國 立 交 通 大 學

電子工程學系 電子研究所碩士班

碩 士 論 文

射頻金氧半場效電晶體之元件佈局對高頻特性與低頻雜訊之影響以應用於射頻與類比電路

RF MOSFET Layout Effect on High Frequency Characteristics and Low Frequency Noise for RF and Analog Circuit Applications

研 究 生：古智友

指導教授：郭治群 博士

中華民國 九十九年 九月

射頻金氧半場效電晶體之元件佈局對高頻特性與低頻雜訊之影響以應用於射頻與類比電路

RF MOSFET Layout Effect on High Frequency Characteristics and Low Frequency Noise for RF and Analog Circuit Applications

研究生： 古智友　　**Student : Chih-You Ku**

指導教授：郭治群 博士　　**Advisor : Dr. Jyh-Chyurn Guo**

國 立 交 通 大 學

電子工程系　電子研究所

碩 士 論 文

A Thesis

Submitted to Department of Electronics Engineering & Institute of Electronics

College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of Master

In

Electronics Engineering

September 2010

Hsinchu, Taiwan, Republic of China

中華民國 九十九年 九月

射頻金氧半場效電晶體之元件佈局對高頻特性與低頻雜訊之影響以應用於射頻與類比電路

學生：古智友　　指導教授：郭治群 博士

國立交通大學　電子工程學系　電子研究所碩士班

摘要

本論文主要是探討元件佈局造成的 STI 應力，和經由 STI 製程頂邊所產生的圓角 TCR (Top Corner Rounding)與隨之而來的△W 效應對直流 I-V 特性、電導(G_m)，以及載子遷移率(μ_{eff})的影響。同時我們也分析此兩種機制對元件低頻雜訊與高頻特性所產生的影響。為了提高的 STI 垂直方向的應力，我們設計了狹窄型 OD 和多重 OD MOSFETs 兩種元件，其概念是由標準多指元件(Standard multi-finger MOSFETs)延伸而來的。而維繞式 MOSFET 是我們所設計的另一種全新結構，它在從垂直方向完全不受到 STI 應力影響，關鍵就是在於它完全在 width 方向避開了的 STI/OD 界面，並藉此改善低頻雜訊。 但是在某些元件的量測結果顯示，我們無法從 STI 的應力解釋在極端狹窄的元件的反而有更低的低頻雜訊。因次我們從 STI 頂邊 TCR 導致的ΔW 的角度去分析，推導出一 semi-empirical 模型，能夠模擬以預測 width 微縮對μ_{eff} 與 G_m之影響，藉由這些模型可以準確的求出有效的通道寬度 W_{eff}。 STI 應力愈強將導致介面缺陷電荷(N_{it})的增加，但是此效應確遠不及ΔW 明顯，對於狹窄型元件來說是一項優點。但不幸的是狹窄型元件的截止頻率還是會因為 C_{gg} 增加而下降，而且無法經由去寄生的方式改善。

在低頻雜訊與高頻特性之間有著權衡的 RF 元件佈局設計考量。因此基板接點的元件設計是一個有趣的課題，我們設計各種基板接點結構的 4-port 測試元件並將基板端獨

立接出來，將我們提出新的 body network 模型萃取元件參數如基板電阻，它影響高頻元件特性與低頻雜訊特性。量測結果顯示愈大的基板電阻會產生愈大的低頻雜訊，我們可以藉由多重環狀佈局的方式設計基板接點形狀，藉此大量降低基板電阻，但同樣地此舉也會使閘極至基板接點間的雜散電容增加而降低截止頻率。

RF MOSFET Layout Effect on High Frequency Characteristics and Low Frequency Noise for RF and Analog Circuit Applications

Student：Chih-You Ku **Advisor：Dr. Jyh-Chyurn Guo**

Department of Electronics Engineering Institute of Electronics

National Chiao Tung University

Abstract

The impact of layout dependent STI stress and △W from STI top corner rounding (TCR) on I-V characteristics, transconductance (G_m), effective mobility (μ_{eff}), low frequency noise (LFN), and high frequency performance has been investigated in this thesis. Narrow-OD and multi-OD MOSFETs are two structures derived from standard multi-finger MOSFET for enhancing transverse stress $\sigma_{\perp}$. Donut MOSFET is another structure created to keep free from $\sigma_{\perp}$, due to the elimination of STI/OD boundary along the channel width direction.

The compressive stress from STI cannot explain the lower LFN in extremely narrow devices. STI top TCR induced ΔW is identified as an important factor responsible for the increase of G_m and the reduction of LFN with width scaling to nanoscale regime. A semi-empirical model was derived to simulate μ_{eff} degradation from STI stress and the increase of effective width (W_{eff}) from ΔW. The proposed model can accurately predict width scaling effect on G_m based on a trade-off between μ_{eff} and W_{eff}. The enhanced STI stress may lead to an increase of interface traps density (N_{it}) but the influence is relatively minor and overcome by W_{eff} effect. Unfortunately, the extremely narrow devices suffer f_T degradation due to an increase of C_{gg}, which cannot be eliminated even through an improved open

deembedding. The trade-off between LFN and high frequency performance provides an important layout guideline for analog and RF circuit design.

Body contact layout effect on LFN and high frequency performance is one more interesting topic of research in this thesis. Four-port test structures were implemented to accommodate 4-terminal MOSFETs with separate body terminal and a new body network model has been developed to simulate the body contact layout and body biases effects. The measurement result reveals that the higher body resistance will lead to the worse LFN characteristic. We can significantly reduce the body resistance by applying multi-ring body contacts, but this may contribute larger parasitic capacitance from poly gate to body contacts and hence lead to lower f_T.

誌謝

首先，我要感謝我的指導教授--郭治群教授。在研究方法及態度上的指導，不斷地替學生尋找研究資源，並且嚴格要求學生完成許多基礎工作。在這過程中，除了建立許多研究領域相關的能力，在國際學術會議及期刊發表論文，也體會到許多待人接物的觀念，相信將會成為我未來工作上或者研究上的準則。

此外，還要感謝 NDL 的研究員—黃國威博士，在研究設備上的支持，讓我能夠接觸並學習到高頻量測設備。也感謝 RFTC 的工程師們，邱佳松、林書毓、蕭治華在量測實驗上的協助以及建議，讓我能夠順利完成量測。

感謝 LAB635 中的所有成員們，國良學長、冠旭學長、依修學長、仁嘉學長在研究上的幫忙，也感謝敬文、弈岑、唯倫、智翔、德昌的陪伴 ，讓我在實驗室的生活更豐富有趣。

最後，我要感謝一直在背後支持我的家人，他們在我求學與成長的路上付出許多。

Contents

Figure Captions

Chapter 2

Chapter 3

Chapter 5

Table Captions

Chapter2

Chapter3

Chapter5

Chapter 1

Introduction

1.1 Research Motivation

Recent advances in the CMOS technology, which possesses the advantage of high integration and low power, is finding more and more important applications in the area of analog, mixed mode, and RF circuits **[1,2]**. RF integrated circuit design requires accurate device models up to the operating frequency, especially when the operating frequency of the circuit approaches the cutoff frequency (f_T) or maximum oscillation frequency (f_{MAX}).

As devices scaled down into the nanometer-scale regime, the stress induced from shallow trench isolation (STI) process and its impact on the device performance becomes significant [3]. The longitudinal stress effect has been widely investigated; however, relatively fewer studies were focused on the transverse stress effects. Besides, STI stresses, either longitudinal or transverse, exhibits high sensitivity to device layouts, which may significantly impact the analog circuit performance. Meanwhile, the increase of effective, namely ΔW from STI top corner rounding (TCR) emerges as an important factor besides STI stress to affect DC characteristics (drain current and trans-conductance), low frequency noise, and high frequency performance.

The multi-finger transistors have been widely as the standard layout in RF circuits to enhance f_{MAX} and reduce RF noise by minimizing the parasitic gate resistance. In this thesis, three kinds of layouts, namely doughnut (donut), narrow-OD, and multi-OD MOSFETs have been designed to investigate layout dependence of STI stress, particularly the transverse stress. A comprehensive characterization was carried out to investigate the transverse stress effect on mobility, current, low frequency noise (LFN), and high frequency performance. Mullti-finger MOSFETs with different body contact layouts is one more research subject in this thesis.

Four-port test structures adopting five kinds of body contact layouts, denoted as multi-ring, U-shape, L-shape, parallel, and perpendicular stripe have been designed to explore the influence on basic device parameters like body resistance and junction capacitances, and the proliferated impact on low frequency noise and high frequency performance.

1.2 Thesis Organization

This thesis presents our research results of RF MOSFET layout effect on low frequency noise and high frequency characteristics, with major purpose for RF and analog circuit applications. The main effect of our focus is the mechanical stress introduced from shallow trench isolation (STI), which increases with technology scaling and becomes dramatically significant in nanometer scale devices. In this thesis, three types of new MOSFET layout, namely narrow-OD, multi-OD, and donut devices are proposed to modulate the transverse stress from STI.

