Electron Devices*, vol. 39,

no. 7, pp. 1694–1703, 1992.

[1.15] K. Tanaka, K. Takeuchi, and M. Hane, ”Practical FinFET designe considering GIDL for LSTP (Low Standby Power) devices,” in *IEDM Tech. Dig.*, 2005, pp. 980–983.

[1.16] Scalable CMOS (SCMOS) Design Rules (Based on *MOSIS* design: http://www.mosis.com/)

[1.17] T. Skotnicki, G. Merckel, and T. Pedron, “The voltage-doping transformation: A new approach to the modeling of MOSFET short-channel effects,” *IEEE Electron Device Lett.*, vol. 9, no. 3, pp. 109–112, 1988.

[1.18] T. Skotnicki, C. F.-Beranger, C. Gallon, F. Boeuf, S. Monfray, F. Payet, A. Pouydebasque, M. Szczap, A. Farcy, F. Arnaud, S. Clerc, M. Sellier, A. Cathignol, J.-P. Schoellkopf, E. Perea, R. Ferrant, and H. Mingam, “Innovative Materials, Devices, and CMOS Technologies for Low-Power Mobile Multimedia,” *IEEE Trans. Electron Devices*, vol. 55, no. 1, pp. 96–130, 2008.

[1.19] *MASTAR4 Code and User Guide*. Available: http://www.itrs.net/models.html

[1.20] S. Thompson, P. Packan, T. Ghani, M. Stettler, M. Alavi, I. Post, S. Tyagi, S. Ahmed, S. Yang, and M.Bohr, ”Source/Drain extension scaling for 0.1mm and below channel length MOSFETs,” in *Proc. Symp. VLSI Tech.*, 1998, pp. 132–133.

[1.21] M. Lundstrom and Z. Ren, “Essential physics of carrier transport in nanoscale MOSFETs,” *IEEE Trans. Electron Devices*, vol. 49, no. 1, pp. 133–141, 2002.

[1.22] S. M. Sze and Kwok K. Ng, *Physics of Semiconductor Devices*, 3rd edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 2007.

[1.23] M. Saitoh and K. Uchida, “Universal relationship between low-field mobility and high-field carrier velocity in high-κ and SiO_2 gate dielectric MOSFETs,” in *IEDM Tech. Dig.*, 2006, pp. 261–264.

[1.24] K. Natori, “Ballistic metal–oxide–semiconductor field effect transistor,” *J. Appl. Phys.*, vol. 76, no. 8, pp. 4879–4890, Oct. 1994.

[1.25] S. Takagi, T. Irisawa, T. Tezuka, T. Numata, S. Nakaharai, N. Hirashita, Y. Moriyama, K. Usuda, E. Toyoda, S. Dissanayake, M. Shichijo, R. Nakane, S. Sugahara, M. Takenaka, and N. Sugiyama, “Carrier-Transport-Enhanced Channel

CMOS for Improved Power Consumption and Performance, ” *IEEE Trans. Electron Devices*, vol. 55, no. 1, pp. 21–38, 2008.

[1.26] A. Rahaman and M. S. Lundstrom, “A compact scattering model for the nanoscale double gate MOSFET,” *IEEE Trans. Electron Devices*, vol. 49, no. 3, pp. 481– 489, 2002.

[1.27] K. Tatshumura, M. Goto, S. Kawanaka, and A. Kinoshita, “Correlation between low-field mobility and high-field carrier velocity in quasi-ballistic-transport MISFETs scaled down to L_g=30 nm,” in *IEDM Tech. Dig.*, 2009, pp. 465–468.

[1.28] P. Packan, S. Akbar, M. Armstrong, D. Bergstrom, M. Brazier, H. Deshpande, K. Dev, G. Ding, T. Ghani, O. Golonzka, W. Han, J. He, R. Heussner, R. James, J. Jopling, C. Kenyon, S-H. Lee, M. Liu, S. Lodha, B. Mattis, A. Murthy, L. Neiberg, J. Neirynck, S. Pae, C. Parker, L. Pipes, J. Sebastian, J. Seiple, B. Sell, A. Sharma, S. Sivakumar, B. Song, A. St. Amour, K. Tone, T. Troeger, C. Weber, K. Zhang, Y. Luo, S. Natarajan, “High Performance 32nm Logic Technology Featuring 2nd Generation High-k + Metal Gate Transistors,” in *IEDM Tech. Dig.*, 2009, pp. 659–662.

[1.29] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, “On the Universality of Inversion Layer Mobility in Si MOSFET’s: Part I–Effects of Substrate Impurity Consentration,” *IEEE Trans. Electron Devices*, vol. 41, no. 12, pp. 2357–2362, Dec. 1994.

[1.30] K. Nishinohara, N. Shygyo, and T.Wada, “Effects of microscopic fluctuations in dopant distributions on MOSFET threshold voltage,” *IEEE Trans. Electron Devices*, vol. 39, no. 3, pp. 634–639, Mar. 1992.

[1.31] T. Ernst, R. Ritzenthaler, O. Faynot, and S. Cristoloveanu, “A model of fringing field in short-channel planar and triple-gate SOI MOSFETs,” *IEEE Trans. Electron Devices*, vol. 54, no. 6, pp. 1366–1375, June. 2007.

[1.32] J. P. Coling, “Multi-gate SOI MOSFETs,” *Microelectronic Engineering*, vol. 84, pp. 2071–2076, 2007.

[1.33] R.-H. Yan, A. Ourmazd, and L. F. Lee, “Scaling the Si MOSFET: from bulk to SOI to bulk,” *IEEE Trans. Electron Devices*, vol. 39, no. 7, pp. 1704–1710, July. 1992.

[1.34] C.-W. Lee, S.-R.-N. Yun, C.-G. Yu, J.-T. Park, J.P. Colinge, “Device design guidelines

for nano-scale MuGFETs," *Solid-State Electronics*, vol. 51, pp. 505–510, 2007.

[1.35] J. P. Coling, "Multiple-gate SOI MOSFETs," *Solid-State Electronics*, vol. 48, pp. 897–905, 2004.

[1.36] S. Bangsaruntip, G. M. Cohen, A. Majumdar, and J. W. Sleight, "Univesality of short-channel effects in undoped-body silicon nanowire MOSFETs," *IEEE Electron Device Lett.* vol. 31, no. 9, pp. 903–905, Sept. 2010.

[1.37] S. Bangsaruntip, G.M. Cohen, A. Majumdar, Y. Zhang, S. U. Engelmann, N. C. M. Fuller, L. M. Gignac, S. Mittal, J. S. Newbury, M. Guillorn, T. Barwicz, L. Sekaric, M. M. Frank, and J. W. Sleight, "High performance and highly uniform gate-all-around silicon nanowire MOSFETs with wire size dependent scaling," in *IEDM Tech. Dig.*, 2009, pp. 297–300.

[1.38] M. Li, K.H. Yeo, S.D. Suk, Y.Y. Yeoh, D.-W. Kim, T.Y. Chung, K. S. Oh, and W.-S. Lee, "Sub-10 nm gate-all-around CMOS nanowire transistors on bulk Si substrate," in *VLSI Tech. Dig.*, 2009, pp. 94–95.

[1.39] G. Bidal, F. Boeuf, S. Denorme, N. Loubet, J.L. Huguenin, P. Perreau, D. Fleury, F. Leverd, S. Lagrasta, S. Barnola, T. Salvetat, B. Orlando, R. Beneyton, L. Clement, R. Pantel, S. Monfray, G. Ghibaudo, and T. Skotnicki, "High velocity Si-nanodot: a candidate for SRAM applications at 16 nm node and below," in *VLSI Tech. Dig.*, 2009, pp. 240–241.

[1.40] S. D. Suk, M. Li, Y. Y. Yeoh, K. H. Yeo, K. H. Cho, I. K. Ku, H. Cho, W. J. Jang, D.-W. Kim, D. Park, andW.-S. Lee, "Investigation of nanowire size dependency on TSNWFET," in *IEDM Tech. Dig.*, 2007, pp. 891–894.

[1.41] M. Saitoh, Y. Nakabayashi, H. Itokawa1, M. Murano, I. Mizushima1, K. Uchida, and T. Numata, "Short-Channel Performance and Mobility Analysis of <110>- and <100>-Oriented Tri-Gate Nanowire MOSFETs with Raised Source/Drain Extensions," in *VLSI Tech. Dig.*, 2010, pp. 169–170.

[1.42] J. Chen, T. Saraya, K. Miyaji, K. Shimizu, and T. Hiramoto, "Experimental Study of Mobility in [110]- and [100]-Directed Multiple Silicon Nanowire GAA MOSFETs on (100) SOI," in *VLSI Tech. Dig.*, 2008, pp. 32–33.

[1.43] *Y.* X. Liu, M. Masahara, K. Ishii, T. Tsutsumi, T. Sekigawa, H. Takahima, H.

Yamauchi and E. Suzuki, “Flexible Threshold Voltage FinFETs with Independent Double Gates and an Ideal Rectangular Cross- Section Si-Fin Channel”, in *IEDM Tech. Dig.*, 2003, pp. 986–989.

[1.44] C. Dupré, A. Hubert, S. Bécu, M. Jublot, V. Maffini-Alvaro, C. Vizioz, F. Aussenac, C. Arvet, S. Barnola, J.-M. Hartmann, G. Garnier, F. Allain, J.-P. Colonna, M. Rivoire, L. Baud, S. Pauliac, V. Loup, T. Chevolleau, P. Rivallin, B. Guillaumot, G. Ghibaudo, O. Faynot, T. Ernst, and S. Deleonibus, “ 15nm-diameter 3D Stacked Nanowires with Independent Gate Operation: ΦFET,” in *IEDM Tech. Dig.*, 2008, pp. 749–752.

CHAPTER 2

VERTICALLY- STACKED CHANNEL MOSFETS FABRICATION

CHAPTER 2 CONTENTS

2.1 SILICON-ON-NOTHING TECHNOLOGY

Silicon-On-Nothing (SON) architecture have been proposed allowing extremely thin buried oxides and Silicon films to be fabricated and thereby better resisting to short-channel effects [2.1–2.3]. The process is based on the use of a sacrificial SiGe layer that is selectively removed versus the silicon as shown in Fig. 2.1. The tunnel under the silicon film is then filled with the oxide. Silicon film and buried oxide are defined by epitaxy that opens access to extremely thin films. Extensions are physically limited by the silicon film thickness preserving a good control of shorts channel effects. In the same time, source and drain remain in continuity with the substrate, limiting self-heating and lowering R_{SD}. The use of thin buried oxide allows the control of the fringing field. Moreover, this architecture is co-integrable with bulk one [2.3].

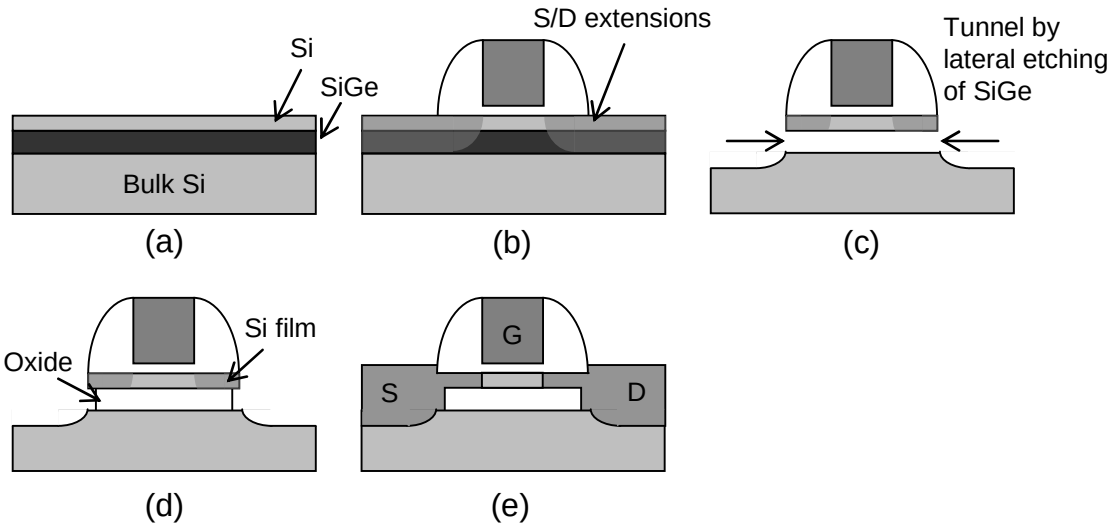


Figure 2.1 Fabrication process of the SON MOSFET: (a) epitaxy of SiGe and Si layers on isolated bulk wafer; (b) conventional CMOS process steps until formation of the nitride spacers; (c) formation of the shallow trenches in the S/D regions and formation of the tunnel under the Si film; (d) filling the tunnel with oxide (optional step); (e) selective epitaxy of S/D regions, implantation and RTA.

Vertically-stacked channel structure has been derived from SON and GAA concepts. The use of sacrificial SiGe layers enables silicon channel to be piled up. In this study, two types of stacked channel MOSFETs were fabricated; one is multi-channel MOSFET, another is vertically-stacked nanowire FET.

2.2 PROCESS STEP OVERVIEW FOR MULTI-CHANNEL MOSFET

The fabrication process of MCFETs, which is based on the principles developed for Silicon-On-Nothing FETs, has been developed by Bernard [2.2–2.4]. Figure 2.2 shows an overview of the MCFET process flow. First, a (25-nm-Si and 30-nm-$Si_{0.8}Ge_{0.2}$) superlattice structure was epitaxially grown on a 20-nm thick SOI substrate (step 1). The grown Si layers will be used as channels for MCFETs. 5nm of high-temperature oxide (HTO) and 80nm of silicon nitride (Si_3N_4) are deposited as a hard mask on the top of the stack. After a gate photolithography, the superlattice is then anisotropically etched down to the bottom SOI layer (step 2). Since L_G is defined at this etching step, a vertical etching profile is an important requirement for this technology in order to suppress the variability of the L_G among the stacked Si channels. To introduce internal spacers, the SiGe layers are partially etched selectively to the Si ones (step 2.1). The depth of the SiGe recess determines the thickness of the future internal spacers. Then, HTO and Si_3N_4 layers are deposited in the cavities with a thickness ratio optimized for the following spacers etch (step 2.2 and 2.3). An anisotropic/isotropic etch sequence of the deposited dielectrics allows one to access the silicon layers for further S/D epitaxy (step 2.4). After the wet cleaning to remove entirely the dielectrics on the sides of silicon layers, crystalline Si S/D was selectively grown with a natural “flat” shape induced by the presence of the internal spacers (step 3). The S/D were then ion-implanted. The channel width was defined by the active-area patterning and etching, giving access to the SiGe layers. The SiGe layers are selectively removed using pure CF_4 at high pressure and low microwave power in a remote plasma tool. The high-*k*/metal gate stack (HfO_2/TiN/N^+ poly-Si) was deposited in the obtained cavities. A

second gate etch is subsequently carried out followed by the formation of external silicon nitride spacers to avoid any short-cut between the gate and the S/D. After the dopant activation anneal, the top of S/D are silicided (with nickel) to reduce the series resistances, followed by a standard back-end-of-the-line (BEOL) process (for contacts and interconnections).

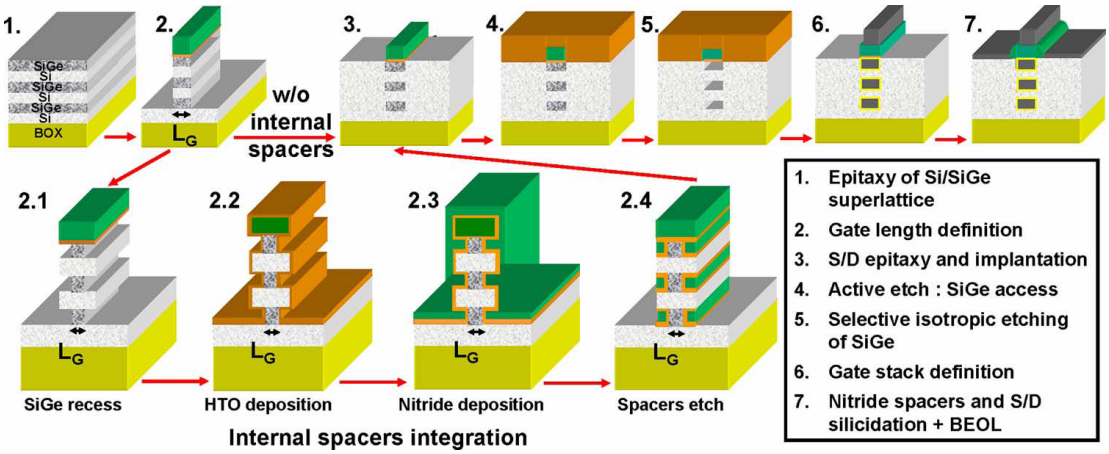


Figure 2.2 MCFET fabrication process overview [2.2].

Figure 2.3 shows the cross-sectional transmission electron microscopy (TEM) images of the fabricated MCFETs along the Si channel and width direction. Here, two SiGe/Si alternating layers, resulting in five Si channels in parallel, were designed for fabrication. Channels 1 and 2 are activated at the top and bottom interfaces of the superior silicon island, whereas channels 3 and 4 have similar formations within the intermediate Si island. Channel 5 is responsible for conduction as the top of the original Si film. The sixth possible channel, which can be activated at the bottom of the Si film by substrate (back-gate) biasing, is not investigated in this work. The resulting channel thickness was 10 nm for all channels. The measured equivalent oxide thickness (EOT) in inversion was ~2.5 nm. The gate lengths of the MCFETs range from 500 to 70 nm.

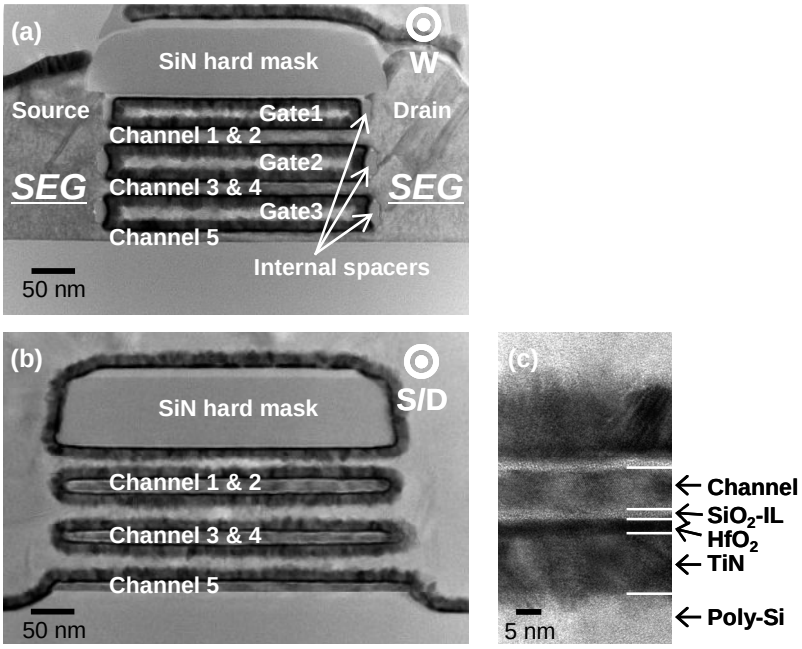


Figure 2.3 Fabricated multi channel FET (MCFET) along (a) channel length and (b) width direction. (c) is the enlarged image of the gate stack.

2.3 PROCESS STEP OVERVIEW FOR VERTICALLY-STACKED NANOWIRE MOSFET

Figure 2.4 illustrates briefly the process flow for making vertically stacked silicon nanowire MOSFET [2.5, 2.6]. First, a (30-nm Si / 30-nm $Si_{0.8}Ge_{0.2}$)x3 superlattice was epitaxially grown on (100) SOI substrate (step1). 5nm of HTO and 40nm of Si_3N_4 are deposited as a hard mask on the top of the stack (step 2). Hybrid DUV/ebeam -lithography and resist trimming were combined to define narrow lines. A damascene process was used: cavities were patterned thanks to anisotropic dry plasma etching of the superlattice (step 3). SiGe layers between Si ones were then etched isotropically (step 4). Optionally, an hydrogen annealing process (at 750 ^{o}C and 20 Torr for 2 min) was applied to obtained circular cross-sectional shape of nanowires (step 5). The obtained cavities were then filled with the gate stack (HfO_2/TiN/N^+ poly-Si) (step 6). The gate length is thus defined by the cavity length. Chemical Mechanical Polishing (CMP) of poly-silicon and the thick HTO hard mask deposition for S/D implantation

were carried out followed by the gate patterning (step 7). After gate etching, S/D implantation, spacers formation, and the dopant activation anneal were performed. The nickel silicide was formed on the top of Source / Drain area. The fabrication ended with a standard BEOL process.

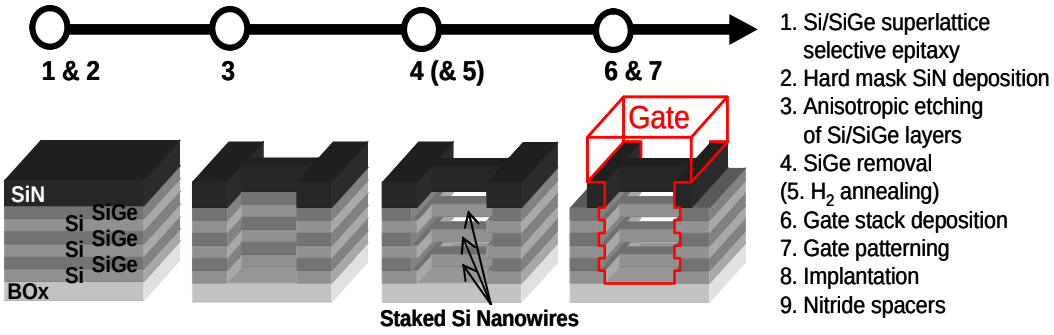


Figure 2.4 Vertically-stacked nanowire MOSFET fabrication process overview [2.5].

Figure 2.5 shows their cross-sectional TEM and SEM images. We successfully fabricated narrow nanowires by using e-beam lithography and isotropic plasma etching of SiGe sacrificial layers. Good uniformity of the nanowires in a 200-mm-wafer has been also achieved as shown in Figure 2.6. The width (W_{NW}) of rectangular shape nanowire ranges from 5 nm up to 30 nm and the height (H_{NW}) is 15 nm as seen in Figure 2.7 and Figure 2.8. The smallest nanowire with 5-nm-W_{NW} was formed by H_2 annealing. Thanks to the vertically-stacked channel structures, large surface gains (W_{eff}/W_{Top}) have been achieved as shown in Table 2.1. Superior on-state currents per surface unit can be achieved in those devices.

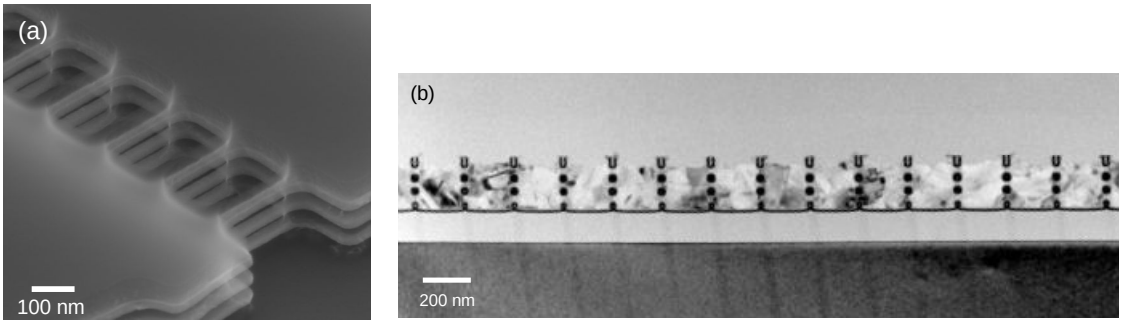


Figure 2.5 Fabricated vertically-stacked silicon nanowires: (a) a top-view SEM image, (b) a cross-section TEM image.

