國立交通大學

電子工程學系　　電子研究所

碩 士 論 文

利用實驗驗證場效電晶體之汲極與源極之遠距庫倫效應

Experimental Evidence for MOSFET S/D Long-Range Coulomb Effects

研究生: 李致葳 Chih-Wei Lee

指導教授: 陳明哲 博士 Prof. Ming-Jer Chen

中華民國 一０一年 十 月

利用實驗驗證場效電晶體之汲極與源極之遠距庫倫效應

Experimental Evidence for MOSFET S/D Long-Range Coulomb Effects

研究生: 李致葳 Chih-Wei Lee

指導教授: 陳明哲 博士 Prof. Ming-Jer Chen

國立交通大學

電子工程學系 電子研究所

碩士論文

A Thesis

Submitted to Department of Electronics Engineering &

Institute of Electronics

College of Electrical and Computer Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Electronics Engineering

October 2012

Hsinchu, Taiwan, Republic of China

中華民國 一0一 年 十 月

利用實驗驗證場效電晶體之汲極與源極之遠距庫倫效應

研究生: 李致葳　　指導教授: 陳明哲 博士

國立交通大學

電子工程學系 電子研究所碩士班

摘要

由近年研究得知，元件尺寸縮減時電子遷移率會伴隨遞減，這也指出了有額外的碰撞機制存在，並且此機制會對下一世代的元件造成很大的影響。因此本篇論文主旨係利用實驗萃取額外遷移率之溫度係數，進而探討N型超短通道場效電晶體下的散射機制。研究內容主要為我們第一次提出實驗證據對於當元件實際通道長度小於 40 奈米會被存在於高濃度的源極與汲極的電漿電子所造成的遠距庫倫散射機制所影響。這一系列的的研究方法為透過載子遷移率的溫度效應以及利用二維模擬器對元件建立的模型來取得重要參數。此外，我們也提供了另一項證據，是我們在大汲極電壓下量測到的轉導值與文獻中考慮遠距庫倫效應下的模擬值相符。

Experimental Evidence for MOSFET S/D Long-Range Coulomb Effects

Student: Chih-Wei Lee Advisor: Dr. Ming-Jer Chen

Department of Electronics Engineering and Institute of Electronics

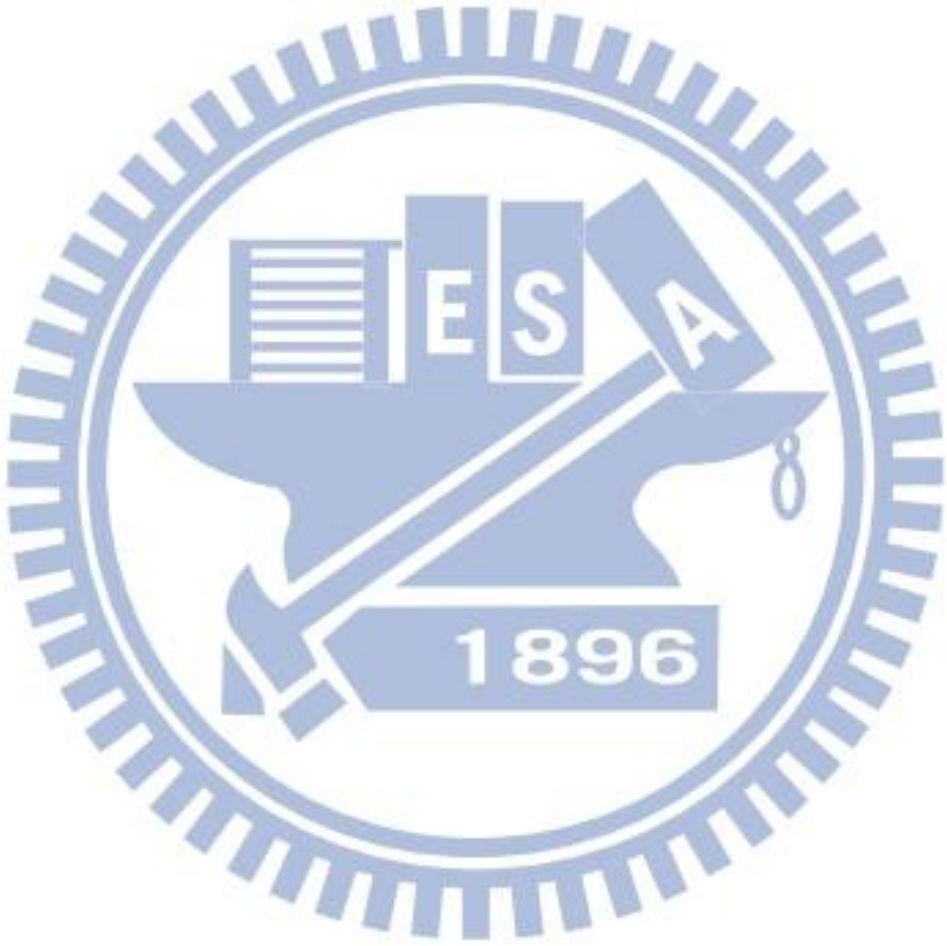
National Chiao Tung University

Abstract

Electron mobility degradation is currently encountered in highly scaled devices. This means that additional scattering mechanisms exist and will become profoundly important in next generation of devices. The aim of this work is to, for first time, present experimental evidence for the existence of long-range Coulomb effects due to plasmons (collective behaviors of fluctuating dipoles) in high-density source/drain (S/D) of MOSFETs, particularly for the metallurgical channel length less than about 40 nm. This is obtained through temperature-dependent mobilities via TCAD-based inverse modeling. Other evidence is further produced in terms of the measured transconductance at high drain voltage, which is comparable with that of sophisticated simulations in the literature taking into account long-range Coulomb interactions.

Acknowledgements

感謝所有曾經幫助我的所有人。



Contents

Chinese Abstract I

Abstract II

Acknowledgements III

Contents IV

Figure Captions VI

Table Captions IX

Chapter 1 Introduction 1

Chapter 2 Experiment 3

2.1 *C-V* Fitting 3

2.2 Measurement Method and Experimentally Assessed Effective Inversion-Layer Mobility 4

Chapter 3 Inverse Modeling 6

3.1 Drift-diffusion Model 6

3.2 Calculation of Inversion Layer Charge Density 8

3.3 Extraction of Parasitic Source/Drain Resistance 9

Chapter 4 Analysis and Discussion 12

4.1 Additional Mobilities 12

4.2 Main Source of Mobility Degradation in Short-channel Device 13

4.3 Evidence of Long-range Coulomb Interactions 15

Chapter 5 Conclusion 16

References 17

Figures 17

Figure Captions

Figure 1.1 Schematic of electron transport under long-range Coulomb interactions with S/D plasmons. 19

Figure 1.2 Flowchart of inverse modeling in this work. 20

Figure 2.1 Temperature-dependent terminal currents at V_d=0.05V versus V_g for L_g=1μm. 21

Figure 2.2 Temperature-dependent terminal currents at V_d=0.05V versus V_g for L_g=65nm. 22

Figure 2.3 Temperature-dependent terminal currents at V_d=0.05V versus V_g for L_g=50nm. 23

Figure 2.4 Comparison of the measured and simulated gate capacitance versus gate voltage. 24

Figure 2. 5 TEM image of the sample. 25

Figure 2.6 The schematic diagram for current flow of nMOSFETs with large gate tunneling current. Besides, I_S<0 and I_d>0. 26

Figure 3.1 Measured and simulated I_{ch} versus V_g at T=292K and V_d=0.05 and 1V. 27

Figure 3.2 Measured and simulated I_{ch} versus V_g at T=330K and $V_{d=}$0.05 and 1V. 28

Figure 3.3 Measured and simulated I_{ch} versus V_g at T=360K and V_d=0.05 and 1V. 29

Figure 3.4 Measured and simulated I_{ch} versus V_g at T=380K and V_d=0.05 and 1V. 30

Figure 3.5 Calibrated 2D simulation structure for L_g=1μm and hence L_m=1μm. W_D is the width of the mid-channel depletion region. 31

Figure 3.6 Calibrated 2D simulation structure for L_g=65nm and hence L_m=48nm. W_D is the width of the mid-channel depletion region. 32

Figure 3.7 Calibrated 2D simulation structure for L_g=50nm and hence L_m=33nm. W_D is the width of the mid-channel depletion region. 33

Figure 3.8 The free electron density under the surface 30nm at one position of the channel region. 34

Figure 3.9 The schematic diagram for inversion layer charge density of metallurgical length L_m=1μ m.....35

Figure 3.10 The schematic diagram for inversion layer charge density of metallurgical length L_m=48nm.....36

Figure 3.11 The schematic diagram for inversion layer charge density of metallurgical length L_m=33nm.....37

Figure 3.12 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 1μm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K. 38

Figure 3.13 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 48nm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K. 40

Figure 3.14 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 33nm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K. 42

Figure 3.15 Simulated N_{inv} versus gate voltage. 44

Figure 3.16 Simulation structure for R_{sd} assessment. 45

Figure 3.17 Simulated R_{sd} versus V_g at T=292K. 46

Figure 3.18 Terminal currents for short-channel device versus V_g at (a)V_b=0V and (b)V_b=-0.4V.....47

Figure 3.19 The extracted R_{sd} by using the experimental method [13]. 48

Figure 3.20 Extracted temperature-dependent effective mobility versus N_{inv} for different L_m with R_{sd} = (a) 100Ω-μm and (b) 120Ω-μm. 49

Figure 4.1 Extracted temperature-dependent additional mobility versus N_{inv} for L_m=48

and 33nm with R_{sd} = (a) 100Ω-μm and (b) 120Ω-μm. 50

Figure 4.2 Extracted additional mobility at N_{inv}=1x$10^{13}$$cm^{-2}$ versus temperature with R_{sd} as a parameter. The power-law coefficient γ is obtained by data fitting. 51

Figure 4.3 Extracted temperature-dependent $\mu_{add,\ extra}$ versus N_{inv} with R_{sd} as a parameter.....52

Figure 4.4 Extracted $\mu_{add,\ extra}$ at N_{inv}=1x$10^{13}$$cm^{-2}$ versus temperature, along with corresponding power-law coefficient γ.....53

Figure 4.5 Simulated conduction-band edge, subband levels, and Fermi level versus position for L_m = 1μm at V_g = (a) 1V, (b) 1.2V, and (c) 1.5V. 54

Figure 4.6 Simulated conduction-band edge, subband levels, and Fermi level versus position for L_m = 48nm at V_g = (a) 1V, (b) 1.2V, and (c) 1.5V.....55

Figure 4.7 Simulated conduction-band edge, subband levels, and Fermi level versus position for L_m = 33nm at V_g = (a) 1V, (b) 1.2V, and (c) 1.5V. 56

Figure 4.8 Comparison of temperature power-law exponent of extracted $\mu_{add,\ extra}$ and simulated bulk phonon-limited mobility, plotted versus N_{inv}. 57

Figure 4.9 Comparison of measured transconductance in this work and simulated ones with and without Coulomb effects [1]. 58

Table Captions

Table 1 The length of the overlap region for different L_{gate}, and the metallurgical channel length ($L_m=L_{gate}-\Delta L$). 59

Table 2 The subband population for different metallurgical channel length at gate voltages of (a) 1V, (b) 1.2V, and (c) 1.5V. 60

Chapter 1

Introduction

It is well known that the inversion layer mobility of nMOSFETs can be limited to three primary scattering mechanisms: (1) Coulomb impurity scattering due to ionized impurity atoms in substrate depletion region; (2) acoustic and optical phonon scattering in inversion channel region; and (3) surface roughness scattering at the SiO_2/Si interface. However, owing to the measured effective mobility degradation in highly scaled devices, this means that additional scattering mechanisms exist. Additional scatterings can be generalized in terms of remote surface roughness, fixed oxide charge, remote Coulomb scattering, short-range Coulomb centers due to halo implant or pockets and/or defects near S/D, and remote phonon scattering due to surface optical (SO) phonons. In spite of so many additional scattering mechanisms, we only need to consider specific scattering mechanisms in short-channel device, which are absent in long-channel device since scattering coming from gate dielectrics above channel should be all the same for both devices.

So far, in the open literature dedicated to short-channel nMOSFETs, the origins of mobility degradation remained controversial. While channel length is shrunk into the region of Thomas-Fermi screening length (like sub-40 nm device), electron interactions near highly doped source/drain (S/D) - channel interface can be viewed as the dynamic screening, resulting in the excitation plasmons emission or absorption, which in turn transfer the momentum to affect the current indirectly according to the published simulation results [1],[2]. In addition, since the channel doping concentration is significantly higher in deep submicron devices compared to long-channel devices due to the presence of halo doping, it is expected to result in higher transversal fields and degraded mobility in short-channel devices [3]. Besides,

the mobility degradation of short-channel device is also attributed to process-induced defects located near the source and drain junctions [4],[5]. Moreover, 2D charge sharing from S/D is also not neglected in short-channel device [3],[6], but it can be excluded in this work because the halo implant is used in our device sample under study.

Of degradation mechanisms mentioned above, the most challenging stems from long-range Coulomb interactions between electrons in channel and plasmons in such high-density regions as gate, source, and drain [1],[7]. However, in the past, the understanding of such long-range effects was primarily through sophisticated simulations [1],[7]. Only recently, we experimentally separated gate-plasmon-limited mobility [8], which can provide evidence for the interaction between channel electrons and gate plasmons.

In this work, we further present experimental evidence for another interaction between channel electrons and S/D plasmons (see Figure 1.1). The methodology is shown in Figure 1.2.

Chapter 2

Experiment

N-channel devices under study were fabricated in a conventional manufacturing process with different channel width (W) and length (L) of W/L=1/1 μm, W/L=1/0.065 μm and W/L=1/0.05 μm. In this process, SiO_2 film was thermally grown on (001) surface, followed by NO annealing. Figure 2.1 to Figure 2.3 show I-V characteristics of device measured at V_d=0.05V for different temperatures (T = 292, 330, 360, and 380K).

2.1 C-V Fitting

We used ICS software to control HP4284 while measuring *C-V* characteristics. The *C-V* measurement was performed on a 10μm×1μm device. Corresponding process parameters can essentially be obtained by fitting the measured gate capacitance versus gate voltage (C_g-V_g) as shown in Figure 2.4. This was realized with the use of a self-consistent Schrödinger and Poisson's equations solver. Two such solvers were cited: one named Schred [9] and the other in the team's work [10]. Obviously, the two sources [9],[10] are consistent each other in the data fitting. Although C_g-V_g data at high gate voltages were seriously distorted due to the use of ultrathin gate oxide where the direct tunneling current is pronouncedly large, the fitting was successfully done in the remaining regions, leading to n^+ polysilicon doping concentration = 1×10^{20} cm^{-3}, gate oxide (SiO_2) physical thickness = 1.27 nm, and p-type substrate doping concentration = 4×10^{17} cm^{-3}. In addition, the TEM of the device also shows gate oxide (SiO_2) thickness of around 1.3nm as shown in Figure 2. 5.

2.2 Measurement Method and Experimentally Assessed Effective Inversion-Layer Mobility

In the measurement method, the conventional inversion-layer mobility is usually extracted according to

$$\mu_{eff}(V_g) = \frac{L}{W} \times \frac{I_d(V_g)}{V_d} \times \frac{I}{qN_{inv}(V_g)} \tag{2.1}$$

$$G_d = \frac{I_d(V_g)}{V_d} \tag{2.2}$$

where G_d is drain conductance. When I_g is small sufficiently, G_d can be presented by $\frac{I_{ch}}{V_d}$, and the channel current (I_{ch}) should be the same as drain current (I_d) and source current (I_s). However, while gate oxides are thin enough to encounter direct tunneling current in long channel device, high gate leakage current would affect the accurate determination of drain conductance (G_d). Hence, the G_d is difficult to define simply by $\frac{I_s}{V_d}$ or $\frac{I_d}{V_d}$.

By following the experimental work by Takagi, et al. [11], we can know that while gate oxide thicknesses are quite thin, the amount of gate current (I_g) affects source current (I_s) and drain current (I_d), giving rise to the opposite sign. The schematic diagram of current flow in MOSFETs with high gate leakage current is shown in Figure 2.6 and I_s and I_d can be written as

$$I_S = I_{channel} + I_{GS} \tag{2.3}$$

$$I_D = I_{channel} - I_{GD} \tag{2.4}$$

where I_{GS} and I_{GD} are the current from the source to the gate and the current from the drain to gate, besides Is<0 and I_d>0. Figure 2.6 illustrated that I_s is larger than the current from the source into the channel (I_{ch}) due to the current flows from source to

gate(I_{GS}), but I_d is smaller than the current from channel to gate (I_{ch}) because the current tunnels from drain into gate (I_{GD}). When V_d is sufficiently small, the I_{GS} must be the same as the I_{GD}, hence the channel current can be defined as $I_{channel} = \frac{I_d + I_s}{2}$. Therefore, inversion layer (channel) mobility is measured by

$$\mu_{eff}(V_g) = \frac{L}{W} \times \frac{(I_d(V_g) + I_s(V_g))}{2V_d} \times \frac{I}{qN_{inv}(V_g)} \tag{2.5}$$

Considering the difference between L_{mask} and the metallurgical channel length L_m, as well as the issue about the parasitic source/drain resistance (R_{sd}), Eq.(2.5) could be defined more accurately by

$$\mu_{eff}(V_g) = \frac{1}{(R_{ch})} \times \frac{(L_{mask} - \Delta L)}{qN_{inv}(V_g)} \tag{2.6}$$

$$R_{ch} = \frac{2V_{ds}}{I_d + I_s} - R_{sd} \tag{2.7}$$

where L_{mask} is the gate length on the polysilicon etch mask: $L_{mask} \approx L_{gate}$. $L_m = L_{mask} - \Delta L$, where ΔL is the length of the overlap region between the source/drain implant straggle and diffusion. Table 1 shows the overlap length for different L_{gate}. We will discuss in next chapter the N_{inv} and R_{sd} extraction.