At first, an overview on low frequency noise theory and measurement methods is presented in chapter 2. Then, the layout dependence of STI stress and its effect on device parameters like drain current (I_{DS}), transcondutance (G_m), and effective mobility (μ_{eff}) are described in chapter 3. The proposed new MOSFETs with layouts of narrow-OD, multi-OD, and donut were fabricated in 90 nm low leakage CMOS process. An extensive device characterization has been carried out through I-V, C-V, S-parameters, and charge pumping (CP) current measurements. Note that the C-V characterization was performed based on S-parameters measurement, which can solve the problems of conventional C-V measurement, such as the parasitic capacitances from pads, interconnection lines, and substrate coupling and gate leakage induced abnormal C-V fall-off. However, one more parasitic capacitance arising from 3-D gate fringing effect cannot be removed from the open deembedding method implemented in S-parameters characterization. Thus, 3-D RLC simulator like Raphael was employed in this thesis to calculate the fringing capacitances contributed from gate sidewalls and finger ends.

Based on the mentioned characterization techniques, STI top corner rounding (TCR) is identified as a key factor affecting I-V, G_m, gate capacitances, LFN, and high frequency performance, etc.

In chapter 4, MOSFET layouts effect on LFN will be presented, with a comprehensive characterization on all of the device structures such as standard multi-finger, narrow-OD, multi-OD, and donut MOSFETs, and covering both NMOS and PMOS. The layout effect can be examined through an extensive comparison between different device structures. Also, the interface trap density N_{it} extracted by CP method is taken as one of key parameters responsible for flicker noise, i.e. LFN under the condition that number fluctuation model is the dominant mechanism, tentatively for NMOS. Two more topics to be covered in chapter 4 are the body contact layouts effect and strain effect on LFN. In this study, totally five body contact layouts, such as multi-ring, U-shape, L-shape, parallel and perpendicular stripes are designed for an investigation on their effect on LFN. A review of strain effect will be classified as stress from strain engineering in 65nm high speed CMOS process and STI stress from new MOSFET layouts introduced in chapter 3. Chapter 5 will focus on MOSFET layout effect on high frequency characteristics, such as f_T, f_{MAX}, and RF noise parameters. An extensive characterization will cover different device structures and also the dependence of frequency and bias. Note that two-port or four-port deembedding is indispensable in high frequency characterization for parasitic RLC extraction and elimination, which is necessary for an accurate extraction of intrinsic device performance under high frequency operation. As for body contact layout effect, four-port test structure is required to accommodate 4-terminal MOSFETs with separate body terminal and a new body network model has been developed to simulate the body contact layout and body biases effects. In the end, chapter 6 concludes this thesis with a summary and suggestions for future work.

Chapter 2

Low Frequency Noise Theory and Measurement Method

2.1 An Overview of CMOS Technology and Low Frequency Noise

It has been well known that noise exists as a fundamental problem in semiconductor devices and electronic circuits. In electronic devices, noise appears as random fluctuations in current or voltage around their DC level, due to fluctuations in carriers transport through the conduction channel. The impact from noise on the desired signal may cause failure of electronic circuits operation when the noise power becomes too large to keep sufficient signal to noise ratio (SNR). Furthermore, the mentioned problem becomes increasingly tough in miniaturized devices accompanied with supply voltage scaling and squeezed dynamic range. Attributed to the nature of random fluctuations, noise is generally characterized by probability density function (PDF) based on multiple measurements over time. The time average of the measured noise currents or noise voltages approaches zero when integrated long enough and provides no useful information; instead, the square quantities namely power spectral density (PSD) is used to characterize noise. The PSD is measured with a spectrum analyzer or dynamic signal analyzer, which will be described in section 2.2.

The experimental indicates that the noise spectral density increases with decreasing frequency at lower frequencies and becomes white thereafter. The corner frequency between the frequency dependent noise and white noise is typically from few Hz up to MHz range and may vary with device types, device dimensions, and bias conditions, etc. The physical mechanism behind the white noise source is well known. However, the excess noise at low frequencies, namely low frequency noise (LFN) brings many questions with lot of debates and open up an interesting research area. Note that the LFN is also known as flicker noise or 1/f noise due to the fact that the frequency dependence sometimes approaches 1/f.

The aggressive advancement of CMOS technology into nanoscale regime in recent five years has driven transistor gate delay to below 10 ps and intrinsic cut-off frequency (f_T) well above 100 GHz **[4]**. According to International Technology Roadmap for Semiconductor (ITRS) **[4]**, the demand on devices and interconnection lines scaling in the area of analog and RF is not as stringent as that for high speed logic CMOS. However, much more stringent criterion on the flicker noise at low frequencies (i.e., LFN) and thermal noise at high frequencies is required and specified for analog and RF circuit design, as shown in Table 2.1(a) and (b) for near term and long term, respectively **[4]**. The yellow blocks marked for 2011~2014 represent that the manufacturing solutions are known but not yet optimized. Unfortunately, the red blocks marked for 2015 and thereafter highlight the problems for which the manufacturable solutions do not exist today. Obviously, the 1/f noise, i.e. LFN emerges as a big roadblock from 2015 and on. As compared with bipolar transistors, CMOS transistors generally suffer higher flicker noise (LFN), due to the nature of surface channel conduction. As a result, the solutions to suppressing LFN become more challenging to CMOS technology, even though CMOS is superior in terms of scalability, high integration, low cost, and low standby power. To facilitate the extension of CMOS technology into RF and analog domain, an in-depth study on LFN in MOSFETs and its dependence on layouts and geometry scaling becomes critically important and is selected as one of research topics in this thesis. It has been recognized that LFN in transistors introduces particular problem in analog and RF circuits like voltage controlled oscillators (VCOs) and mixers. For VCOs, the LFN is upconverted to phase noise at small frequency offsets from the carrier frequency and then sets the ultimate separation limitation to two channels **[5-8]**. Fig. 2.1 illustrates schematically phase noise spectrum and different physical origins. As for mixers, the LFN originated from transistors may lead to severe degradation of SNR **[9,10]**. This can be understood if the signal is translated to very low frequency domain where the flicker noise may dominate and

overwhelm the signal, particularly worse for low voltage operation. The impact from LFN on mixer may be relieved by using long channel transistors; however the penalty of to be paid is the degradation of transconductance and circuit speed.

Table 2.1(a) RF and Analog Mixed-Signal CMOS Technology Requirements – Near term **[4]**

Year of Production	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*	*2015*
DRAM ½ Pitch (nm) (contacted)	*65*	*57*	*50*	*45*	*40*	*35*	*32*	*28*	*25*
Performance RF/Analog [1]									
Supply voltage (V) [2]	1.2	1.1	1.1	1	1	1	1	0.95	0.85
T_{ox} (nm) [2]	2	1.9	1.6	1.5	1.4	1.3	1.2	1.1	1.2
Gate Length (nm) [2]	53	45	37	32	28	25	22	20	18
g_m/g_{ds} at $5 \cdot L_{min\text{-}digital}$ [3]	32	30	30	30	30	30	30	30	30
1/f-noise ($\mu V^2 \cdot \mu m^2/Hz$) [4]	160	140	100	90	80	70	60	50	60
σV_{th} matching ($mV \cdot \mu m$) [5]	6	6	5	5	5	5	5	5	5
I_{ds} ($\mu A/\mu m$) [6]	13	11	9	8	7	6	6	5	4
Peak F_t (GHz) [7]	170	200	240	280	320	360	400	440	490
Peak F_{max} (GHz) [8]	200	240	290	340	390	440	510	560	630
NF_{min} (dB) [9]	0.25	0.22	0.2	<0.2	<0.2	<0.2	<0.2	<0.2	<0.2
Precision Analog/RF Driver [1]									
Supply voltage (V)	2.5	2.5	2.5	1.8	1.8	1.8	1.8	1.8	1.8
T_{ox} (nm) [10]	5	5	5	3	3	3	3	3	3
Gate Length (nm) [10]	250	250	250	180	180	180	180	180	180
g_m/g_{ds} at $10 \cdot L_{min\text{-}digital}$ [11]	220	220	220	160	160	160	160	160	160
1/f Noise ($\mu V^2 \cdot \mu m^2/Hz$) [4]	500	500	500	180	180	180	180	180	180
σV_{th} matching ($mV \cdot \mu m$) [5]	9	9	9	6	6	6	6	6	6
Peak F_t (GHz) [7]	40	40	40	50	50	50	50	50	50
Peak F_{max} (GHz) [8]	70	70	70	90	90	90	90	90	90
									switch to DG device
CMOS NFET [1 HP CMOS lag 2 yrs]	-								
V_{dd}: Power Supply Voltage (V) [13]	1.1	1.1	1.1	1	1	1	0.95	0.9	0.9
EOT: Equivalent Oxide Thickness (Å) [13]	12	11	11	9	7.5	6.5	5.5	5	6
L_g: Physical L_{gate} for High Performance logic (nm) [13]	32	28	25	22	20	18	16	14	13
Peak F_t (GHz) [7]	280	320	360	400	440	490	550	630	670
Peak F_{max} (GHz) [8]	340	390	440	510	560	630	710	820	880
NF_{min} (dB) at 24GHz[14]	2	1.8	1.6	1.4	1.3	1.2	1.1	1	0.9
NF_{min} (dB) at 60GHz[14]	5.1	4.5	4.0	3.6	3.3	3.0	2.7	2.4	2.3
									switch to DG device

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

Table 2.1(b) RF and Analog Mixed-Signal CMOS Technology Requirements – Long term **[4]**