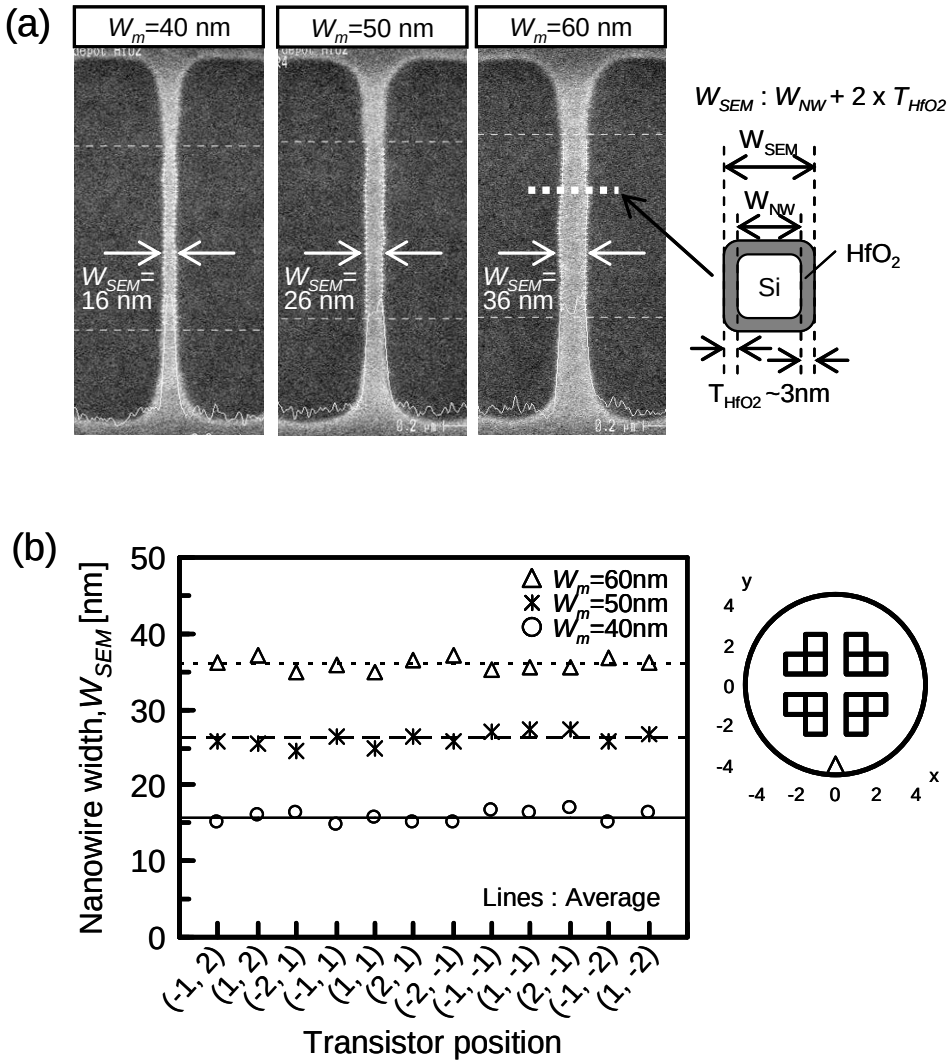


Figure 2.6 (a) Top-view SEM images of silicon nanowires after HfO_2 deposition with width W_{SEM}= 16, 26, and 36 nm. (b) Variation of nanowire width in a 200-nm-wafer. The variations are less than +/- 1.5 nm. The thickness of HfO_2 on side walls (3nm x 2) is included in the values of W_{SEM}. Wm is the mask width.

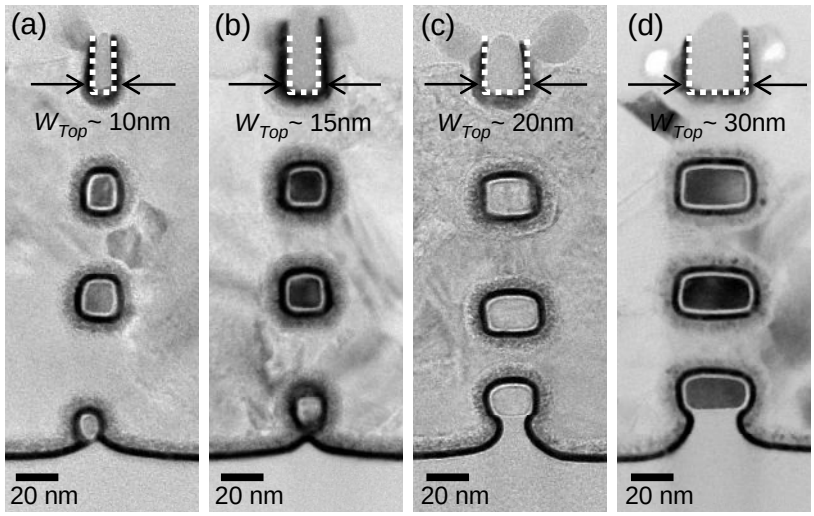


Figure 2.7 Cross-sectional TEM images of vertically-stacked silicon nanowire MOSFET with top-view width W_{Top}= 10, 15, 20 and 30 nm.

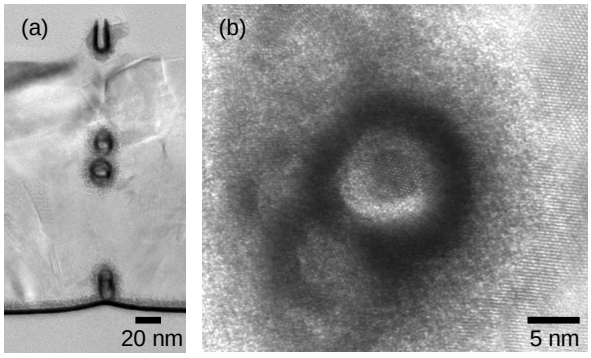
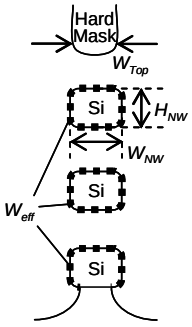


Figure 2.8 (a) Cross-sectional TEM image of vertically-stacked silicon nanowire MOSFET with top-view width W_{Top}= 5 nm. (b) Enlarged image of 5-nm-diameter nanowire.

Table 2.1 Nanowire width with various definitions and surface gain factor W_{eff}/W_{Top}.

W_m [nm]	40	45	50	60
W_{Top} [nm]	10	15	20	30
Max. W_{NW}/H_{NW} [nm]	11/14	15/14	21/15	32/15
W_{eff} [nm]	102	130	171	227
W_{eff}/W_{Top}	10.2	8.7	8.5	7.6



The NWs are [110]-oriented and horizontally arrayed with 50 or 10 parallel wires. The physical wire lengths (L_{NW}) are in the 42 – 607 nm range as shown in Figure 2.9. Effective gate length (L_{eff}) and source/drain resistance (R_{SD}) were extracted thanks to the Y-function-based technique as shown in Chapter 3. Differences between L_{NW} and L_{eff} were less than 10 nm. This means that the source/drain implantation and activation annealing are well-controlled. The resulting R_{SD} are 159 Ω.µm for NMOS and 161 Ω.µm for PMOS.

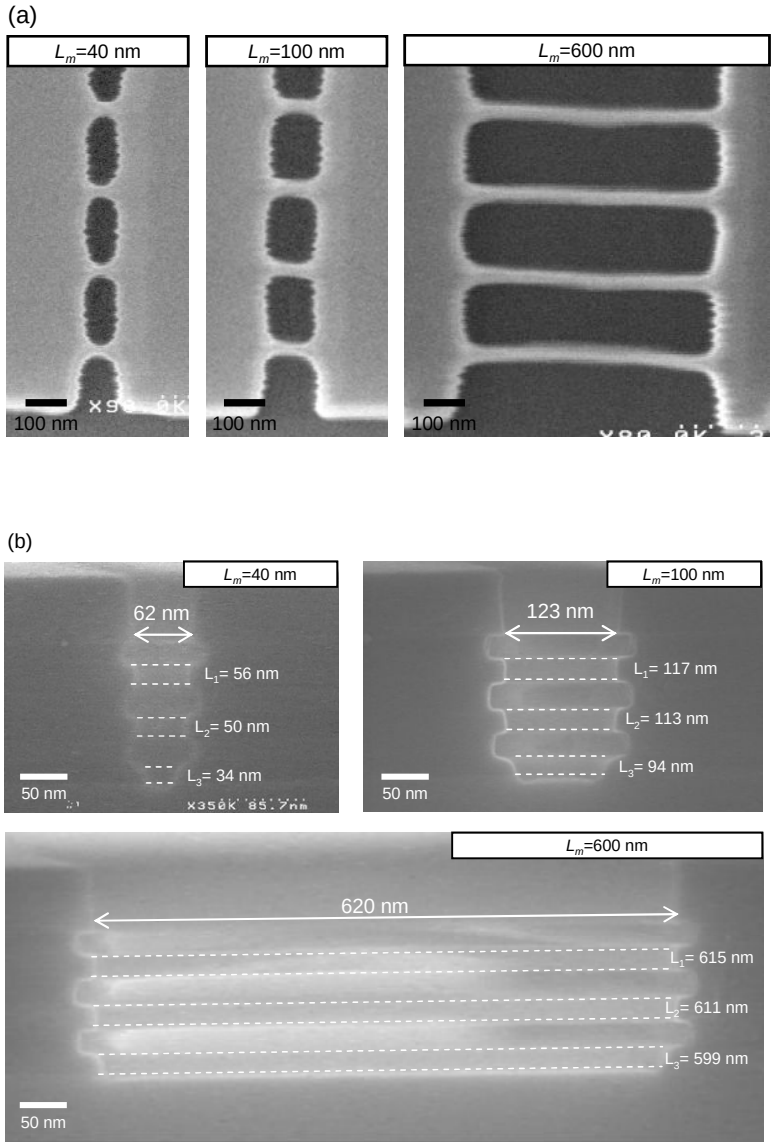


Figure 2.9 Top-view (a) and cross-section (b) of SEM images of vertically-stacked silicon nanowire MOSFET with mask length L_m= 40, 100, and 600 nm.

2.4 KEY STEPS

In the processes of two type of vertically-stacked channel MOSFETs, the following technologies were commonly used.

2.4.1 SiGe Epitaxy and Etching

■ **$Si_{0.8}Ge_{0.2}$/Si superlattice epitaxial growth**

$Si_{0.8}Ge_{0.2}$/Si epitaxial growth is processed in the Epi Centura Reduced-Pressure Chemical Vapor Deposition (RP-CVD) industrial cluster tool [2.7]. The Si layers are grown at 700°C and the SiGe ones at 650°C. Such a low growth temperature enables to grow quite thick layers without any elastic relaxation of the strain through the formation of surface undulations [2.1]. Pure dichlorosilane (SiH_2Cl_2) is used as the source of Si and germane (GeH_4) diluted at 2% in H_2 as the source of Ge. The number of vertically aligned channels is determined by that of the grown SiGe/Si layers and is practically limited by the possible superlattice relaxation due to the compressively strained SiGe layers.

■ **Anisotropic plasma etching of $Si_{0.8}Ge_{0.2}$/Si superlattice**

Hybrid lithography with Deep UV (248nm wavelength) (equipment: ASM 300) and ebeam (equipment: LEICA VB6HR) was used to define the initial photo-resist patterns. The whole stack with Si_3N_4/SiO_2 Hard Mask (HM) is etched in an applied material reactor (Applied Material Centura tool). The etching process is divided into five steps as seen in Table 2.1.

Table 2.2 Process description of anisotropic etching of SiGe/Si superlattice

Step description	Plasma chemistry	Step type
Resist trimming (optional)	Cl_2/O_2	Adjusted Time
Resist cure (optional)	HBr	Time= 60s
Oxide hard mask etch	CF_4/CH_2F_2/He	Time=15s
Hard Mask etch	$CF_4/CH_2F_2/O_2$/He	Endpoint
Si/SiGe multilayer main etch	Cl_2/HBr/O_2	Endpoint + over etch

2.4.2 High-*k*/Metal Gate Stacks

High-κ dielectrics and metal gate are strongly required to achieve small EOT with low gate leakage current density. TiN/HfO_2 has been chosen as a gate stack for both n- and p-MOS. 3 nm of HfO_2 and 10 nm of TiN are deposited in the cavities by uniform deposition methods of Atomic Layer Deposition (ALD) and Chemical Vapor Deposition (CVD), respectively. The thick N^+ polysilicon layers (CVD) are used to fill in the cavities and connect the vertically-stacked gates. Figure 2.10 shows cross-sectional TEM image of silicon nanowire with 3nm-thick-HfO_2 and 10nm-thick-TiN gate stack. A SiO_2-like interfacial layer was observed. This layer is grown by the thermal process after the gate deposition. The resulting equivalent oxide thickness (EOT) is ~1.7 nm.

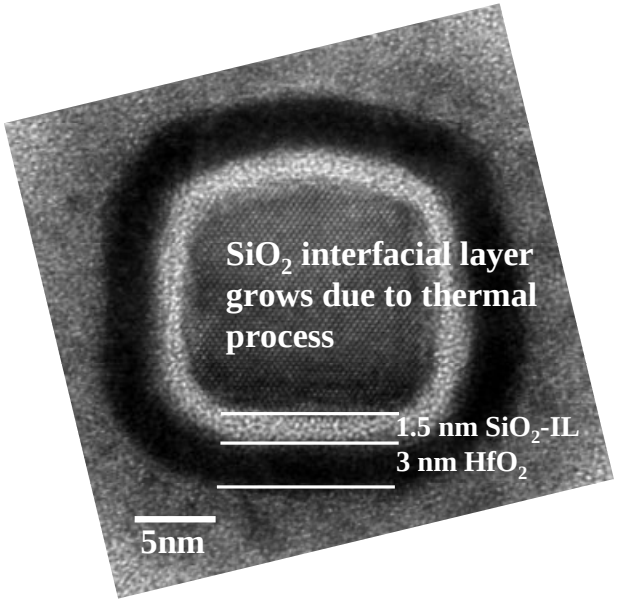


Figure 2.10 Cross-sectional TEM image of vertically-stacked silicon nanowire with high-κ/metal gate.

2.5 CONCLUSIONS

In this chapter, the fabrication process of the vertically-stacked channel MOSFETs were described. The use of sacrificial SiGe layers enables silicon channel to be piled up. The gate stacks were uniformly surrounded owing to ALD and CVD process.

2.6 REFERENCES

[2.1] A.J. Pidduck, D.J. Robbins, A.G. Cullis, W.Y. Leong, A.M. Pitt, "Evolution of surface morphology and strain during SiGe epitaxy", *Thin Solid Films*, Vol. 222, Issue 1-2, pp. 78-84, 1992.

[2.2] E. Bernard, T. Ernst, B. Guillaumot, N. Vulliet, T.C. Lim, O. Rozeau, F. Danneville, T. Skotnicki, S. Deleonibus, O. Faynot, "Novel internal spacers 50nm gate length Multi-Channel MOSFET (MCFET) with TiN/HfO2 gate stack", *IEEE Electron Device Letters*, 2009.

[2.3] E. Bernard, T. Ernst, B. Guillaumot, N. Vulliet, P. Coronel, T. Skotnicki, S. Deleonibus, O. Faynot, "Multi-Channel Field Effect Transistors (MCFET) - Part I: Electrical performances and current gain analysis", to be published in *IEEE Transactions on Electron Devices*, 2009.

[2.4] E. Bernard, T. Ernst, B. Guillaumot, N. Vulliet, P. Coronel, T. Skotnicki, S. Deleonibus, O. Faynot, "Multi-Channel MOSFET (MCFET) structures - Part II : Analysis of the gate stack and the series resistances influence on the performances", *IEEE Transactions on Electronic Devices*, 2009.

[2.5] C. Dupré, A. Hubert, S. Bécu, M. Jublot, V. Maffini-Alvaro, C. Vizioz, F. Aussenac, C. Arvet, S. Barnola, J.-M. Hartmann, G. Garnier, F. Allain, J.-P. Colonna, M. Rivoire, L. Baud, S. Pauliac, V. Loup, T. Chevolleau, P. Rivallin, B. Guillaumot, G. Ghibaudo, O. Faynot, T. Ernst, and S. Deleonibus, "15nm-diameter 3D Stacked Nanowires with Independent Gate Operation: ΦFET," in *IEDM Tech. Dig.*, 2008, pp. 749–752.

[2.6] K. Tachi, M. Casse, D. Jang, C. Dupré, A. Hubert, N. Vulliet, V. Maffini-Alvaro, C. Vizioz, C. Carabasse, V. Delaye, J. M. Hartmann, G. Ghibaudo, H. Iwai, S. Cristoloveanu, O. Faynot, and T. Ernst, "Relationship between mobility and high-*k* interface properties in advanced Si and SiGe nanowires," in *IEDM Tech. Dig.*, pp.313–316, 2009, Baltimore, US.

[2.7] J.M. Hartmann, M. Py, P.H. Morel, T. Ernst, N. Vulliet, B. Previtali and J.P. Barnes, "Cyclic Deposition / Etch processes for the formation of Si raised sources and

drains in advanced MOSFETs," *ECS Trans.*, Vol. 33, No. 6, pp. 391–407, 2010.

CHAPTER 3

ELECTRICAL CHARACTRIZATION METHODS

CHAPTER 3 CONTENTS

3.1 INTRODUCTION

In order to analyze the performances of the MCFET architecture, it was necessary to extract its intrinsic parameters such as the threshold voltage, the carrier mobility or the series resistances. In the literature, a lot of extraction methods have been proposed: “Shift&Ratio”, “Mc Larty”, “Hamer”, “Y function” or “split-CV” (only for mobility extraction). They are all based on the exploitation of the I_D-V_G characteristics. For our extractions, we preferentially used the Y function [3.1] and split-CV methods [3.2, 3.3]. Those latter are detailed in the following.

3.2 Y-FUNCTION METHOD

This method is based on the combined exploitation of the I_D-V_G and g_m-V_G characteristics. The MOSFET parameter extraction is performed within the strong inversion regime of the MOSFET linear region and, therefore, relies on the well known drain current expression given by equation (3.1):

$$I_D = \frac{W}{L}\mu_0 . C_{ox} . \frac{(V_G - V_T).V_D}{1 + \theta_1 .(V_G - V_T)} \tag{3.1}$$

where W and L are the channel width and channel length, V_T is the threshold voltage, C_{ox} is the gate oxide capacitance, μ_0 is the low field mobility and θ_1 is the first mobility reduction coefficient which takes into account the influence of the series resistances (3.2).

$$\theta_1 = \theta_{1,0} + \frac{W}{L}\mu_0 . C_{ox} R_{SD} \tag{3.2}$$

where $\theta_{1,0}$ is the intrinsic mobility reduction factor.

The basic idea consists in constructing a function which is independent of the series resistances R_{SD}, i.e. which does not contain θ_1 (because R_{SD} strongly degrade the drain current especially in short channel devices). This can be achieved by dividing the current expression by the square root of the transconductance (3.3):

$$Y = \frac{I_D}{\sqrt{g_m}} = \sqrt{\beta . V_D} . (V_G - V_T) \tag{3.3}$$

where g_m and β are respectively the transconductance and gain factor of the transistor. Their expressions are given by equations (3.4) and (3.5) respectively:

$$g_m = \left. \frac{\partial I_D}{\partial V_G} \right|_{V_D} = \beta . \frac{1}{[1 + \theta_1 . (V_G - V_T)]^2} \tag{3.4}$$

$$\beta = \frac{W}{L} . \mu_0 . C_{ox} = \frac{{S_Y}^2}{V_D} \tag{3.5}$$

where S_Y is the Y function slope.

In the strong inversion regime, the Y function varies linearly with the gate voltage (Figure 3.1). The threshold voltage can then be obtained by the extrapolation of the linear part of the curve (intercept with the X-axis).

On the other hand, the low field mobility μ_0 can be extracted from the Y function slope as expressed in equation (3.6):

$$\mu_0 = \frac{{S_Y}^2}{C_{ox} . V_D} \frac{L}{W} \tag{3.6}$$

For that, we only need to know the oxide capacitance C_{ox}, which is extracted by capacitance measurements (C_{GC} (V_G)), and the effective gate length and width, L_{eff} and W. According to equation (3.7), the effective gate length can be obtained by plotting $1/\beta$ as a function of the gate length and by taking the value for which $1/\beta$ is null (Figure 3.2).

$$\frac{1}{\beta} = \frac{L - \Delta L}{W . \mu_0 . C_{ox}} \tag{3.7}$$

But this method supposes that the difference ΔL between the effective gate length L_{eff} and the mask gate length L_m is constant, which is not always true. In order to circumvent this problem, another method based on the Shift&Ratio procedure has been set up by *Cretu et al* [1.4]. This latter is based on the variation of the Y function

between a short channel transistor and a long channel one (this method is not explained here). It should be noted that the Y function method allows a separate determination of the threshold voltage and the mobility.

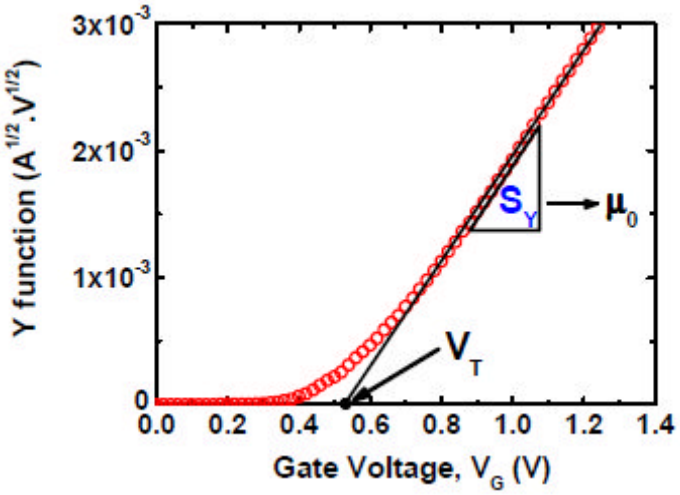


Figure 3.1 Y- function as a function of the gate voltage.

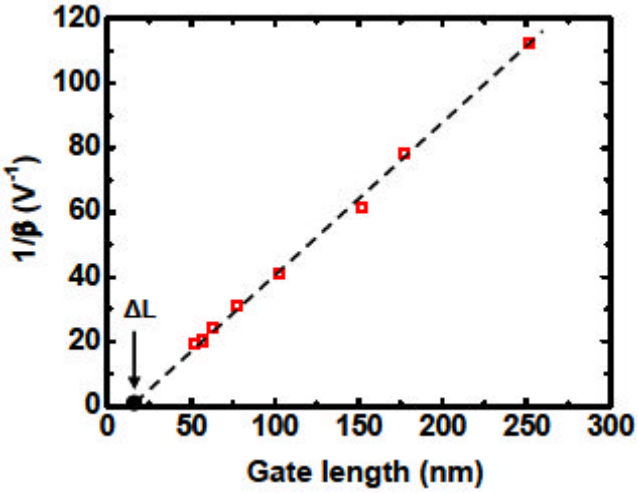


Figure 3.2 Extraction of the effective gate length.

After the extraction of the threshold voltage, we can calculate the mobility reduction factor θ_{eff} defined by equation (3.8):

$$\theta_{eff} = \frac{S_Y^{\,2}}{I_D} - \frac{1}{V_G - V_T} \tag{3.8}$$

where θ_2 is the second mobility reduction coefficient taking into account the surface roughness.

Indeed, at high V_G ($V_G >> V_T$), the mobility is described by two attenuation factors (3.9):

$$\mu_{eff} = \frac{\mu_0}{1+\theta_1.(V_G - V_T)+\theta_2.(V_G - V_T)^2} \tag{3.9}$$

For $V_G >> V_T$, θ_{eff} is proportional to the gate voltage as shown by equation (3.10):

$$\theta_{eff}(V_G >> V_T) = \theta_1 + \theta_2.(V_G - V_T) \tag{3.10}$$

So, by plotting θ_{eff} as a function of the gate voltage, we can extract θ_1 (intercept with the Y-axis) and θ_2 (slope). Note that a low value for θ_2 indicates a good Si/SiO_2 interface (low surface roughness). From θ_1, we can finally extract the series resistances. To do this, θ_1 is plotted as a function of β. The R_{SD} are then extracted from the slope of this curve (Figure 3.3).

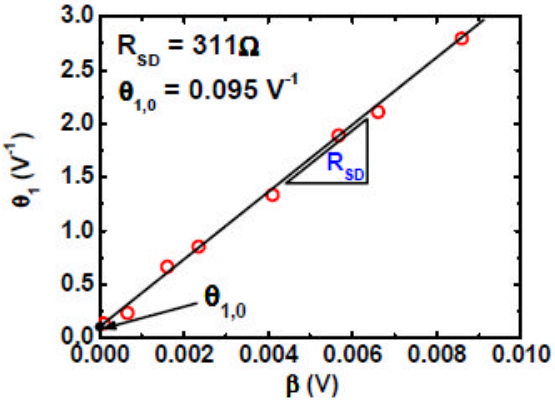


Figure 3.3 Extraction of the series resistances for 40nm to 600nm gate lengths MCFET devices (W = 500nm).

The accuracy of the extraction can be improved by introducing the second mobility reduction coefficient θ_2 in the Y function (3.11) and by recalculating all the parameters.

$$Y_{new} = Y \cdot \sqrt{1-\theta_2 \cdot (V_G - V_T)^2} \tag{3.11}$$

This correction is only necessary if $\theta_2 \approx 1/(V_G - V_T)^2$.

At the end of the extraction, we can compare the theoretical linear drain current calculated by using the extracted parameters and the experimental one. For that, we use the expression of the linear current given by equation (3.12).

$$I_D = \frac{W}{L_{eff}} \mu_0 . C_{ox} . \frac{(V_G - V_T).V_D}{1 + \theta_1 . (V_G - V_T) + \theta_2 . (V_G - V_T)^2} \tag{3.12}$$

Figure 3.4 shows the comparison the calculated curves with the measured data. As we can see, an excellent fit is obtained between the model and the experiments. The model is then well adapted even for short channel MOSFET devices (L_G < 100nm).

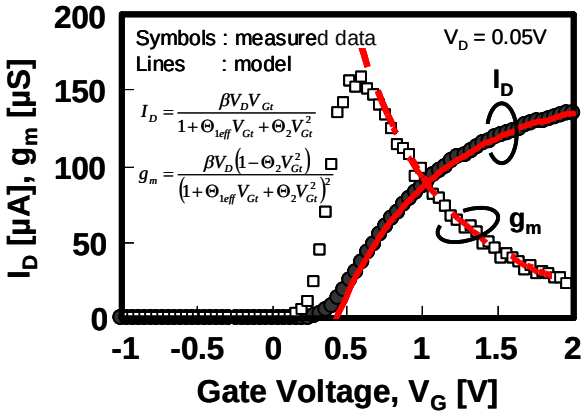


Figure 3.4 Measured and modeled I_{Dlin}-V_G and g_{mlin}-V_G characteristics of TiN/HfO$_2$ n-MCFETs. Gate length and width are 70 nm and 350 nm, respectively.

3.3 SPLIT *C–V* METHOD

This method is based on the combination of two capacitance measurements and one current-voltage measurement in order to obtain the effective mobility μ_{eff} as a function of the inversion charge Q_{inv} or the effective electric field E_{eff}.