Chapter 3

Inverse Modeling

In this section, we make use of a TCAD tool named Sentaurus to reproduce I_{ch}-V_g, especially for the subthreshold conduction. This procedure is so-called inverse modeling. First, we use the parameters which is got from *C-V* fitting: n^+ polysilicon doping concentration N_{poly} (= 1×10^{20} cm^{-3}), gate oxide (SiO_2) physical thickness t_{ox} (= 1.27 nm), and p-type substrate doping concentration P_{sub} (= 4×10^{17} cm^{-3}). Second, we add two extra peak doping impurities: the source/drain extension N_{sde} and the halo implant P_{halo}. Remarkably, a fairly good fitting was achieved for different gate lengths, different temperatures, and different drain voltages, all with the same source/drain extension doping concentration N_{sde} (= 4.95×10^{20} cm^{-3}) and halo implant P_{halo} (= 2.5×10^{19} cm^{-3}). Figure 3.1 to Figure 3.4 show I_{ch}-V_g fitting results. I_{ch}-V_g calibration leads to device doping profiles for L_m=1μm, 48nm and 33nm as shown in Figure 3.5 to Figure 3.7. Corresponding extension overlap and hence the metallurgical channel length L_m can be drawn.

TCAD further delivers the inversion-layer charge density qN_{inv} as well as the source/drain series resistance R_{sd}. The method we extracted those parameters will be illustrated clearly as follows. Also, we will introduce the physical model (drift-diffusion) we used in above calibration.

3.1 Drift-diffusion Model

The drift-diffusion model is widely used for the simulation of carrier transport in semiconductors and is defined by the Poisson equation and continuity equations. The

three governing equations for charge transport in semiconductor devices are Poisson equation and the electron and hole continuity equations. Poisson equation is:

$$\nabla \bullet \varepsilon \nabla \phi = -q(p - n + N_D - N_A) - \rho_{trap} \tag{3.1}$$

Where ε is the permittivity, q is the elementary electronic charge, n and p are the electron and hole densities, N_D is the concentration of ionized donors, N_A is the concentration of ionized acceptors, and ρ_{trap} is the charge density contributed by traps and fixed charges.

The electron and hole continuity equations are

$$\nabla \bullet \vec{j}_n = qR_{net} + q\frac{\partial n}{\partial t} \tag{3.2}$$

$$-\nabla \bullet \vec{j}_p = qR_{net} + q\frac{\partial p}{\partial t} \tag{3.3}$$

where R_{net} is the net electron–hole recombination rate, $\vec{j}_n$ is the electron current density, and $\vec{j}_p$ is the hole current density.

Combining equation 3.1~ equation 3.3, we can derive current densities for electrons and holes as given by:

$$\vec{j}_n = -nq\mu_n \nabla \Phi_n \tag{3.4}$$

$$\vec{j}_p = -pq\mu_p \nabla \Phi_p \tag{3.5}$$

where μ_n and μ_p are the electron and hole mobilities, and Φ_n and Φ_p are the electron and hole quasi-Fermi potentials, respectively.

3.2 Calculation of Inversion Layer Charge Density

In this section, we will introduce the method we calculated inversion layer density from TCAD calibration model. However, we want to explain why we don't use the same method as we used in extraction for additional mobility from gate plasmons in long-channel device [8]. From Figure 3.5 to Figure 3.7, we can observe that when channel length is shrunk, the substrate doping concentration under the SiO_2/Si interface is obviously larger. This is due to the halo implant P_{halo} in short-channel device. The altered substrate doping concentration may influence the inversion layer density directly and further bring about mobility degradation [3]. Consequently, 1D simulator [10] is not enough in this work. Instead, we use 2D simulator TCAD [12] to calculate inversion charge layer density involving the halo implant effect.

First, we use the calibration model to see the free electron density in vertical direction as shown in Figure 3.8. Then we integrate the electron density vertically under the oxide/substrate interface to 30nm deep, which is the quantum confinement region. So, the area under the e-density curve in Figure 3.8 is the N_{inv} at this x position. Then we cut L_m=1μm, 48nm and 33nm devices into many pieces to do the integration, and in order not to contact the source/drain extension in the calculation, we only counted the channel between them. The region we made the calculation is between two dashed lines which is depicted in Figure 3.9 to Figure 3.11. After that, we did the same calculation for different gate voltages from 0V to 1.8V and temperatures from 292K to 380K. Corresponding inversion layer charge density results are shown in Figure 3.12 for L_m=1μm, Figure 3.13 for L_m=48nm and Figure 3.14.for L_m=33nm. Since the N_{inv} for short-channel device is not a constant value, unlike the long-channel device. We further did an average for the inversion layer charge density, i.e. we

integrated the N_{inv} along the channel and then divided the channel length. Finally, we get the N_{inv} for L_m=1μm ,48nm and 33nm versus gate voltage as shown in Figure 3.15, which is the N_{inv} mentioned in the mobility calculation in Euation 2.6.

3.3 Extraction of Parasitic Source/Drain Resistance

Since the parasitic resistance cannot be neglected in short channel device, we will explain the method we use in this work. First, we used the calibration model from TCAD simulation as shown in Figure 3.5 to Figure 3.7. Then, we only retained the drain region and constructed a metal contact beneath the oxide/substrate interface to 1 and 3nm deep. For this metal contact, one side connects the device and the other side is grounded as shown in Figure 3.16. After that, device is operated at drain voltage of 0.05V. Hence, we can get the current flow into the metal contact when we applied gate voltage. Therefore, we can derive R_{sd} as below:

$$\frac{R_{sd}}{2} = \frac{V}{I} . \tag{3.6}$$

Figure 3.17 shows the R_{sd} under various bias conditions from 0.8V to 1.8V with the values of around 100~110Ω -μ m.

Besides, we also provide another calculation method for R_{sd} extraction. The method of extracting parasitic source/drain resistance (R_{sd}) is well described elsewhere [13]. The R_{sd} can be derived as follows. For the intrinsic MOSFET operated in linear region, drain current can be expressed as

$$I_d = \frac{C_{ox} W \mu}{L_m}(V_{gs} - V_{th} - \frac{\alpha}{2}V_{ds})(V_{ds} - R_{sd} I_d) \tag{3.7}$$

The constant mobility can be achieved with two sets of bias conditions:($V_{gs}^{V_{bs1}}, V_{th}^{V_{bs1}}$) and ($V_{gs}^{V_{bs2}}, V_{th}^{V_{bs2}}$), where $V_{gs}^{V_{bs2}} = V_{gs}^{V_{bs1}} + (1/\eta_{ana} - 1)(V_{th}^{V_{bs1}} - V_{th}^{V_{bs2}})$. We can get R_{sd} as below

$$R_{sd} = (\frac{B}{I_d^{V_{bs2}}} - \frac{A}{I_d^{V_{bs1}}}) \frac{\eta_{ana} V_{ds}}{(V_{th}^{V_{bs1}} - V_{th}^{V_{bs2}})} \tag{3.8}$$

where $A = V_{gs}^{V_{bs1}} - V_{th}^{V_{bs1}} - 0.5\alpha V_{ds}$, and

$$B = V_{gs}^{V_{bs1}} + (\frac{1}{\eta_{ana}} - 1)V_{th}^{V_{bs1}} - \frac{V_{th}^{V_{bs2}}}{\eta_{ana}} - 0.5\alpha V_{ds} .$$

In modern MOSFET with thin oxide, α is close to 1, and V_{bs1} and V_{bs2} represent two different back biases. Hence, mobility μ_{ana} and E_{eff_ana} can be analytical written as

$$\mu_{ana} = \frac{L_m}{WC_{ox}} \frac{I_d}{(V_{gs} - V_{th} - 0.5\alpha V_{ds})(V_{ds} - R_{sd} I_d)} \tag{3.9}$$

$$E_{eff\_ana} = \frac{V_{gs} + (\frac{1}{\eta_{ana}} - 1)V_{th} - \frac{1}{\eta_{ana}}V_{FB} - \frac{2}{\eta_{ana}}\Psi_B}{\frac{3}{\eta_{ana}}T_{OX}} \tag{3.10}$$

While minimizing the error between μ_{ana} and $\mu_{universal}$ as described in [13]. Herein, η_{ana} is approximated through iteration to minimize its error. Figure 3.18 shows the I-V curves for short-channel devices where V_{bs1} =0V and V_{bs2} = -0.4V. The result for extracting R_{sd} value is derived in Figure 3.19, and its value is about 123Ω -μm for

η_{ana} = 0.33, which is reasonable and consistent with the value (η_{ana} ≈1/2~1/3) used in [13]. Also, the R_{sd} value obtained by using this method is reasonably consistent with the extracted one (R_{sd} =120Ω-μm) in [14]. Thus, we use R_{sd} =100~120Ω-μm to calculate the mobility for short-channel device.

According to inverse modeling, the extension overlap and hence the metallurgical channel length L_m are easy to obtain. Combining the inversion layer charge density N_{inv} and source/drain resistance R_{sd}, the measured effective mobility (μ_{eff}) are obtained as demonstrated in Figure 3.20 versus inversion layer density N_{inv} for different L_m with R_{sd} of 100 to 120Ω-μm as parameters.

Chapter 4

Analysis and Discussion

With the inverse modeling in Chapter 3, we extracted the mobility with R_{sd} =100~120Ω -μ m for different metallurgical channel length L_m of 1μ m, 48nm and 33nm. In Figure 3.20 we found the mobility degradation in short-channel device. Due to the controversial mechanism for the mobility degradation in short-channel device, we want to use the temperature-dependent experimental method as in [8] to determine the main source of degradation. First, we extracted additional mobility $\mu_{add}(48\ nm)$ and $\mu_{add}(33\ nm)$ of short-channel device with respect to long-channel one ($L_{m=}$ 1μ m), and further obtained their temperature dependencies at high N_{inv} (= 1x10^{13} cm^{-2}). Besides, we also extracted the extra additional mobility $\mu_{add,extra}$ and decided the main source of this scattering mechanism. Moreover, we further provided an important evidence to verify our hypothesis about this additional scattering mechanism.

4.1 Additional Mobilities

While comparing measured effective electron mobility (μ_{eff}) between the long-channel and short-channel device shown in Figure 3.20, additional scattering in short-channel device can be dealt with using Matthiessen's rule which essentially is valid under high N_{inv} as follows:

$$\frac{1}{\mu_{add}(SC)} = \frac{1}{\mu_{eff}(SC)} - \frac{1}{\mu_{eff}(LC)} \quad (4.1)$$

where SC and LC mean short-channel and long-channel, respectively. Figure 4.1 shows the extracted temperature-dependent additional mobility versus N_{inv} for L_m=48 and 33nm with R_{sd} = 100 to 120 Ω-μm. Note that the devices under study are all characterized by the same oxide thickness, so it is expected that scattering coming from above channel should be the same. It can be seen that additional mobility for L_m = 48 nm $\mu_{add}(48\,nm)$ increases with N_{inv}, whereas for L_m = 33nm $\mu_{add}(33\,nm)$ exhibits a saturating trend which dominates in both cases of R_{sd}. We further extracted additional mobility at N_{inv}=1x10^{13}cm^{-2} versus temperature with R_{sd} as a parameter. In addition, we also provided a temperature-coefficient γ to clarify the temperature-oriented trend, there is a relationship between the additional mobility and temperature as follows:

$$\mu_{add} \propto T^{\gamma} \quad (4.2)$$

Apparently, corresponding temperature dependencies are also opposite to each other: γ = 0.37~0.55 for L_m = 48nm and γ = -0.13~-0.17 for L_m = 33nm as shown in Figure 4.2 for varying R_{sd}.

4.2 Main Source of Mobility Degradation in Short-Channel Device

Several mechanisms are considered to be involved in the process of mobility

degradation for short-channel device. Among them, S/D plasmons, bandgap narrowing, short-range Coulomb scattering due to halo implant or pockets and defects near S/D are believed to be the most probable culprits to degrade the performance [1]-[5]. Additional mobility of L_m = 48 nm increases with temperature, consistent with the result of [3] about mobility degradation in short-channel device, indicating that the responsible mechanisms are those of short-range Coulomb centers (due to halo implant or pockets [3] and/or defects near S/D [4],[5]).

As to L_m = 33 nm, it exhibits a slight decrease with temperature (Figure 4.2), suggesting other mechanisms. The corresponding mobility component can be written as:

$$\frac{1}{\mu_{add,\ extra}} = \frac{1}{\mu_{add}(33\ nm)} - \frac{1}{\mu_{add}(48\ nm)} \quad (4.3)$$

The resulting $\mu_{add,extra}$ is plotted in Figure 4.3 versus N_{inv}. And in Figure 4.4 we present the power-law relation of $\mu_{add,extra}$ versus temperature at N_{inv} = 1x10^{13} cm^{-2}. Surprisingly, temperature-dependent $\mu_{add,extra}$ in Figure 4.4 is satisfactorily close to that of gate-plasmon-limited ones [8]. The corresponding γ has a value of –0.69 to -0.83. Here, we have to rule out the mechanism of 2D charge sharing from S/D [3],[6]. The reasons are twofold. First, although the width W_D of simulated substrate depletion region increases with decreasing L_m (Figure 3.5 to Figure 3.7), which may act as a signature of 2D charge sharing, the subband separation does not appear to decrease but increase (see Table 2), greatly contrary to [3],[6]. Here, we consider two lowest subbands as an estimation of population distribution versus energy. Meanwhile, this calculation is done for population for 1nm below surface. Figure 4.5 to Figure 4.7 show the subband energy below surface 1nm for gate voltage from 1.0V~1.5V. This obvious difference is simply because in our work, the halo implant is used and its

effect is enhanced with decreasing L_m. Second, we conducted bulk-phonon-limited mobility simulation using the simulation program [7],[10]. The resulting γ lies at a value from -1.5 to -1.6, which is much more negative than that of $\mu_{add,extra}$ (Figure 4.8). It is therefore argued that only for L_m less than about 40 nm can long-range Coulomb effects become noticeable.

4.3 Evidence of Long-Range Coulomb Interactions

Finally, in Figure 4.9 we quote simulated transconductance at V_d = 1.0 V and V_g of 0.75 and 1.0 V above threshold [1] versus L_m, for comparison with measured transconductance at V_d = 0.8 and 1.0 V in this work. Evidently, the dimension of our devices, which is carefully chosen to meet the criterion, lies across the activation point of long-range Coulomb effects. With above evidence, we further argue that the long-range Coulomb interactions would be the main factor for performance degradation in ultra-short devices.

Chapter 5

Conclusion

In this work, TCAD-based inverse modeling has been carried out with aim reconstruct the process parameters. Consequently, interesting and useful results have been created. First, the overlap region for short-channel device can be accurately determined. Also the halo implant P_{halo}, which has a significant impact on the substrate doping concentration and further affect the inversion layer charge density, is also solved by calculating the free electron density in the channel region from 2D simulation. Moreover, the parasitic source/drain resistance (R_{sd}) is also extracted from the calibration model. In addition, we also provide experimental method to estimate the R_{sd} for the comprehensive analysis.

Second, the resulting temperature power-law exponent (γ) as extracted from our experimentally-determined additional mobility data points out that the long-range Coulomb interactions exist in the metallurgical channel length less than about 40 nm. Thus, experimental evidence of long-range Coulomb interactions has been drawn. Furthermore, underlying physical origins have all been distinguished for short channel device. Therefore, long-range Coulomb effect, which is not to be ignored in ultra-short devices, has been for the first time experimentally corroborated in the device samples under study.

References

[1] M. V. Fischetti , S. Jin , T.-W. Tang , P. Asbeck , Y. Taur , S. E. Laux , M. Rodwell and N. Sano, “Scaling MOSFETs to 10 nm: Coulomb effects, source starvation, and virtual source model,” *J. Comput. Electron.*, vol. 8, p.60 , 2009.

[2] M. V. Fischetti and S.E. Laux, “Long-range Coulomb interactions in small Si devices. Part Ⅰ: Performance and reliability,” *J. Appl. Phys.*, vol. 89, no. 2, pp. 1232-1248, January 2001.

[3] K. Rim, S. Naeasimha, M. Longstreet, A. Mocuta, and J. Cai, “Low field Mobility characteristics of sub-100 nm unstrained and strained si MOSFETs,” in *IEDM Tech. Dig.* , pp. 43-46, 2002.