Year of Production	*2016*	*2017*	*2018*	*2019*	*2020*	*2021*	*2022*
DRAM ½ Pitch (nm) (contacted)	*22*	*20*	*18*	*16*	*14*	*13*	*11*
Performance RF/Analog [1]							
Supply voltage (V) [2]	0.8	0.8	0.8	0.8	0.75	0.75	0.7
T_{ox} (nm) [2]	1.1	1.1	1	1	0.9	0.9	0.8
Gate Length (nm) [2]	16	14	13	12	11	10	10
g_m/g_{ds} at 5·$L_{min\text{-}digital}$ [3]	30	30	30	30	30	30	30
1/f-noise ($\mu V^2 \cdot \mu m^2/Hz$) [4]	50	50	40	40	30	30	30
σ V_{th} matching (mV·μm) [5]	4	4	4	4	3	4	5
I_{dx} (μA/μm) [6]	4	3	3	3	2	2	2
Peak F_t (GHz) [7]	550	630	670	730	790	870	870
Peak F_{max} (GHz) [8]	710	820	880	960	1050	1160	1160
NF_{min} (dB) [9]	<0.2	<0.2	<0.2	<0.2	<0.2	<0.2	<0.2
Precision Analog/RF Driver [1]							
Supply voltage (V)	1.8	1.8	1.8	1.5	1.5	1.5	1.5
T_{ox} (nm) [10]	3	3	3	2.6	2.6	2.6	2.6
Gate Length (nm) [10]	180	180	180	130	130	130	130
g_m/g_{ds} at 10·$L_{min\text{-}digital}$ [11]	160	160	160	110	110	110	110
1/f Noise ($\mu V^2 \cdot \mu m^2/Hz$) [4]	180	180	180	135	135	135	135
σ V_{th} matching (mV·μm) [5]	6	6	6	5	5	5	5
Peak F_t (GHz) [7]	50	50	50	70	70	70	70
Peak F_{max} (GHz) [8]	90	90	90	120	120	120	120
	switch to DG device						
CMOS NFET [1 HP CMOS lag 2 yrs]							
V_{dd}: Power Supply Voltage (V) [13]	0.9	0.8	0.8	0.7	0.7	0.7	0.65
EOT: Equivalent Oxide Thickness (Å) [13]	6	6	5.5	5.5	5.5	5	5
L_g: Physical L_{gate} for High Performance logic (nm) [13]	11	10	9	8	7	6	5.5
Peak F_t (GHz) [7]	790	870	960	1080	1220	1420	1550
Peak F_{max} (GHz) [8]	1050	1160	1300	1470	1690	1990	2180
NF_{min} (dB) at 24GHz[14]	0.8	0.7	0.6	0.6	0.5	0.4	0.4
NF_{min} (dB) at 60GHz[14]	2.0	1.8	1.6	1.4	1.2	1.0	0.9
	switch to DG device						

Manufacturable solutions exist, and are being optimized
Manufacturable solutions are known
Interim solutions are known
Manufacturable solutions are NOT known

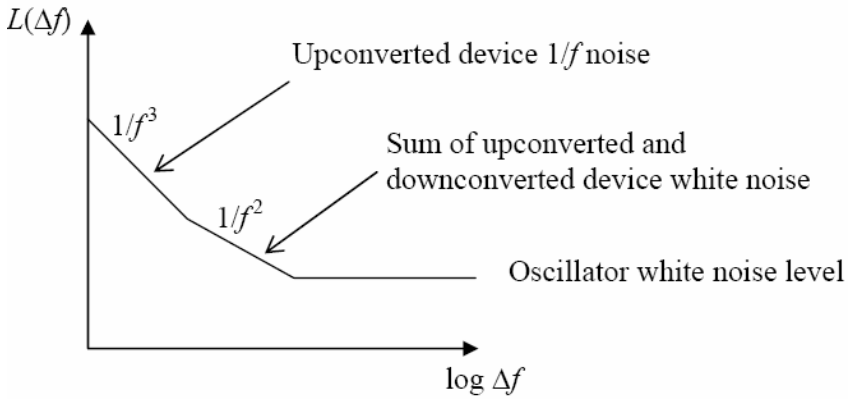


Fig.2.1 Schematical illustration of phase noise spectrum in VCO

2.2 Low Frequency Noise Theory

In the past several decades, the origins and physical mechanisms underlying flicker noise remain an open question, with lot of debates and arguments in the experimental results and modeling to match the measurement. Number fluctuation model and mobility fluctuation model appear as two most popular mechanisms to explain and predict the measured flicker noise **[11,12]**. In 1957, McWhorter published a flicker noise model based on quantum mechanical tunneling transitions of electrons between the gate oxide and channel **[11]**. In practice, the tunneling time varies exponentially with distance, and it is assumed that trap density is uniform in both energy and distance from the channel interface to extract the time constants for generating flicker noise. The McWhorter model, namely number fluctuation model may be useful due to its simplicity and good agreement with experimental, particularly for n-channel MOSFETs **[12,13]**. However, the mobility fluctuation model appears to better explain the flicker noise measured from p-channel MOSFETs **[14,15]**. As compared to surface channel MOSFETs, buried-channel MOSFETs or bipolar junction transistors (BJT) demonstrate significantly lower flicker noise **[16-21]**. The published results are in favor of the number fluctuation model that the flicker noise is originated from the traps in the oxide or at oxide/channel interface. However, the surface carrier mobility is reduced compared to the bulk value due to additional surface scattering (acoustic phonon and surface roughness), which has an impact on the mobility fluctuation. Hooge mobility noise **[22]**, which is sensitive to the crystalline quality, can be employed to explain the higher flicker noise for surface channel devices in which the carriers are in close proximity to the gate oxide and may suffer aggravated mobility fluctuation. In the following, the number fluctuation and mobility fluctuation models will be described in more detail.

2.2.1 Number Fluctuation Theory [11]

The physical mechanism underlying the number fluctuation noise is the interaction between the channel carriers and slow traps in the gate oxide, which is illustrated in Fig. 2.2. The dynamic exchange of carriers between the gate oxide and channel causes a fluctuation in the surface potential (ψ_S) and then gives rise to fluctuations in the inversion carrier density δQ_{inv}. This in turn leads to noise in the drain current. Note that δQ_{inv} (the fluctuation in the inversion carrier density) can occur even without a current flowing the channel and the channel current is only used to sense the fluctuations. The mathematical formulas for expressing number fluctuation model in different operation regions are provided as follows

In weak inversion region

$$\frac{S_{I_{DS}}}{I_{DS}^{\ 2}} = \frac{\lambda N_t}{WLC_{ox}^{\ 2}k_BT}\left(\frac{q^2}{m}\right)^2 f^{-\gamma},\ \ m = 1+\frac{C_D + C_{it}}{C_{ox}} \tag{2.1}$$

In strong inversion region

$$linear\ \ region : g_m = \frac{W}{L}C_{ox}\mu_{eff}V_{DS} \Rightarrow \left(\frac{g_m}{I_{DS}}\right)^2 = \left(\frac{1}{V_{GS}-V_T}\right)^2 \tag{2.2}$$

$$\frac{S_{I_{DS}}}{I_{DS}^{\ 2}} = \frac{q^2k_BT\lambda N_t}{f^{\gamma}WLC_{ox}^{\ 2}}\left(\frac{g_m}{I_{DS}}\right)^2 = \frac{q^2k_BT\lambda N_t}{f^{\gamma}}\frac{WC_{ox}\mu_{eff}^{\ 2}V_{DS}^{\ 2}}{L^3}\frac{1}{I_{DS}^{\ 2}} \tag{2.3}$$

$$\text{saturation } region : g_m = WC_{ox}\upsilon_{sat} \Rightarrow \left(\frac{g_m}{I_{DS}}\right)^2 = \left(\frac{1}{V_{GS}-V_T}\right)^2 \tag{2.4}$$

$$\frac{S_{I_{DS}}}{I_{DS}^{\ 2}} = S_{V_{FB}}\left(\frac{1}{V_{GS}-V_T}\right)^2 = \frac{q^2k_BT\lambda N_t}{f^{\gamma}WLC_{ox}^{\ 2}}\left(\frac{1}{V_{GS}-V_T}\right)^2 \tag{2.5}$$

N_t : *the density of traps at quasi - Fermi level*

The frequency dependence with the exponent γ may deviate from 1 under the condition that the trap density N_t is not uniform in depth. For the case when the trap density near the gate oxide/channel interface is higher than that in the interior of the gate oxide, γ tends to be

smaller than 1. For the opposite case, γ may become larger than 1. As for the bias dependence predict by the number fluctuation model, the normalized drain current noise S_{IDS}/I_{DS}^2 varies with approximately as $1/I_{DS}^2$ or $1/(V_{GS}-V_T)^2$ in strong inversion region given by (2.2)~(2.5) while is nearly independent of bias in weak inversion region, shown in (2.1). In this work, the LFN in terms of S_{IDS}/I_{DS}^2 measured from n-channel MOSFETs just follows number fluctuation model and varied with I_{DS} according to the relationship of $1/I_{DS}^2$.

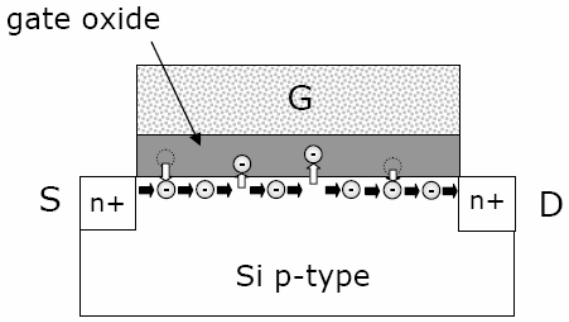
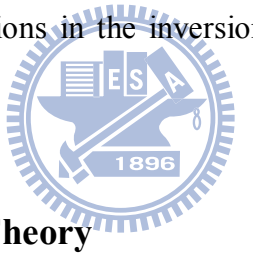


Fig.2.2 Schematical illustration of electrons in the channel of MOSFET moving in and out of the traps giving rising to fluctuations in the inversion carrier density and thereby the drain current.