In a first time, the gate-to-channel capacitance C_{GC} is measured for different gate voltages. The inversion charge Q_{inv} at a given gate voltage V_G is then obtained by integrating this capacitance until the whished gate voltage (3.13).

$$Q_{inv}(V_G) = \frac{1}{L_{eff} \cdot W_{eff}} \cdot \int_{-\infty}^{V_0} C_{GC}(V_G).dV_G \tag{3.13}$$

In a second time, the gate-to-substrate capacitance C_{GB} is measured and integrated so as to obtain the depletion charge (3.14).

$$Q_{dep}(V_G) = \frac{1}{L_{eff} \cdot W_{eff}} \cdot \int_{V_{fb}}^{V_G} C_{GB}(V_G).dV_G \tag{3.14}$$

The effective field is then calculated according to equation (2.15).

$$E_{eff} = \frac{\eta \cdot Q_{inv} + Q_{dep}}{\varepsilon_{Si}} \tag{3.15}$$

where η is an empiric parameter equal to 1/2 for electrons and to 1/3 for holes.

Note that on SOI substrate C_{GB} measurement cannot be performed due to the presence of the BOX. But for ultra-thin films transistors on SOI with undoped channels and thick BOX (which is the case of our transistors), the depletion charge is negligible compared to the inversion charge and the effective field is solely determined from the inversion charge.

Finally, the effective mobility is obtained by dividing the drain current by the inversion charge (3.16):

$$\mu_{eff} = \frac{L_{eff}}{W_{eff}} \frac{I_D}{Q_{inv}.V_D} = \frac{I_D.L_{eff}^{\ 2}}{Q_{invS}.V_D} \tag{3.16}$$

where $Q_{invS} = Q_{inv}.W_{eff}.L_{eff}$ is the measured inversion charge. As a consequence, the effective mobility is only dependant of the effective gate length.

The split C-V method is applicable to short channel devices [3.3], only if: the drain current is corrected by the series resistances, the capacitances are corrected from the parasitic capacitances and the effective gate length is correctly extracted.

The correction of the drain current by the R_{SD} is given by equation (3.17):

$$I_{D0} = \frac{I_D}{1 - \frac{R_S.I_D}{V_D}} \tag{3.17}$$

where I_D is the measured current and I_{D0} the corrected current.

On the other hand, L_{eff} is extracted thanks to the measurement of C_{GC} for a short channel device with a gate length L_{eff_short} and for a long channel device with a gate

length L_{eff_long} (considered equal to the mask gate length). The effective gate length of the short channel device is then calculated according to equation (3.18).

$$L_{eff\_short} = \frac{C_{inv0}^{short}}{C_{inv0}^{long}} \cdot L_{eff\_long} \tag{3.18}$$

where $C_{inv0}=C_{GC}- C_p$ is the maximum of the gate-to-channel capacitance corrected by the parasitic capacitances C_p (overlap and fringing capacitances). Finally, the inversion charge is calculated by using C_{GC0} (instead of C_{CG}) which is the gate-to-channel capacitance corrected from the parasitic capacitances.

In order to accurately evaluate the transport properties, the effective mobility was extracted by split C-V technique with parasitic capacitance and R_{SD} corrections. Figure 3.5 shows the validity of the extraction. A good agreement when compared to other techniques was achieved.

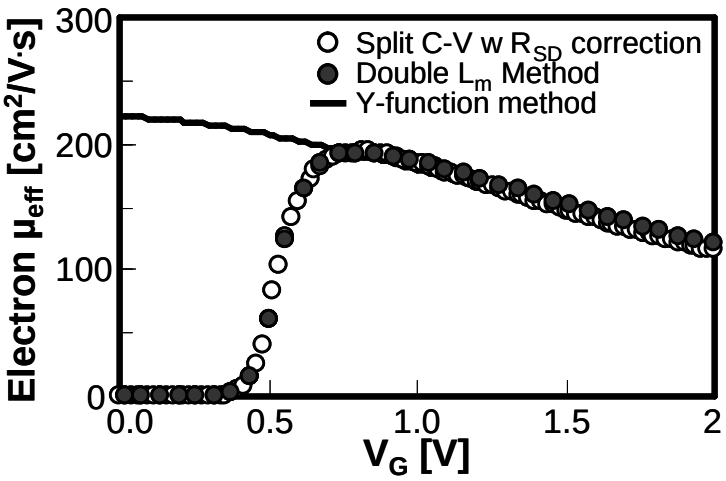


Figure 3.5 Comparison of effective mobility extracted by split C-V, double L_m method, and from parameters extracted by Y-function method. The measured device is the stacked SNWTs with W_{NW}=15 nm and L_{eff}=242 nm. The device with L_{eff}=592 nm was also used for double L_m method.

3.4 CONCLUSIONS

In this chapter, the two electrical characterization methods were described. The calculated *I–V* curve by Y-function method was well fitted to the measured data. This enables to precisely extract device parameters. Finally, good agreement of effective mobility when compared between Y-function method and split C-V technique was achieved.

3.5 REFERENCES

[3.1] G. Ghibaudo, "New Method for the Extraction of MOSFET Parameters", *Electronics Letters*, Vol. 24, Issue 9, pp. 543-545, 1988.

[3.2] J. Koomen, "Investigation of the MOST channel conductance in weak inversion", *Solid-State Electronics*, Vol. 16, pp.801-810, 1973.

[3.3] K. Romanjek, F. Andrieu, T. Ernst, G. Ghibaudo, "Improved Split C-V Method for Effective Mobility Extraction in sub-0.1-Rm Si MOSFETs", *IEEE Electron Device Letters*, Vol. 25, Issue 8, pp. 583-585, 2004.

CHAPTER 4

SOURCE/DRAIN DOPING TEQUNIQUES FOR VERTICALLY-STACKED CHANNEL STRUCTURE

CHAPTER 4 CONTENTS

4.1 INTRODUCTION

The reduction of parasitic access resistances serially connected at source and drain region of MOSFETS are one of the challenging technology to meet the performance required in the roadmap as discussed in Chapter 1. The voltage drops at the resistances reduces the applied voltages at the drain and gate electrodes in the transistors, resulting in the decrease in the overdrive voltage to lower the on-current. These resistances also cause degradation in the time constant, commonly refereed as RC delay, to lower the switching speed. Therefore, as the channel resistance reduces with the scaling in the gate length, the parasitic series resistance should be further reduced not to increase its proportion in the total resistance at on-state.

MOSFETs with 3 dimensional channels (e.g. Fin FETs) severely suffer from the parasitic access resistance with scaling, as the cross-sectional area at the entrance from the source to channel becomes. Therefore, novel processes dedicated for 3 dimensional FETs should be implemented; to modify the shape of the source and drain region to reduce the current density or to further reduce the resistivity of the wiring. One example of the former process is selective epitaxy growth process to elevate the source and drain region to reduce the current density. Since the growth is selectively done only on Si surfaces, the source and drain regions can be modified without any short circuit to other electrodes. One of the concerns is the excess growth of the source and drain regions results in the increase in the gate to drain capacitances, to lower the switching speed. The later process includes recoil ion-implantation, plasma doping and refractory metal silicide technologies; those are aggressively under research.

As with vertically stacked multi channel devices [4.1–4.3], the later approach to reduce the resistivity of the source and drain region should be done as the source and drain regions are already grown to access to the channels. Therefore, a further increase in the doping concentration with higher activation ratio at source and drain regions are mandatory. In terms of the gate to drain capacitances, the use of oxide spacers with lower k-values between gate and drain regions are reported to be effective and a reduction in the intrinsic *CV/I* delay by 39 % has been achieved [4.1].

In this chapter, a novel process to decrease the resistivity of the source and drain regions of MCFET is presented using *in situ* doped selective epitaxial growth in combination with conventional ion-implantation (I/I) as a novel process for source and drain formation. The effect of source and drain formation process on the series resistance and carrier mobility will be discussed through electrical characteristics.

4.2 EXPERIMENTAL CONDITIONS

Three types of MCFETs with different S/D doping schemes, listed in Table 4.1, were investigated. For control MCFET samples, arsenic ions were implanted to the un-doped selective epitaxial growth (SEG) S/D (with a thickness of 200 nm) with a dose of 10^{15} cm^{-2} of at 50 keV for *n*-MCFETs. For *p*-MCFETS, boron fluoride ions were implanted with a dose of 10^{15} cm^{-2} at 40 keV. For sample A, the source and drain regions were *in situ* doped during the SEG step. The dopant atoms (~2 x 10^{19} cm^{-3}) for the *in situ* doped SEG were phosphorus and boron for *n*- and *p*-MCFETs, respectively. For sample B, in situ doped SEG was combined with I/I.

Table 4. 1 Doping scheme

	Un-doped SEG	*In situ* doped SEG	Ion implantation (I/I)
Control	✓		✓
Process A		✓	
Process B		✓	✓

4.3 ELECTRICAL CHARACTERISTICS

4.3.1 *I–V* Characteristics

Drive current (at $|V_G\text{-}V_T|$=0.9 V and $|V_D|$=1.2 V) against standby current (at $|V_G\text{-}V_T|$=0.3 V and $|V_D|$=1.2 V) for MCFETs with a L_G = 70 nm is shown in Figure 4.1. The currents were normalized by a top-view width. A large enhancement in the drive

current with process A and B respectively, can be obtained with *in situ* doped SEG process compared to the control MCFET. This improvement applies for both *n*- and *p*-MCFETs. The mean drive currents of 2.4 and 1.2 mA/μm are achieved for *n*- and *p*-MCFETs with process B, respectively. These high current densities are due to the in situ doped SEG combined with I/I and 3-D configuration of vertically-stacked channels.

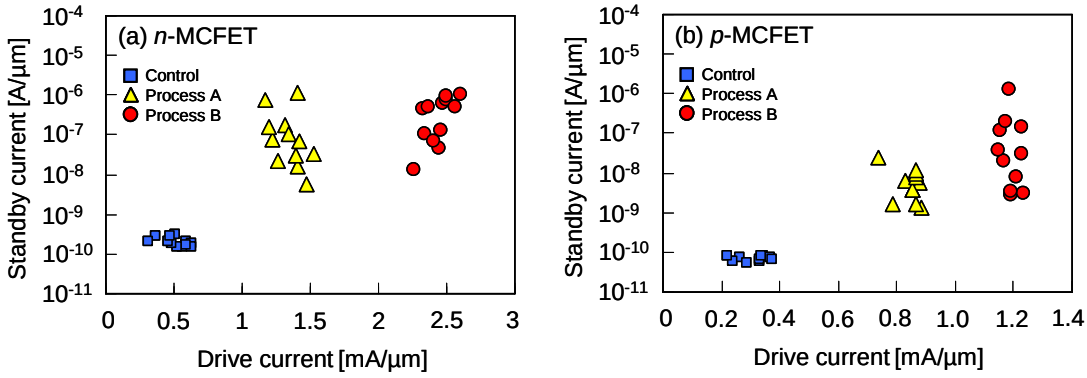


Figure 4.1 On-off relations of (a) n- and (b) p-MCFETs.

Although the variability in the drive current for n- and p-MCFET showed little dependence within the same process conditions, the standby currents tend to scatter and the mean value slightly increases with process A and B. The reason of the variability increase in the standby current might be originated from the residual defect in the *in situ* doped SEG. These defects, commonly observed in highly doped silicon, produce generation centers which degrade the on/off junction property.

The threshold voltage dependence on the L_G is shown in Figure 4.2. Enhanced roll-off properties were observed with sample A and B, where in situ doped SEG was adopted. This suggests the influence of the dopant diffusion to lower the abruptness of the junction.

The dependence of the on-current per a unit width (μm) on the L_G, for both n- and p-MCFETs, is shown in Figure 4.3. Here, the total width (W_{TOT}) of the channels was used for the normalization. The I_{ON} (L_G) dependence clearly demonstrates the scaling achieved with *in situ* SEG process compared to those with the control MCFETs. Especially one can see the merits of process B, which enables further enhancement in

the I_{ON} when L_G is scaled down below 100 nm.

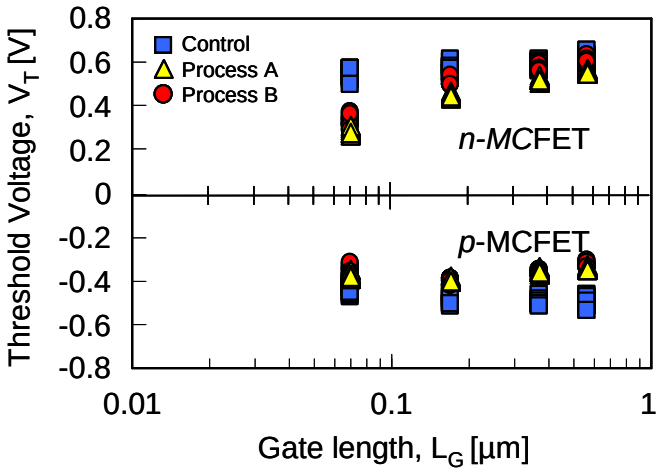


Figure 4.2 Threshold voltage roll-off characteristics of MCFETs with L_G scaling.

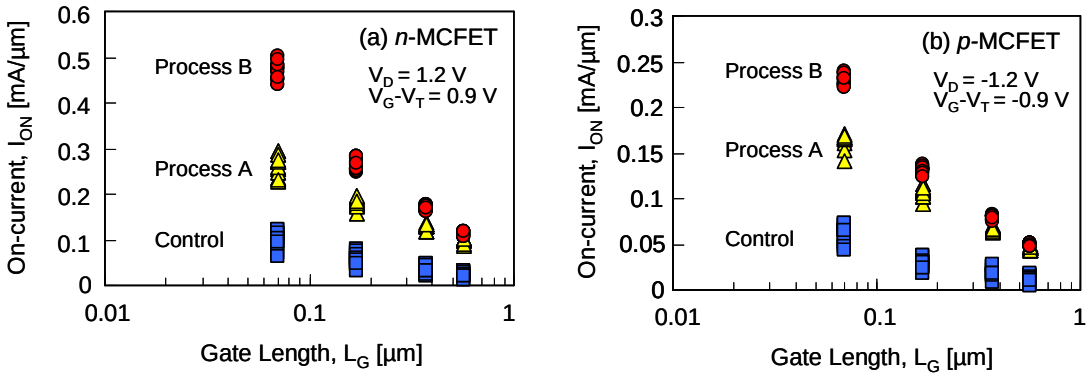


Figure 4.3 On-current dependency on the gate length for (a) n- and (b) p- MCFETs.

4.3.2 Source/Drain Series Resistance Evaluation

The access resistance values are evaluated through electrical measurements of the transistors by Y-function based methods (see Chapter 3). Figure 4.4 shows typical modeling results of the transconducance of multi channel FETs with different L_G. One can confirm a fairly nice modeling with the above equations.

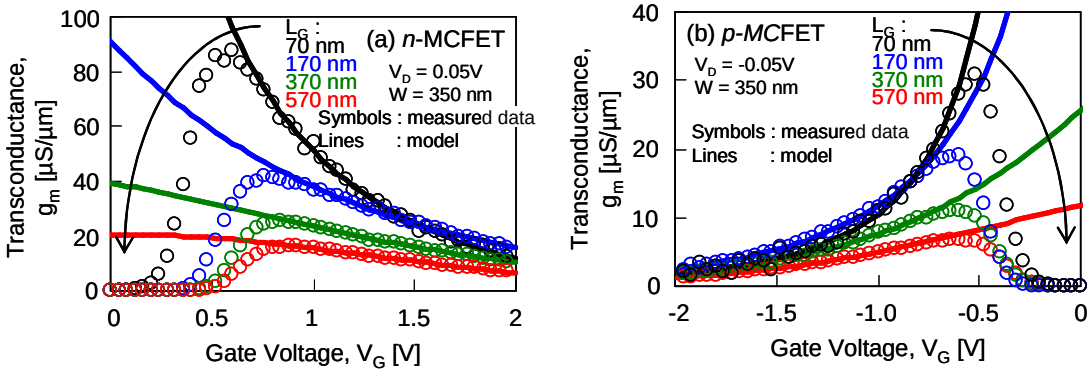


Figure 4.4 Transconductance of the (a) n- and (b) p-MCFETs. Solid lines represent the fitted model.

R_{SD} can be extracted from the slope in the relation between Θ_{1eff} and β as is shown in Figure 4.5. The good linearity in Θ_{1eff} vs. β curves, for all the L_G down to 70 nm, indicates that R_{SD} is identical among different L_G. The extracted values of R_{SD} normalized for a unit channel width (μm), for the control sample and MCFETs with process A and B were 4.5, 3.6 and 0.4 kΩ-μm for *n*-MCFET and 5.2, 3.2 and 1.2 kΩ-μm for *p*-MCFET, respectively. Note the remarkable R_{SD} reduction which has been successfully obtained by combining in situ doped SEG with I/I for both n- and p-MCFETs (process B). The relatively large R_{SD} in the control sample may be attributed to un-optimized doping profile in un-doped SEG, so that ion implantation with multiple acceleration energy appears to be necessary.

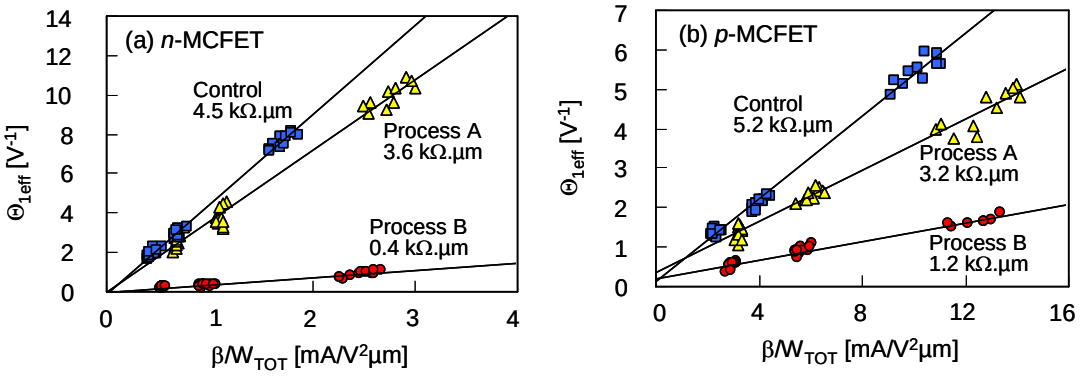


Figure 4.5 Θ_{1eff} vs. β curves of the fabricated (a) n- and (b) p-MCFETs.

On the other hand, *in situ* doped SEG process facilitates and improves the processing of vertically aligned MCFETs S/D-channel junctions. The additional I/I into in situ doped SEG further reduces the R_{SD} by 90 % and 60% for n- and p-MCFET, respectively. The origin of the reduction is still unclear, however, for n-MCFET the better positioning and uniformity of the doped S/D and the high solubility of implanted arsenic atoms in silicon compared to that of phosphorus atoms may increase the activation rate, leading to a decrease in R_{SD}.

4.3.3 Carrier Mobility Evaluation

Besides the R_{SD} values, the effect of junction formation process affects the effective mobility μ_{eff} in the channel, in particular when neutral defects are formed during I/I steps. A degraded μ_{eff} is commonly observed for small L_G. We here investigate the μ_{eff} of the MCFETs and compared the influence of the junction formation process.

The μ_{eff} was extracted by a modified split C-V method in order to exclude the parasitic capacitance, which may become dominant in scaled L_G. Figure 4.6 shows examples of gate-to-channel capacitances C_{GC} (V_G) curves of the MCFETs with process A and B, indicating that the parasitic capacitances were identical among different L_G, so that the inversion carrier density can be extracted.

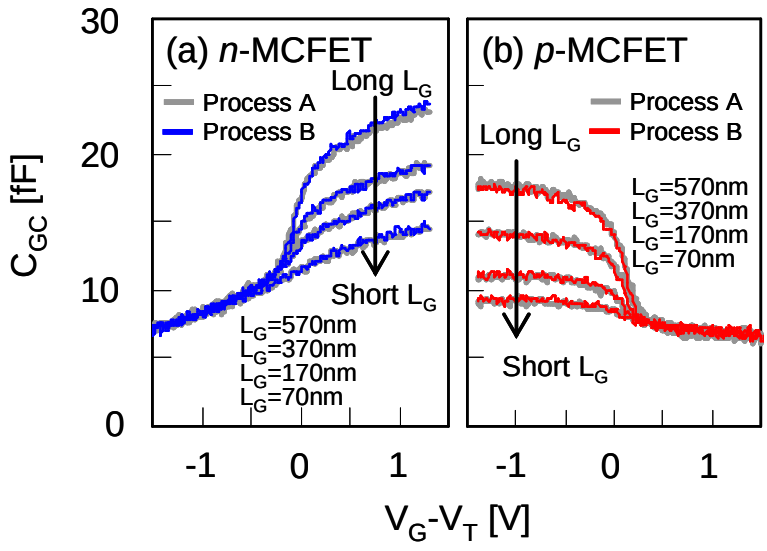


Figure 4.6 C_{GC} characteristics with various L_G for (a) n- and (b) p-MCFETs.

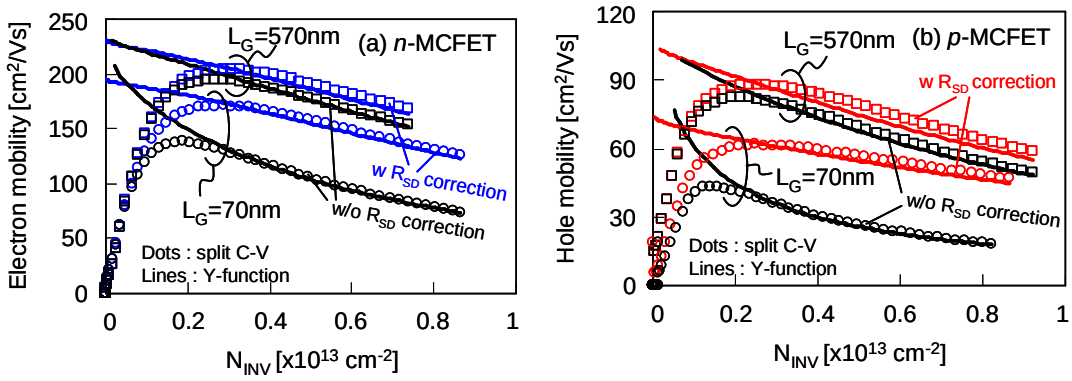


Figure 4.7 μ_{eff} *of the MCFETS with different* L_G *of 570 and 70 nm.*

Additionally, R_{SD} correction for μ_{eff} extraction is necessary as the output conductance is strongly degraded by the parasitic resistance. Figure 4.7 (a) and (b) show the μ_{eff} of *n*- and *p*-MCFETs of L_G=70 and 570 nm with process B with and without RSD correction, respectively. The contribution of R_{SD} on the μ_{eff} is obviously more prominent when the L_G is scaled, as the ratio of the R_{SD} to the total resistance increases. The μ_{eff} calculated by Y-function method agrees well with the μ_{eff} extracted by split C–V method in the region of high carrier density, confirming the correctness of the extracted μ_{eff}. Process A results in slightly higher mobility values, as illustrated in Figure 4.7.

One can observe a degraded μ_{eff} with smaller L_G for both *n*- and *p*-MCFETs even after R_{SD} correction. This fact implies the existence of another mechanism degrading the μ_{eff} in the direction of channel. Figure 4.8 compares the low-field mobility μ_0 values in MCFETs with process B and C. A distinct degradation in the μ_0 with L_G scaling can be observed for both *n*- and *p*-MCFETs. Using the model proposed by Bidal *et al.*[4.4], a L_G-dependent limiting mobility, $\mu_0(L_G)$, can be postulated as

$$\frac{1}{\mu_0(L)} = \frac{1}{\mu_{\max}} + \frac{\alpha_\mu}{L_G}, \qquad (3.7)$$

where $\mu_{\max}$ and α_μ are the maximum mobility in long channel MCFETs and the mobility degradation factor, respectively. Using this model with the two fitting parameters, μ_0 at further scaled L_G can be predicted as shown in Figure 4.8. Although no μ_0 difference is expected at L_G of 10 nm, the mean μmax showed smaller values with process B compared to those with process A, suggesting the influence of I/I.

On the other hand, the α_μ values were found to be similar in the rage of 0.09~0.10 and 0.22~0.24 nm-Vs/cm^2 for *n*- and *p*-MCFETs, respectively. Therefore, the degraded μ_{eff} in the low carrier density region can be considered to be originated by Coulomb scattering from the dopant atoms diffusion from S/D regions induced by I/I B-type process. Indeed the V_T roll-off characteristics on L_G shown in Figure 4.2 revealed slightly changed characteristics for both *n*- and *p*-MCFETs, supporting the possibility of diffusion of dopants. Moreover, the threading dislocation patterns observed in the TEM images shown in Figure 2.3 suggest an enhanced diffusion of dopant atoms, typically reported as 100 times higher, through the defects which reduces the abruptness of channel and S/D regions [4.5].

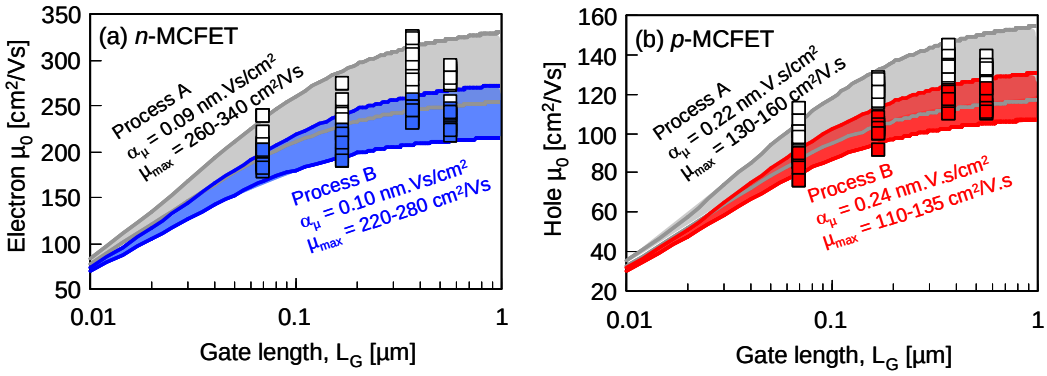


Figure 4.8 Estimated μ_0 on L_G scaling.