[4] Antoine Cros, Krunoslav Romanjek, Dominique Fleury, Samuel Harrison, Robin Cerutti, Philippe Coronel, Benjamin Dumont, Arnaud Pouydebasque, Romain Wacquez, Blandine Duriez, Romain Gwoziecki, Frederic Boeuf, Hugues Brut, Gerard Ghibaudo and Thomas Skotnicki, “Unexpected mobility degradation for very short devices: A new challenge for CMOS scaling,” in *IEDM Tech. Dig.*, pp. 663-666, 2006.

[5] Vincent Barral, Thierry Poiroux, Daniela Munteanu, Jean-Luc Autran, and Simon Deleonibus, ‘Experimental investigation on the quasi-ballistic transport: part II—backscattering coefficient extraction and link with the mobility,” *IEEE Trans. Electron Devices*, vol. 56, no. 3, pp. 420-430, March. 2009.

[6] P.Packan, S.Cea, H.Deshpande, T.Ghani, M.Giles, O.Golonzka, M.Hattendorf, R.Kotlyar, K.Kuhn, A.Murthy, P.Ranade, L.Shifren, C.Weber and K.Zawadzki, “High performance Hi-K + metal gate strain enhanced transistors on (110) Silicon,” in *IEDM Tech. Dig.*, pp. 63-66, 2008.

[7] M. V. Fischetti, “Long-range Coulomb interactions in small Si devices Part II.

Effective electron mobility in thin-oxide structures," *J. Appl.Phys.*, vol. 89, no. 2, pp. 1232–1250, Jan. 2001.

[8] Ming-Jer Chen, Li-Ming Chang, Shin-Jiun Kuang, Chih-Wei Lee, Shang-Hsun Hsieh, Chi-An Wang, Sou-Chi Chang, and Chien-Chih Lee, "Temperature-oriented mobility measurement and simulation to assess surface roughness in ultrathin-gate-oxide (~1 nm) nMOSFETs and Its TEM evidence," *IEEE Trans. Electron Devices*, vol. 59, no. 4, pp. 949-955, April. 2012.

[9] Schred, http://nanohub.org/resources/schred.

[10] M. J. Chen, C. C. Lee, and K. H. Cheng, "Hole effective masses as a booster of self-consistent six-band k • p simulation in inversion layers of pMOSFETs," *IEEE Trans. Electron Devices*, vol. 58, pp. 931-937, April 2011.

[11] S. Takagi and M. Takayanagi, "Experimental evidence of inversion-layer mobility lowering in ultrathin gate oxide metal-oxide-semiconductor field-effect-transistors with direct tunneling current," *Jpn. J. Appl. Phys.*, vol. 41, pt. 1, no. 4B, pp. 2348-2352, Apr. 2002.

[12] TCAD. http://www.synopsys.com/Tools/TCAD/Pages/default.aspx.

[13] D.W. Lin, M. L. Cheng, S.W.Wang, C. C.Wu, and M. J. Chen, "A novel method of MOSFET series resistance extraction featuring constant mobility criteria and mobility universality," *IEEE Trans. Electron Devices*, vol. 57, no. 4, pp. 890–897, Apr. 2010.

[14] K. Romanjek, F. Andrieu, T. Ernst, and G. Ghibaudo, "Improved split C-V method for effective mobility extraction in sub-0.1-μm Si MOSFETs," *IEEE Electron Devices Letters*, vol. 25, no. 8, pp. 583-585, Aug. 2004.

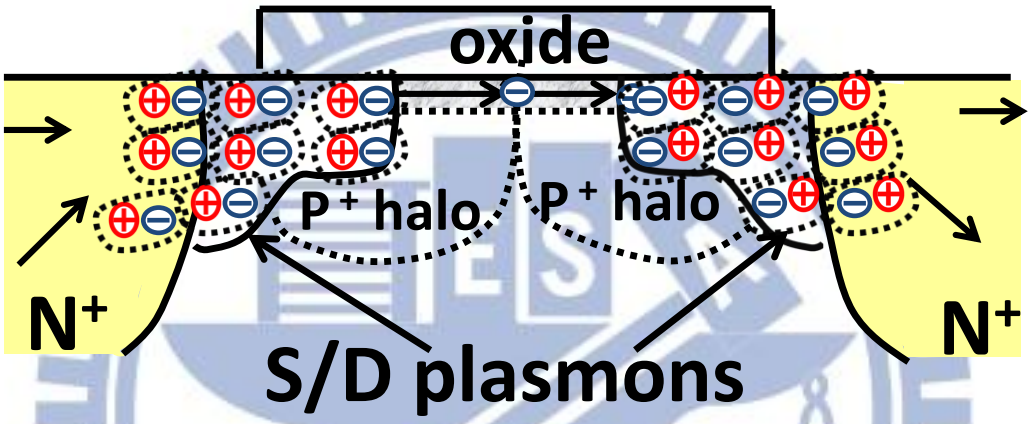


Figure 1.1 Schematic of electron transport under long-range Coulomb interactions with S/D plasmons.

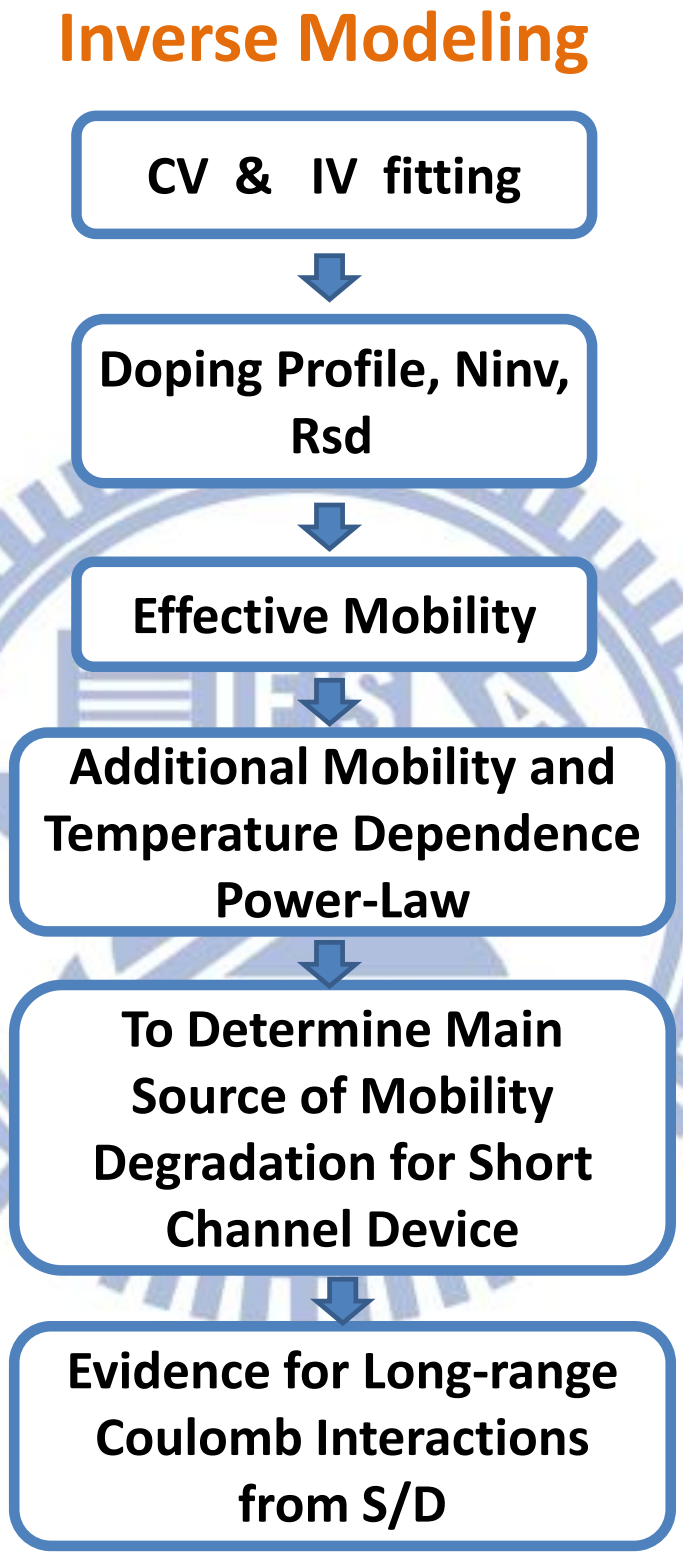


Figure 1.2 Flowchart of inverse modeling in this work.

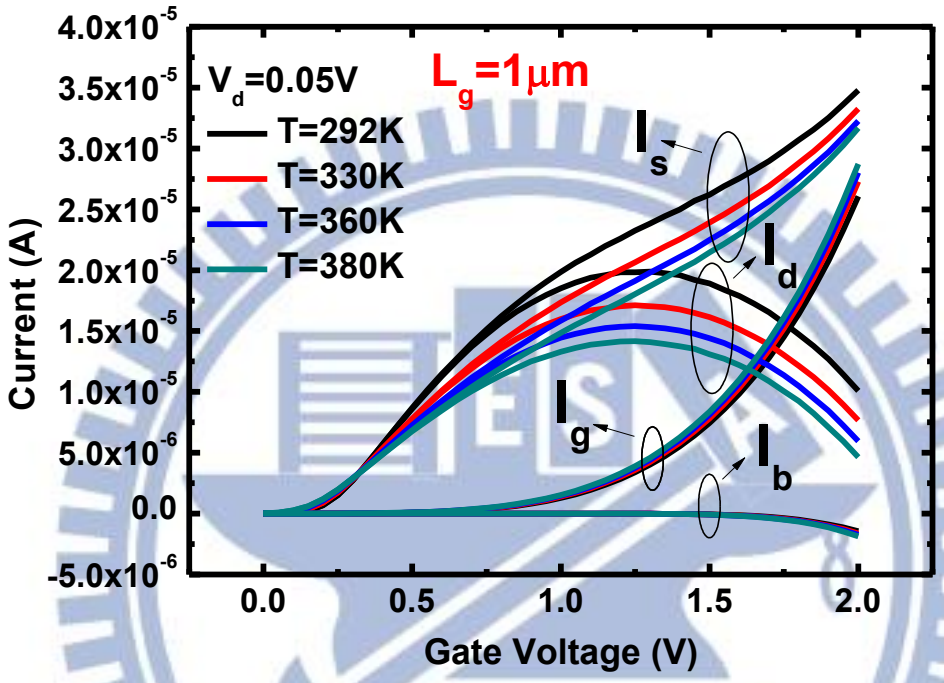


Figure 2.1 Temperature-dependent terminal currents at V_d=0.05V versus V_g for L_g=1μm.

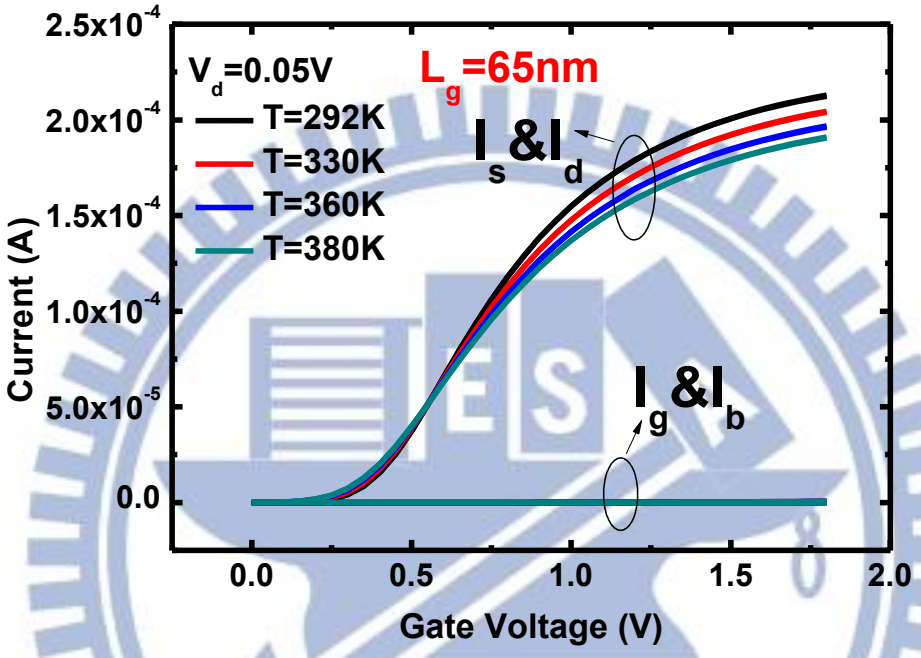


Figure 2.2 Temperature-dependent terminal currents at V_d=0.05V versus V_g for L_g=65nm.

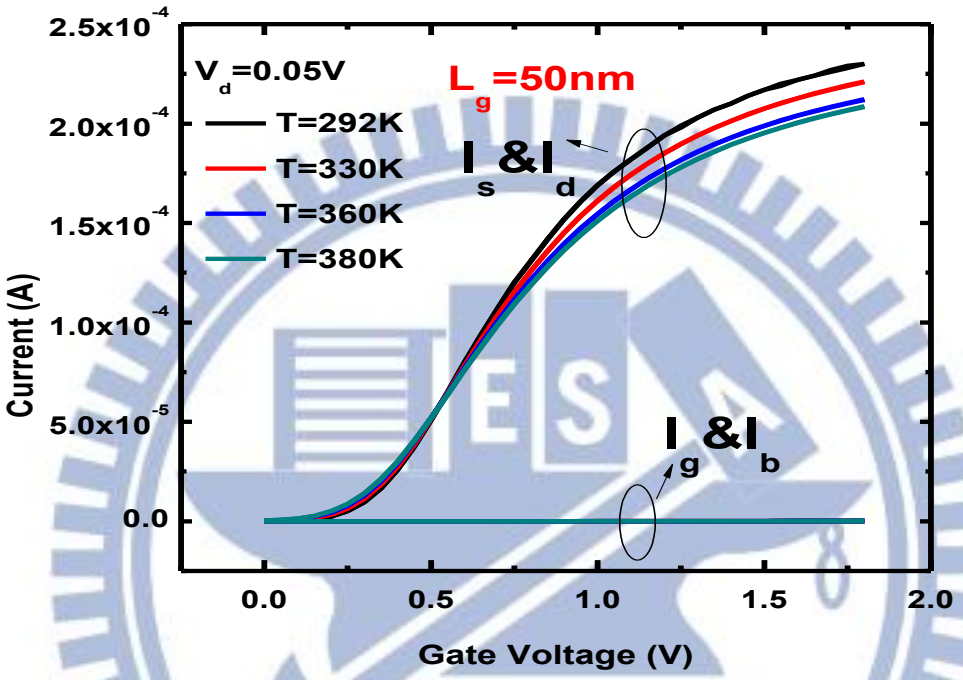


Figure 2.3 Temperature-dependent terminal currents at V_d=0.05V versus V_g for L_g=50nm.

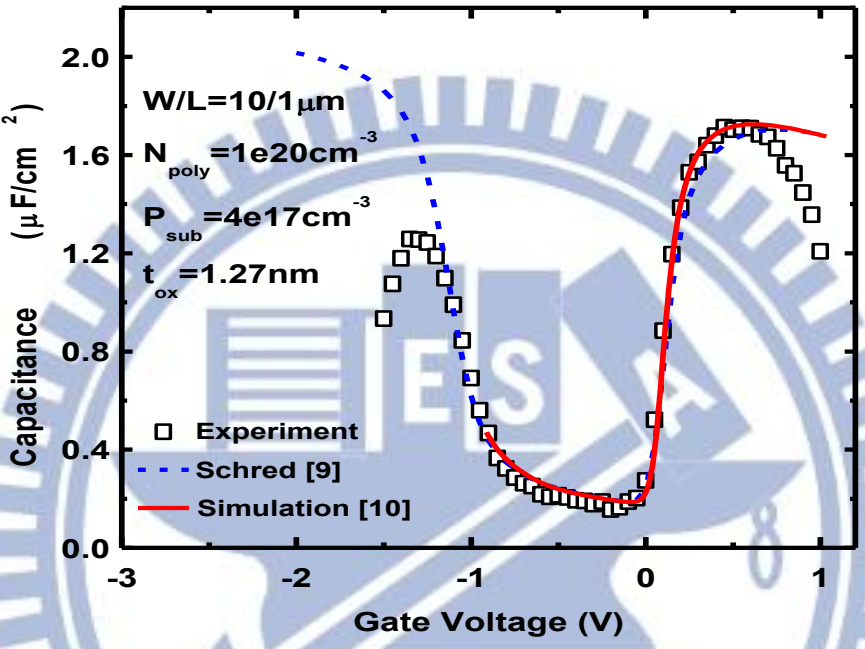


Figure 2.4 Comparison of the measured and simulated gate capacitance versus gate voltage.