2.2.2 Mobility Fluctuation Theory

Mobility fluctuation is another mechanism, which can contribute flicker noise. The mobility fluctuation model was first proposed by F.N. Hooge with an empirical formula given for the resistance fluctuation [23]. According to the Hooge empirical formula, the drain current noise generated by fluctuation in the channel carrier mobility can be written as (2.6).

$$\frac{S_{I_{DS}}}{I_{DS}^{\ 2}} = \frac{q\alpha_H}{fWLQ_{inv}} \tag{2.6}$$

where α_H is a dimensionless parameter and referred as Hooge parameter. The typical values of α_H range between 10^{-3} and 10^{-6} for surface channel transistors. α_H may be down to 10^{-7} for buried channel transistors like N^+ gate pMOSFETs and even lower to the order of 10^{-8} for JFETs. Note that phonon scattering was proposed as the primary source generating mobility

fluctuation noise **[22]**. The effective mobility μ_{eff} of the channel carriers is determined by different scattering mechanisms, which vary in different ways with the effective normal field $E_{\perp eff}$ as a function of inversion carriers density Q_{INV} and body depletion charge Q_B. As a result, α_H is not only dependent on technology but also on the bias conditions. In general, each scattering process generates mobility fluctuation noise with the amount given by each respective Hooge parameter, denoted as α_{Hj}. Assume the scattering processes are independent of each other and then Matthiessen's rule can be applied as follows

$$\frac{1}{\mu_{eff}} = \sum_j \frac{1}{\mu_j} \tag{2.7}$$

The fluctuations in different scattering processes are assumed independent. Then the variation applied to (2.7) can lead to

$$\frac{\Delta\mu_{eff}}{\mu_{eff}^{\ 2}} = \sum_j \frac{\Delta\mu_j}{\mu_j^{\ 2}} \tag{2.8}$$

The power spectral density can be derived as

$$\frac{S_{I_{DS}}}{I_{DS}^{\ 2}} = \sum_j \left(\frac{\mu_{eff}}{\mu_j}\right)^2 \frac{S_{\mu_j}}{\mu_j^{\ 2}} \tag{2.9}$$

$$\frac{S_{\mu_j}}{\mu_j^{\ 2}} = \frac{q\alpha_{H,j}}{fWLQ_{inv}} \tag{2.10}$$

$$\frac{q\alpha_H}{fWLQ_{inv}} = \sum_j \left(\frac{\mu_{eff}}{\mu_j}\right)^2 \frac{q\alpha_{Hj}}{fWLQ_{inv}} \Rightarrow \alpha_H = \sum_j \left(\frac{\mu_{eff}}{\mu_j}\right)^2 \alpha_{H,j} \tag{2.11}$$

It can be understood from (2.11) that α_H varies with biases due to the bias dependent factor $\left(\mu_{eff}/\mu_j\right)^2$. The total drain current noise is evaluated by adding the noise contribution from each channel segment derived for linear region as follows.

$$\frac{S_{I_{DS}}}{I_{DS}^{\ 2}} = \frac{q\alpha_H}{fWL^2}\int_0^{V_{DS}} \frac{W\mu_{eff}}{I_{DS}} dV = \frac{q\alpha_H\mu_{eff}}{fL^2 I_{DS}} V_{DS} \tag{2.12}$$

$$\frac{S_{I_{DS}}}{{I_{DS}}^2} = \frac{q\alpha_H}{f \cdot WLC_{ox}} \frac{1}{(V_{GS} - V_T)} = \frac{q\alpha_H}{f \cdot WLC_{ox}} \frac{1}{V_{GT}} \tag{2.13}$$

The drain current noise contributed from mobility fluctuation as shown in (2.13) predicts that the larger α_H and shorter channel length (L) will lead to higher flicker noise whereas the wider channel width (W) and higher gate overdrive V_{GT} can help reduce the flicker noise. The thinner gate oxide thickness is another beneficial factor, which can suppress flicker noise. In our experimental, the mobility fluctuation model can provide much better fit to the flicker noise measured from p-channel MOSFETs than number fluctuation model.

2.3 Low Frequency Noise Measurement System

Three flicker noise measurement systems with different equipment configurations are employed in this thesis for characterization of low frequency noise (LFN) in MOSFETs with different layouts. The conventional one is Celestry 9812B Noisepro system, which is supported by NDL and has been commonly used in transistor flicker noise characterization but limited to measurement using DC probes. The other two are new systems built up in our RF Lab. – one is SR570 system and the other one is model-5184 system. The major difference from Celestry 9812B is that these two LFN systems can support both RF and DC test structures, with RF GSG probes and DC probes, respectively. **SR570** system shown in **Fig. 2.3** is equipped with SR570 LNA (trans-resistances amplifier), Agilent DSA 35670 (dynamic signal analyzer), and HP4142B (DC power supply). The model-5184 LNA shown in **Fig. 2.4** contains 5184 LNA (voltage amplifier), PA14A1 for ultra low noise DC source, low pass filter, Agilent 35670, and HP4142B. Note that Agilent ICCAP is adopted for measurement auto-control and data collection.

1) SR570 Low Frequency Noise Measurement System

Fig. 2.3 illustrates SR570 LFN measurement system in which the SR570 LNA was a

trans-resistances amplifier produced by Stanford Research Systems. Essentially, LNA acts as a key element in this flicker noise measurement system through which it can provide the second stage low noise gain to reduce system noise and offer sufficiently large signal amplitude to dynamic signal analyzer (DSA) without increasing the undesired noises. In practice, the SR570 LNA can provide an offset current source at the DUT output (i.e., drain terminal of MOSFETs), which serves as an input to the amplifier. The DUT output current can be modulated by the low noise offset current source, so as not to drive the amplifier into saturation. The SR570 is also equipped with a filter with different options, such as low-pass, high –pass or bandpass filter. In our applications, we typically set bandpass filter to select the frequency range of our interest, i.e. 1 ~ 100k Hz. The HP4142B DC power supply was controlled by ICCAP to provide the gate bias for MOSFET through a low pass filter in order to remove the power supply's noise within the frequency range of interest that is above 1Hz. In other words, the supply voltage through the low pass filter is nearly a pure DC voltage source under the condition that the signal above 1Hz was removed by the low pass filter. The LNA output is connected to DSA, which can perform dynamic signal ($V/\sqrt{Hz}$) measurement and analysis. All the equipments such as 4142B, SR570 LNA, and Agilent DSA 35670 are integrated together and the measurement can be carried out through the control of ICCAP . Through a dynamic signal analysis done by DSA, we can obtain the measured noise from ICCAP. In general, we have to make a simple calculation to recover the original noise measured from DUT output stage. The measured noise is transformed to current spectrum density with a unit of A^2/Hz.

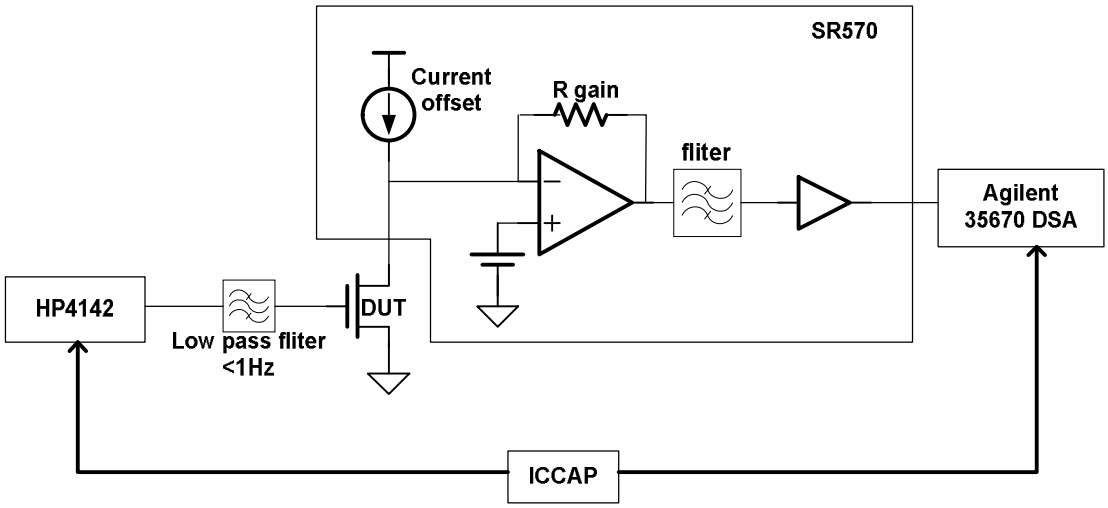


Fig.2.3 Schematic of SR570 flicker noise measurement system setup with SR570 LNA (trans-resistance amplifier), Agilent DSA 35670 (dynamic signal analyzer), and HP4142B (DC power supply)

Unfortunately, The measurement system adopting SR570 LNA has a stringent current limitation of 5 mA. For advanced multi-finger RF MOSFETs, which sometimes have large total width for sufficient current drivability and transconductance, this current limitation generally restricts the DUTs to limited operation conditions, such as linear or subthreshold regions. To solve the mentioned problem, we proposed a new system configuration by using model-5184 LNA to overcome the current limitation. Unlike SR570, model 5184 LNA is a kind of voltage amplifier with a separated output bias source. In the following, we will have a discussion on this new LFN measurement system.