The values of the α_μ and μ_{max} for MCFETs are summarized with reported values for planar FETs in Figure 4.9. The MCFETs with in situ SEG process show 2 and 3 times higher values in α_μ than the ballistic limit ($\alpha_{\mu,bal}$), for *n*- and *p*-MCFETs, respectively. This deviation indicates the presence of Coulomb scattering in the channel which can also be inferred from the lower μ_{max} value compared to other devices. The message is that, although in situ doped SEG process is useful for R_{SD} reduction, further process optimization in the dopant activation step is needed in order to suppress the diffusion of dopant atoms in to the channel.

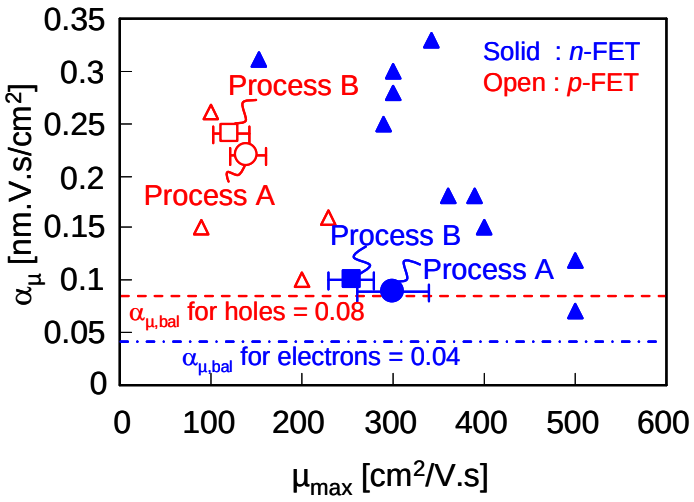


Figure 4.9 Summary of the extracted α_μ and μ_{max}.

4.3.4 Gate Length Scaling

Using the extracted α_μ and μ_{max}, one can estimate the μ_0 of the MCFETs with further scaled L_G. From Figure 4.9, the μ_0 with L_G scaling can be estimated. The μ_0 trend showed further reduction at scaled L_G, even considering the variability. The difference of the μ_0 becomes smaller between the process A and B. This estimation indicates that the electrical influence of the diffused dopants from the source and drain region can be neglected as large portion of ballistic limited mobility dominates at these L_G regions.

Here we examine the global MCFETs down-scaling, including short-channel effects. The I_{OFF} behavior, when normalized by a common threshold voltage V_T, reflects the subthreshold properties such as DIBL and SS (not the V_T roll-off, however). Figure 4.10 shows the I_{ON}-I_{OFF} characteristics of n- and p-MCFET with several L_G and W. The currents were normalized by the total channel surface W_{total} (W_{total} = W x 5ch. + T_{Si} x 6side-ch.). When reducing L_G down to 70 nm, I_{OFF} increases progressively due to the enhanced DIBL, while the SS value remains constant at ~70 mV/decade for nFET and ~75 mV/decade for pFET. MCFETs with L_G smaller than 70 nm have degraded subthreshold properties without the expected I_{ON} enhancement. This kind of degradation can be suppressed by adding the lateral gates electrostatic control through *W* down-scaling [4.6]. Drive current gains of 14 % for nFET and 20 % for pFET were

observed when *W* was reduced from 350 nm down to 100 nm. We measured in the meantime an improved mobility μ_0 of 11 % and 18 % for nFET and pFET, respectively. This suggests that *W* down-scaling may reduce the L_G dependent mobility degradation. Additional investigation is needed in order to obtain a physical explanation of this phenomenon which is compatible with volume inversion.

Lastly, we present the scaled MCFETs characteristics with the process C. Figure 4.11 shows I_D-V_G and I_D-V_D characteristics associated to 50-nm-L_G 80-nm-*W* nMCFET and 40-nm-L_G and 70-nm-*W* pMCFET. We obtained extremely high I_{ON}-currents of 4.1 mA/µm for nFET and a record 2.7 mA/µm for pFET at V_{DD} = 1.2 V. These values are obtained thanks to the 3D configuration of the vertically stacked channels and the enhanced impact of lateral conduction with small gate width. However, I_{OFF} is still high due to the non-optimized threshold voltage on those samples. When normalized at V_{OFF}+V_{DD}, the I_{ON}-currents are 3.3 mA/µm for nFET and 2.0 mA/µm for pFET. When normalized by W_{total}, the I_{ON}-currents at V_{OFF}+V_{DD} for n- and p-FET are 538 µA/µm and 396µA/µm, respectively. These normalized I_{ON} values are comparable to planar fully depleted – SOIFETs when using the same (unoptimized) gate stack [4.7].

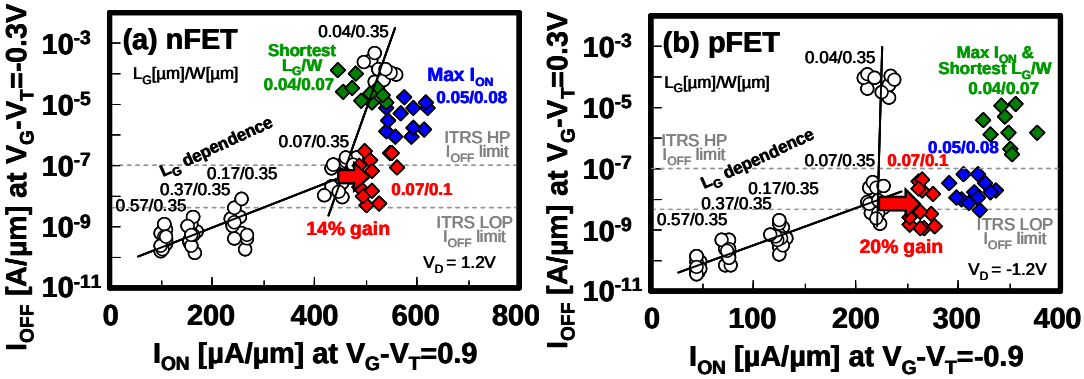


Figure 4.10 I_{ON}-I_{OFF} characteristics with several channel sizes for nMCFET (a) and pFET (b).

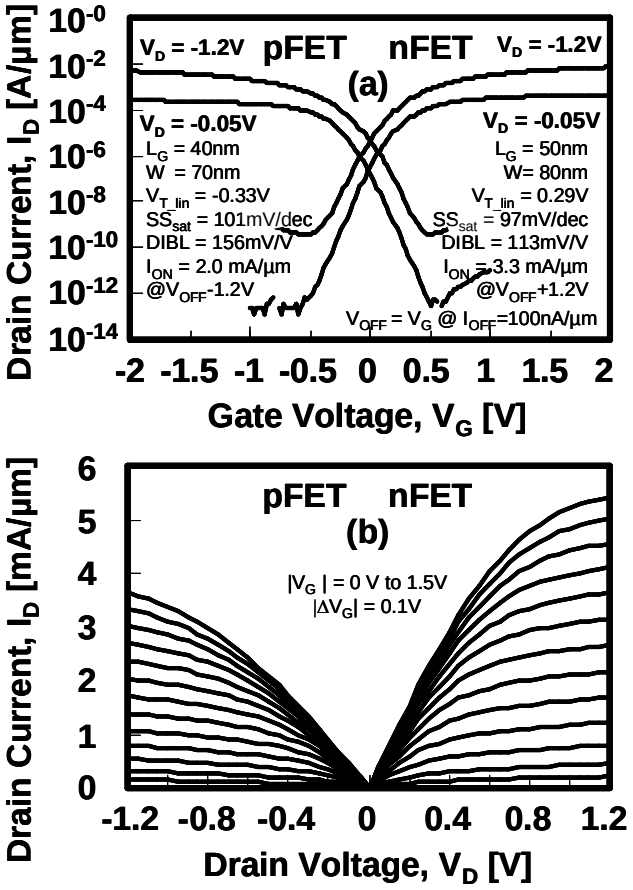


Figure 4.11 I_D-V_G (a) and I_D-V_D (b) characteristics for the scaled MCFETs.

4.4 CONCLUSIONS

The influence of in situ doped SEG source and drain has been examined for vertically aligned MCFETs. A large enhancement, by a factor of 2 in the drive current, can be obtained when in situ doped SEG process is adopted. Detailed parameter extraction from the electrical measurements shows that the R_{SD} values can be reduced by 90 and 75% for *n*- and *p*-MCFETs, respectively, when in situ doped SEG is reinforced by adding ion implantation. On the other hand, V_T roll-off characteristics and the effective mobility behavior are slightly degraded, especially when ion implantation is combined to the SEG process. Mobility analysis has revealed an increase in the Coulomb scattering with L_G scaling, indicating the diffusion of dopant atoms from S/D regions. These results indicate an avenue to further improve the performance by optimizing the S/D activation annealing step.

4.5 REFERENCES

[4.1] E. Bernard T. Ernst, B. Guillaumot, N. Vulliet, T. C. Lim, O. Rozeau. F. Danneville, P. Coronel, T. Skotnicki, S. Deleonibus, O. Faynot, First Internal Spacers' Introduction in Record High I_{ON}/I_{OFF} TiN/HfO_2 Gate Multichannel MOSFET Satisfying Both High-Performance and Low Stanby Power Requirements, *IEEE Elec. Dev. Lett.*, 30 (2) (2009) pp. 148–151.

[4.2] S.-Y. Lee, E.-J. Yoon, D.-S. Shin, S.-M. Kim, S.-D. Suk, M.-S. Kim, D.-W. Kim, D. Park, K. Kim and B.-il Ryu, Sub-25nm Single-Metal Gate CMOS Multi-Bridge-Channel MOSFET (MBCFET) for High Performance and Low Power Application, Tech. Dig. Symp. VLSI Tech. (2005) pp. 154–155.

[4.3] K. Tachi, N. Vulliet, S. Barraud, B. Guillaumot, V. Maffini-Alvaro, C. Vizioz, C. Arvet, Y. Campidelli, P. Gautier, J.M. Hartmann, T. Skotnicki, S. Cristoloveanu, H. Iwai, O. Faynot and T. Ernst, "3D Source/Drain Doping Optimization in Multi-Channel MOSFET," 40th European Solid-State Device Research Conference (ESSDERC), pp. 368, 2010, Spain.

[4.4] G. Bidal, D. Fleury, G. Ghibaudo, F. Boeuf, and T. Skotnicki, Guidelines for MOSFET Device Optimization acounting for L-dependent Mobility Degradation, *IEEE Silicon Nanoelectronics Workshop*, (2009) pp. 5–6.

[4.5] K. Inoue, F. Yano, A. Nishida, T. Tsunomura, T. Toyama, Y. Nagai and M. Hasegawa, Three dimensional characterization of dopant distribution in polycrystalline silicon by laser-assisted atom probe, Appl. Phys. Lett. 93 (2008) 133507.

[4.6] E. Bernard *et al.*, "Novel integration process and performances analysis of Low STandby Power (LSTP) 3D multi-channel CMOSFET (MCFET) on SOI with metal / high-K gate stack," in *VLSI Symp. Tech. Dig.*, 2008, pp. 16-17, 17–19.

[4.7] F. Andrieu *et al.*, "Comparative scalability og PVD and CVD TiN on HfO_2 as a metal gate stack for FDSOI cMOSFET down to 25 nm gate length and width," in *IEDM Tech. Dig., 2006*, pp. 641–644.

CHAPTER 5

CARRIER TRANSPORT PROPERTIES OF VERTICALLY- STACKED NANOWIRE MOSFETS

CHAPTER 5 CONTENTS

5.1 INTRODUCTION

To achieve devices with both high speed and low power consumption for future LSI applications, GAA SNWTs are one of the promising candidates because of their strong short-channel effect immunity. Moreover, to increase the drive current per unit area with the higher density for integration, vertical stacking of NWs enables the use more available silicon surface per device. Recently, short channel GAA-SNWTs have been successfully fabricated with diameter of less than 10 nm using several top-down CMOS compatible processes; they successfully suppress the short-channel effects. On the other hand, transport property degradation in SNWTs was also reported by several groups. However, the mobility behavior when the width is reduced has been remained unclear. Carrier transport in SNW is commonly discussed in terms of two main mechanisms; one is one-dimension (1-D) transport model, and the other is a facet-dominated transport model. The former can be adapted to sub-10 nm diameter, and the latter to larger one. From the fabrication process viewpoints, (dispersion, yield rate), a larger diameter is production friendly, provided that the short-channel effects under aggressively scaled gate length are suppressed.

In this chapter, carrier transport limiting components for vertically-stacked nanowire MOSFETs will be discussed to obtain better performance with suppressing short channel effects.

5.2 ELECTRICAL CHARACTERISTICS

5.2.1 *I–V* Characteristics

The measured I_D-V_D (Figure 5.1) and I_{DS}-V_{GS} (Figure 5.2) characteristics for vertically-stacked 15 nm width SNWTs with 32 nm effective gate length (L_{eff}) for NMOS and 42 nm for PMOS show well-behaved characteristics. I_D-V_G curves exhibit an excellent subthreshold slope (64 mV/dec for NMOS and 74 mV/dec for PMOS) and very low Drain Induced Barrier Lowering (32mV/V for NMOS and 62 mV/V for PMOS). On-currents I_{ON} (normalized by total circumference) of 840 μA/μm and 540

μA/μm with I_{OFF} of 4 nA/μm and 96 nA/μm are obtained for NMOS and PMOS, respectively. Comparable results were obtained in FDSOI FET. When the currents are normalized by top-view width, the I_{ON} for NMOS is 7.2 mA/μm for NMOS and 4.7 mA/μm for PMOS, showing the interest of 3-D devices to increase current density for a given layout. These extremely high currents are due to the vertically stacked structure. The devices showed excellent SCEs immunity as seen in Figure 5.3–5.6. Figure 5.7 shows drain currents as a function of L_{eff}. Gate length scaling is still effective down to sub-50nm L_{eff}.

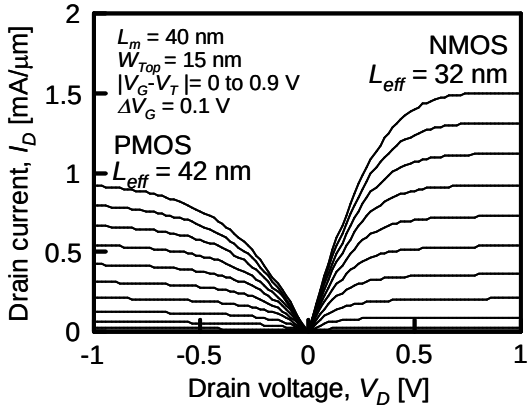


Figure 5.1 I_D-V_D characteristics of vertically-stacked silicon nanowire n- and p-MOSFET with sub-50-nm-L_{eff} and 15-nm-W_{Top}.

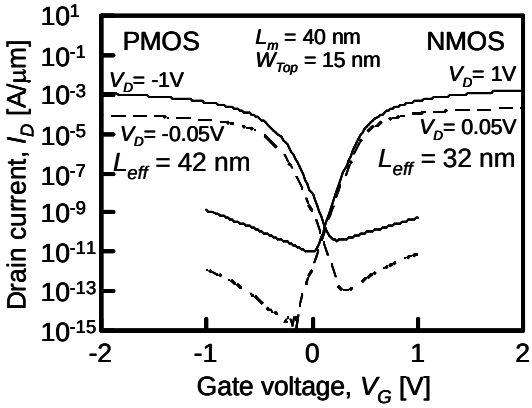


Figure 5.2 I_D-V_G characteristics of vertically-stacked silicon nanowire n- and p-MOSFET with sub-50-nm-L_{eff} and 15-nm-W_{Top}.

Table 5.1 Device parameters for vertically-stacked silicon nanowire n- and p-MOSFET with sub-50-nm-L_{eff} and 15-nm-W_{Top}. The on-currents I_{ON} are extracted at V_G-V_T = 0.7 and -0.7 V for n- and p-MOSFETs, respectively. The off-currents I_{OFF} are extracted at V_G-V_T = -0.3 and 0.3 V for n- and p-MOSFETs, respectively.

	NMOS	PMOS
NW cross-section	3-level-stacking	
Max W_{NW}/H_{NW}	15nm/14nm	
L_m [nm]	40	
L_{eff} [nm]	32	42
EOT [nm]	1.7	1.7
V_{DD} [V]	1	1
I_{ON}/W_{eff} [mA/mm]	840	540
I_{ON}/W_{Top} [mA/mm]	7.2	4.7
I_{ON}/I_{OFF}	~2x10^5	~6x10^3
V_{Tsat} [V]	0.50	-0.37
DIBL [mV/V]	32	63
S.S.$_{sat}$ [mV/dec]	64	73

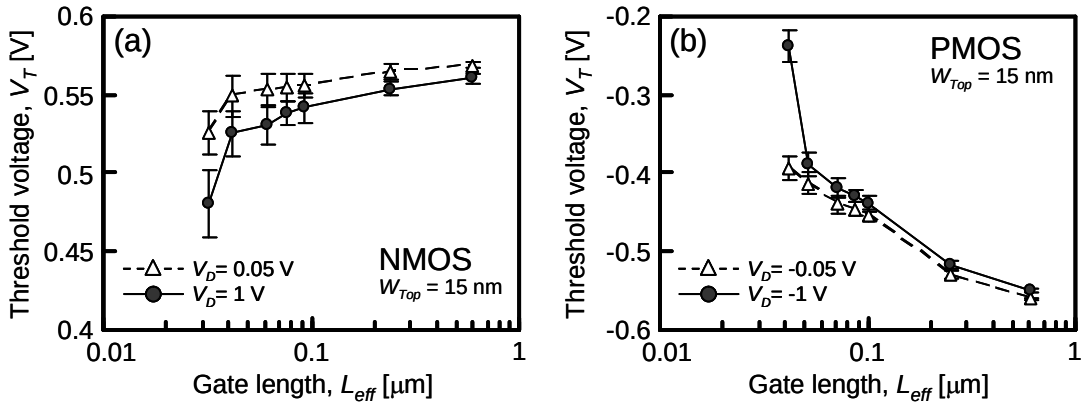


Figure 5.3 Threshold voltage as a function of effective gate length for vertically-stacked silicon nanowire n- and p-MOSFET with 15-nm-W_{Top}.

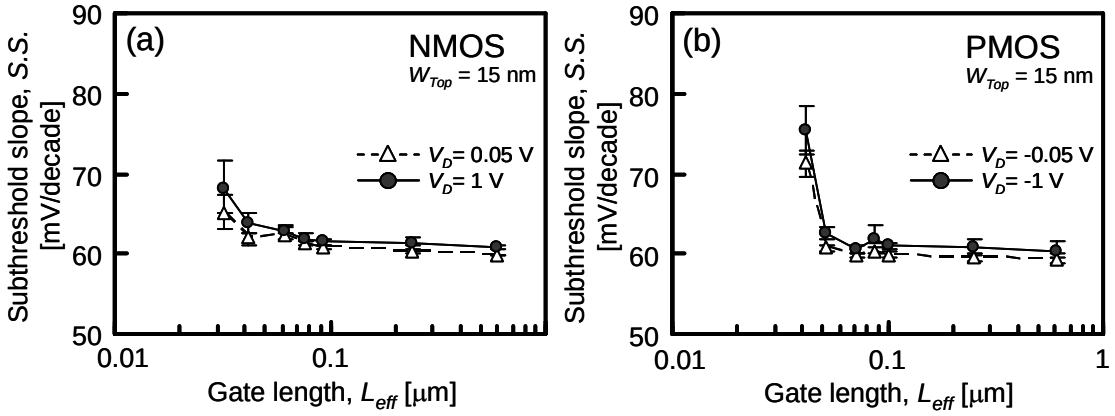


Figure 5.4 Subthreshold slope as a function of effective gate length for vertically-stacked silicon nanowire n- and p-MOSFET with 15-nm-W_{Top}.

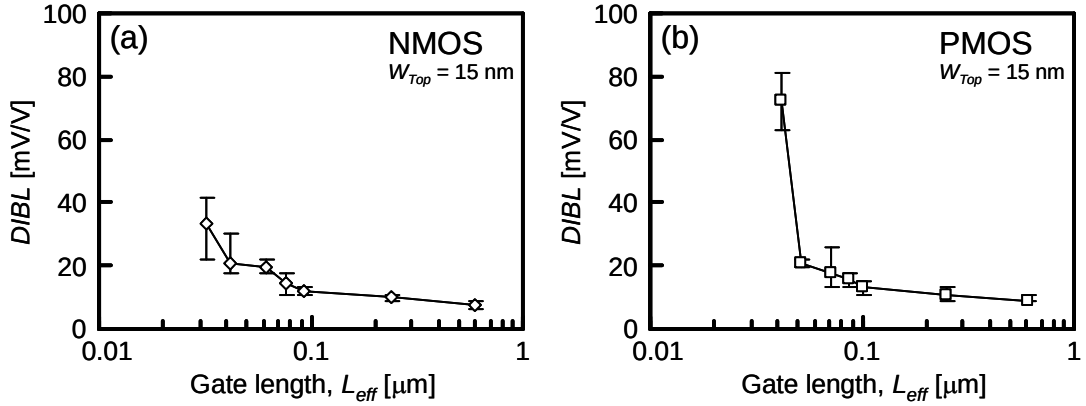


Figure 5.5 Drain-induced barrier lowering (DIBL) as a function of effective gate length for vertically-stacked silicon nanowire n- and p-MOSFET with 15-nm-W_{Top}.

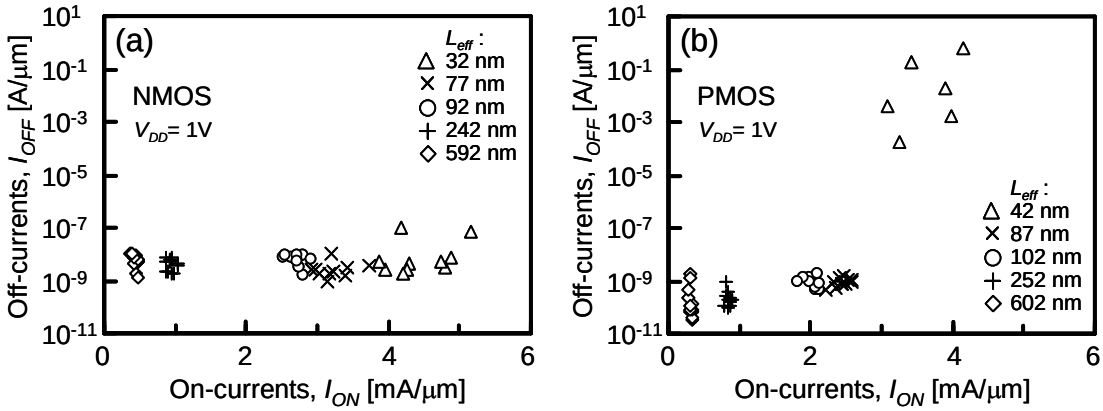


Figure 5.6 I_{ON}-I_{OFF} characteristics of vertically-stacked silicon nanowire n- and p-MOSFET with 15-nm-W_{Top}.

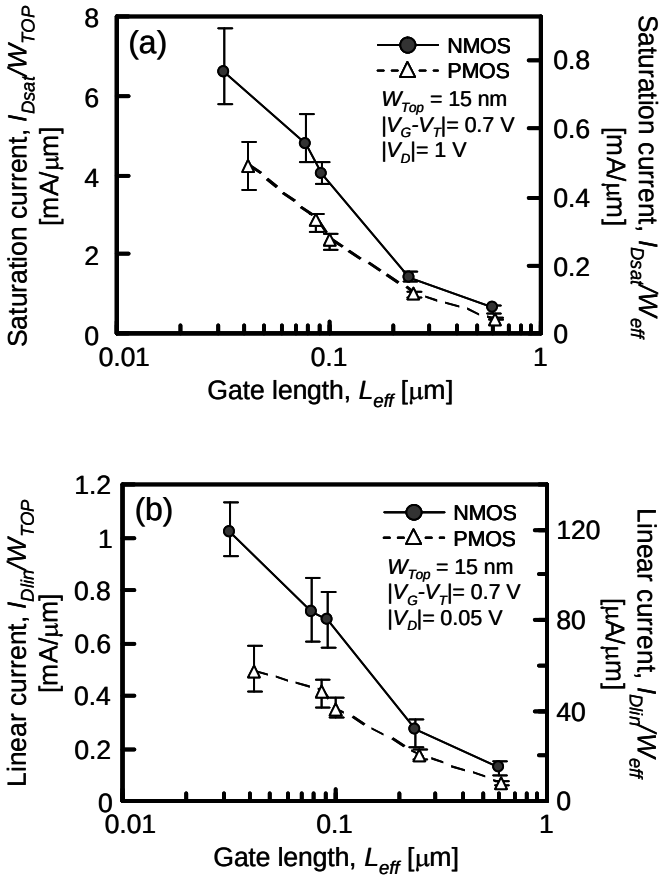


Figure 5.7 Saturation current density (a) and linear current density (b) as a function of effective gate length for vertically-stacked silicon nanowire n- and p-MOSFET with 15-nm-WTop. The currents are normalized by top-view width WTop (left y-axis) and effective total width Weff (right y-axis).