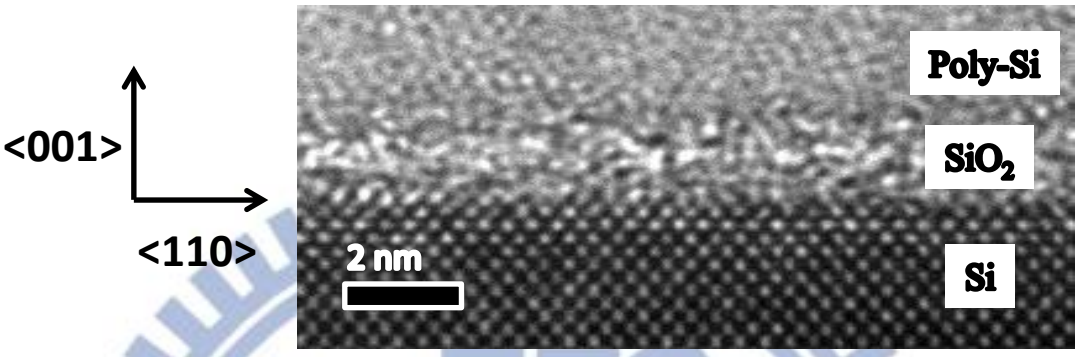


Figure 2. 5 TEM image of the sample.

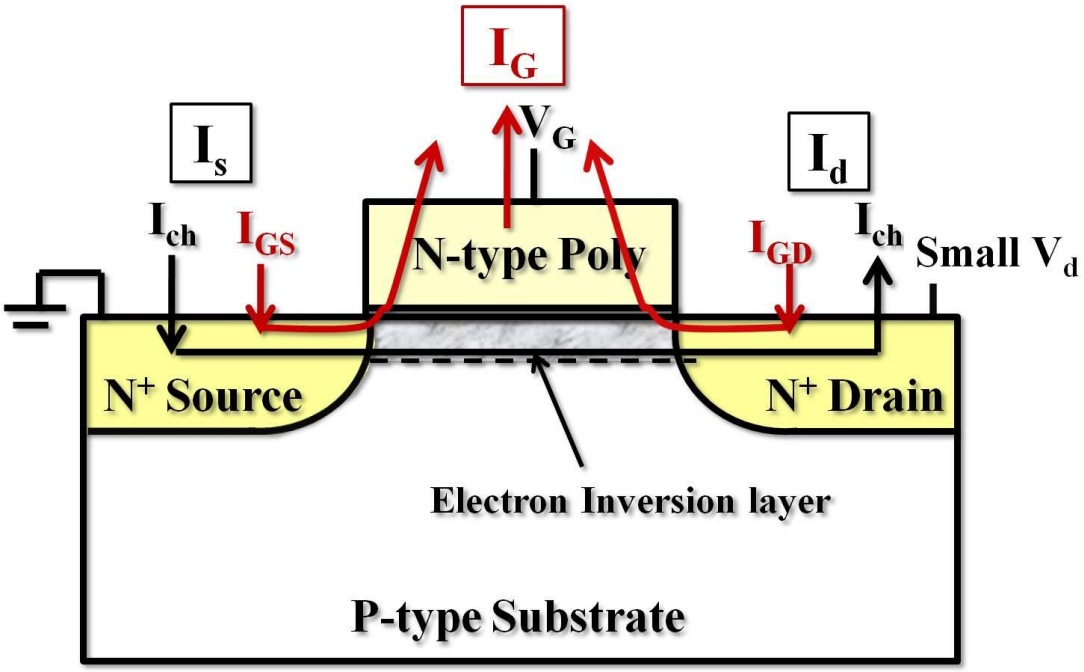


Figure 2.6 The schematic diagram for current flow of nMOSFETs with large gate tunneling current. Besides, $I_S<0$ and $I_d>0$.

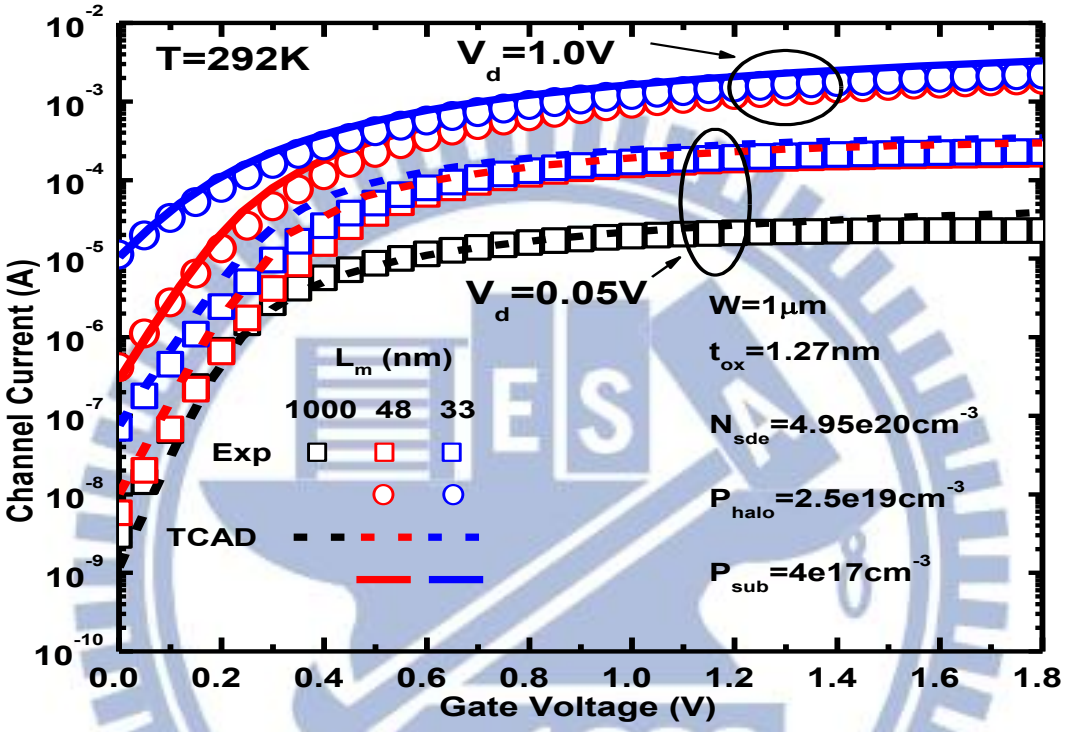


Figure 3.1 Measured and simulated I_{ch} versus V_g at T=292K and V_d=0.05 and 1V.

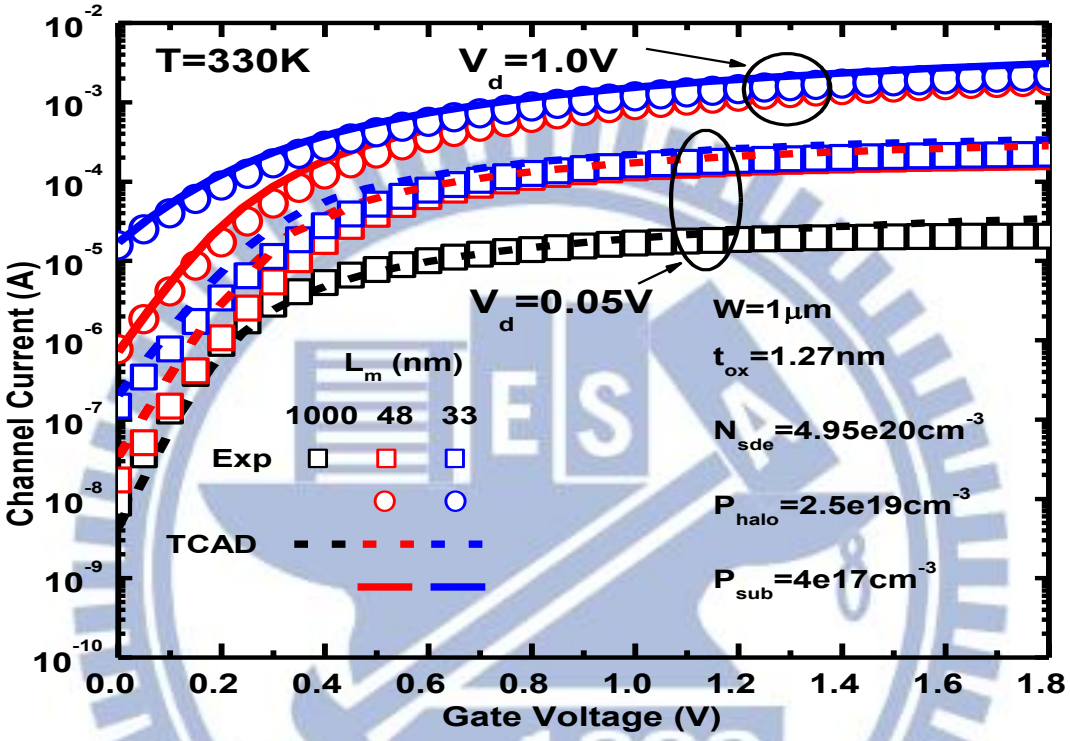


Figure 3.2 Measured and simulated I_{ch} versus V_g at T=330K and $V_{d=}$0.05 and 1V.

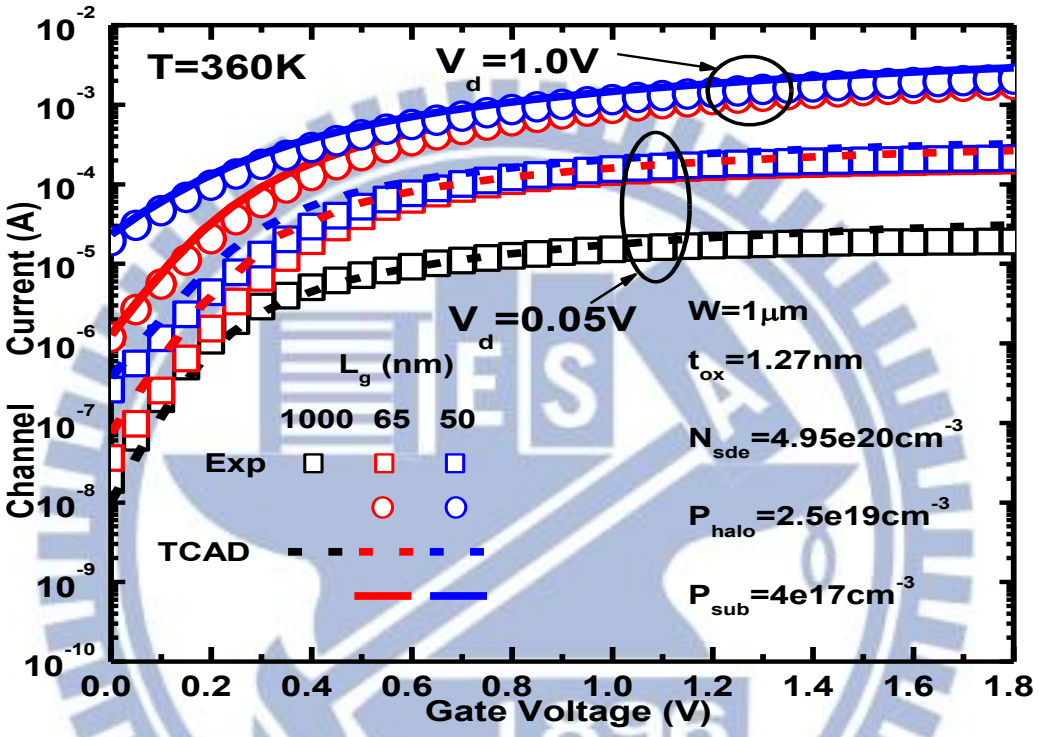


Figure 3.3 Measured and simulated I_{ch} versus V_g at T=360K and V_d=0.05 and 1V.

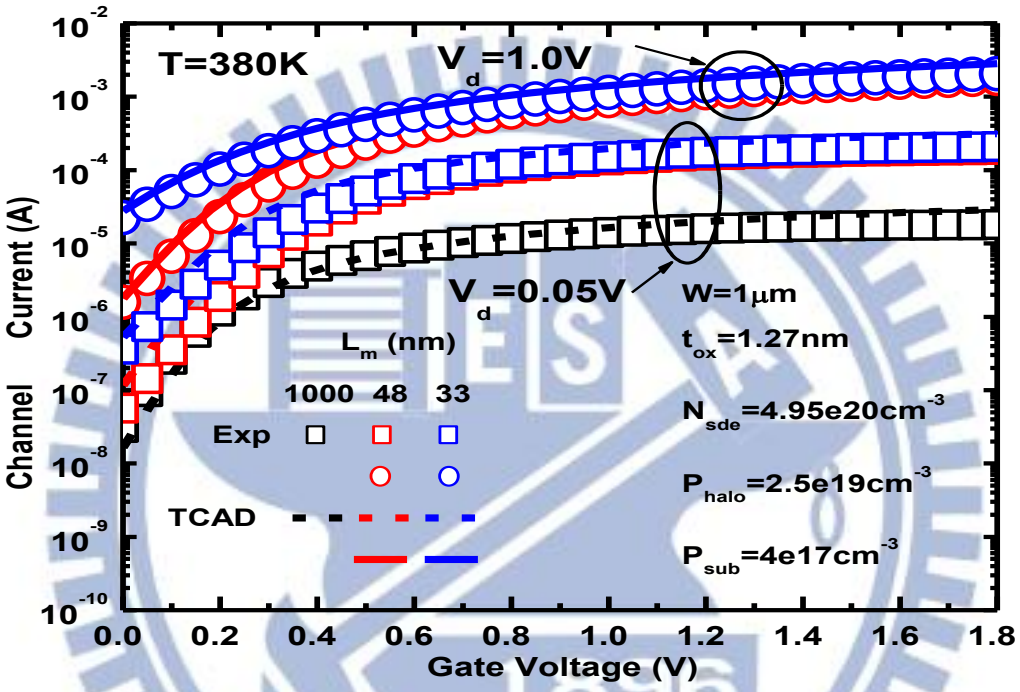


Figure 3.4 Measured and simulated I_{ch} versus V_g at T=380K and V_d=0.05 and 1V.

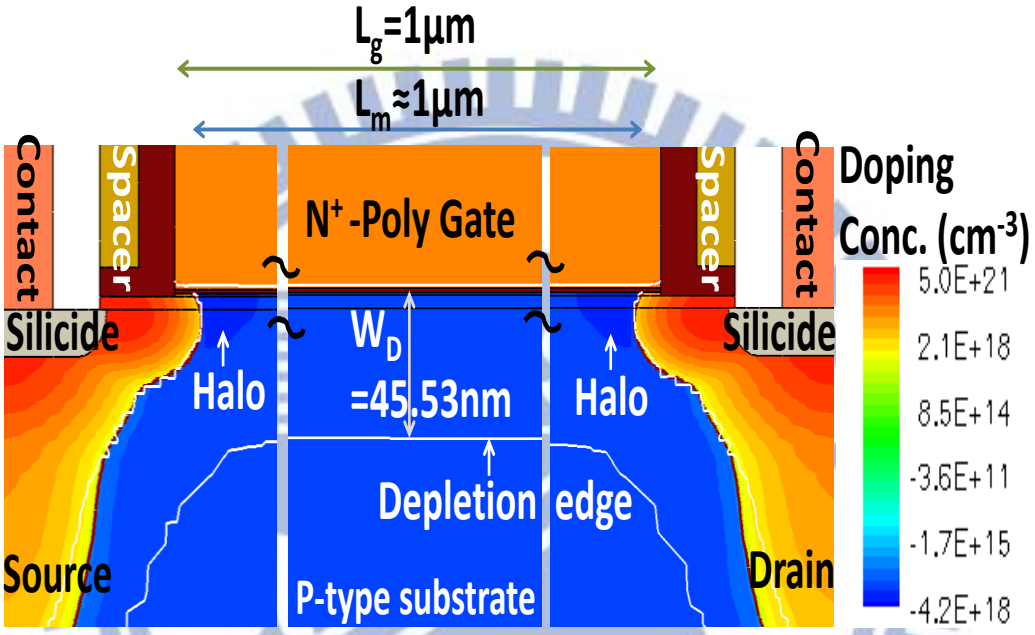


Figure 3.5 Calibrated 2D simulation structure for L_g=1μm and hence L_m=1μm. W_D is the width of the mid-channel depletion region.

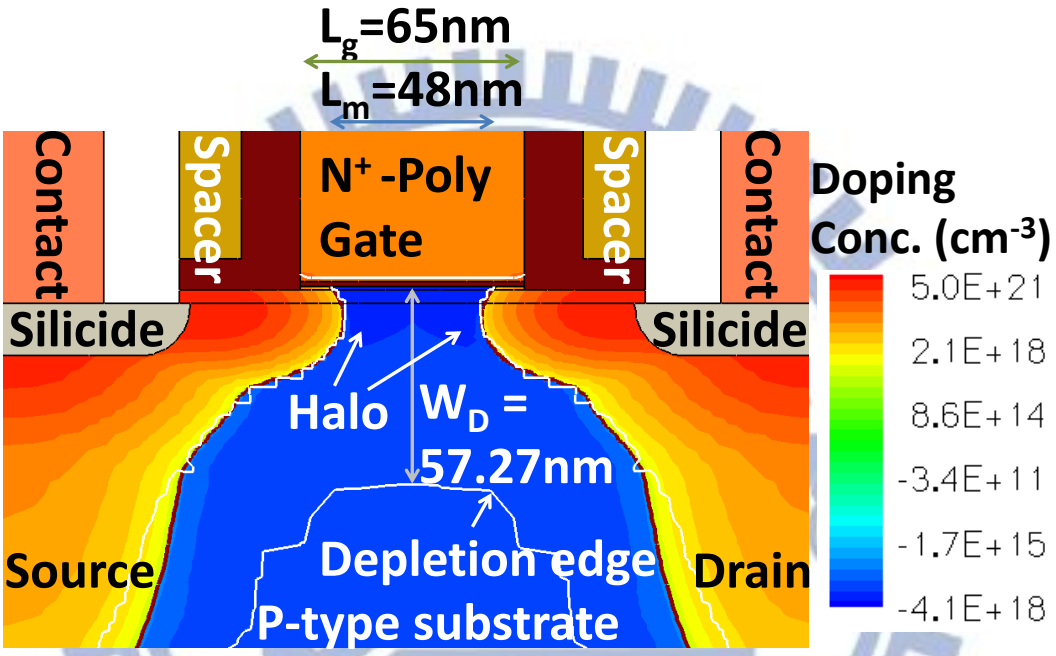


Figure 3.6 Calibrated 2D simulation structure for L_g=65nm and hence L_m=48nm. W_D is the width of the mid-channel depletion region.