2) **New Configuration with Model 5184 LNA and ultra low noise DC source (PA14A1)**

An appropriate operation of SR570 LNA requires that the gain setting resistor should be lower than the output resistance of DUT (r_{out}). As for the new solution using model 5184 LNA (voltage amplifier), the voltage gain and LNA noise are independent of the output resistance of DUT. In this way, the LNA noise is independent of the DUT and thus we can obtain a system noise level from measurement without DUT. The noise floor of the system can be pushed to 10^{-20} (V^2/Hz), which is well below that of DUT in the frequency range of our interest. However, one major drawback of this voltage-mode LNA is that it does not include a

DUT output biasing source, thus reducing the system integration level. Therefore, a reliable and ultra-low noise DC source at DUT output is required to minimize the extra noise contributed from the system. Battery is one of low noise sources, which can effectively reduce the noise generated from power supply. Besides the simple setup using batteries, we propose a new solution using an ultra low noise DC source (PA14A1) instead of batteries. A complete system configuration incorporating batteries or low noise DC source (PA14A1) is illustrated in **Fig. 2.4.** In practice**,** we implement on a PCB (printed circuit board) the metal lines in three paths for connecting LNA (model 5184), low noise source (PA14A1), and DUT drain terminal as shown in **Fig. 2.5 (a)** and **(b)**.

In Fig.2.5(b), the ultra low noise DC source (PA14A1) was used. The advantage of using PA14A1 is that it can achieve the desired output voltage but a simple setup with batteries shown in Fig. 2.5(a) cannot meet the requirement.

In conclusion, we can obtain almost the same measurement results under the same bias condition by using SR570 LNA and model 5184 LNA either with batteries or PA14A1. Besides the mentioned achievement, we have extended this work to 4-terminal MOSFET flicker noise measurement in both NMOS and PMOS.

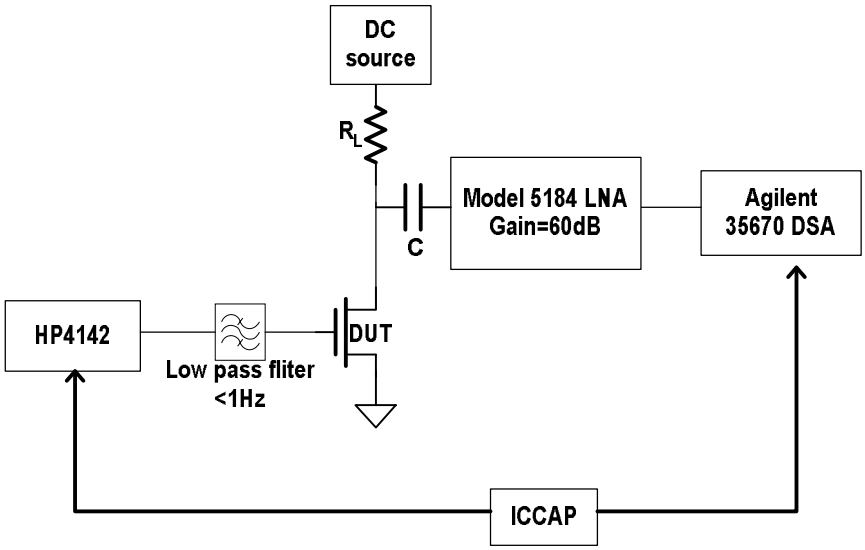
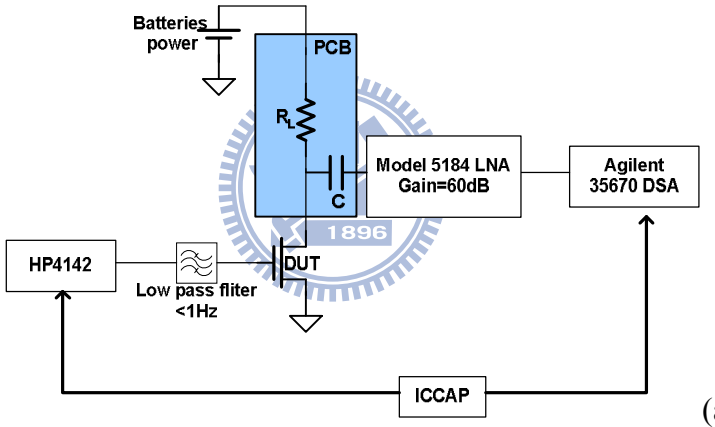
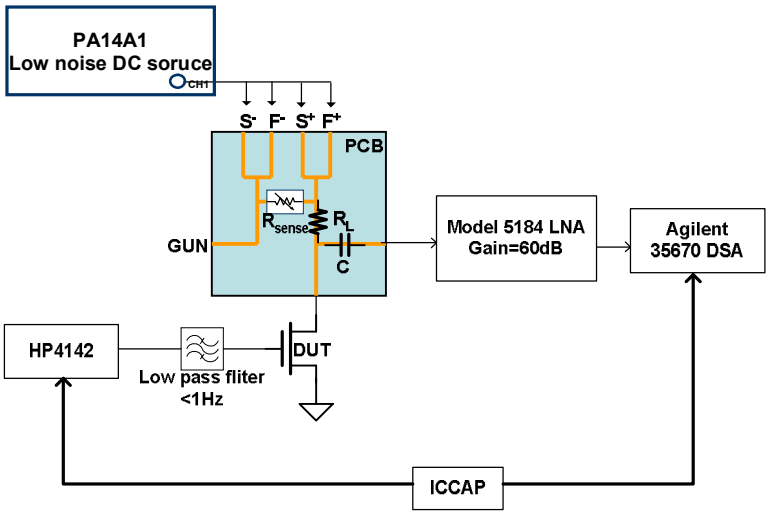


Fig.2.4 Schematic of low frequency noise measurement system consisting of model 5184 LNA (voltage amplifier), HP4142B for DC supply, and Agilent 35670 for dynamic signal analyzer



(a)



(b)

Fig.2.5 New solution for using model 5184 LNA (a) Battery (b) ultra low noise DC source (PA14A1) as a power supply

3) Celestry 9812B NoisePro System

Fig. 2.6 illustrate the configuration of Celestry 9812B system, later named as ProPLUS NoisePro system, which consists of 9812B analyzer (hardware), NoisePro software with automatic control, and built-in 1/f noise SPICE model parameter extraction routine to provides off-the-shelf 1/f noise system.

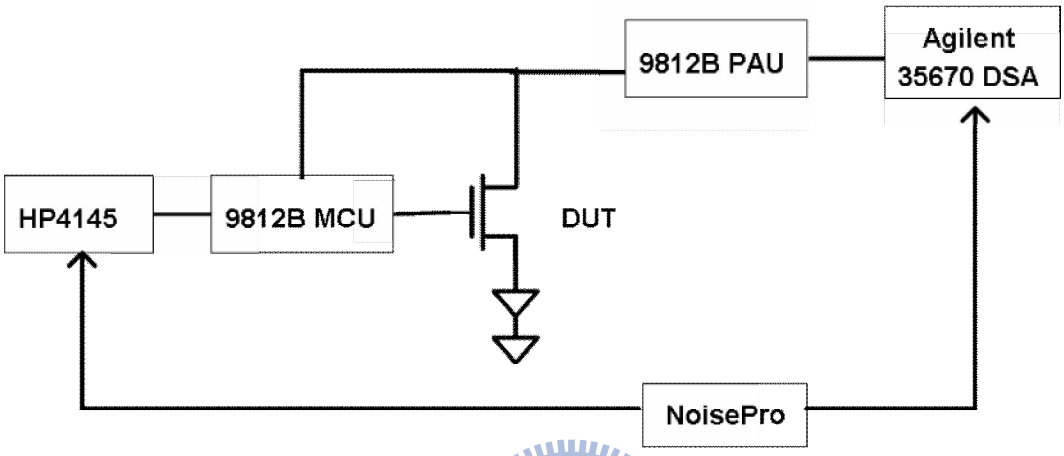


Fig.2.6 Schematic illustration of Celestry 9812B Noisepro system consisting of HP4145 for DC supply, 9812B for pre-amplifier, and Agilent 35670 for dynamic signal analyzer.

Celestry 9812B system can support noise characterization and model parameter extraction for MOSFETs, BJTs, diodes and resistors, based on either public or proprietary models. This system offers V-mode and I-mode biasing modes with accuracy specification of user options. Detachable pre-amplifier unit can be placed closer to DUTs or in the measurement chamber, which greatly reduces the system noise floor. The battery operated current amplifier provides as low as $0.5pA/(Hz)^{1/2}$ @ 1 KHz floor noise and voltage amplifier provides $1nV/(Hz)^{1/2}$ @ 1 KHz floor noise.

When proceeding a system calibration, NoisePro will automatically perform a leakage testing, and the system will apply voltage of 1-Volt and measure the leakage current. The measured values of leakage and load resistors will be saved in a file to be used in the noise measurement. There are seven load resistors used in a noise amplifier. Based on the model

and version of noise amplifier, their values differ. **Table 2.2** lists the default values of load resistors **(R_{load})** used in noise amplifiers and those of filter resistors in load circuit. If 9812B is selected, user needs to choose the type of amplifier : current or voltage amplifier, internal or external amplifier. If external amplifier is selected, the user needs to provide method of converting the A, B, C values.