5.2.2 Transport Limiting Velocity

In thermal equilibrium conditions, carriers can be injected from the source reservoir to the channel with a thermal velocity. The part (r) of injected carriers can be elastically backscattered towards the source, whereas the ($1-r$) part propagates towards the drain. Therefore, the effective source injection velocity (v_{inj}) is smaller than thermal velocity. In the linear regime (small lateral field), the injection velocity, resulting from forwarded and backscattered fluxes and given by the ratio $(1-r)/(1+r)$, will be the only limitation of the total drain current, this being called the quasi-ballistic transport. However for high lateral electric field, saturation velocity v_{sat} resulting from optical phonon-electron interactions may constitute a stronger limitation than injection velocity.

Therefore, whatever the conduction regime (linear or on-state), there exists a certain limiting velocity that can be expressed as: $v_{lim} = min\ (v_{sat}, v_{inj})$. This simple reasoning gives ground for a unification of all transport mechanisms, within one universal and continuous drain current model that is a kind of Matthiessen's rule. Since v_{sat} and v_{inj} have very close values, it is difficult to identify the true limiting mechanism. Fortunately, the temperature dependences of v_{sat} and v_{inj} are opposed, and thus they can reveal the limiting mechanism [5.1].

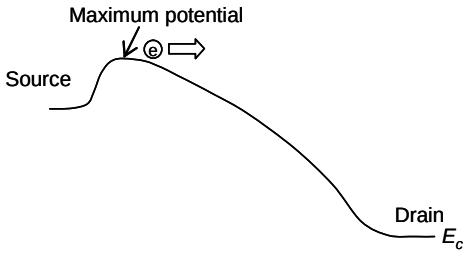


Figure 5.8 Schematic image of conduction band

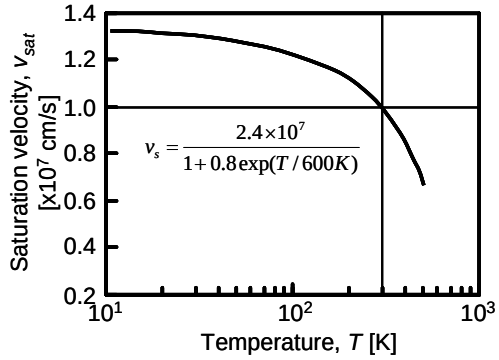


Figure 5.9 Temperature dependence of saturated electron drift velocity [5.2].

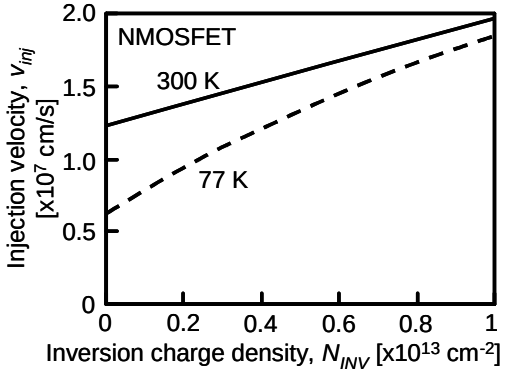


Figure 5.10 Injection velocity of a NMOSFET on (100) plane as a function of inversion charge density at 300 and 77 K. [5.3].

Low field mobility μ_0 is degraded with decreasing L_{eff} as shown in Figure 5.11. One of the possible reasons is the ballistic motion of carriers [5.4]. The part of injected carriers can ballistically reach drain. Indeed, the mobility experimentally extracted can be limited by ballistic transport. Thanks to the temperature dependence of both saturation velocity and injection velocity, the nature of the transport can be evidenced by plotting the temperature dependence of the limiting velocity. The temperature dependence of I_{DS}-V_{GS} curves for the SNWT with L_{eff} of 32 nm is shown in Figure 5.12. The temperature range is from 5 to 300 K. The threshold voltage decreases with temperature, while the sub-threshold slope increases. These changes in V_T and *SS* with temperature are mainly due to band gap changes and are consistent with the theory. Figure 5.13 shows temperature dependence of v_{lim} for NMOS. It is clear that the vertically-stacked SNWTs are almost exclusively v_{sat} limited. This result implies high backscattering rate and the presence of strong scattering components.

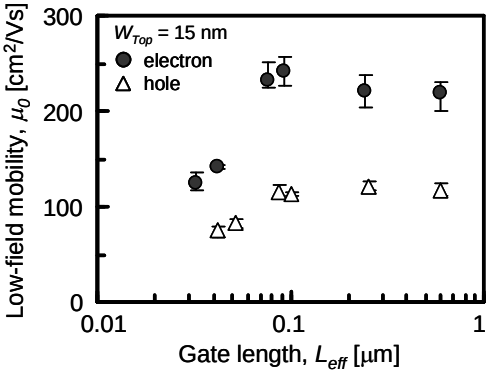


Figure 5.11 Low-field mobility μ_0 as a function of effective gate length for vertically-stacked silicon nanowire n- and p-MOSFET with 15-nm-W_{Top}.

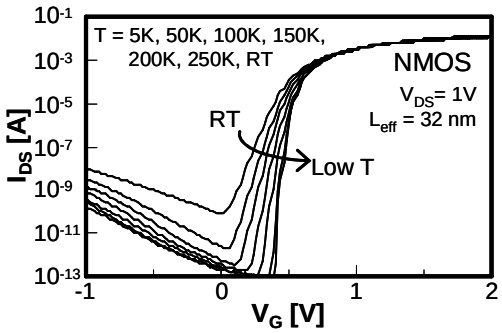


Figure 5.12 Temperature dependence of I_D-V_G characteristics for vertically-stacked SNWTs.

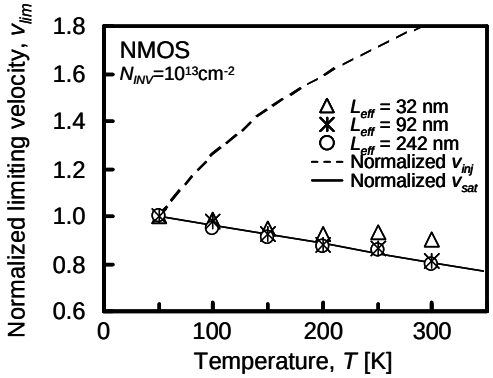


Figure 5.13 Temperature dependence of extracted limiting velocity v_{lim}. Theoretical dependence of saturation velocity v_{sat} from [5.5] and injection velocity v_{inj} from [5.6] are given.

5.2.3 Carrier Mobility Evaluation

Figure 5.14 shows possible carrier transport limiting factors intrinsically for silicon nanowire transistor. In general, the carrier mobility in two-dimensional (2-D) transport for planar FET strongly depends on its surface orientations due to the effective mass difference. The electron mobility on (100)-surface is about two times as high as that on (110)-surface, while the hole mobility is opposite. The rectangular nanowires directed to <110> have two oriented surfaces, that is, (100)-surface for the top and bottom channels and (110)-surface for the side channels. In that case, as shrinking the wire width, that is, the (100)-surface, it is expected that the electron mobility decreases, while the hole mobility increase as shown in Figures 5.14 (b) and 5.15. Moreover, in silicon nanowire with less than 10 nm in diameter, carrier transport will become one-dimension (1-D). In that case, it is expected that carrier limiting components show different behaviors from two-dimensional transport as shown in Figure 5.14 (c). In addition, as nanowire width decreases, transport property at the corners for rectangular shaped nanowires becomes dominant. The carriers at the corners could possibly behave like one-dimensional transport depending radius of curvature as shown in Figure 5.14 (d) and 5.16. As the results of those carrier transport limiting factors, the experimentally extracted mobility is difficult to understand if the degradation is due to the extrinsic causes. In this study, mobility limiting components in silicon nanowires will be investigated in detail by observing temperature dependence of the mobility.

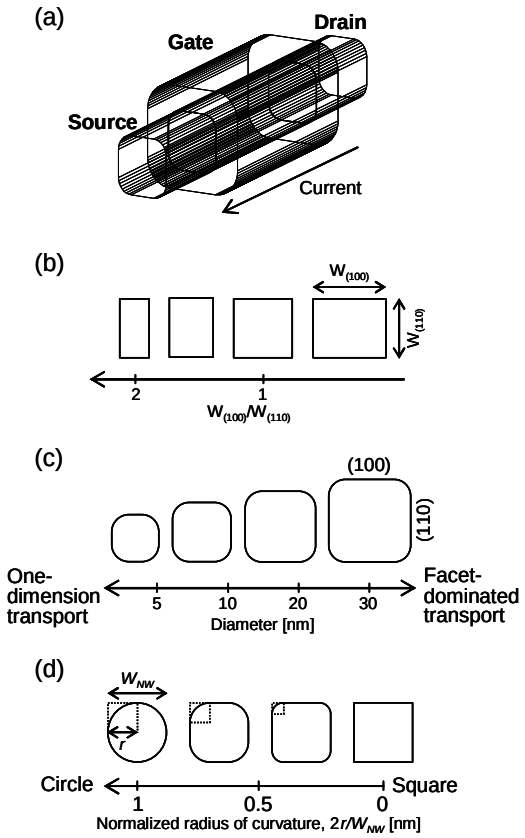


Figure 5.14 Schematic image of nanowire MOSFET (a) and its cross-section: (b) varying ratio of (100) width to (110) width, (c) varying size, (d) varying radius of curvature at the corners.

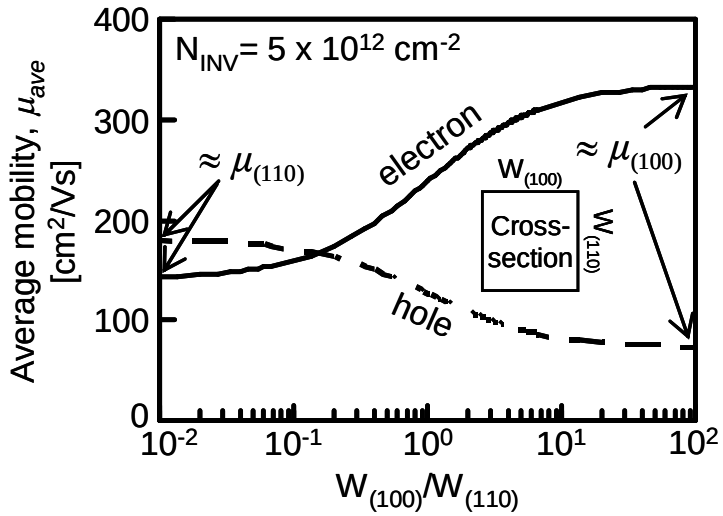


Figure 5.15 Nanowire mobility as a function of $W_{(100)}$-to- $W_{(110)}$ ratio. The average mobility is mave=($\mu_{(100)}W_{(100)}$+ $\mu_{(100)}$ $W_{(100)}$)/($W_{(100)}$+$W_{(110)}$), where $\mu_{(100)}$ and $\mu_{(110)}$ are the mobilities on (100) and (110) surfaces, respectively.

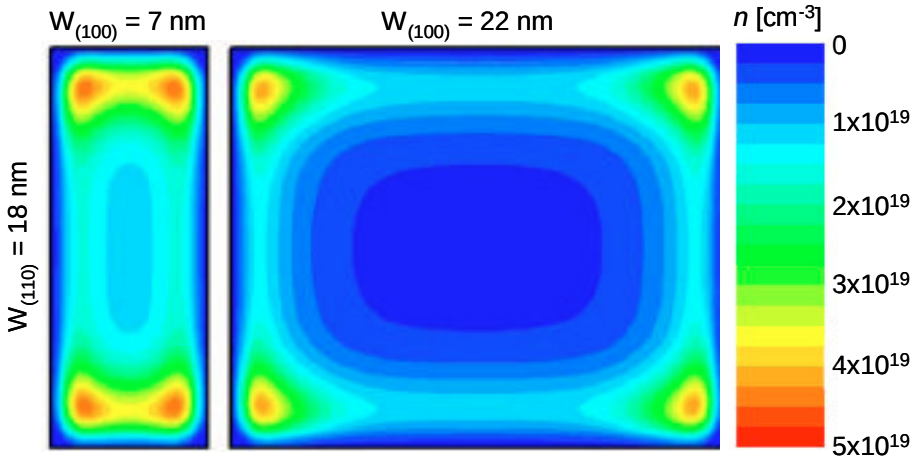


Figure 5.16 Calculated profiles of electron density across the cross section of silicon nanowires with $W_{(110)}$=18 nm, and $W_{(100)}$= 7 and 22 nm [5.7].

Since the backscattering rate is directly related to the mobility, it is important to examine the temperature dependence of mobility to quantify the contribution of each scattering mechanism. Figures 5.17 show effective mobility for electrons in vertically stacked SNWTs and fully-depleted SOI FET as a function of inversion charge density at different temperatures. It is clear that the electron mobility dependence on temperature for the SNWT is much lower than that for FDSOI. In addition, the electron mobility dependence on temperature is also much lower than that for hole. In general, mobility in MOSFETs is limited by three scattering components; coulomb, phonon, and surface roughness as shown in Figure 5.19 (b). The coulomb-(μ_{cb}) and phonon-(μ_{ph}) limited mobilities have negative and positive contribution at low temperature, respectively. Meanwhile, the surface roughness-limited mobility (μ_{sr}) does not depend on temperature. At low temperature, mobility is limited by only the coulomb scattering only at low N_{inv} and only by the surface-roughness scattering only at high N_{inv}. Figure 5.19 (a) shows effective mobility for electron and hole at high N_{inv} as a function of temperature. The μ_{sr} values can be extracted by extrapolating the mobility at 10 K. The μ_{sr} values for vertically-stacked SNWTs are much smaller than that for FDSOI. Phonon limited mobility can be extracted by using Matthiessen's rule and the extracted μ_{sr}. Figure 5.21 shows comparison of mobility limiting components at high N_{INV}. It is clear

that electron effective mobility is strongly limited by surface roughness scattering, while hole mobility is limited by both surface roughness and phonon scattering at high N_{INV}. In Figure 5.22, electron mobility at low N_{inv} is degraded at lower temperatures. This indicates that electron mobility at low N_{inv} in vertically-stacked SNWTs is limited by coulomb scattering. On the other hand, hole mobility increase at low temperature. This means that phonon scattering is dominant at low N_{INV}. Mobility for vertically-stacked SNWTs is degraded due to surface roughness. The surface roughness could be one of the key factors to obtain high performance.

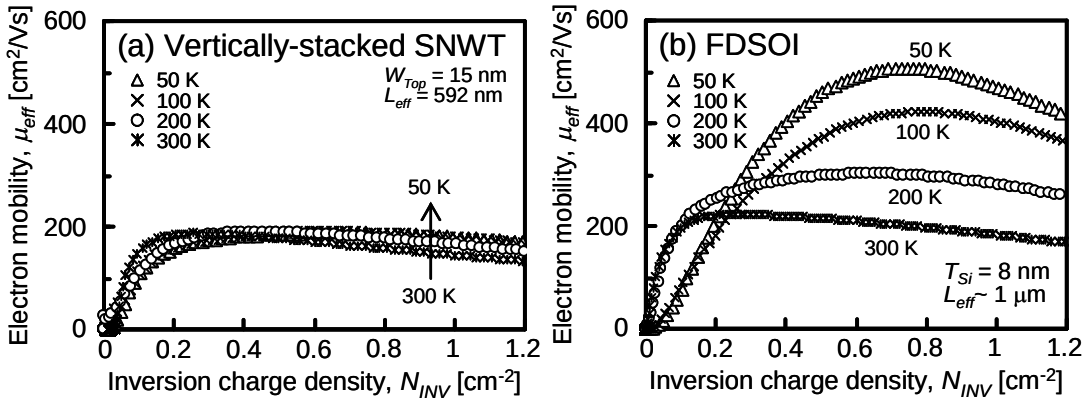


Figure 5.17 Temperature dependence of effective electron mobility in vertically-stacked silicon nanowire FET (a) and in fully-depleted SOI-FET (b).

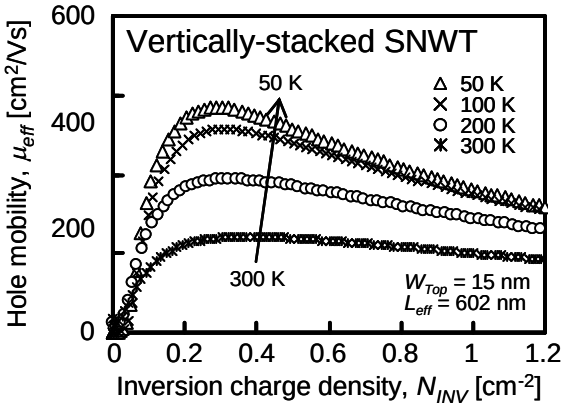


Figure 5.18 Temperature dependence of effective hole mobility in vertically-stacked silicon nanowire FET.

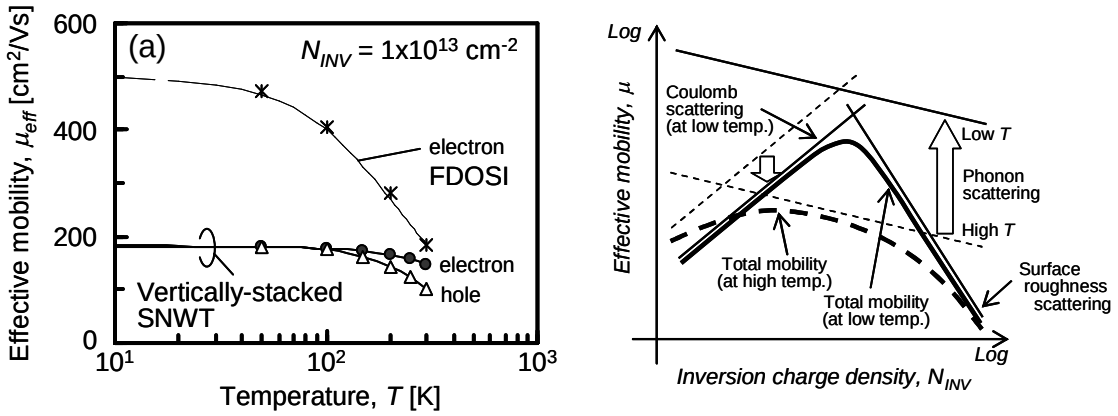


Figure 5.19 (a) Temperature dependence of effective mobility at high inversion charge density (N_{INV}=10^{13}cm^{-2}). (b) Schematic diagram of mobility limiting components at low temperature.

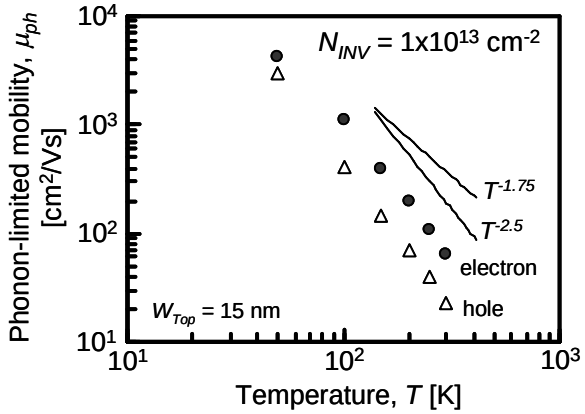


Figure 5.20 Temperature dependence of phonon-limited mobility at high inversion charge density (N_{INV}=10^{13}cm^{-2}) for vertically-stacked silicon nanowire MOSFETs.

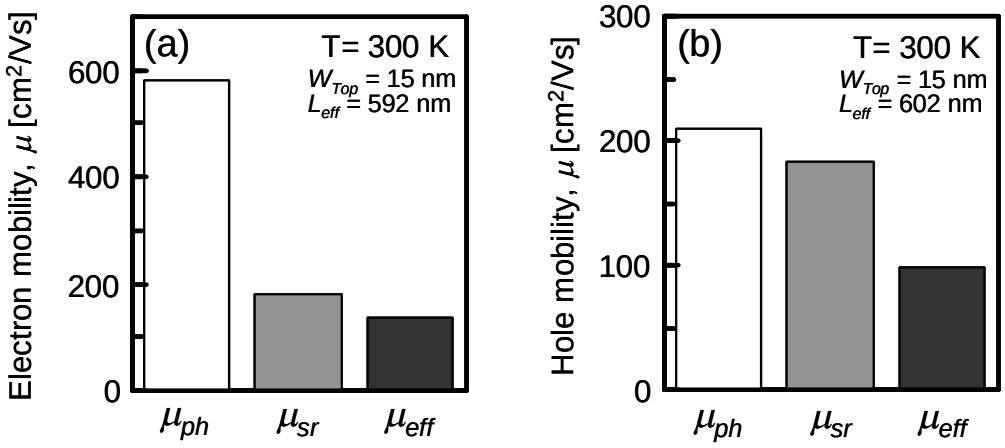


Figure 5.21 Mobility limiting components for electron (a) and hole for vertically-stacked silicon nanowire MOSFETs at 300 K.

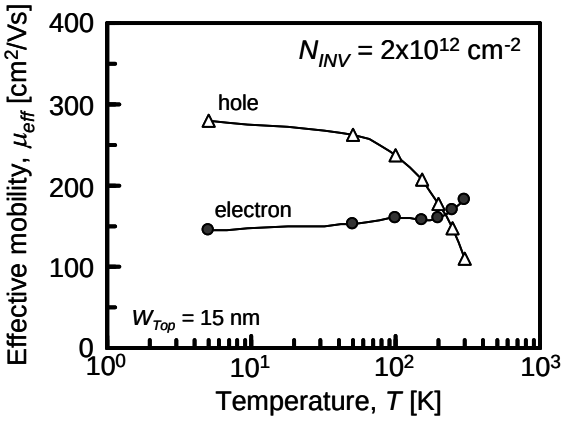


Figure 5.22 Temperature dependence of effective mobility at low inversion charge density (N_{INV}=2x10^{12} cm^{-2}).

5.3 IMPACT ON PLASMA ETCHING OF SILICON-GIRMANIUM SACRIFICIAL LAYERS

One of the possible reasons of degraded surface-roughness limited mobility is damage due to the selective SiGe dry and isotropic etching. According to results reported by C. Dupre *et. al.* [5.8], Si planar NMOS transistors have been roughened using an isotropic plasma etching, similar to the one used for SON technologies. AFM was used to evaluate the RMS of damaged Si surfaces as shown in Figure5.23 and 5.24. In this section, impact on plasma etching of SiGe sacrificial layers for vertically-stacked SNWTs is investigated. To compare between with and without the plasma etching, one-level SNWTs were fabricated without SiGe sacrificial layer.

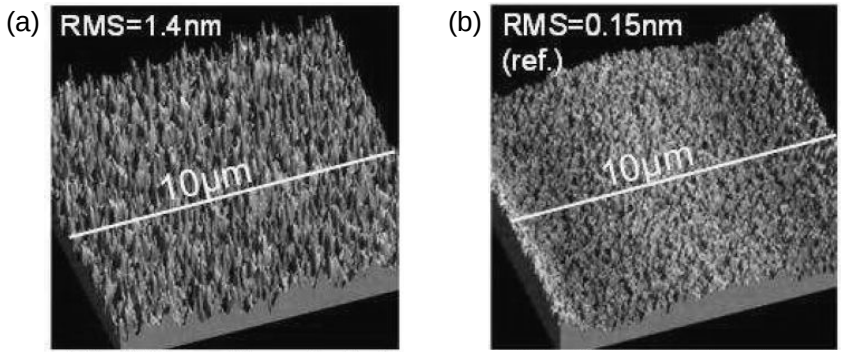


Figure 5.23 Atomic force microscopy images of silicon surface after isotropic SiGe dry etching (a) and the reference sample without the etching (b).

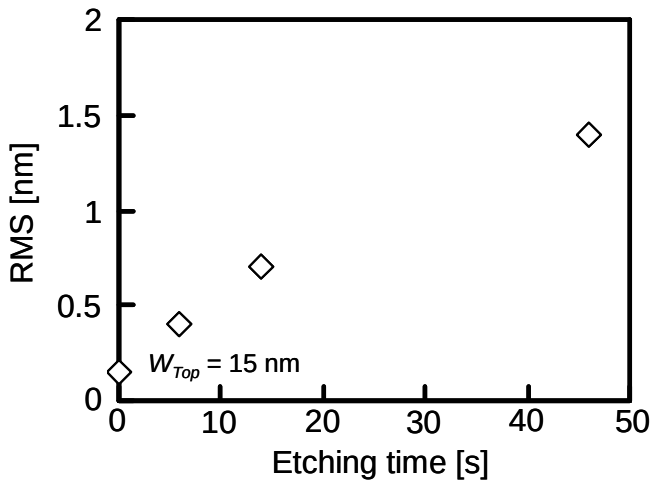


Figure 5.24 Root mean square (RMS) values as a function of isotropic SiGe dry etching.

5.3.1 One-Leveled Nanowire MOSFET Fabrication

In order to investigate the impact of the selective SiGe isotropic etching, one-level nanowire MOSFETs as reference are fabricated without SiGe epitaxy and etching process as shown in Figure 5.25. The NW diameter is controllable down to 5 nm by self limited oxidation [5.9] while keeping regularly arrayed NWs as shown in Figure 5.26 (a).

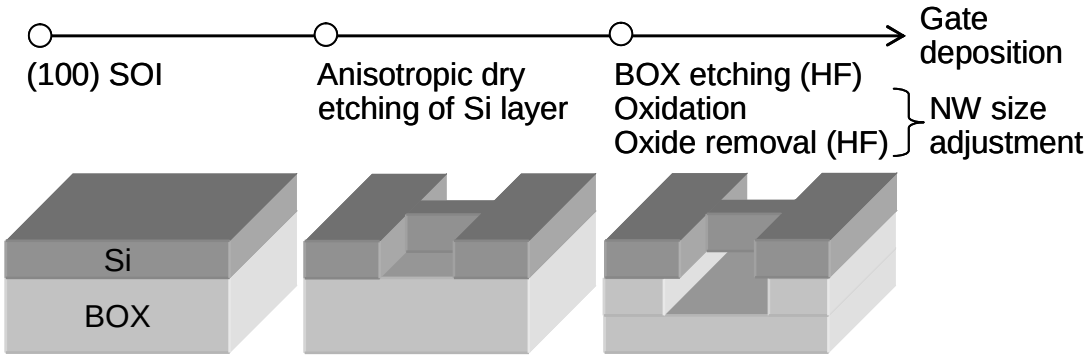


Figure 5.25 Brief process flow of 1-level silicon nanowire MOSFET.