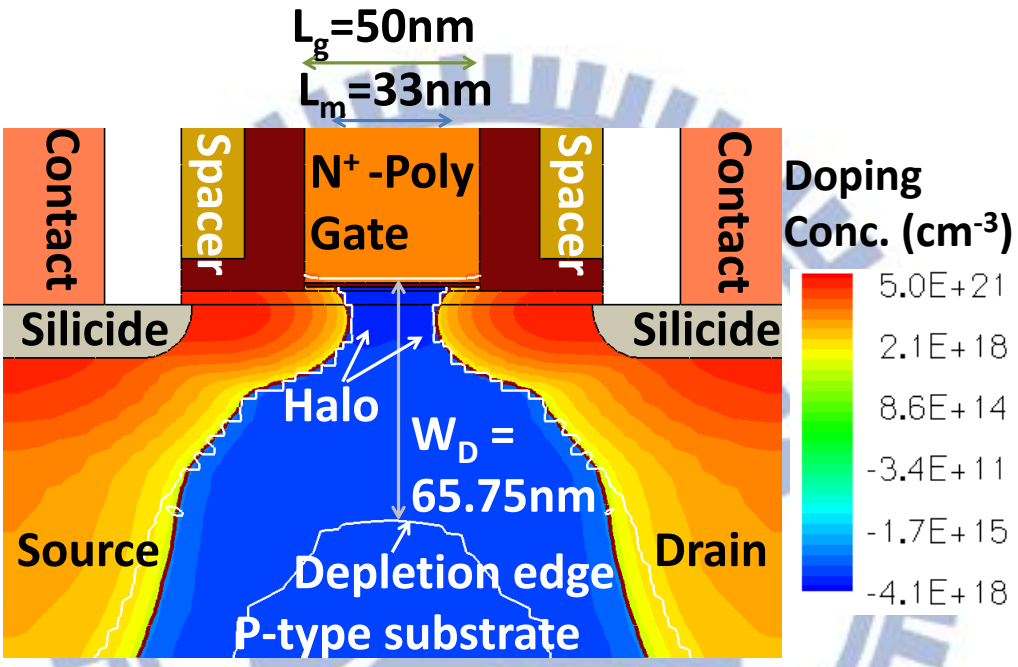


Figure 3.7 Calibrated 2D simulation structure for L_g=50nm and hence L_m=33nm. W_D is the width of the mid-channel depletion region.

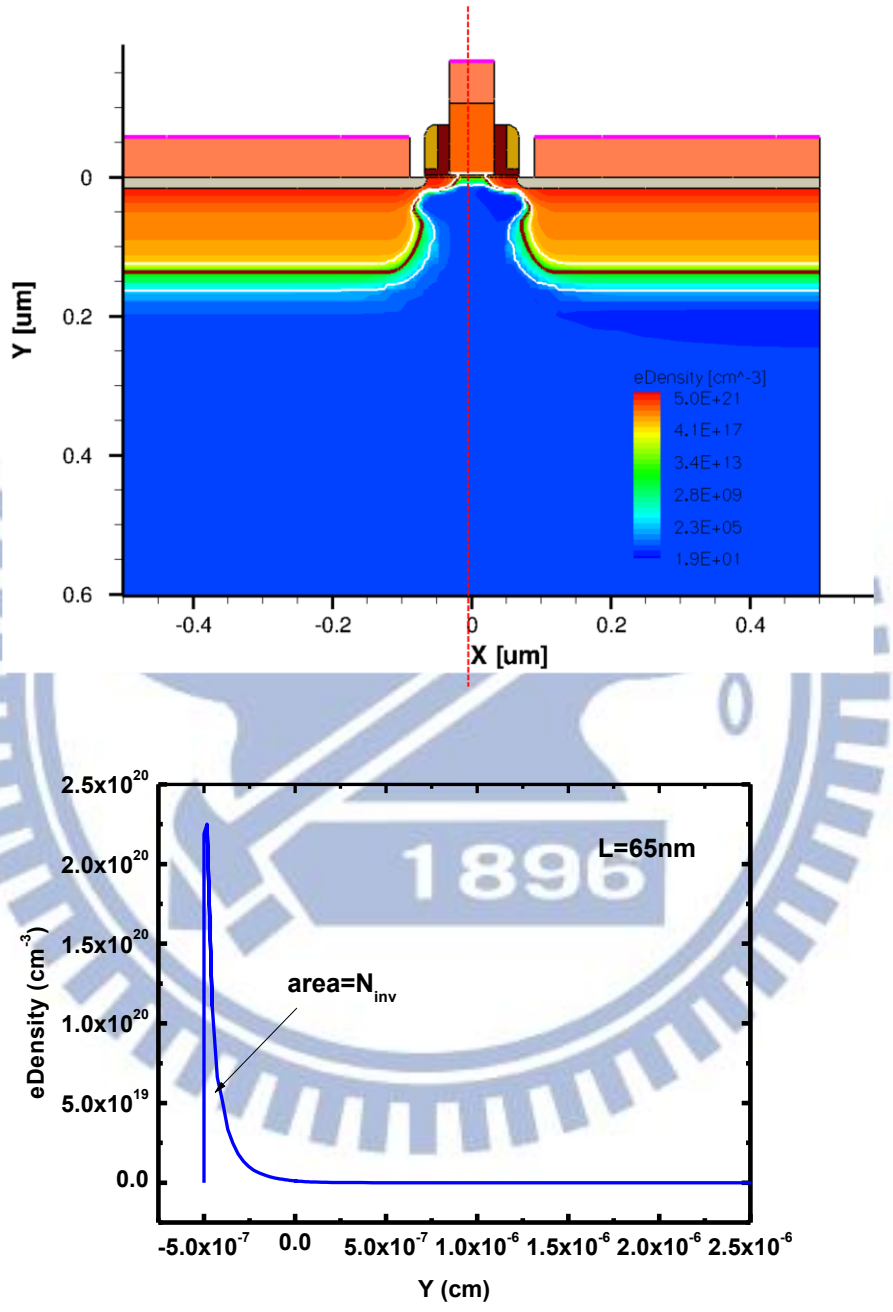


Figure 3.8 The free electron density under the surface 30nm at one position of the channel region.

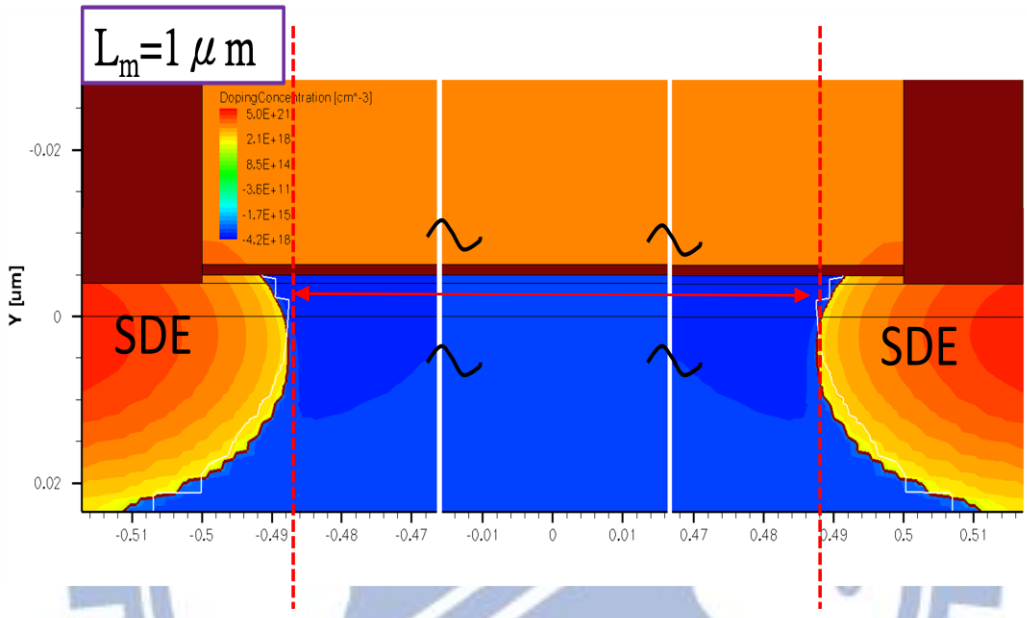


Figure 3.9 The schematic diagram for inversion layer charge density of metallurgical length L_m=1μ m.

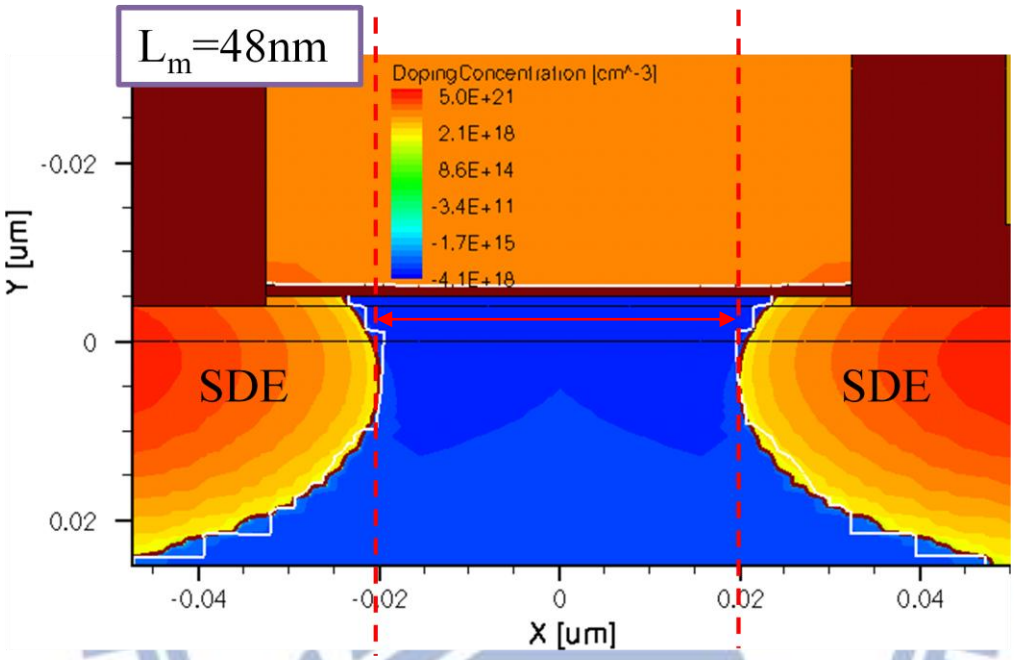


Figure 3.10 The schematic diagram for inversion layer charge density of metallurgical length L_m=48nm.

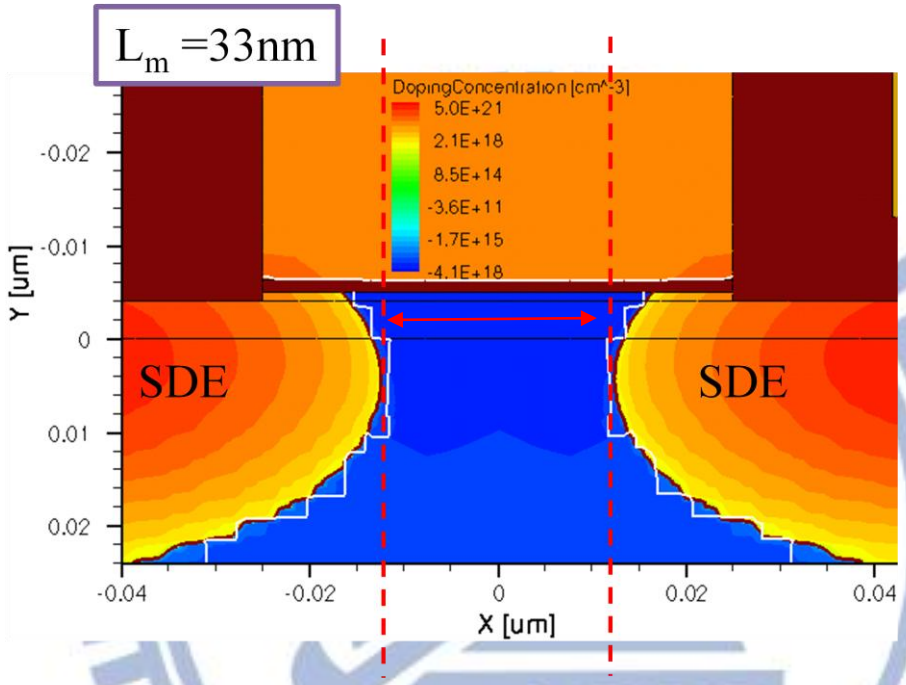
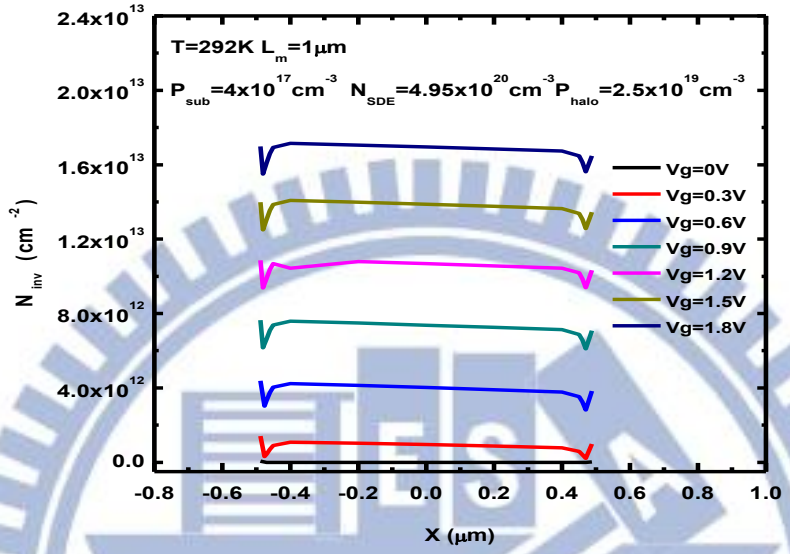
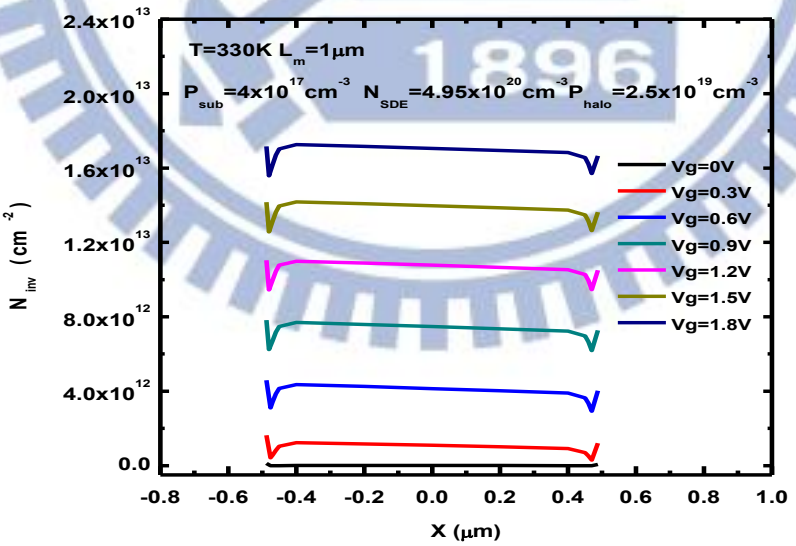


Figure 3.11 The schematic diagram for inversion layer charge density of metallurgical length L_m=33nm.