Table 2.2 Load resistors and filter resistors used in 9812B pre-amplifier

9812B load R (Ω)		9812B Filter Rd (Ω)	
R_{d1}	100	Short	300
R_{d2}	333	Medium	1K
R_{d3}	1k	Long	3K
R_{d4}	3.3k		
R_{d5}	10k		
R_{d6}	33k		
R_{d7}	100k		

2.4 Low Frequency Noise Data Analysis Method

As mentioned previously, an extensive noise characterization indicates that low frequency noise (LFN) in nMOSFETs follows the number fluctuation model whereas that in pMOSFET is in better agreement with mobility fluctuation model. For MOSFETs, the measured LFN is generally expressed as normalized PSD (power spectral density) of drain current noise, denoted as S_{ID}/I_{DS}^2 to identify which model is the dominant one in the devices under test. According to the normalized PSD derived for number fluctuation model for subthreshold, linear, and saturation regions in (2.1), (2.3), and (2.5), respectively, the S_{ID}/I_{DS}^2 is proportional to $1/I_{DS}^2$ and increases linearly with oxide trap density N_t. On the other hand for the normalized PSD governed by mobility fluctuation model, expressed in (2.12) and (2.13) as a function of I_{DS} and V_{GT}, respectively, the S_{ID}/I_{DS}^2 is proportional to $1/I_{DS}$ or $1/ V_{GT}$ and independent of the oxide trap density N_t. The bias dependence suggests that increasing gate

overdrive V_{GT} can suppress LFN. Sometimes, the experimental present V_{GS} dependence stronger than what predicted by model. McWhorter's carrier trapping model may provide one of mechanisms responsible for the stronger V_{GS} dependence. The oxide trap density N_t is known to be a function of the potential energy across the band gap of semiconductor substrate. When V_{GS} is varied, Fermi level of the channel moves accordingly. If the number of particular traps that cause flicker noise drastically changes along with the V_{GS}, the noise amplitude could be a stronger function of the bias. Another factor that we should take into account is the fact that carrier mobility is also a function of V_{GT}. Regarding device scaling effect, both models reveal dependence of device dimensions in terms of $1/WLC_{ox}^2$, which means the channel length and width scaling will lead to increase of S_{ID}/I_{DS}^2 but the oxide thickness T_{OX} scaling can help reduce LFN in terms of S_{ID}/I_{DS}^2. Note that W, L, C_{ox}, μ_{eff}, α_H, and V_{GT} appear as key parameters in LFN models and how to accurately extract or determine the mentioned parameters becomes critical for LFN data analysis and diagnosis. In LFN models, W and L represent effective channel width and effective channel length, rather than drawn dimensions on layout or physical dimensions from optical measurement. The accurate extraction of effective channel dimensions brings additional challenges for miniaturized MOSFETs with short channel and narrow width to sub-100 nm regimes in 90nm technology and below. STI top corner rounding (TCR) effect on effective channel width will be described in chapter 3. As for $\mathbf{C_{ox}}$, it represents the gate capacitance density per unit area, under strong inversion condition in which $\mathbf{C_{ox}}$ is contributed from three major capacitances in series – gate oxide capacitance, gate depletion capacitance, and inversion channel capacitance. Again, an accurate extraction of C_{ox} becomes difficult in miniaturized MOSFETs, due to several complicated factors like deviations in effective channel length and width, gate depletion effect, inversion channel quantization effect, and more importantly the parasitic capacitances from pads, interconnection lines, and substrate. A large device with channel length and width large

enough to minimize or even eliminate the deviation of effective channel dimensions is adopted to solve this problem. Furthermore, a dedicated open deembedding to the bottom metal, i.e. M1 is necessary to approach an accurate extraction of C_{ox}.

The accuracy or deviation of mentioned basic device parameters will have direct impact on the determination of μ_{eff} and α_H. Furthermore, both μ_{eff} and α_H are a function of scattering mechanisms and biases. The bias dependence of μ_{eff} and α_H comes from electrical field effect on carriers transport and their scattering through the conduction channel. Note that the gate bias dependence manifests normal field effect and is expressed in terms of gate overdrive $V_{GT}=V_{GS}-V_T$ in which V_T variations from processes, geometries, profiles, and biases can be taken into account. As shown in **Fig. 2.7**, the measured LFN in terms of S_{ID}/I_{DS}^2 decrease with increasing $|V_{GT}|$ for both NMOS and PMOS. **Fig. 2.8 (a)** and **(b)** illustrate S_{ID}/I_{DS}^2 versus V_{GT} for NMOS and PMOS, respectively and reveal the difference of V_{GT} dependence between NMOS and PMOS. The results can be explained by **(2.3)** and **(2.5)** for NMOS governed by number fluctuation, and **(2.13)** for PMOS dominated by mobility fluctuation.

Frequency dependence is one more key factor for LFN analysis and modeling. Assuming an ideal condition that trap density N_t is a constant, the frequency dependence of flicker noise will follow 1/f. However, in most of real cases, the exponent γ in frequency dependence may deviate from 1 due to the fact that the trap density N_t is not uniform in depth. For the case when the trap density near the gate oxide/channel interface is higher than that in the interior of the gate oxide, γ tends to be smaller than 1. For the opposite case, γ may become larger than 1. Referring to **Fig. 2.7**, the frequency dependence of PMOS (Fig.2.7(b)) follows 1/f very consistently but that of NMOS (Fig. 2.7(a)) reveals certain deviation from 1/f , with the exponent $\gamma < 1$.

It has been a conventional belief from analog circuit design using N^+ poly gate technology that PMOS has the advantage of lower flicker noise (LFN) than NMOS. It may be a truth for

buried channel PMOS, attributed to deeper channel and less surface scattering. However, the advancement of CMOS process has driven phase-out of N^+ poly gate technology and deployment of dual gate technology, i.e. N^+/P^+ poly gate for N/P MOSFET, since 0.25 μm technology node and beyond. The dual gate technology can realize surface channel for both NMOS and PMOS and help push V_T scaling for low voltage operation. In this way, the advantage of PMOS over NMOS, in terms of flicker noise can no longer be maintained. **Fig. 2.9** makes a comparison of measured LFN (S_{ID}/I_{DS}^2) between NMOS and PMOS under the same bias, i.e. $|V_{GT}|$=0.3V and $|V_{DS}|$=0.05V. As can be seen clearly that S_{ID}/I_{DS}^2 of PMOS is higher than NMOS by around one order at very low frequency (<10 Hz) and the difference decreases at higher frequency. The results suggest that conventional design using PMOS in VCO for low phase noise is no longer valid. A detailed analysis of LFN to investigate the mechanisms responsible for NMOS and PMOS and the optimization design for suppressing LFN become an important task.

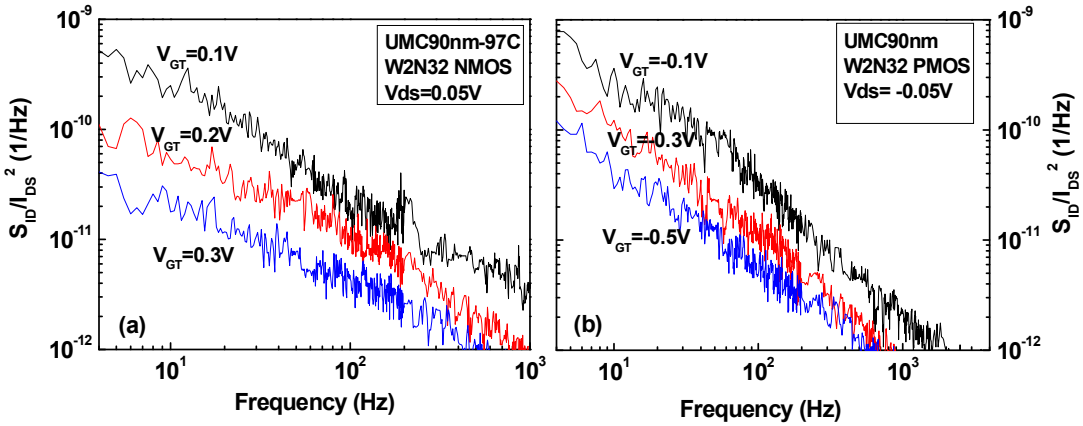


Fig.2.7 Normalized PSD of drain current noise S_{ID}/I_{DS}^2 measured from (a) NMOS and (b) PMOS with increasing frequency, under various V_{GT} ($|V_{GT}|$=0.1, 0.3, and 0.5V). NMOS and PMOS fabricated in 90nm low leakage process.