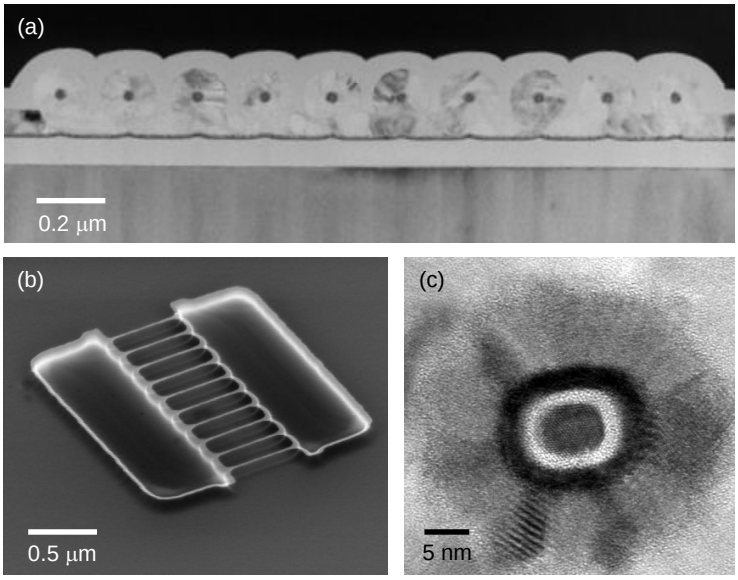


Figure 5.26 Gate-all-around 1-level silicon nanowire MOSFETs fabricated without SiGe epitaxy and selective etching.

5.3.2 Carrier Mobility Evaluation

Figure 5.27 shows effective mobility comparisons between vertically-stacked and 1-level SNWTs with 15 nm of W_{NW} at 300 K and 5 K. The 1-level SNWTs show higher mobility than vertically-stacked ones. The higher effective mobility at high inversion charge density in 1-level SNWTs means less surface-roughness scattering as shown in Figure 5.28–5.30. Moreover, in the case of 1-level SNWTs, coulomb scattering is less dominant as shown in Figure 5.31. The reason why stronger coulomb scattering is higher in stacked SNWTs may be the degraded interface quality with high-k because of the use of SiGe sacrificial layers. Additional surface treatments may thus be needed.

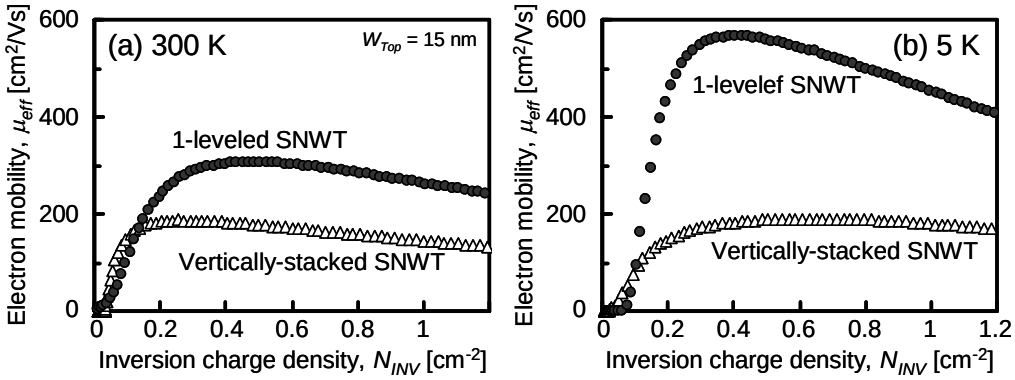


Figure 5.27 Electron mobility comparisons between 1-leveled and vertically-stacked silicon nanowire MOSFET at 300 K (a) and 5 K (b).

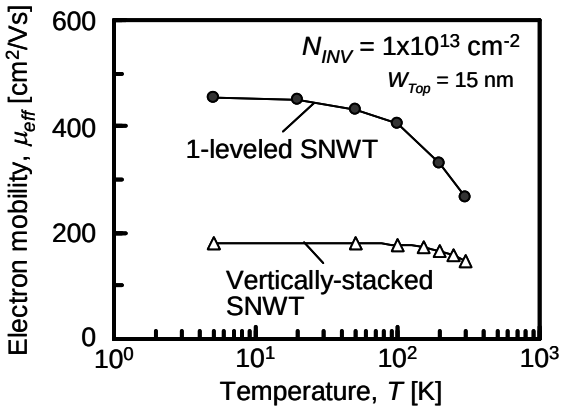


Figure 5.28 Temperature dependence of effective mobility at high inversion charge density (N_{INV}=10^{13}cm^{-2}) for 1-leveled and vertically-stacked silicon nanowire MOSFET.

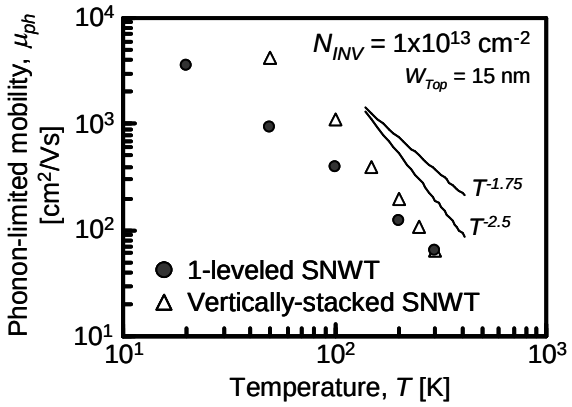


Figure 5.29 Temperature dependence of phonon-limited mobility at high inversion charge density (N_{INV}=10^{13}cm^{-2}) for 1-leveled and vertically-stacked silicon nanowire MOSFETs.

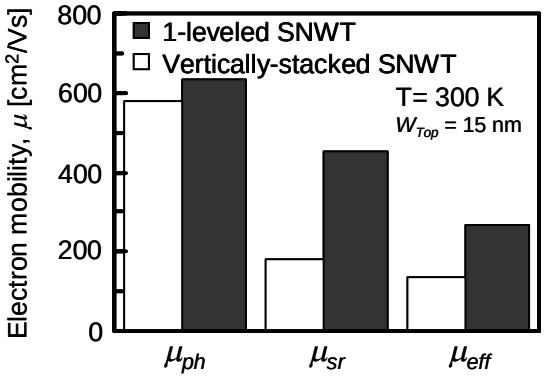


Figure 5.30 Mobility limiting components comparison at high inversion charge density between 1-leveled and vertically-stacked silicon nanowire MOSFETs at 300 K.

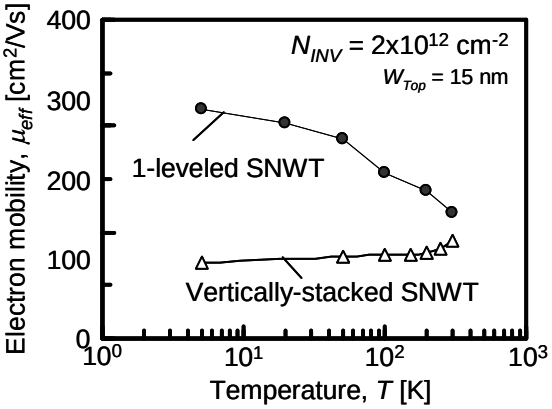


Figure 5.31 Temperature dependence of effective mobility at low inversion charge density (N_{INV}=2x10^{12} cm^{-2}).

5.4 EFFECT OF HYDROGEN ANNEALING

Hydrogen annealing can, however, be used intentionally for three-dimensional profile transformation by rounding sharp corners while diminishing the surface roughness and keeping the active layer crystalline [5.10, 5.11]. In this section, a mobility study was performed in order to highlight the impact of hydrogen annealing on etched surfaces.

5.4.1 Cross-Sectional Shape

Hydrogen annealing (750°C, 2min) was performed on 15 nm wide Si nanowires (Figure 5.32 (a)). Si nanowires were rounded thanks to this specific process (Figure 5.32 (b)).

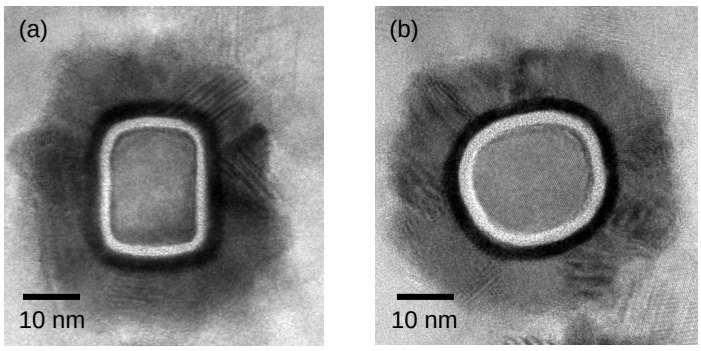


Figure 5.32 Cross-sectional TEM images of silicon nanowire (a) without and (b) with hydrogen annealing at 750 ^{o}C for two minutes.

5.4.2 Carrier Mobility Evaluation

First, the impact of H_2 annealing on NW surface quality was investigated by using 1-leveled SNWTs. Figure 5.33–5.36 show the comparisons of electron effective mobility between with and without hydrogen annealing. It is clear that effective mobility in H_2-annealed SNWTs is more degraded by coulomb scattering, while surface roughness is improved. Figure 5.37 shows a mobility comparison between with and without hydrogen annealing for vertically-stacked SNWTs. A circular shape formed by hydrogen annealing leads to mobility degradation at low inversion charge density (N_{inv}). Improvement of μ_{eff} at high N_{inv} is however observed for circular NWs because their surface roughness is reduced by the H_2 annealing as well as 1-leveled SNWTs.

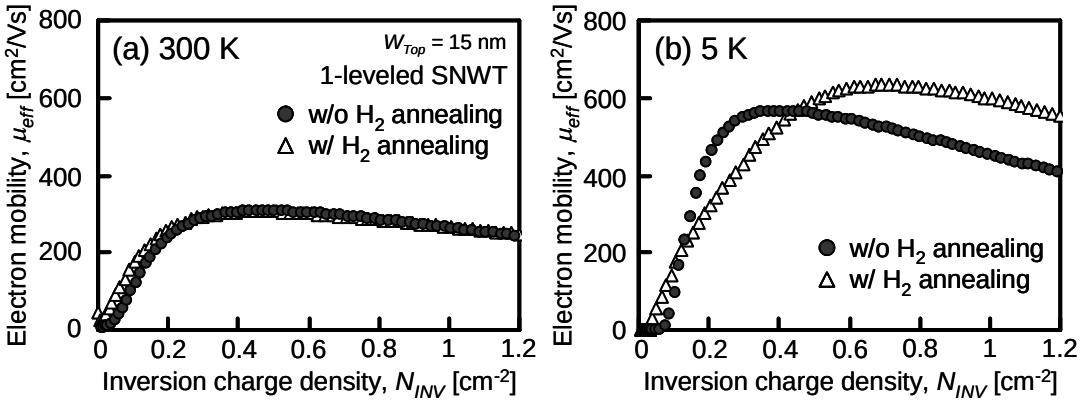


Figure 5.33 Electron mobility comparison of 1-leveled silicon nanowire MOSFETs between with and without hydrogen annealing. The measurement temperatures are 300 K (a) and 5 K (b).

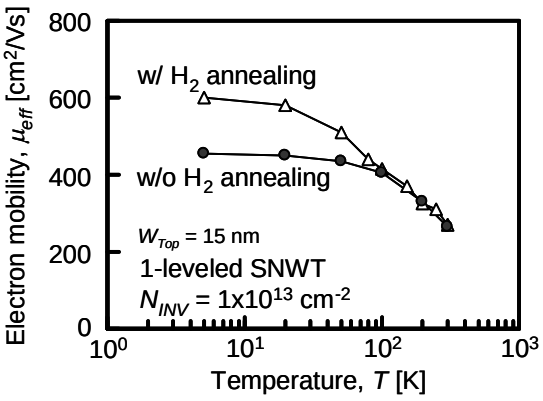


Figure 5.34 Temperature dependence of effective mobility at high inversion charge density (N_{INV}=10^{13}cm^{-2}) for 1-leveled nanowire MOSFET with and without H_2 anneal.

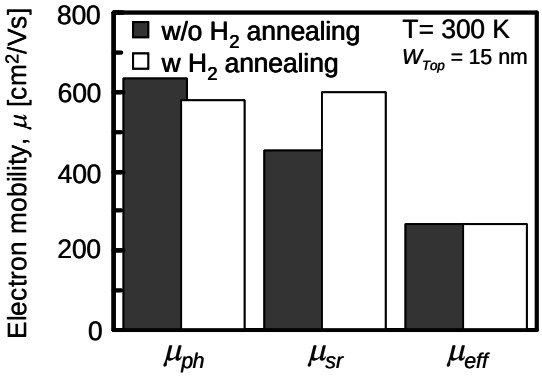


Figure 5.35 Mobility limiting components comparison at high inversion charge density and 300 K for 1-leveled silicon nanowire MOSFET between with and without hydrogen annealing.

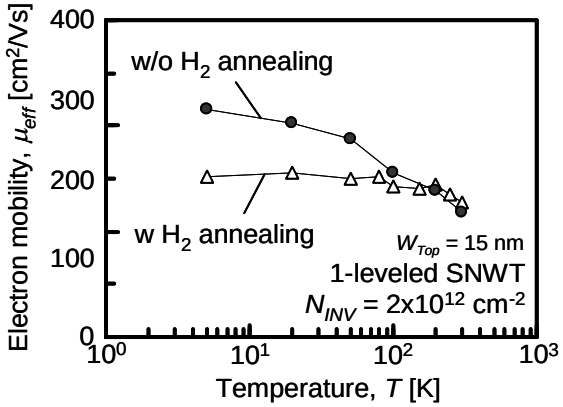


Figure 5.36 Temperature dependence of effective mobility at low inversion charge density (N_{INV}=2x10^{12} cm^{-2}).

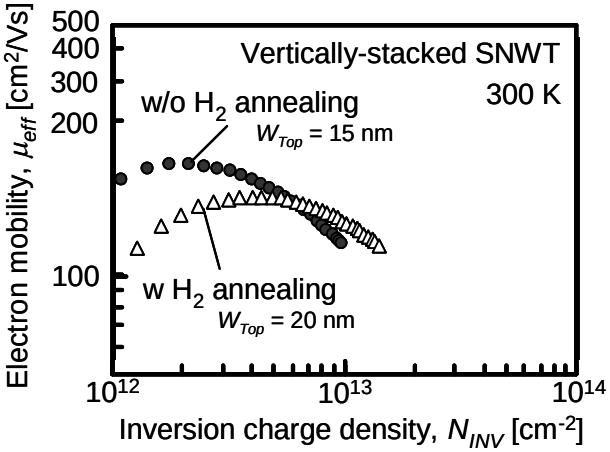


Figure 5.37 Electron mobility comparison of vertically-stacked silicon nanowire MOSFETs between with and without hydrogen annealing.

5.4.3 Interface Trap density

Mobility at low N_{inv} is mainly limited by (remote) coulomb scattering due to interface and oxide charges and/or interface dipoles between high-*k* and interfacial layer in the case of high-k/metal gate stack. To evaluate the interface quality, the interface trap density (D_{it}) have been quantified by adapting the charge pumping method with gated-diode structures. Figure 5.38 shows the charge pumping current (I_{cp}) in Si NWs which exhibits a typical "hat" shape. The peak I_{cp}-freqency (f) plots shows a good linearity in Figure 5.39. Circular NWs have roughly 3 times higher mean D_{it} values than rectangular ones. Figure 5.40 show the energy profiles of D_{it} (obtained by temperature dependence of Si band gap). At both the upper and lower regions of the gap, the D_{it} of circular NWs is higher, leading to higher mean D_{it} values.

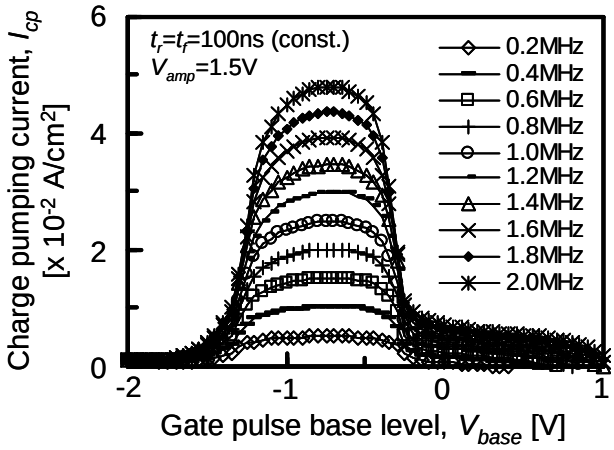


Figure 5.38 Charge pumping currents I_{cp} obtained base voltage sweep on nanowire gated-diode with L_G = 240 nm and W_{NW}/H_{NW}= 20 nm/15 nm. The currents are normalized by W_{eff} obtained from TEM images.

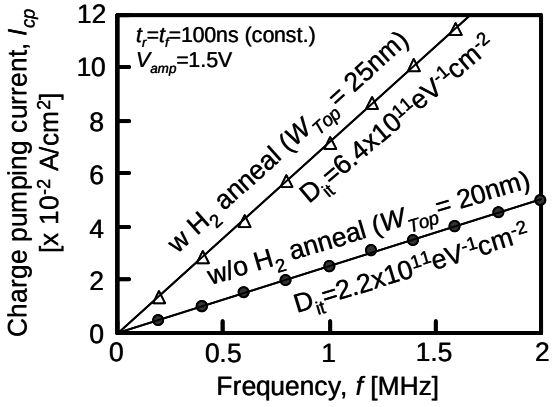


Figure 5.39 Charge pumping currents I_{cp} as a function of frequency f.

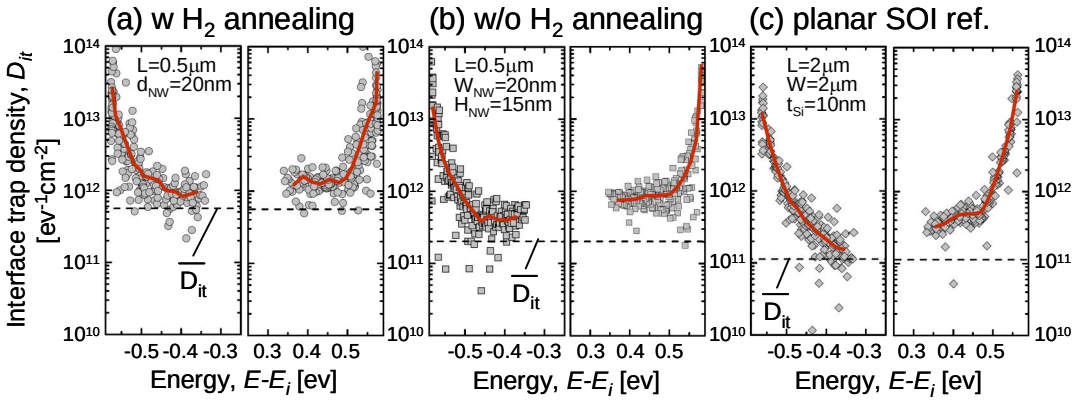


Figure 5.40 Interface trap density as a function of energy for vertically-stacked nanowires with (a) and without (b) hydrogen annealing, and planar SOI devices (c) with the same gate stack (3 nm HfO_2 ALD/10 nm TiN CVD). The profile is obtained by scanning temperature from 300 K down to 25 K by 25 K steps. The bold line represents the mean value of $D_{it}(E)$. The dashed line is the directly measured mean value of interface trap density over the full energy range at 300 K which evidence the lower density of interface traps in the middle of the gap.

5.5 SILICON-GIRMANIUM NANOWIRE MOSFET

5.5.1 Device Fabrication Process

The fabrication process of vertically-stacked SiGe nanowire FETs was changed from Si ones as the following steps. SOI (001) wafers were used for Si and compressively (c)-strained SiGe NWs. Tensile-strained (1.3 GPa) SOI (001) wafers were used for un-strained SiGe NWs, respectively. After anisotropic etching of $Si/Si_{0.8}Ge_{0.2}$ superlattices, isotropic etching of Si layers between $Si_{0.8}Ge_{0.2}$ layers was performed to obtain the suspended $Si_{0.8}Ge_{0.2}$ nanowires. In order to achieve better interface quality, a 2 nm-thick-Si cap was grown at 650 ^{o}C on the SiGe NWs.

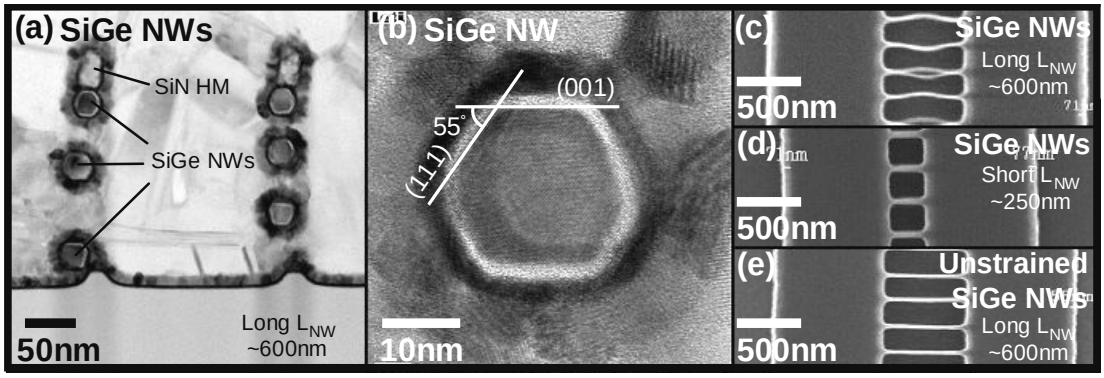


Figure 5.41 (a) Cross-sectional TEM micrographs of 3D-stacked compressively(c)-strained SiGe NWTs, (b) enlarged images of c-strained SiGe NW, (c) top view of bended c-strained SiGe NWs with L_{NW}=600nm, (d) top view of c-strained SiGe NWs with L_{NW}=250nm, and (e) top view of un-strained SiGe NWs with L_{NW}=600nm. Short length SiGe NWs are straight, this whatever their strain state.

5.5.2 *I–V* Characteristics

C-strained SiGe and un-strained SiGe NWs were evaluated in order to boost pFET performances. Figure 5.42 shows I_{ON}/I_{OFF} characteristics of Si, c-strained and un-strained SiGe NWs. The currents are normalized by the number of wires. Both the SiGe NWTs showed larger off-current than SNWTs. This is due to the lower V_T for the SiGe NWTs as shown in Figure 5.43. The c-strained SiGe NWTs show higher on-current. However the best I_{ON}/I_{OFF} performance is obtained for Si NWs.

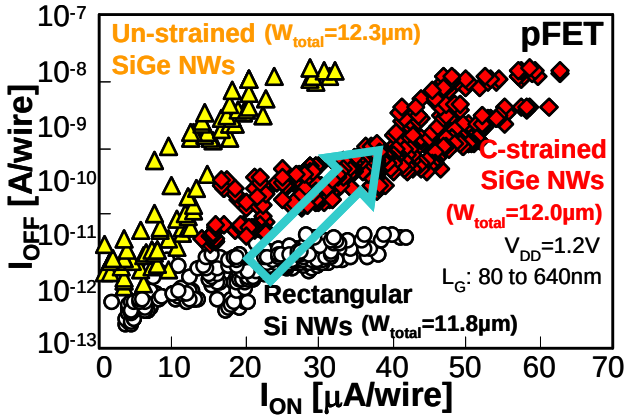


Figure 5.42 I_{ON}/I_{OFF} characteristics of Si, c-strained and un-strained SiGe NWs normalized by the number of wires. The total NW surface W_{total} is estimated from the cross-sectional TEM images. The W_{NW} of all NWs is ~20nm.

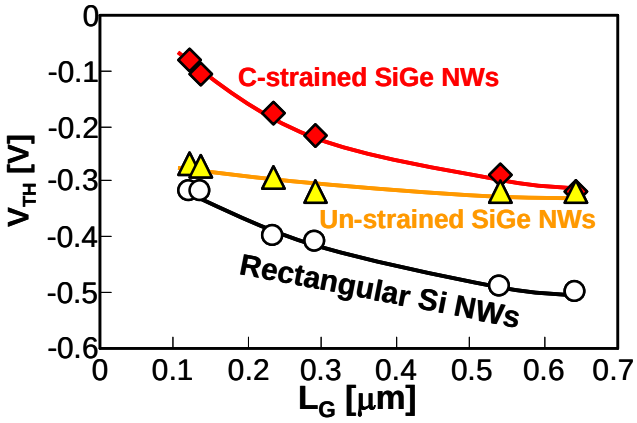


Figure 5.43 Threshold voltage of Si, c-strained and un-strained SiGe NWs as a function of gate length. The W_{NW} of all NWs are ~20nm.

5.5.3 Carrier Mobility Evaluation

Figure 5.44 shows a mobility comparison. The large enhancement of mobility was obtained in the c-strained SiGe NWTs compared with un-starained ones. This can be due to the compressive strain effect. However, in comparison with SNWTs, a small impact on mobility was observed. The c-strained SiGe NWTs have higher μ_{eff} at high N_{inv} lead to a larger I_{ON} current than for Si NWTs. The hexagonal cross section of SiGe NWs with (111) sidewalls could also contribute to mobility degradation as shown in Figure 5.44.