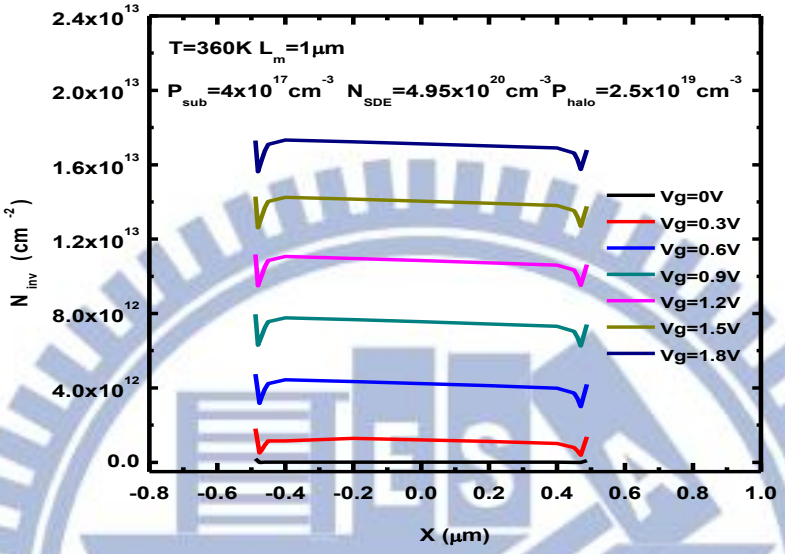


(a)

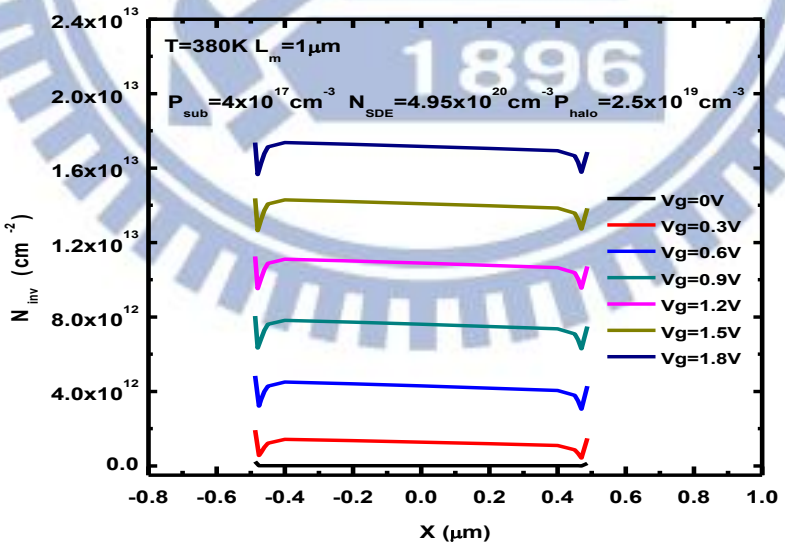


(b)

Figure 3.12 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 1μm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K.

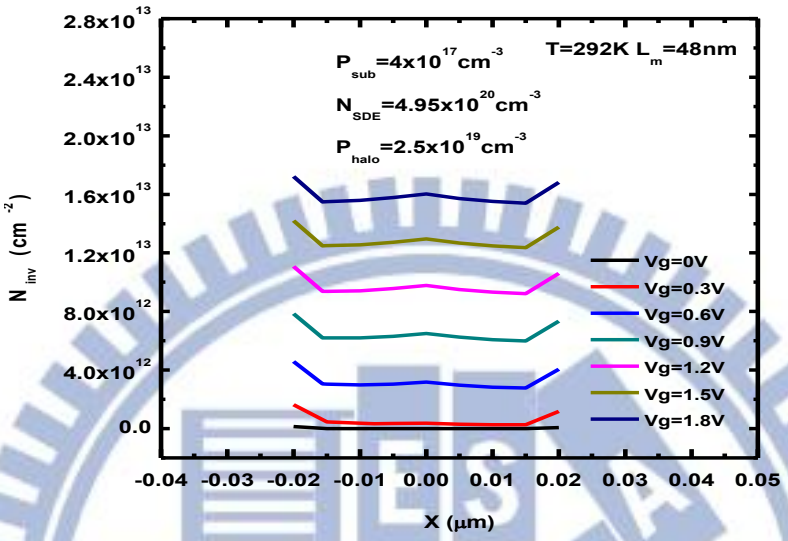


(c)

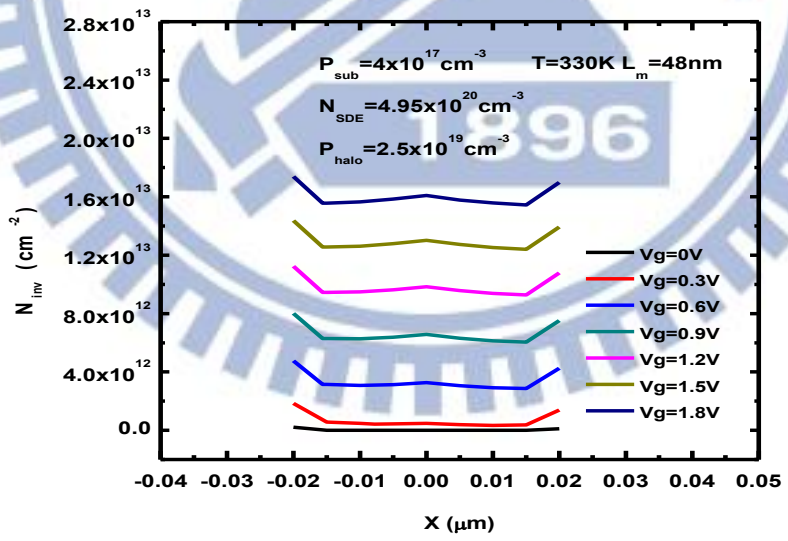


(d)

Figure 3.12 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 1μm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K.

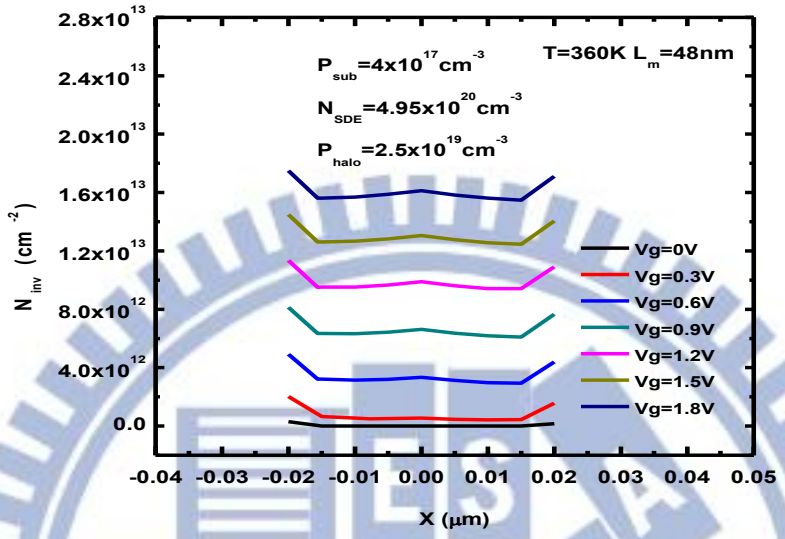


(a)

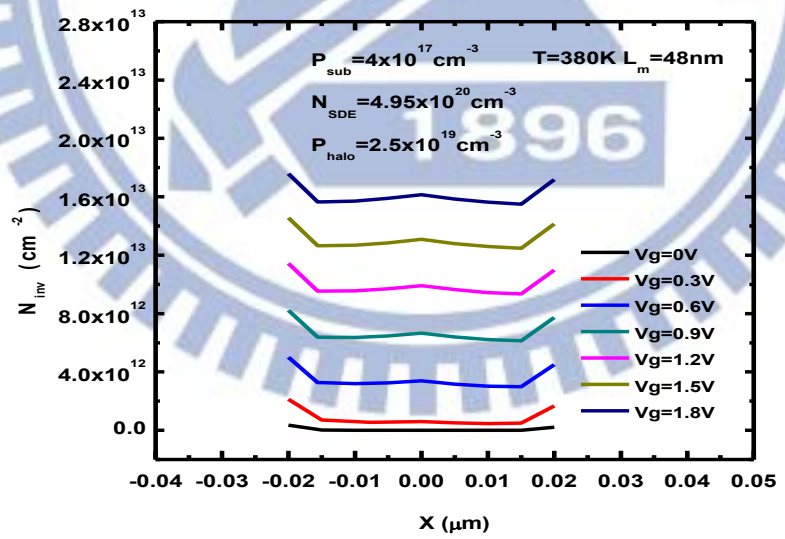


(b)

Figure 3.13 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 48nm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K.

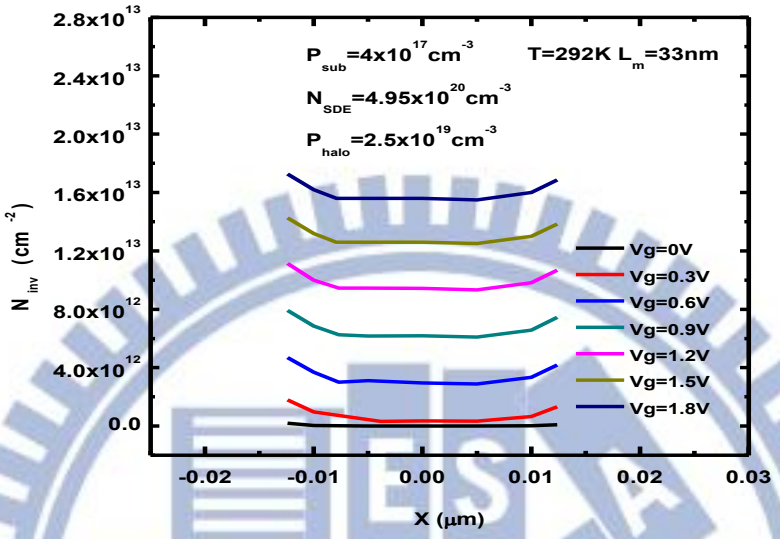


(c)

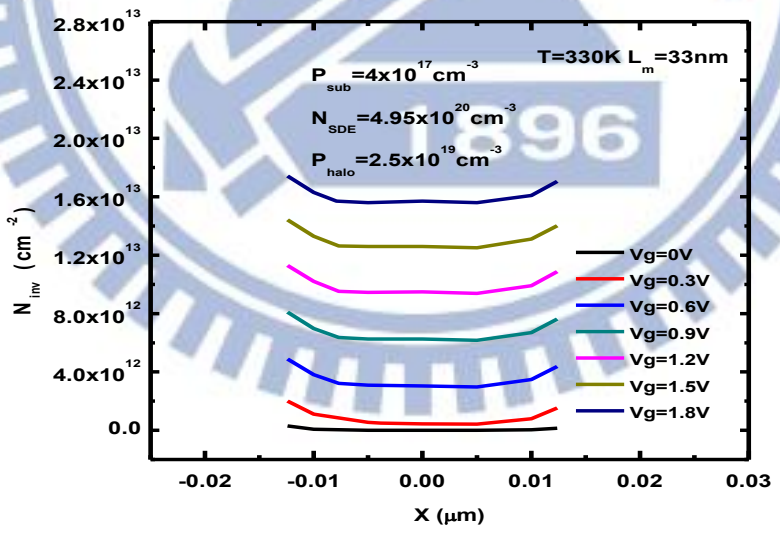


(d)

Figure 3.13 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 48nm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K.

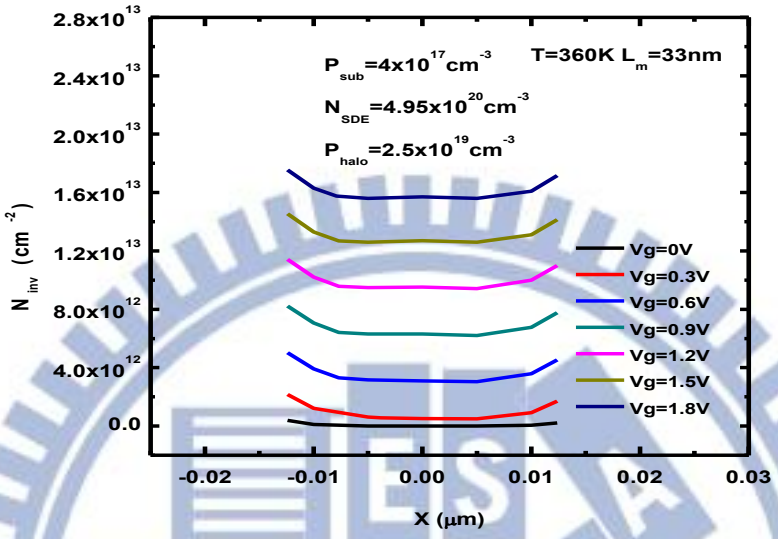


(a)

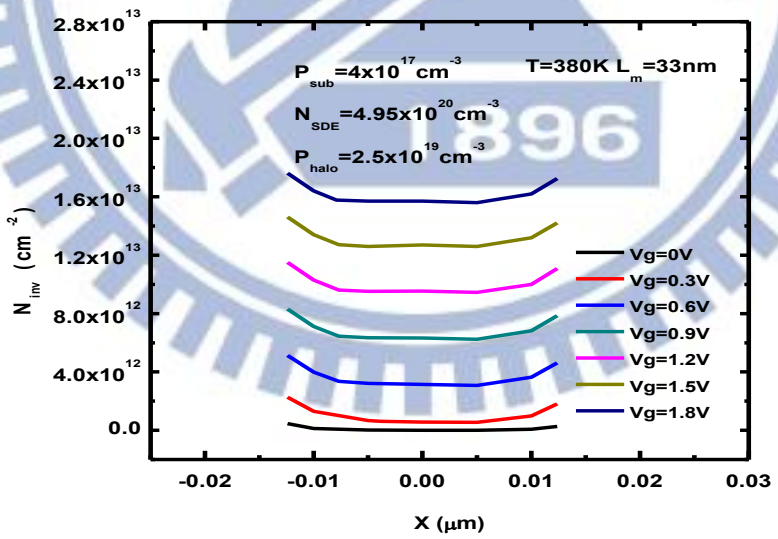


(b)

Figure 3.14 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 33nm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K.



(c)



(d)

Figure 3.14 The calculated N_{inv} along the channel direction under interface 30nm for L_m = 33nm at T= (a) 292K, (b) 330K, (c) 360K, and (d) 380K.

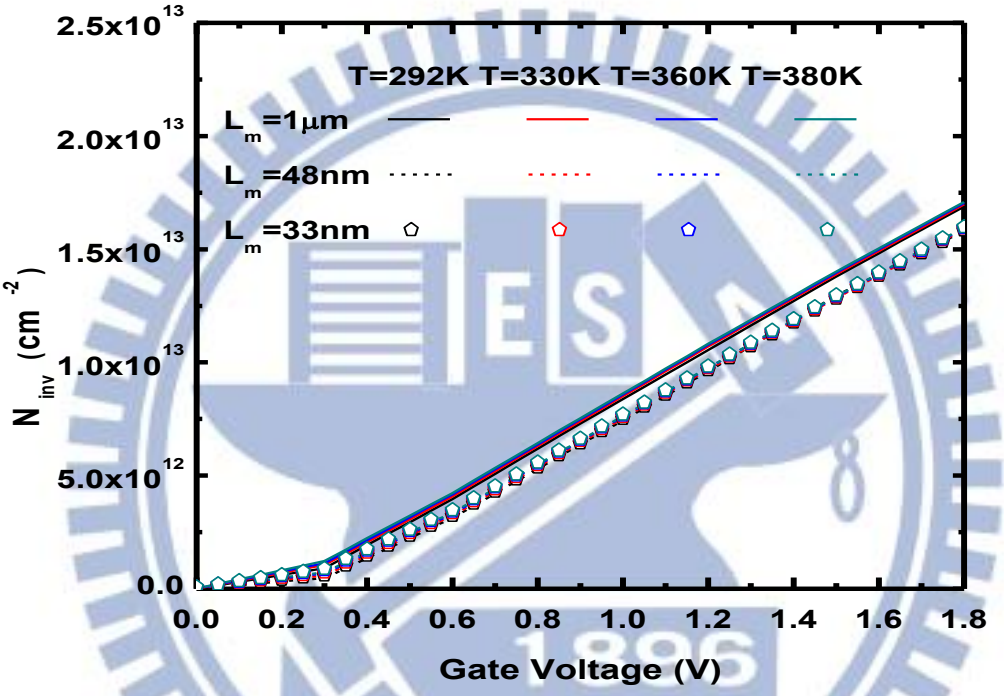


Figure 3.15 Simulated N_{inv} versus gate voltage.

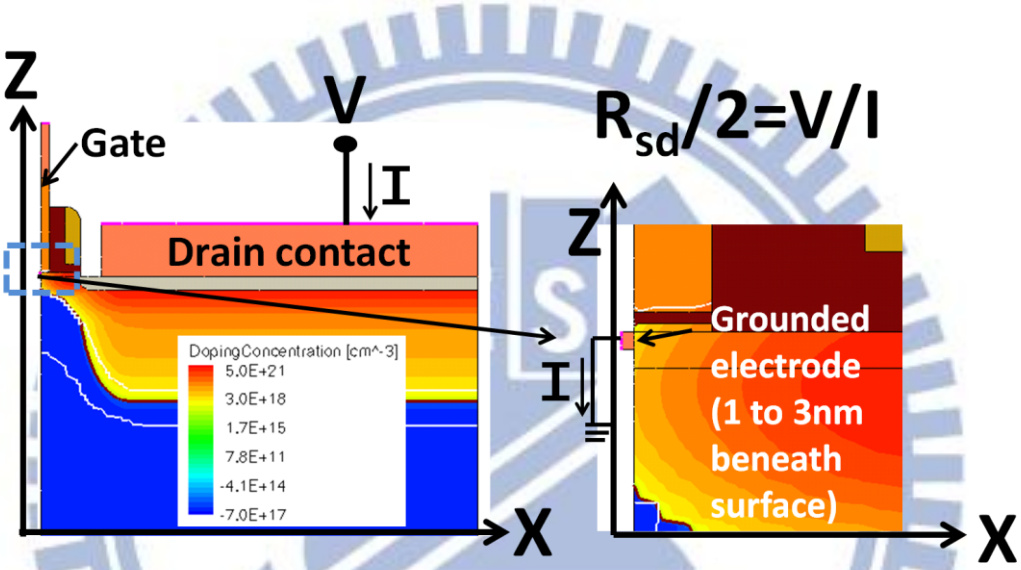


Figure 3.16 Simulation structure for R_{sd} assessment.