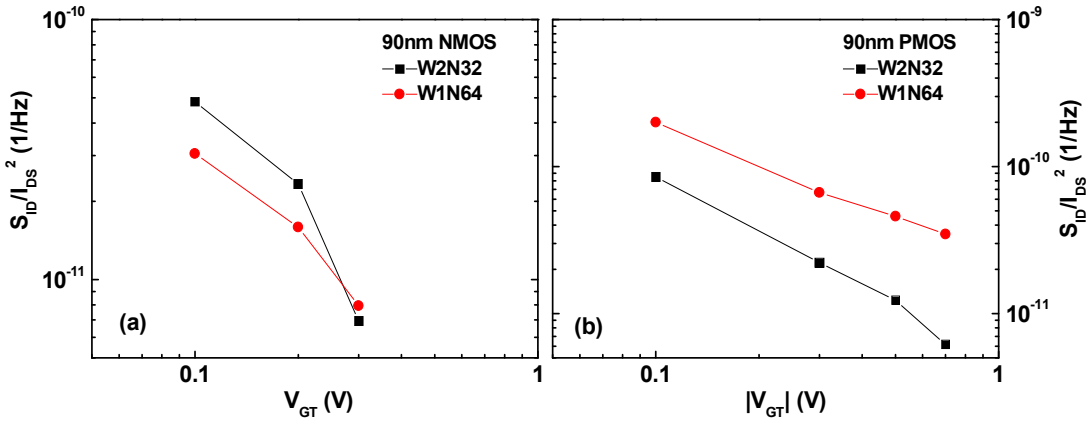


Fig.2.8 Normalized PSD of drain current noise S_{ID}/I_{DS}^2 measured from (a) NMOS and (b) PMOS under increasing $|V_{GT}|$ and fixed $|V_{DS}|$=0.05V.

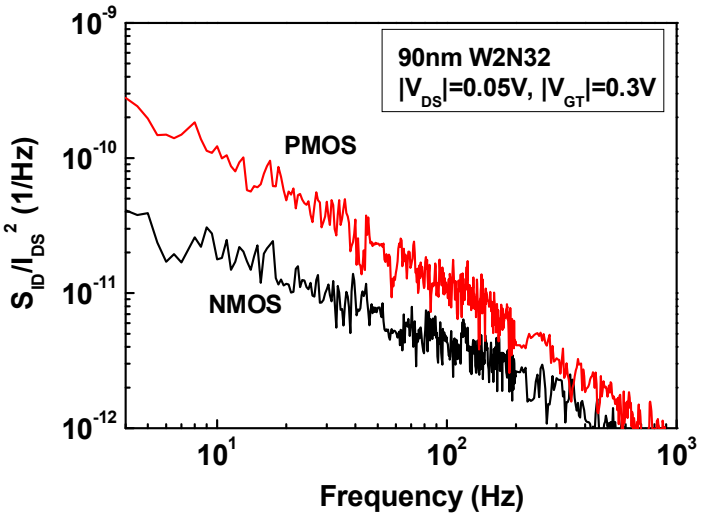


Fig.2.9 Normalized PSD of drain current noise S_{ID}/I_{DS}^2 versus frequency, measured from (a) NMOS and (b) PMOS under $|V_{GT}|$=0.3V and $|V_{DS}|$=0.05V. NMOS and PMOS were fabricated in 90nm low leakage CMOS process.

Chapter 3

Analysis of Layout Effect on STI Stress and Device Parameters

3.1 STI stress Mechanics in RF MOSFET

Shallow trench isolation (STI) emerged to replace LOCOS as the standard isolation technology in CMOS process at 0.25 μm node and below **[24,25].** The technology transition proves itself the success that STI enables continuous scaling of active and isolation regions to far beyond 0.25μm node, and realize higher integration level for IC design. Also, STI brings the advantages like reduced sidewall capacitance for AC performance and improved surface planarity for high integration **[26]**. In spite of the mentioned advantages, STI indeed introduces some other impacts, such as mechanical stress from refilled trench or known as Length of Oxide Definition (LOD) stress-effect **[27,28]**. The STI stress may have a significant influence on MOSFETs' electrical characteristics, such as threshold voltage (V_T), carrier effective mobility (μ_{eff}), transconductance (G_m), gate speed, cut-off frequency (f_T), and low frequency noise, etc. The mentioned impacts always increase with devices scaling and becomes the key factors of consideration in MOSFETs layout for integrated circuit design, particularly for RF and analog circuits of our focus in this thesis.

The STI module implemented in standard CMOS processes from 0.25μm to 65nm nodes generally introduce compressive stress, either along or transverse to the channel length, i.e. the current flow direction. The stress along the channel length is defined as longitudinal stress $\sigma_{//}$, and that transverse to the channel length is denoted as transverse stress $\sigma_{\perp}$. The reverse type of stress, namely tensile stress will introduce its influence in the opposite direction with that of compressive stress. The impact from STI stress on device characteristics is strongly dependent on the device types and orientations. **Table 3.1** summarizes the stress favorable for mobility enhancement and reveals fundamental differences between NMOS and PMOS. For NMOS, tensile stress, either $\sigma_{//}$ or $\sigma_{\perp}$ can

improve μ_{eff}. As for PMOS, compressive stress in $\sigma_{//}$ or tensile stress in $\sigma_{\perp}$ is the right one for μ_{eff} enhancement **[29]**. The results indicate that compressive stress in transverse direction, i.e. $\sigma_{\perp}$ along the channel width always leads to mobility degradation in both NMOS and PMOS. It means that channel width scaling will result in G_m degradation due to μ_{eff} degradation under the condition that STI stress is the only one factor influencing μ_{eff}. However, the experimental from a comprehensive coverage of narrow device layouts reveals an interesting result that an increase of effective channel width, namely ΔW from STI top corner rounding (TCR) emerges as an anti-factor, which will trade off with μ_{eff} degradation from STI stress in determining G_m. The details will be described in sections 3.4 and 3.5.

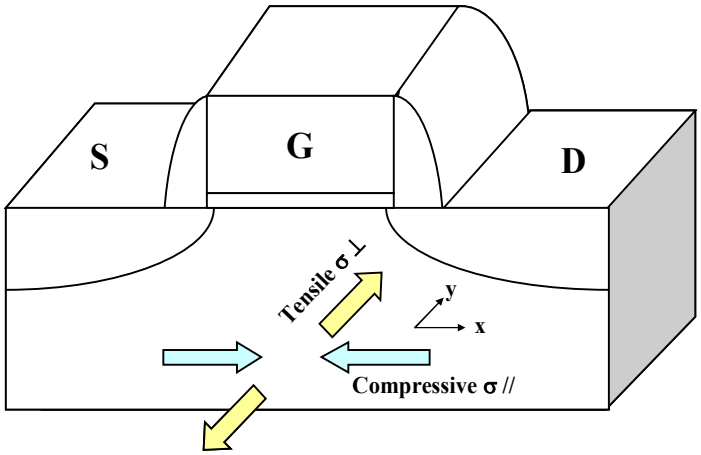


Figure 3.1 Schematic views of stress types and orientations

Table 3.1 Stress favorable for mobility enhancement in NMOS and PMOS along longitudinal and transverse directions **[29]**

Directions	Stress favorable for mobility enhancement	
	NMOS	PMOS
Longitudinal ($\sigma_{//}$)	Tensile	Compressive
Transverse ($\sigma_{\perp}$)	Tensile	Tensile

3.2 MOSFET Layouts for STI stress Modulation (Multi-finger, Multi-OD, and Doughnut Structures)

In this work, MOSFETs with various layouts were fabricated in 90nm low leakage CMOS process with nitrided oxide of target physical thickness at 2.2nm. The gate length drawn on the layout is 90nm, i.e. L_{drawn}=90nm and the total channel width W_{tot} are specified at 32 μm or 64 μm. To investigate STI transverse stress effect, two new layouts derived from multi-finger MOSFET, namely narrow-OD and multi-OD MOSFETs with extremely narrow widths were designed and implemented. Note that OD means oxide diffusion, which is equivalent to active area, generally denoted as AA.

Fig. 3.2 displays narrow-OD MOSFET layouts in which gate finger numbers (N_F) and finger width (W_F) are varied simultaneously to keep $W_F \times N_F = W_{tot}$. In this work, three splits of $W_F \times N_F$, namely W2N16 ($W_F \times N_F$=2μm×16), W1N32 ($W_F \times N_F$=1μm×32), and W05N62 ($W_F \times N_F$=0.5μm×64) corresponding to W_{tot}=32μm were fabricated, as shown in Fig. **3.2**(a)~(c). The multi-OD MOSFETs shown in Fig.3.3 represent multiple OD fingers with simultaneously varied OD finger width (W_{OD}) and OD finger number (N_{OD}) under a specified finger width, which is $W_F = W_{OD} \times N_{OD}$. As shown in Fig. 3.3(a)~(c), three splits of layout, namely OD1 ($W_{OD} \times N_{OD}$=2μm×1), OD8 ($W_{OD} \times N_{OD}$=0.25μm×8) and OD16 ($W_{OD} \times N_{OD}$=0.125μm×16) are designed corresponding to W_F=2μm and N_F is fixed at 16. Note that the poly-gate edge to OD edge distance along the direction of channel length is fixed at 0.5 μm for both narrow-OD and multi-OD MOSFETs.