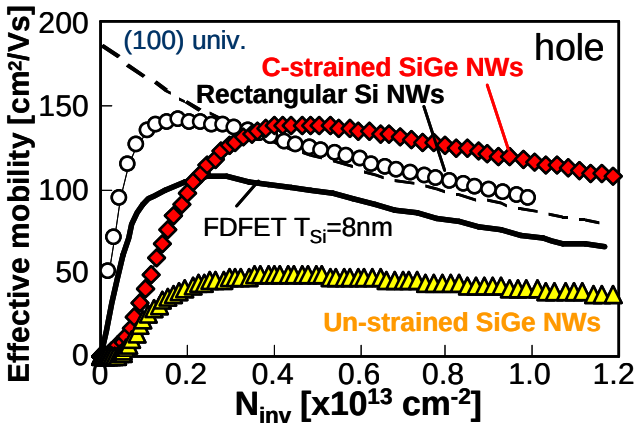


Figure 5.44 Effective hole mobility of Si, c-strained and un-strained SiGe NWs. The W_{NW} of all NWs are ~20nm.

5.5.4 Noise Measurement

Low-frequency noise measurements performed on the NWs between 10 Hz and 10 kHz at V_{DS}=50 mV, show an oxide trap density (N_t) for SiGe NWs 3.5 times larger than for Si NWs (Figure 5.45). This higher trap density may reduce the mobility.

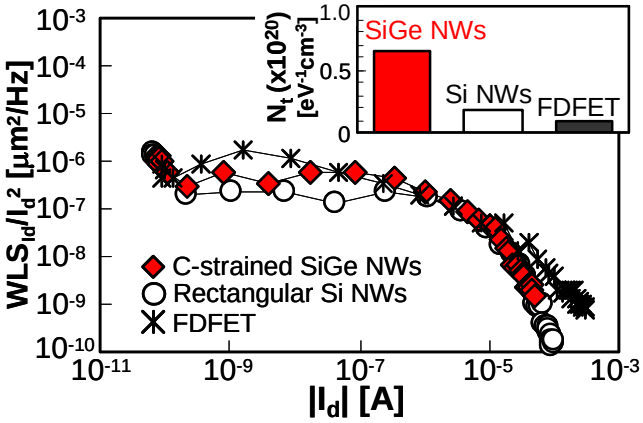


Figure 5.45 Low-frequency noise of Si and c-strained SiGe NWs. Inserted figure is a comparison of oxide trap density (N_t). L_G and W_{NW} are ~290nm and ~20nm, respectively.

5.6 CONCLUSIONS

In this chapter, the electrical characteristics of vertically-stacked SNWTs have been investigated. Vertically-stacked nanowire structure can achieve extremely high on-currents per given layout surface with good short-channel effects immunity. This result is expected to achieve high integration and low power consumption. On the other hand, in terms of its performance, the optimisation of short-channel CMOS nanowire drive current will have to take into account specific effects. In particular, the use of SiGe sacrificial layer to make vertically-stacked channels cause the large mobility degradation due to the surface roughness, resulting from the damage of plasma etching. This result can evidence the poor ballisticity in the short channel SNWTs.

The hydrogen annealing can improve the surface-roughness limited mobility a little. Charge pumping measurements, however, revealed that circular-shaped SNWTs, which are formed by hydrogen annealing, have a higher D_{it} than rectangular ones, leading to low-field mobility degradation. This high D_{it} might be caused by the

continuously-varying surface orientation. The resulting additional coulomb scattering could partly explain the quite low mobility in 5 nm diameter SNWTs together with the already known transport limitations in NWs.

The vertically-stacked SiGe NWTs have been also investigated. Compressively-strained SiGe showed slightly higher mobility than Si ones. One of the possible reasons of the small mobility enhancement is the higher trap density for SiGe nanowires. Additionally, the hexagonal cross section of SiGe NWs with (111) sidewalls could also contribute to mobility degradation

5.7 REFERENCES

[5.1] D. Fleury, G. Bidal, A. Cros, F. Boeuf, T. Skotnicki and G. Ghibaudo, "New Experimental Insight into Ballisticity of Transport in Strained Bulk MOSFETs*" VLSI Tech. Dig*., pp. 16-17, 2009.

[5.2] *C. Jacoboni, C. Canali, G Ottaviani, and A. A. Quaranta, "A Review of Some Charge Transport Properties of Silicon," Solid-state Electron., 20,77 (1977).*

[5.3] *K. Natori, "Ballistic Metal-Oxide-Semiconductor Field Effect Transistor," J. Appl. Phys., 76,4879 (1994).*

[5.4] M. S. Shur, "Low Ballistic Mobility in Submicron HEMTs," *IEEE Electron Device Lett.*, vol. 23, no. 9, pp. 511–513, Sep. 2002.

[5.5] S. M. Sze and Kwok K. Ng, *Physics of Semiconductor Devices*, 3rd edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 2007.

[5.6] V. Barral, T. Poiroux, F. Rochette, M. Vinet, S. Barraud, O. Faynot, L. Tosti, F. Andrieu, M. Casse, B. Previtali, E. Bernard, R. Ritzenthaler, P. Grosgeorges, G. Lecarval, D. Munteanu, J-L. Autran, S. Deleonibus, "Will Strain Be Useful for 10nm Quasi-Ballistic FDSOI Devices? An Experimental Study", *VLSI Tech. Dig.*, pp.128-129, 2007.

[5.7] *L. Sekaric, O. Gunawan, A. Majumdar, X. H. Liu, D. Weinstein, and J. W. Sleight, "Size-dependent modulation of carrier mobility in top-down fabricated silicon nanowires." Appl. Phys. Lett., 95, 023113, 2009.*

[5.8] C. Dupre, T. Ernst, S. Borel, Y. Morand, S. Descombes, B. Guillaumot, X. Garros, S. Becu, X. Mescot, G. Ghibaudo, S. Deleonibus, “Impact of Isotropic Plasma Etching on Channel Si Surface Roughness Measured by AFM and on NMOS Inversion Layer Mobility”, *Proceedings of ULIS*, pp. 133-136, 2008.

[5.9] A. Hubert et al., ECS Trans, 13(1), p.195, 2008.

[5.10] M.-C. M. Lee and M. C. Wu, “Thermal Annealing in Hydrogen for 3-D Profile Transformation on Silicon-on-Insulator and Sidewall Roughness Reduction,” Journal of Microelectromechanical systems. Vol. **15**, no. 2, pp.338–343, 2006.

[5.11] E. Dornel, T. Ernst, J. C. Barbe, J. M. Hartmann, V. Delaye, F. Aussenac, C. Vizioz, S. Borel, V. Maffini-Alvaro, C. Isheden, and J. Foucher. “Hydrogen annealing of arrays of planar and vertically stacked Si nanowires”, *Applied Physics Letters*, vol. 91, 233502, 2007.

CHAPTER 6

THRESHOLD VOLTAGE CONTROL OF VERTICALLY-STACKED NANOWIRE MOSFETS

CHAPTER 6 CONTENTS

6.1 INTRODUCTION

In previous chapter, we discussed the possibility of vertically-stacked SNWTs in terms of R_{SD} resistance, transport properties, and short channel effect immunity. Another issue for the vertically-stacked SNWTs is how to control the threshold voltage. For various CMOS applications such as HP, LOP, and LSTP, it is important to achieve flexible threshold voltage in a transistor. In this capter, the possibility of the flexible threshold voltage for vertically-stacked SNWT will be discussed. In particular, the flexibility and the short-channel effects immunity are investigated by numerical simulations.

6.1.1 Threshold Voltage Control by Independent-Gate FinFET

In usual case, the threshold voltages are mainly controlled by the gate work function and channel doping level. However, its control by the high-k/metal gate is difficult and complex because of the sensitivity of process conditions and the necessity of the dual metal and/or dual high-κ. Furthermore, the use of channel dope technique yields large variations in a wafer due to dopants fluctuations. As one of the solutions for three-dimensional devices, independent-gate FinFETs (IG-FinFETs) have been proposed and demonstrated with excellent experimental results of threshold voltage control by the second gate and the synchronized driving mode operation by the double gates [6.1, 6.2]. The independent gates have been successfully fabricated by using a chemical–mechanical-polishing process or an etch-back process [6.1–6.4] as shown in Figure 6.1.

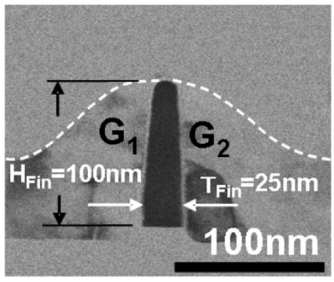


Figure 6.1 Cross-sectional TEM image of the independent-gate FinFET fabricated by the resist etch back process [6.4].

6.1.2 Vertically-Stacked Nanowire Transistor with Independent Gates

Vertically-stacked SNWTs with independent gates have been experimentally demonstrated by C. Dupre *et al.* [6.5]. The device has internal spacers between the nanowires as shown in Figure 6.2, named ΦFET since its shape is similar to the Greek letter Φ. Figure 6.3 summarizes the fabrication process of Φ-FET. First, Reduced Pressure- Chemical Vapor Deposition (RP-CVD) was used to epitaxially grow (25nm-Si /25nm-SiGe)x4 superlattice on SOI wafers (Fig.5.3, step 1). A SiN hard-mask was then deposited. After an hybrid DUV/e-beam lithography, the resist was trimmed to define narrow lines (fin width; W_{Si}~30nm). Then, the exposed Si and (Si/SiGe)x4 areas were trenched by an anisotropic dry plasma etching (Fig.5.3, step 2). The same RIE reactor was used to remove the SiGe isotropically using a CF_4 + O_2 chemistry in order to liberate the suspended Si-nanowires. Then, HTO and SiN were deposited. The partitions between the stacked nanowires were formed by internal spacer obtained by anisotropic and isotropic etchings of SiN selectively to HTO (Fig.5.3, steps 2.1 and 2.2). After chemical cleaning of the channel surface, a HfO_2 / TiN / Poly-Si gate stack was deposited. The gate stack over the SiN hard mask was removed by Chemical Mechanical Polishing (CMP) (Fig.5.3, step 3). After the gate etching, Source/Drain implantations and spacer formation, dopant atoms were activated and the top of S/D regions were silicided. The fabrication ended with a standard Back-End Of Line (BEOL) process. Figure 6.4 shows the cross-sectional TEM pictures of Φ-FET.

This structure is expected to obtain flexible threshold voltage with keeping better short channel effects immunity due to their partially surrounding gates. Figure 6.5 and 5.6 shows the electrical results. Threshold voltage shift have been demonstrated due to a coupling effect between the two gates for ΦFET. Moreover, ΦFET's I_{OFF} currents are 2-decade lower than IG-FinFET ones thanks to an improved electrostatic control.

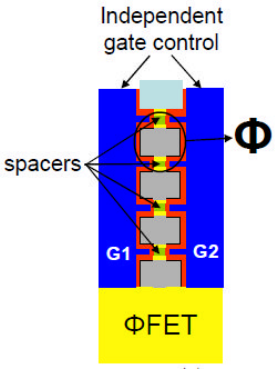


Figure 6.2 Φ-FET scheme.

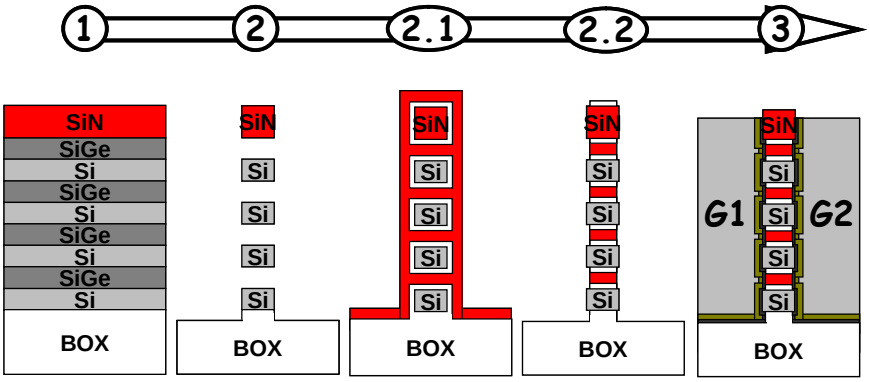


Figure 6.3 Schematic fabrication sequence of Φ-FET.

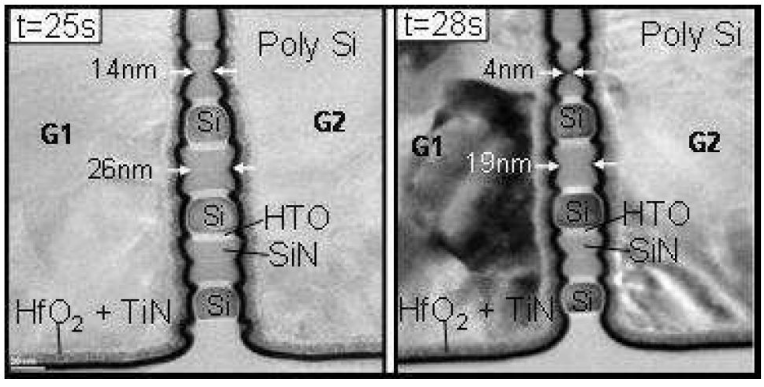


Figure 6.4 Cross-sectional TEM pictures of Φ-FET (3 stacked nanowires). Left: 25s SiN isotropic etching Right: 28s SiN isotropic etching.

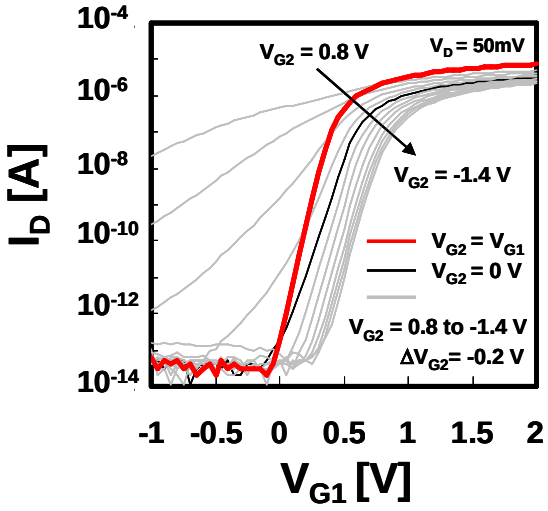


Figure 6.5 Experimental I_d-V_{g1} characteristics at various V_{g2} for n-channel Φ-FET. The gate length and channel width are 550 nm and 25 nm, respectively.

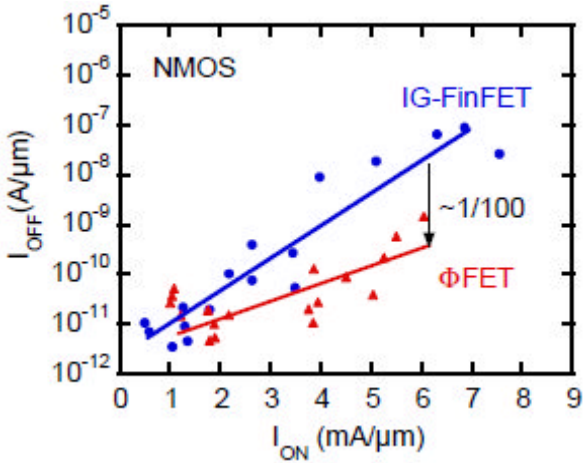


Figure 6.6 I_{on}-I_{off} characteristics comparison between Φ-FET and IG-FinFET.

6.2 OPTIMIZATION OF DEVICE DIMENTIONS

Although the Φ-FETs have been successfully demonstrated, it is necessary to optimize the structure in detail to obtain flexible threshold voltage with better short channel effects immunity. First, we will go through the simulation details. Figure 6.7 shows the simulated Φ-FET structure. To extract DIBL associated to the given structure, the currents in the subthreshold regime are determined by using FlexPDF software as

the following procedure.

The source term is detailed as

$$S = \frac{q}{\varepsilon}(N_a - N_d + n - p), \tag{6.1}$$

where N_a is the acceptor density (=10^{15} cm^{-3}), N_d the donor density (N_d=n_i^2/N_a with n_i the intrinsic carrier density). Here, N_d is negligible due to p-type Si. n and p are given as a function of the potential V:

$$n = N_d \exp\left(\frac{qV}{kT}\right), \tag{6.2}$$

$$p = N_d \exp\left(-\frac{qV}{kT}\right). \tag{6.3}$$

The subthreshold current for long-channel MOSFET is given by:

$$I_{Dlong} = \frac{1}{L}\frac{kT}{q}\mu_{eff} Q_{is}(1 - e^{\frac{qV_D}{kT}}), \tag{6.4}$$

where Q_{is} is the inversion charge per surface on the source side which is deduced by integrating n from the equation (6.2) in the Si volume. The short-channel current thus can be easily rewritten owing to a correction factor C.F [6.6]

$$I_{Dshort} = \frac{I_{Dlong}}{C.F.}, \tag{6.5}$$

$$C.F = \frac{1}{t_{Si} W_{Si} L}\int \exp\left(\frac{-q(V_{short} - V_{long})}{kT}\right), \tag{6.6}$$

where V_{short} and V_{long} are the potentials in short and long channels, respectively. The boundary conditions for a simple nanowire structure as shown in Figure 6.8 are adapted to Φ-FET structure as following: V_G = 0.6 V on the gate oxide and the built-in voltage (V_{BI}) as each nanowire end as

$$V_{BI} = \frac{kT}{q}\ln\left(\frac{N_a N_d}{n_i^2}\right). \tag{6.7}$$

The long-channel case differs from the short-channel case only by applying Neumann boundary conditions: the gradient of the potential (electric field) is forced to zero at each nanowire end. Once the potential distribution is simulated, we used it to extract the

short-channel effects through the natural length λ [6.7],

$$\lambda = \sqrt{\frac{\varepsilon_{Si}}{\varepsilon_{ox}} \frac{t_{Si} t_{ox}}{\eta}}, \tag{6.8}$$

where η is the empirical parameter which varies as a function of the gate configuration: planar (η=1), double gate (η =2), tri-gate (η =3), and gate-all-around (η ≈4). In the long-channel, the potential is constant in the middle of the channel, while in the short-channel, the potential is parabolic as shown in Figure 6.9. In terms of subthreshold currents, *n* and *p* terms can be removed from equation (6.1). Then the equation (6.1) can be rewritten as follows

$$S = \frac{q}{\varepsilon} N_a, \tag{6.9}$$

A quantum correction was introduced to have zero-charge at the Si/SiO_2 interface [6.8]. The poisson equation (6.10) thus is solved:

$$\nabla^2 V = S. \tag{6.10}$$

Figure 6.10 shows the simulated subthreshold currents. From the currents, we extracted the threshold voltage and the DIBL. The threshold voltage was extracted using the current constant method: $V_T = V_G$ at $I_D = 10^{-7} L/W$.

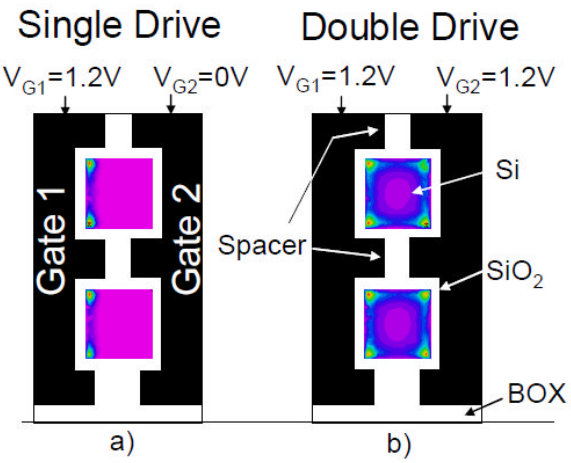


Figure 6.7 Simulated inversion charge density in a ΦFET for (a) one gate activated (single drive mode) and (b) two gates activated (double drive mode).

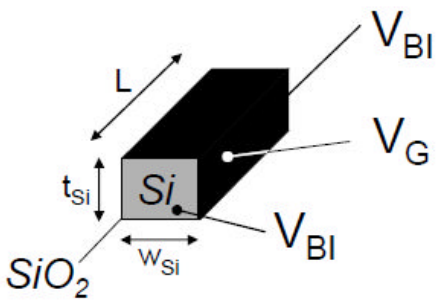


Figure 6.8 Schmatic illustration of a SNWT with boundary conditions.

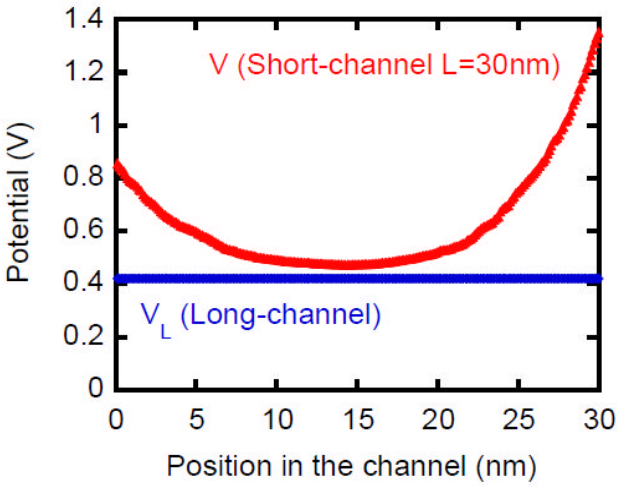


Figure 6.9 Potential along the channel for a long and short-channel transistor.

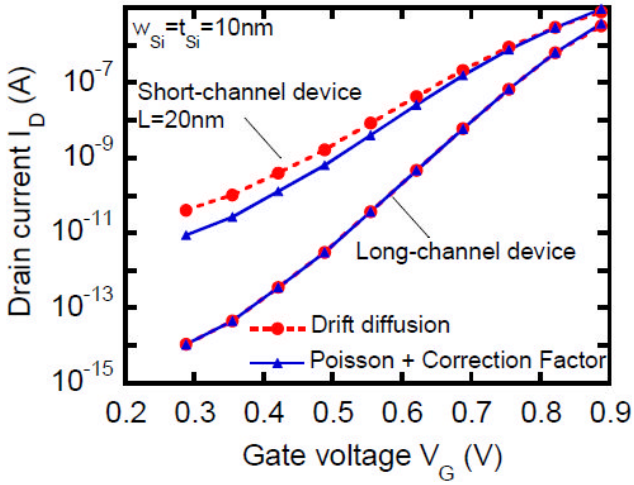


Figure 6.10 Simplified Poisson equation resolution for long-channel with C.F. for short channel is compared to the drift-diffusion model.

As the Φ-FET structure is perfectly symmetric, the coupling factor (α) can be written as the threshold voltage sensitivity [6.9]:

$$\alpha \equiv \frac{\Delta V_{T1}}{\Delta V_{G2}} \approx \frac{C_{Si}}{C_{ox} + C_{Si} + C_{it}}, \tag{6.11}$$

where V_{T1} is the threshold voltage on the side of the gate 1, and V_{G2} is the gate 2 voltage. The coupling factor has been extracted on 2D simulations with long channel. We checked that the coupling is not degraded for shorter gate lengths. For a long-channel, the coupling was evaluated at 0.37 while it is equal to 0.33 at L=10 nm.

From here, the simulation results are discussed. To understand a relationship between flexibility of the threshold voltage and short channel effects immunity, DIBL– α characteristics are plotted as a function of silicon width (Figure 6.11), silicon thickness (Figure 6.12), and spacer width (Figure 6.13). DIBL decreases and α increases when the Si width is reduced as shown in Figure 6.11. For a given W_{Si}, DIBL and α both decrease when changing from an IG-FinFET architecture to a Φ-FET. The coupling decrease is understandable considering the rounded gates shape decreasing the gate coupling compared to the IG-FinFET's straight gates. The gate shape immunity to coupling effects has already been studied [6.10]. The lateral gates can screen narrow silicon body from the other gate influence as shown in Figure 6.14. The coupling decrease is understandable considering the rounded gates shape decreasing the gate coupling compared to the IG-FinFET's straight gates. The t_{Si} variation impact is shown on Figure 6.12. Decreasing the Si thickness improves the architecture electrostatics. The gate coupling is enhanced for the t_{Si} highest value, reaching the coupling value of IG-FinFET at t_{Si} = h_{Si} = 200nm. The spacer width W_{sp} impact is shown on Figure 6.13. For W_{sp} = 29nm, the spacers are aligned with the Si. This structure is equivalent to an IG-FinFET but with oxide and nitride alternatively with Si. Replacing Si by nitride or oxide decreases slightly the DIBL. wsp has to be adjusted to have a low DIBL and a satisfying α value. For $W_{sp}/W_{Si} \approx 0.5$, a good DIBL–coupling trade-off is obtained.

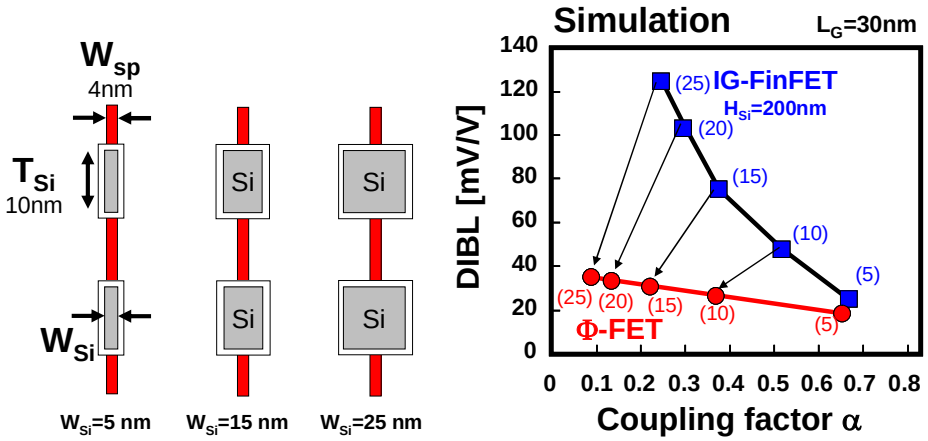


Figure 6.11 DIBL versus coupling factor: Silicon width (W_{Si}) dependence.