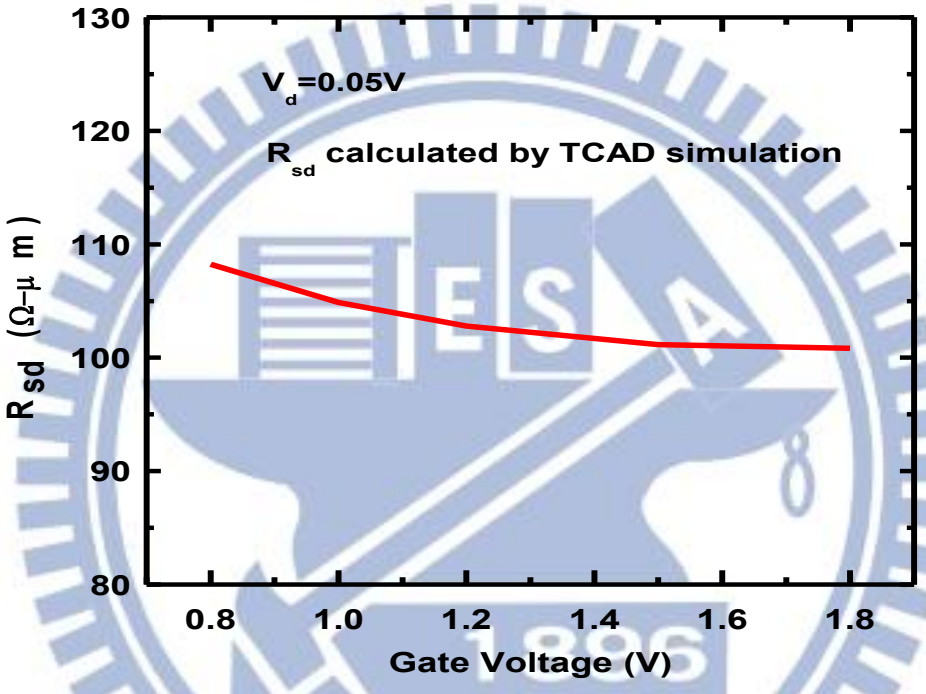
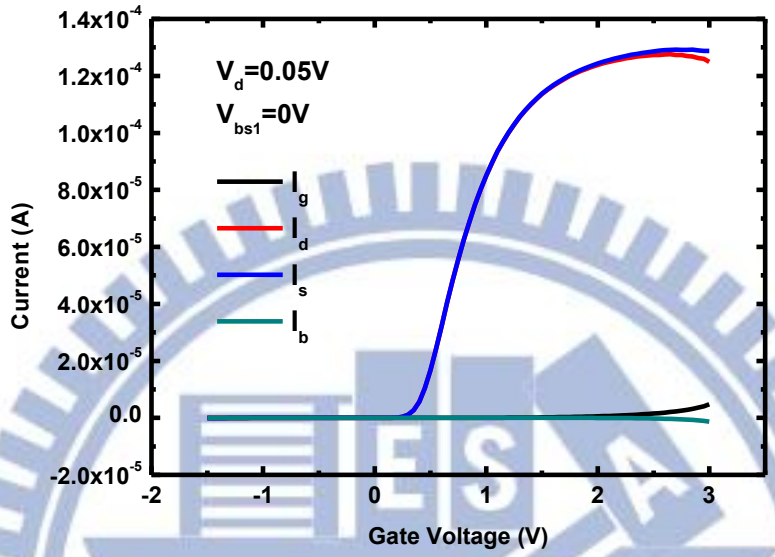
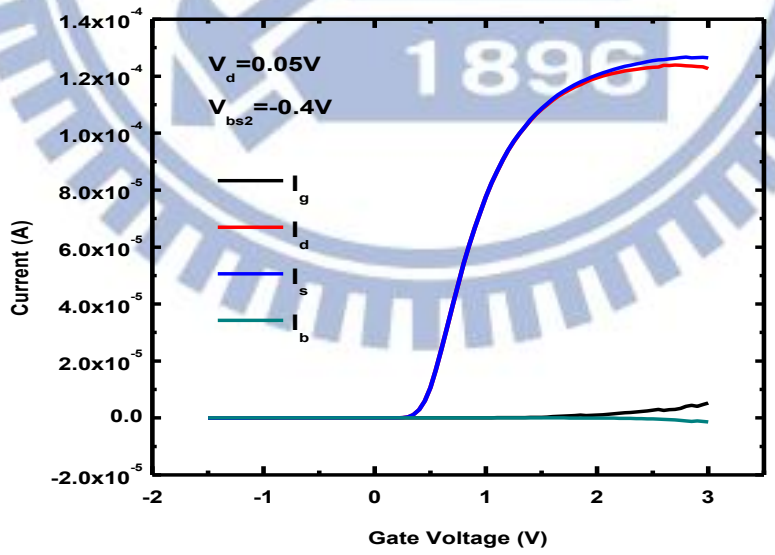


Figure 3.17 Simulated R_{sd} versus V_g at T=292K.



(a)



(b)

Figure 3.18 Terminal currents for short-channel device versus V_g at (a)V_b=0V and (b)V_b=-0.4V.

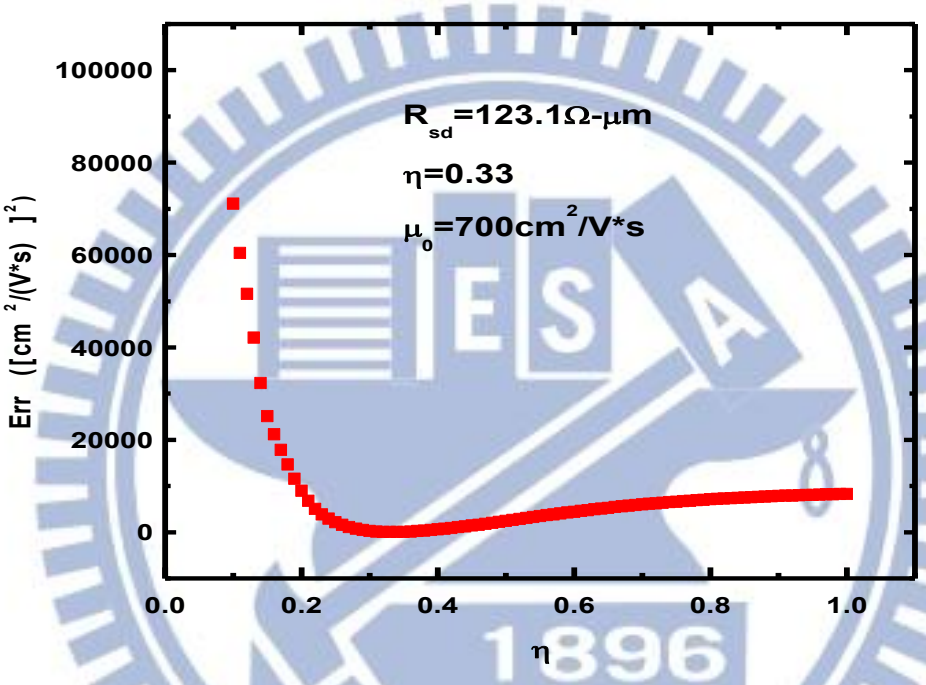
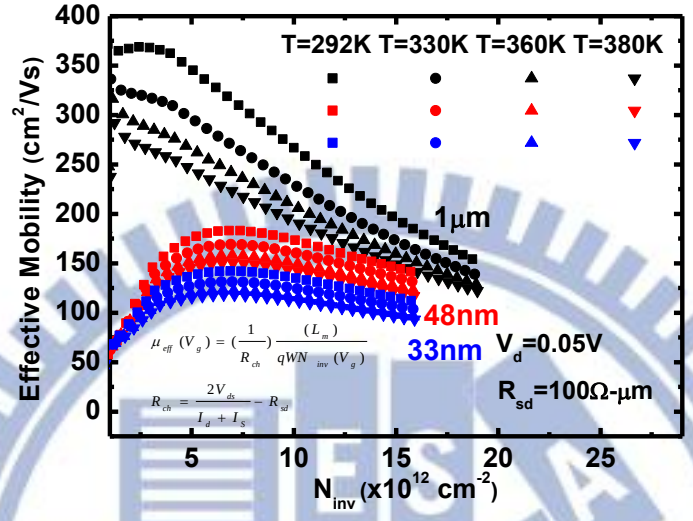
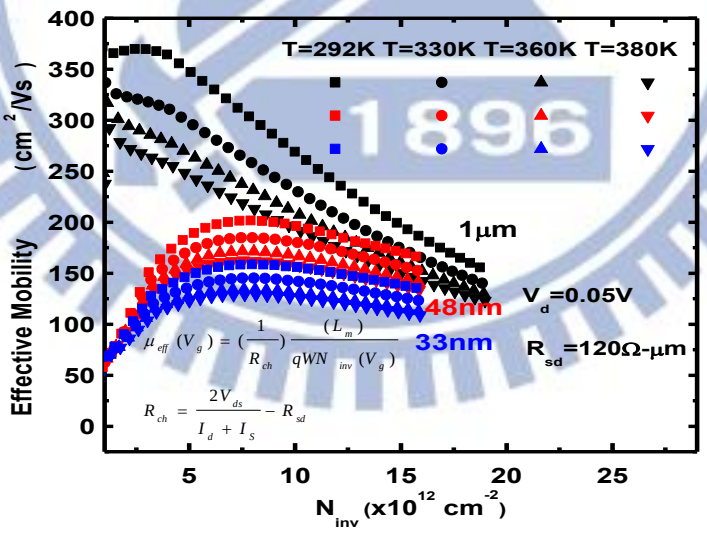


Figure 3.19 The extracted R_{sd} by using the experimental method [13].

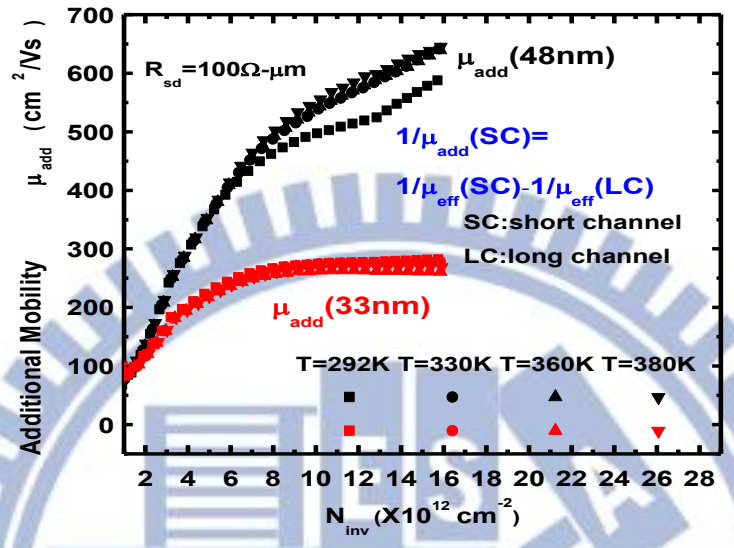


(a)

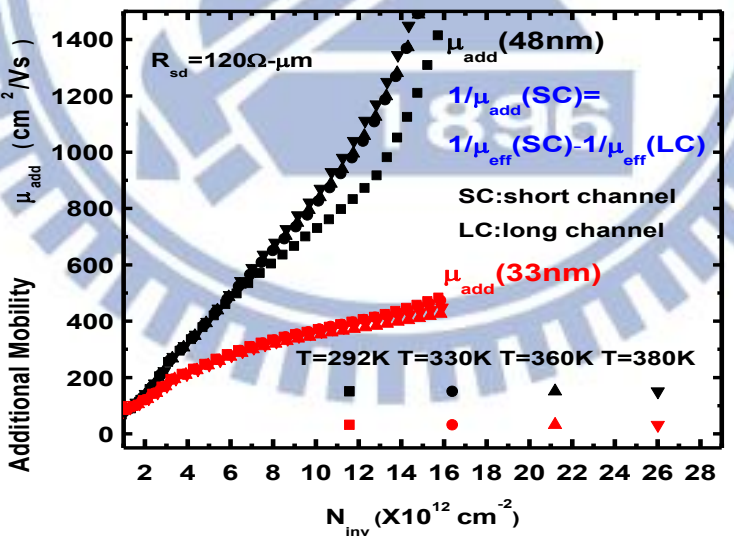


(b)

Figure 3.20 Extracted temperature-dependent effective mobility versus N_{inv} for different L_m with R_{sd} = (a) 100Ω-μm and (b) 120Ω-μm.



(a)



(b)

Figure 4.1 Extracted temperature-dependent additional mobility versus N_{inv} for L_m=48 and 33nm with R_{sd} = (a) 100Ω-μm and (b) 120Ω-μm.

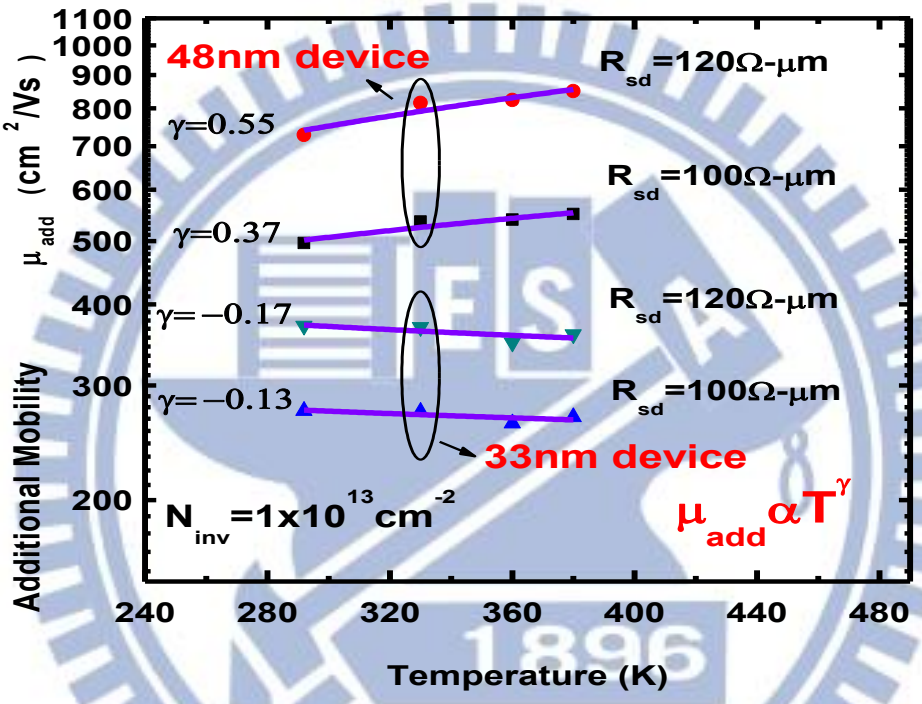


Figure 4.2 Extracted additional mobility at N_{inv}=1x10^{13}cm^{-2} versus temperature with R_{sd} as a parameter. The power-law coefficient γ is obtained by data fitting.

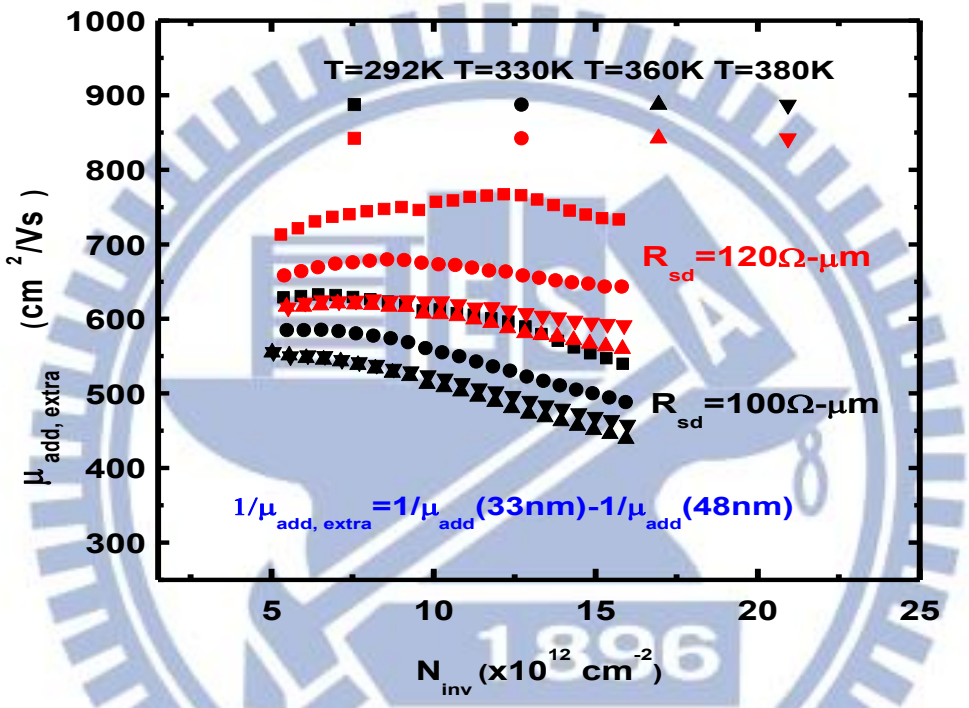


Figure 4.3 Extracted temperature-dependent $\mu_{add,\ extra}$ versus N_{inv} with R_{sd} as a parameter.

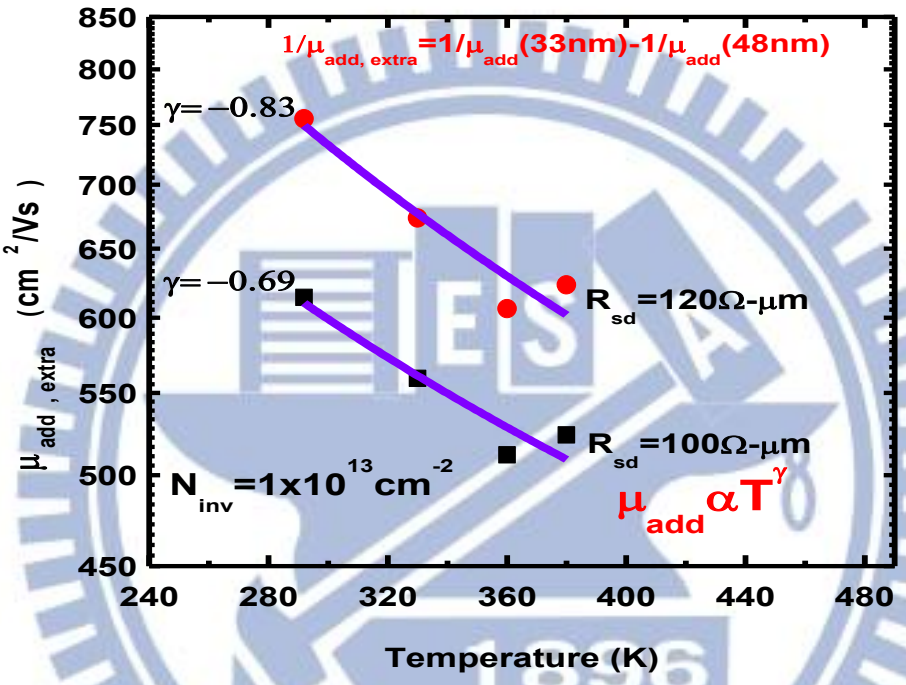
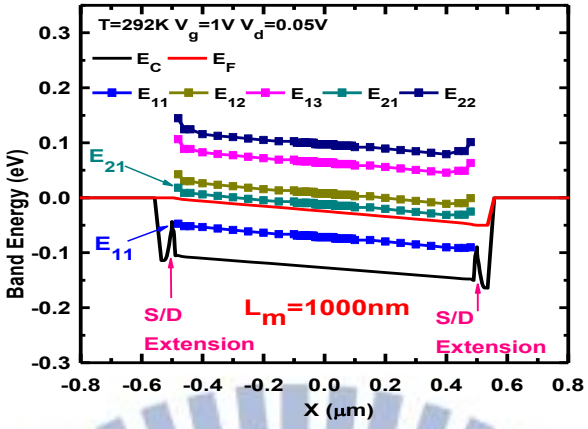
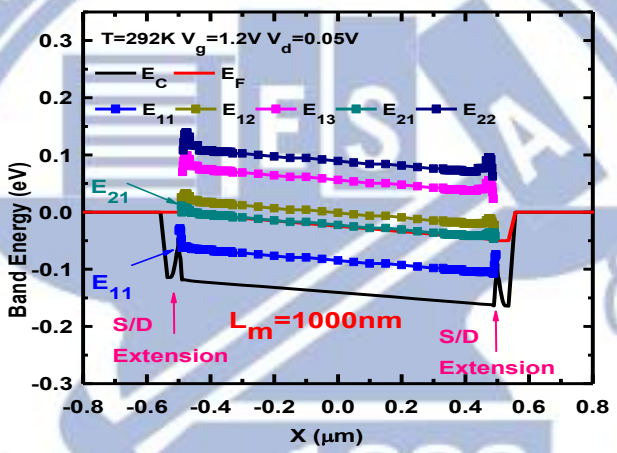


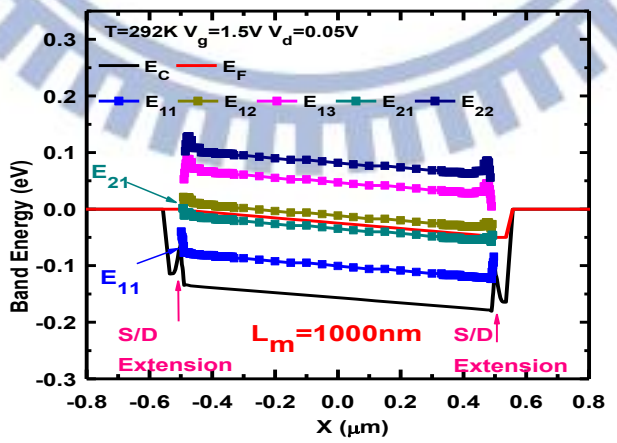
Figure 4.4 Extracted $\mu_{add,\ extra}$ at $N_{inv}=1x10^{13}cm^{-2}$ versus temperature, along with corresponding power-law coefficient γ.



(a)

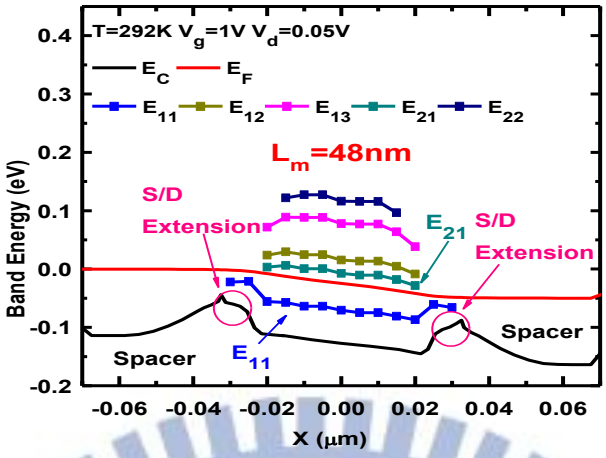


(b)

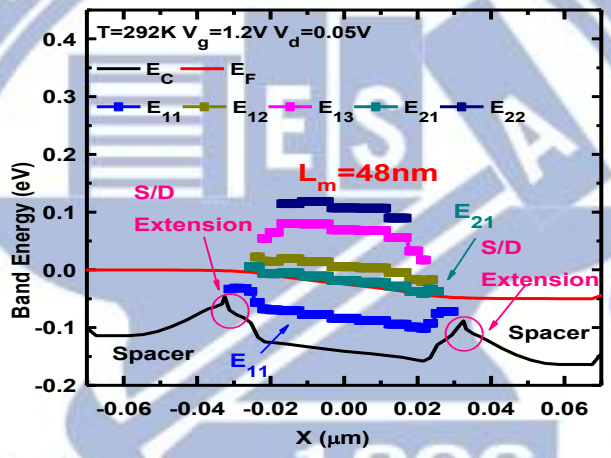


(c)

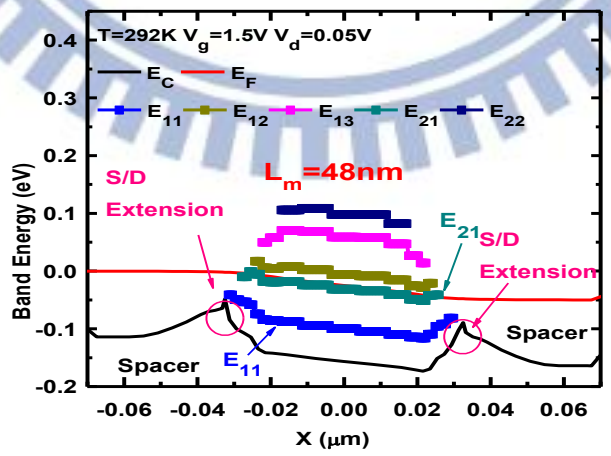
Figure 4.5 Simulated conduction-band edge, subband levels, and Fermi level versus position for L_m = 1μm at V_g = (a) 1V, (b) 1.2V, and (c) 1.5V.



(a)

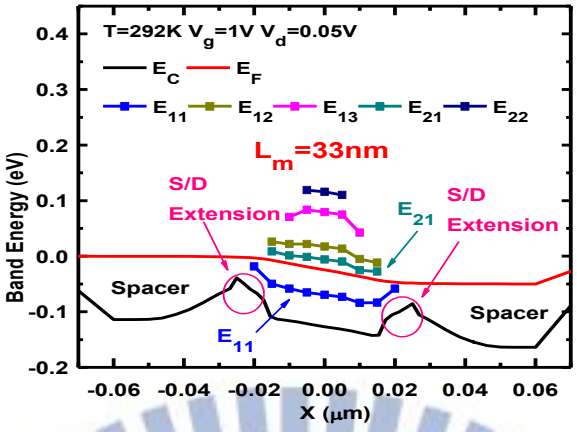


(b)

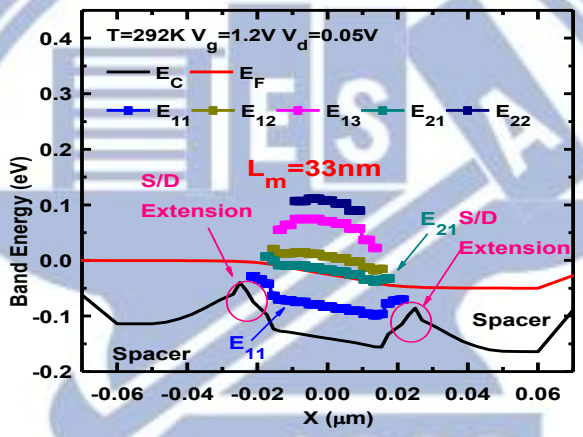


(c)

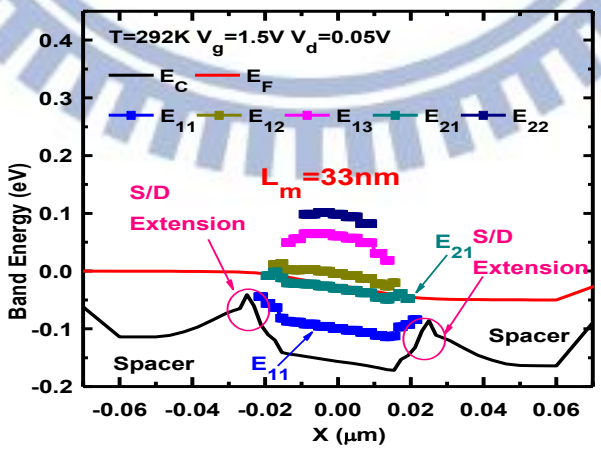
Figure 4.6 Simulated conduction-band edge, subband levels, and Fermi level versus position for L_m = 48nm at V_g = (a) 1V, (b) 1.2V, and (c) 1.5V.



(a)



(b)



(c)

Figure 4.7 Simulated conduction-band edge, subband levels, and Fermi level versus position for L_m = 33nm at V_g = (a) 1V, (b) 1.2V, and (c) 1.5V.

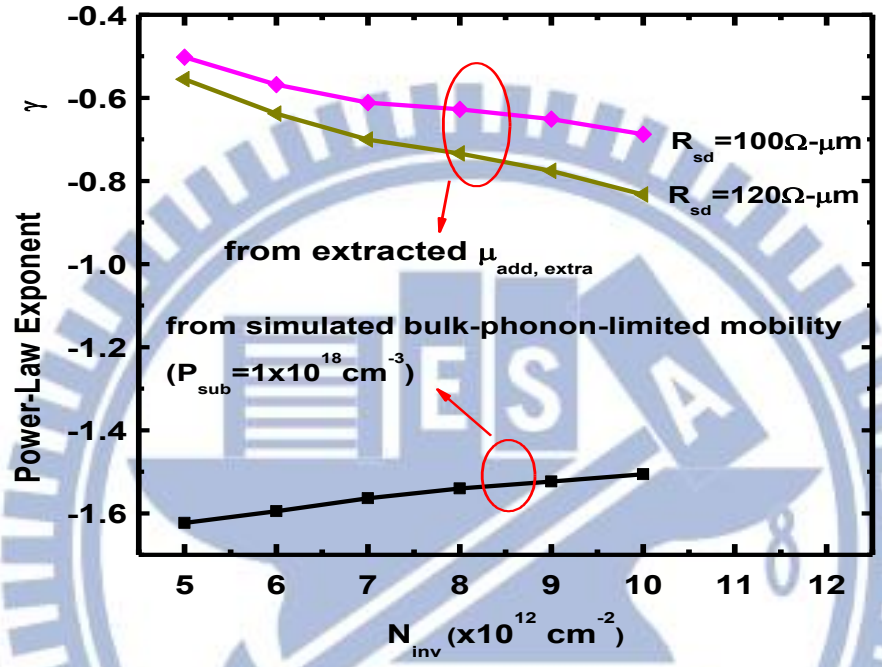


Figure 4.8 Comparison of temperature power-law exponent of extracted $\mu_{add,\ extra}$ and simulated bulk phonon-limited mobility, plotted versus N_{inv}.

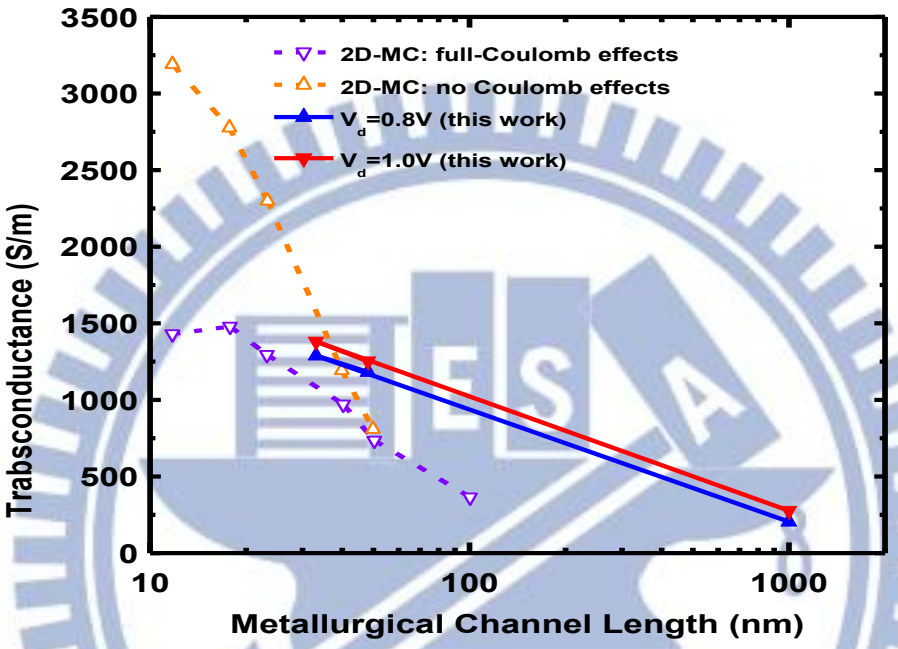


Figure 4.9 Comparison of measured transconductance in this work and simulated ones with and without Coulomb effects [1].

	L_{gate}=1μm	L_{gate}=65nm	L_{gate}=50nm
ΔL	16.776nm	16.9nm	16.89nm
L_m	1μm	48nm	33nm

Table 1 The length of the overlap region for different L_{gate}, and the metallurgical channel length (L_m=L_{gate}-ΔL).

Gate Voltage V_g=1V			
Gate Length	L_m=1μm	L_m=48nm	L_m=33nm
E_{11} occupancy	0.49	0.52	0.53
E_{21} occupancy	0.51	0.48	0.47

(a)

Gate Voltage V_g=1.2V			
Gate Length	L_m=1μm	L_m=48nm	L_m=33nm
E_{11} occupancy	0.46	0.49	0.50
E_{21} occupancy	0.54	0.51	0.50

(b)

Gate Voltage V_g=1.5V			
Gate Length	L_m=1μm	L_m=48nm	L_m=33nm
E_{11} occupancy	0.43	0.45	0.46
E_{21} occupancy	0.57	0.55	0.54

(c)

Table 2 The subband population for different metallurgical channel length at gate voltages of (a) 1V, (b) 1.2V, and (c) 1.5V.