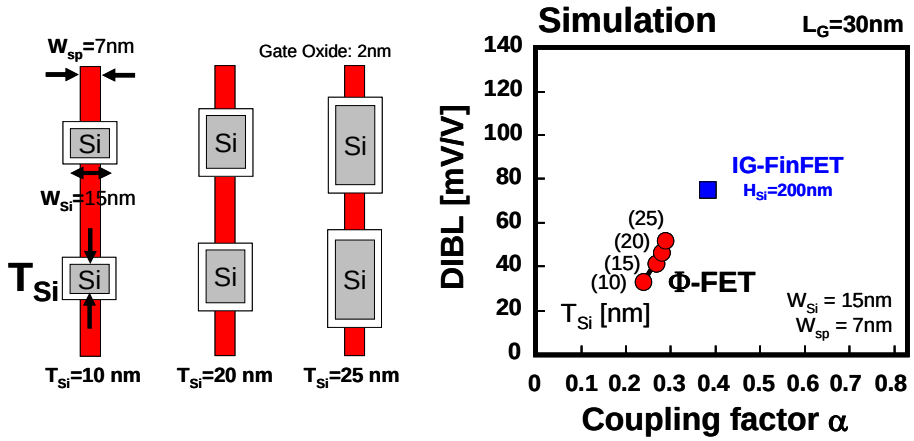


Figure 6.12 DIBL versus coupling factor: Spacer width (T_{Si}) dependence.

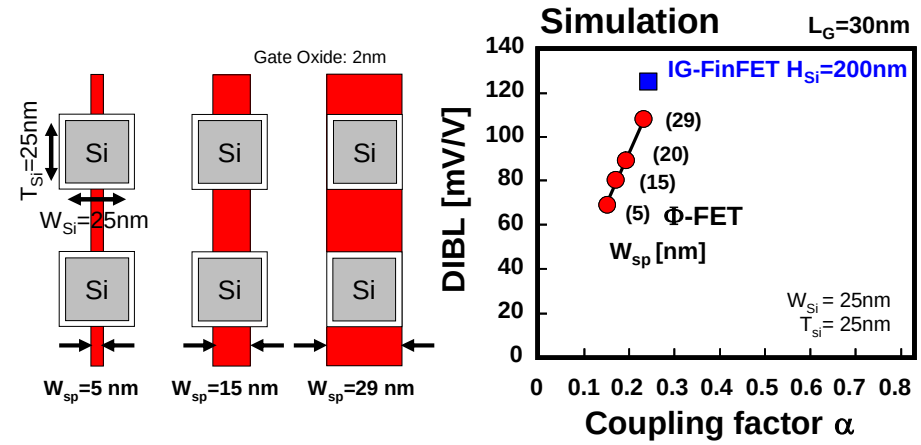


Figure 6.13 DIBL versus coupling factor: Spacer width (W_{sp}) dependence.

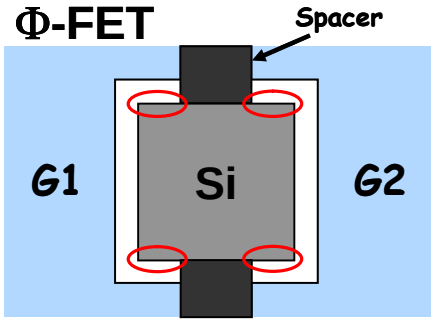


Figure 6.14 Lateral gates can screen narrow silicon body from the other gate influence.

6.3 CONCLUSIONS

In this chapter, vertically-stacked SNWTs architectures with independent gate operation (Φ-FET) have been evaluated by simulation in terms of the flexibility of the threshold voltage and the short channel effects immunity. Φ-FET structure can achieve better short-channel immunity than IG-FinFET owing to the partially surrounded gate structure. On the other hand, the lateral gates screen narrow silicon body from the other gate influence. This causes the degradation of coupling factor. However, the change of cross-sectional dimensions can make the threshold voltage flexible.

The proposed architectures can provide solutions for future technological nodes. It enable to achieve extremely high integration density and low leakage currents with multi-threshold voltage.

6.4 REFERENCES

[6.1] *Y.* X. Liu, M. Masahara, K. Ishii, T. Tsutsumi, T. Sekigawa, H. Takahima, H. Yamauchi and E. Suzuki, "Flexible Threshold Voltage FinFETs with Independent Double Gates and an Ideal Rectangular Cross- Section Si-Fin Channel", *IEEE International Electron Devices Meeting IEDM*, p. 986, 2003.

[6.2] *Y.* X. Liu, M. Masahara, K. Ishii, T. Sekigawa, H. Takahima, H. Yamauchi and E. Suzuki, "A Highly Threshold Voltage-Controllable 4T FinFET with an

8.5-nm-Thick Si-Fin Channel", *IEEE Electron Device Letter*, vol. 25, no. 7, pp. 510–512, 2004.

[6.3] D. Fried, J. Duster, and K. Kornegay, "Improved Independent Gate N-Type FinFET: Fabrication and Characterization", *IEEE Electron Device Letters*, vol. 24, no. 9, p. 592, 2003.

[6.4] K. Endo, Y. Ishikawa, Y. Liu, K. Ishii, T. Matsukawa, S. O'uchi, M. Masahara, E. Sugimata, J. Tsukada, H. Yamauchi, and E.Suzuki, "Four-Terminal FinFETs Fabricated Using an Etch-Back Gate Separation", *IEEE Transactions on Nanotechnology*, vol. 6, no. 2, p. 201,2007.

[6.5] C. Dupré, A. Hubert, S. Bécu, M. Jublot, V. Maffini-Alvaro, C. Vizioz, F. Aussenac, C. Arvet, S. Barnola, J.-M. Hartmann, G. Garnier, F. Allain, J.-P. Colonna, M. Rivoire, L. Baud, S. Pauliac, V. Loup, T. Chevolleau, P. Rivallin, B. Guillaumot, G. Ghibaudo, O. Faynot, T. Ernst, and S. Deleonibus, " 15nm-diameter 3D Stacked Nanowires with Independent Gate Operation: FET," in *IEDM Tech. Dig.*, 2008, pp. 749–752.

[6.6] A. Tsormpatzoglou, C. A. Dimitriadis, R. Clerc, Q. Rafhay, G. Pananakakis and G. Ghibaudo, "Semi-Analytical Modeling of Short-Channel Effects in Si and Ge Symmetrical Double-Gate MOSFETs", *IEEE Transactions on Electron Devices*, vol. 54, no. 8, p. 1943, 2007.

[6.7] R.-H. Yan, A. Ourmazd, K. Lee, "Scaling the Si MOSFET: From Bulk to SO1 to Bulk", *IEEE Transactions on Electron Devices*, vol. 39, no. 7, p. 1704, 1992.

[6.8] W. Hansch, Th. Vogelsang, R. Kircher, M. Orlowski, "Carrier transport near the Si/SiO2 interface of a MOSFET", *Solid-State Electronics*, Vol. 32, no.10, p.839, 1989.

[6.9] H.K. Lim and J. G. Fossum, "Threshold Voltage of Thin-Film Silicon-on-Insulator (Sol) MOSFET's, *IEEE Transactions on Electron Devices*, vol. 30, no. 10, p. 1244, 1983.

[6.10] R. Ritzenthaler, S. Cristoloveanu, O. Faynot, C. Jahan, A. Kuriyama, L. Brevard, S. Deleonibus, "Lateral coupling and immunity to substrate effect in sFET devices", *Solid-State Electronics*, vol. 50, p. 558, 2006.

CHAPTER 7

CONCLUSIONS

CHAPTER 7 CONTENTS

6.1 SUMMARY

In this thesis, to achieve both high speed and low power consumption with high integration for future LSI applications, vertically-stacked silicon nanowire MOSFETs (SNWTs) have been experimentally investigated as one of the possible solutions to various problems related to the CMOS scaling explained in Chapter 1.

The contributions of this work can be divided into three main subjects:

(1) Study of source/drain series resistance for thick source/drain region in vertically-stacked channel MOSFETs (Chapter 4),
(2) Study of carrier transport in vertically-stacked SNWTs (Chapter 5),
(3) Study of threshold voltage controllability for vertically-stacked SNWTs with separated gates (Chapter 6).

In Chapter 2, the device fabrication process was described. Vertically-stacked channel MOSFETs have been successfully fabricated by adapting Silicon-On-Nothing technology with sacrificial SiGe layers.

In Chapter 3, the electrical characterization methods was described. In order to analyze the performances of the fabricated devices, the intrinsic parameters extraction methods, Y-function method and Split C–V, were detailed in the latter half.

In Chapter 4, the influence of *in situ* doped SEG source/drain has been examined for vertically-stacked channel MOSFETs. A large enhancement, by a factor of 2 in the drive current, can be obtained when *in situ* doped SEG process is adopted. Detailed parameter extraction from the electrical measurements shows that the R_{SD} values can be reduced by 90 and 75% for *n*- and *p*-FETs, respectively, when in situ doped SEG is reinforced by adding ion implantation. On the other hand, V_T roll-off characteristics and the effective mobility behavior are slightly degraded, especially when ion implantation is combined to the SEG process. Mobility analysis has revealed an increase in the

Coulomb scattering with L_G scaling, indicating the diffusion of dopant atoms from S/D regions. These results indicate an avenue to further improve the performance by optimizing the S/D activation annealing step.

In chapter 5, the carrier transport limiting components for vertically-stacked nanowire MOSFETs have been discussed to obtain better performance with suppressing short channel effects. The optimization of drive currents will have to take into account specific effects to vertically-stacked SNWTs. In particular, the use of SiGe sacrificial layer to make vertically-stacked channels cause the large mobility degradation due to the surface roughness, resulting from the damage of plasma etching. This result can evidence the poor ballisticity in the short channel SNWTs.

The hydrogen annealing can improve the surface-roughness limited mobility a little. Charge pumping measurements, however, revealed that circular-shaped SNWTs, which are formed by the annealing, have a higher interface trap density (D_{it}) than rectangular ones, leading to low-field mobility degradation. This high D_{it} might be caused by the continuously-varying surface orientation. The resulting additional coulomb scattering could partly explain the quite low mobility in 5 nm diameter SNWTs together with the already known transport limitations in NWs.

In chapter 6, vertically-stacked SNWTs with independent gates by internal spacers between the nanowires to control V_T, which is named Φ-FETs, have been evaluated. Φ-FETs demonstrated excellent V_T control by inter-gate coupling effects. As the results of numerical simulations to optimize Φ-FETs structures, it have been found that when the spacer width is reduced, the DIBL value can be lowered by a factor of 2 compared to independent-gate FinFETs with the same silicon width. The superior scaling of Φ-FETs with narrow spacer results from a better electrostatic control which also attenuates the inter-gate coupling.

6.2 CONCLUSIONS AND PERSPECTIVE

In this thesis, it have been demonstrated that gate-all-around silicon nanowire structure can dramatically suppress short-channel effects. Moreover, the introduction of internal spacers between the nanowires can control threshold voltage. These technologies enable to achieve ultra-low power consumption.

In order to obtain high speed operation, the carrier transport limiting components in vertically-stacked SNWTs have been investigated in detail. In addition, the study of mechanical stress to the nanowires indicates a guide of mobility enhancement.

Vertically-stacked channel structure has yielded extremely high drive current density per top-viewed channel width compared to planer MOSFETs. However, this high current may not connect to intrinsic delay reduction because the parasitic capacitance increases in proportion to the number of channels. A benefit of the structure is a possibility of ultra-high integration per given layout area, that is, gate width scaling. For SNWTs, the optimization of the integration is strongly limited by horizontal spaces between the nanowires. The use of vertically stacked channel structure without any nanowires in parallel enables to cancel this limitation. To realize this structure, it is necessary to increase the number of channels in vertical direction. To do that, various process developments are needed such as (Si/SiGe) x *n* superlattice formation shown in Figure 7.1 and its etching with vertically straight line edge.

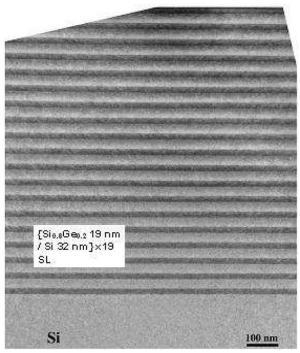


Figure 7.1 Cross-sectional TEM image of the 19 period superlattice with 19 nm $Si_{0.8}Ge_{0.2}$ and 32 nm of Si [fabricated by J.M. Hartmann in CEA-Leti].

Publications and Presentations

Nanowire MOSFETs

[Publications as 1st author]

1. K. Tachi, N, Vulliet, S. Barraud, K. Kakushima, H. Iwai, S. Cristoloveanu and T. Ernst, "Influence of source/drain formation process on resistance and effective mobility for scaled multi-channel MOSFET," Special Issue of Solid-State Electronics, (accepted for publication)

2. K. Tachi, S. Barraud, K. Kakushima, H. Iwai, S. Cristoloveanu and T. Ernst, "Comparison of low-temperature electrical characteristics of gate-all-around nanowire FETs, Fin FETs and fully-depleted SOI FETs," Microelectronics Reliability, (accepted for publication)

[Presentations as 1st author]

3. K. Tachi, M. Cassé, S. Barraud, C. Dupré, A. Hubert, N. Vulliet, M.E. Faivre, C. Vizioz, C. Carabasse, V. Delaye, J.M. Hartmann, H. Iwai, S. Cristoloveanu, O. Faynot and T. Ernst, "Experimental study on carrier transport limiting phenomena in 10 nm width nanowire CMOS transistors," *IEDM Tech. Dig.*, p.784, 2010, San Francisco, US.

4. K. Tachi, N. Vulliet, S. Barraud, B. Guillaumot, V. Maffini-Alvaro, C. Vizioz, C. Arvet, Y. Campidelli, P. Gautier, J.M. Hartmann, T. Skotnicki, S. Cristoloveanu, H. Iwai, O. Faynot and T. Ernst, "3D Source/Drain Doping Optimization in Multi-Channel MOSFET," 40th European Solid-State Device Research Conference (ESSDERC), pp. 368, 2010, Spain.

5. K. Tachi, M. Casse, D. Jang, C. Dupré, A. Hubert, N. Vulliet, V. Maffini-Alvaro, C. Vizioz, C. Carabasse, V. Delaye, J. M. Hartmann, G. Ghibaudo, H. Iwai, S. Cristoloveanu, O. Faynot, and T. Ernst, "Relationship between mobility and

high-*k* interface properties in advanced Si and SiGe nanowires," *IEDM Tech. Dig.*, p.313, 2009, Baltimore, US.

5. K. Tachi, T. Ernst, C. Dupré, A. Hubert, S. Bécu, H. Iwai, S. Cristoloveanu, O. Faynot, " Transport Optimization with Width Dependence of 3D-stacked GAA Silicon Nanowire FET with High-*k*/Metal Gate Stack," *2009 IEEE Silicon Nanoelectronics Workshop*, p.11, 2009, Kyoto, Japan.

6. K. Tachi, T. Ernst, C. Dupré, A. Hubert, S. Bécu, H. Iwai, S. Cristoloveanu, O. Faynot, " F-FET: Electrostatics and influence of spacer thickness," 5th *EUROSOI*, 2009, Chalmers University of Technology, Sweden.

[Publications as co- author]

7. D. Jang, J. W. Lee, **K. Tachi**, L. Montes, T. Ernst, G. T. Kim and G. Ghibaudo, "Low-frequency noise in strained SiGe core-shell nanowire p-channel field effect transistors, " Appl. Phy. Lett. 97, 073505, 2010.

8. M. Casse, **K. Tachi**, S. Thiele and T. Ernst, "Spectroscopic charge pumping in Si nanowire transistors with a high-k/metal gate, " Appl. Phy. Lett. 96, 123506, 2010.

9. T. Ernst, **K. Tachi**, E. Saracco, A. Hubert, C. Dupré, S. Bécu, N. Vulliet, E. Bernard, P. Cherns, V. Maffini-Alvaro (CEA-LETI), J. Damlencourt, C. Vizioz, J. Colonna, C. Bonafos and J. Hartmann, "3D suspended nanowires integration for CMOS and beyond," *Electrochemical Society (ECS) Transaction*, vol. 25, Issue.7 , pp. 471-478, 2009.

[Presentations as co-author]

10. K-I. Na, W. Van Den Daele, L. Pham-Nguyen, M. Bawedin, K-H. Park, J. Wan, **K. Tachi**, S-J. Chang,1 I. Ionica, Y-H. Bae, J.A. Chroboczek, C. Fenouillet-Beranger, T. Ernst, E. Augendre, C. Le Royer, A. Zaslavsky, H. Iwai and S. Cristoloveanu, "(Invited) Selected SOI puzzles and tentative answers, " 6th SemOI workshop & 1st Ukrainian-French SOI Seminar, 2010, Ukraine.

11. A. Hubert, E. Nowak, **K. Tachi**, V. Maffini-Alvaro, C. Vizioz, C. Arvet, J.-P. Colonna, J.-M. Hartmann, V. Loup, L. Baud, S. Pauliac, V. Delaye, C. Carabasse, G. Molas, G. Ghibaudo, B. De Salvo, O. Faynot and T. Ernst, "A stacked SONOS technology, up to 4 levels and 6nm crystalline nanowires, with gate-all-around or independent gates (Φ -Flash), suitable for full 3D integration," *IEDM Tech. Dig.*, p.673, 2009, Baltimore, US.

12. T. Ernst, S. Bécu, P. Cherns, L. Durrafourg, C. Dupré, P. Andreucci, A. Hubert, J.-M. Hartmann, J.-P. Colonna, M. Jublot, G. Delapierre, V. Delaye, **K. Tachi**, E. Saracco, J.-F. Damlencourt, J. Foucher, and O. Faynot, "A 3D stacked nanowire technology - Applications in advanced CMOS and beyond," 7th International Conference on Frontiers of Characterization and Metrology for Nanoelectronics, 2009, US.

13. P. D. Cherns, F. Lorut, S. Beçu, C. Dupré, **K. Tachi**, D. Cooper, A. Chabli, and T. Ernst, "A Study Of Gate-All-Around Transistors By Electron Tomography," AIP Conf. Proc. 1173, 290, 2009.

High-k/Metal Gate

[Publications as 1st author]

14. K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Improvement of interface properties of W/La2O3/Si MOS structure using Al capping layer," *Electrochemical Society (ECS) Transaction* Vol.11 No.4 pp.191-198, 2007.

15. K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Effect of Oxygen for Ultra-Thin La2O3 Film Deposition," *Electrochemical Society (ECS) Transaction 3 (3)*, pp. 425-435, 2006.

[Presentations as 1st author]

16. K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Improvement of interface properties of W/La2O3/Si MOS structure using Al capping layer," *212th Electrochemical Society (ECS) Meeting*, Abs#E4-0752O, October, 2007, Washington, DC.

17. K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Effect of Oxygen for Ultra-Thin La2O3 Film Deposition," *Physics and Technology of High-k Gate Dielectrics 4, 210th Electrochemical Society (ECS) Meeting*, Abs#1128, November, 2006, Cancun, Mexico.

[Publications and Presentations as co-author]

18. K. Kakushima, K. Okamoto, T. Koyanagi, K. Tachi, M. Kouda, T. Kawanago, J. Song,P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Selection of Rare Earth Silicate with SrO Capping for EOT Scaling below 0.5 nm," 39th ESSDERC, pp 403, 2009, Greece.

19. K. Kakushima, K. Tachi, J. Song, S. Sato, H. Nohira, E. Ikenaga, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori, and H. Iwai, " Comprehensive x-ray photoelectron

spectroscopy study on compositional gradient lanthanum silicate film," Journal of Applied Physics, vol. 106, 124903, 2009.

20. T. Koyanagi, K. Tachi, K. Okamoto, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori, and H. Iwai, "Electrical Characterization of La2O3-Gated Metal Oxide Semiconductor Field Effect Transistor with Mg Incorporation," Japanese Journal of Applied Physics, vol. 48, 05DC02, 2009.
21. K. Kakushima, K. Okamoto, K. Tachi, J. Song, S. Sato, T. Kawanago, K. Tsutsui, N. Sugii, P. Ahmet, T. Hattori, and H. Iwai, "Observation of band bending of metal/high-*k* Si capacitor with high energy x-ray photoemission spectroscopy and its application to interface dipole measurement," Journal of Applied Physics, vol. 104, 104908, 2008.

21. M. Kouda, K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Electrical Properties of CeOX /La2O3 Stack as a Gate Dielectric for Advanced MOSFET Technology," *Electrochemical Society (ECS) Transaction*, vol. 16, Issue.5 , pp. 153-160, 2008.

22. K. Kakushima, K. Okamoto, M. Adachi, K. Tachi, S. SatoT. Kawanago, , J. Song, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori, and H. Iwai, "Impact of Thin La2O3 Insertion for HfO2 MOSFET," *Electrochemical Society (ECS) Transaction*, vol. 13, Issue.2 , p. 29, 2008.

23. K. Kakushima, K. Tachi, M. Adachi, K. Okamoto, S. Sato, J. Song, T. Kawanago, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori, and H. Iwai, "Advantage of La2O3 gate dielectric over HfO2 for direct contact and mobility improvment," 38th ESSDERC, p.126, 2008.

24. K. Kakushima, K. Okamoto, K. Tachi, S. Sato, T. Kawanago, J. Song, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori, and H. Iwai, "Further EOT Scaling below 0.4

nm for High-k Gated MOSFET," International Workshop on Dielectric Thin Films for future ULSI devices, p. 9, 2008, Japan.

25. T. Koyanagi, K. Tachi, K. Okamoto, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori, and H. Iwai, "Electrical Characterization of La2O3-Gated MOSFET with Mg Incorporation," International Workshop on Dielectric Thin Films for future ULSI devices, p. 73, 2008, Japan.

26. J. Molina, K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Effects of N2-Based Annealing on the Reliability Characteristics of Tungsten/La2O3/Silicon Capacitors," *J. Electrochem. Soc.*, Volume 154, Issue 5, pp. G110-G116, 2007

27. T. Kawanago, K. Tachi, J. Song, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai "Electrical Characterization of Directly Deposited La-Sc Oxide Complex for Gate Insulator Application," Microelectronic Engineering, vol. 84, issue. 9-10, pp. 2235-2238, 2007.

29. S. Sato, K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii , T. Hattori and H. Iwai "Thermal-stability improvement of LaON thin film formed using nitrogen radicals," Microelectronic Engineering, vol. 84, issue. 9-10, pp. 1894-1897, 2007.

30. S. Sato, K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Thermal Stability of Lanthanum Oxynitride Thin Film Fabricated by in-situ Radical Nirtidation," *54th Japan Society of Applied Physics (JSAP) Spring Meeting*, 29p-ZH-13, 2007, Aoyama Gakuin University, Japan.

31. K. Kakushima, P. Ahmet, J.-A. Ng, J. Molina, H. Sauddin, Y. Kuroki, K. Nakagawa, A. Fukuyama, K. Tachi, Y. Shiino, J. Song, K. Tsutsui, N. Sugii, T.

Hattori and H. Iwai, "Study of La2O3 Gate Dielectric Suitability for Future MIM and MOSFETs", *2006 IEEE Silicon Nanoelectronics Workshop*, 12th June 2006, Hawaii, USA.

32. H. Nohira, T. Matsuda, K. Tachi, Y. Shiino, J. Song, Y. Kuroki, J. Ng, P. Ahmet, K. Kakushima, K. Tsutsui, E. Ikenaga, K. Kobayashi, H. Iwai and T. Hattori, "Effect of Deposition Temperature on Chemical Structure of Lanthanum Oxide/Si Interface Structure," 2006 International Workshop on Dielectric Thin Films for Future ULSI Devices (IWDTF), pp.29-30, November, 2006, Kawasaki, Japan

33. J. Molina, K. Tachi, K. Kakushima, P. Ahmet, K. Tsutsui, N. Sugii, T. Hattori and H. Iwai, "Effects of N2-Based Annealing on the Reliability Characteristics of Tungsten/La2O3/Silicon Capacitors," *Electrochemical Society (ECS) Transaction 3 (3)*, pp. 233, 2006.

34. H. Nohira, T. Matsuda, K. Tachi, Y. Shiino, J. Song, Y. Kuroki, J. Ng, P. Ahmet, K. Kakushima, K. Tsutsui, E. Ikenaga, K. Kobayashi, H. Iwai and T. Hattori, "Effect of Deposition Temperature on Chemical Structure of Lanthanum Oxide/Si Interface Structure," *Electrochemical Society (ECS) Transaction 3 (2)*, pp. 169, 2006.

35. H. Nohira, T. Matsuda, K. Tachi, Y. Shiino, J. Song, Y. Kuroki, J. Ng, P. Ahmet, K. Kakushima, K. Tsutsui, E. Ikenaga, K. Kobayashi, H. Iwai and T. Hattori, "Thermal Stability of LaOX/Si Interfacial Transition Layer", *53rd Japan Society of Applied Physics (JSAP) Spring Meeting*, 23p-V-2, 2006, Musashi Institute of Technology, Japan.

36. K. Nakajima, K. Kimura, K. Tachi, K. Kakushima and H. Iwai, "Hydrogen Depth Profiling in La2O3/Si(001) by High-resolution Erastic Recoil Detection",

53rd Japan Society of Applied Physics (JSAP) Spring Meeting, 23p-V-5, 2006, Musashi Institute of Technology, Japan.

Awards

2010 IEEE EDS Japan Chapter Student Award

2009 IEEE EDS Japan Chapter Student Award