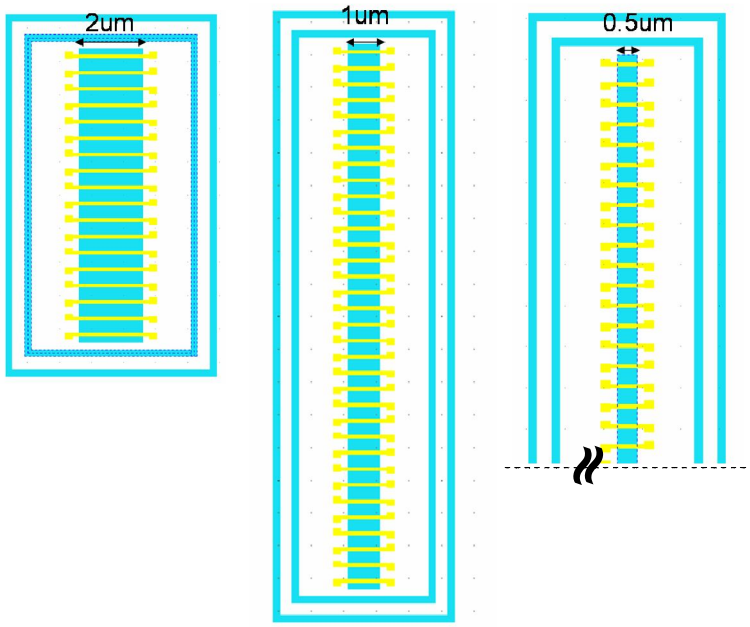


Fig.3.2 Schematics of narrow-OD MOSFETs with three layouts (a) W2N16 (W_F×N_F=2μm×16) (b) W1N32 (W_F×N_F=1μm×32) (c) W05N62 (W_F×N_F=0.5μm×64) corresponding to W_{tot}=32μm

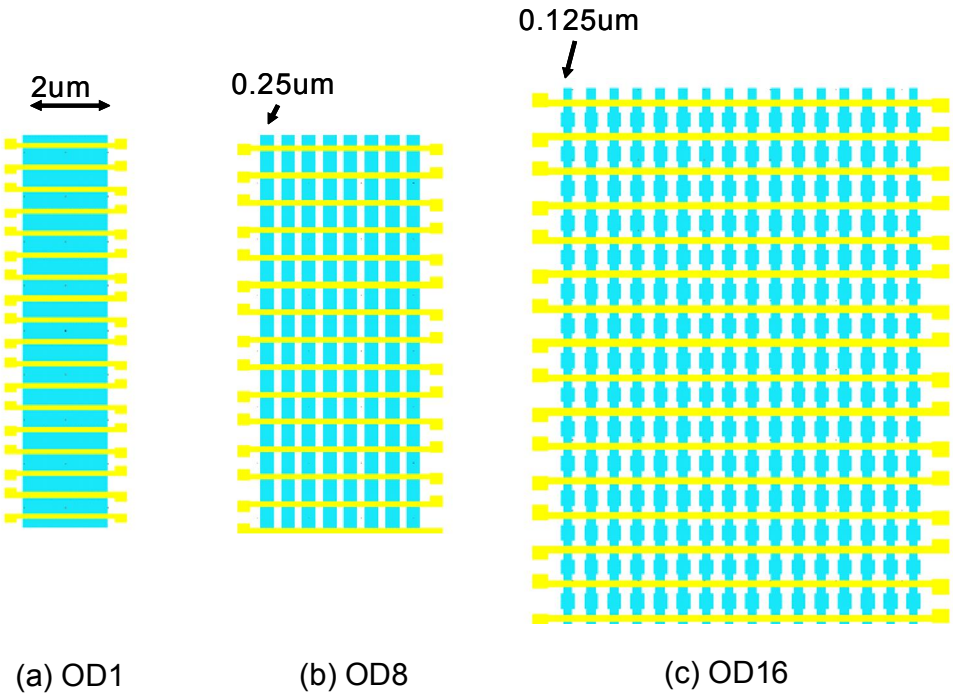


Fig.3.3 Schematics of multi-OD MOSFETs with three layouts (a) OD1 (W_{OD} × N_{OD}=2μm×1) (b) OD8 (W_{OD} × N_{OD}=0.25μm×8) (c) OD16 (W_{OD}×N_{OD}=0.125μm×16)

Besides the multi-finger MOSFETs, a new MOSFET layout, namely doughnut (donut) is proposed to create devices free from STI transverse stress σ_{∞}, along the channel width. As shown in **Fig. 3.4**, donut MOSFETs are constructed as 4-side polygons in which the corners contribute very little to the channel current **[30]**. In this thesis, donut devices with two layout dimensions were implemented. In **Fig. 3.4(a)**, D1S1 represents donut MOSFET in which the space from poly gate to STI edge follows the minimum rule, i.e. 0.3μm, to maximize the compressive stress from STI along the channel (i.e., longitudinal stress $\sigma_{//}$). Meanwhile, D10S10 shown in **Fig. 3.4(b)** denotes donut MOSFET with 10 times larger space between poly gate and STI edge, i.e. 3μm, intentionally to relax $\sigma_{//}$ from STI.

In the following section, an extensive characterization has been performed on both NMOS and PMOS devices to explore the STI stress effect on channel current (I_{DS}), transconductance (G_m) and effective mobility (μ_{eff}).

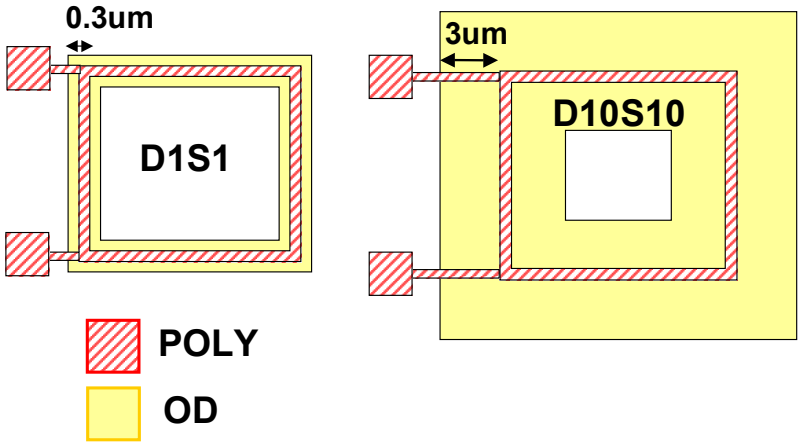


Fig.3.4 A brief layout of donut MOSFET with two major layers, such as active region (OD) and poly gate (PO) (a) D1S1 (poly gate to STI edge distance =0.3μm (min. rule) (b) D10S10 (poly gate to STI edge distance =3μm=10x min. rule), 4-side polygons length is 16 μm x4.

3.3 Layout Effect on DC Characteristic

Regarding layout dependent STI stress effect on electrical characteristics of MOFSETs, DC characteristics such as I-V, V_T, and G_m are the most fundamental ones to be verified.

Besides the mentioned DC characteristics, effective mobility μ_{eff} is the most important parameter to be extracted for investigating STI effect on carrier transport. At first, I-V characterization was performed using HP4156B Semiconductor Parameter Analyzer. For the extraction of μ_{eff}, I-V characteristics in linear region and C-V characteristics from high frequency S-parameter measurement have been carried out. Note that an accurate extraction of intrinsic gate capacitance is indispensable for accurate determination of inversion carriers density Q_{inv}, which is required for μ_{eff} extraction.

In the following, the effective mobility μ_{eff} is extracted from linear I-V under very low drain bias V_{DS} ($|V_{DS}|$=0.05V in this work) and Q_{inv} determined by intrinsic gate capacitance from S-parameters through an appropriate deembedding. In this work, S-parameters were measured by Agilent network analyzer E8364B for high frequency characterization up to 40 GHz and AC parameters extraction. Open de-embedding was performed on the measured two-port or four-port S-parameters to remove the parasitic capacitances from the pads as well as interconnection lines and short de-embedding was done to eliminate the parasitic resistances and inductances originated from the metal interconnection lines.

$$I_{DS} = W_{eff} Q_{inv} \mu_{eff} \frac{V_{DS0}}{L_{eff}} \tag{3.1}$$

where, L_{eff} and W_{eff} are the effective channel length and width of the intrinsic channel region; V_{DS0} is the internal drain voltage, which is applied to the intrinsic channel excluding the parasitic drain and source resistances, given by

$$V_{DS0} = V_{DS} - I_{DS}(R_D + R_S) \tag{3.2}$$

Assume that $I_{DS}(R_D+R_S) << V_{DS}$ at V_{DS}= 0.05V to simplify the problem and easily derive μ_{eff} as follows (an example for NMOS, and an opposite sign for all biases for PMOS)

$$\mu_{eff} = \frac{I_{DS}}{V_{DS}} \frac{1}{\frac{W_{eff}}{L_{eff}} Q_{inv}} \tag{3.3}$$

$$Q_{inv} = C_{ox(inv)}(V_{GS} - V_T - \alpha V_{DS})$$
$$\alpha \le \frac{1}{2} \quad for\ MOSFET\ in\ linear\ region \tag{3.4}$$

note that (3.4) is valid under strong inversion condition, i.e.

$$(V_{GS} - V_T - \alpha V_{DS}) > 0$$

$C_{ox(inv)}$ is the intrinsic gate capacitance density per unit area, under strong inversion condition, which can be calculated by intrinsic Y-parameters ($Im(Y_{11})$) from S-parameters after an open deembedding to the bottom metal, i.e. M1

$$C_{ox(inv)} = \frac{C_{gg,\text{int}}}{W_{eff} L_g} \cong \frac{C_{gg,\text{int}}}{WL} \tag{3.5}$$

For a long-channel and wide width MOSFET

$$C_{gg,\text{int}} \cong C_{gg(DUT,OM1)} = \frac{\text{Im}(Y_{11}^{DUT,OM1})}{\omega} \tag{3.6}$$

$$\left[Y_{ij}^{DUT,OM1}\right] = \left[Y_{ij}^{mea}\right] - \left[Y_{ij}^{openM1}\right] \tag{3.7}$$

Note that $C_{ox(inv)}$ is determined from a large device in which W and L are large enough to make the critical dimension loss negligibly small compared to the drawn dimension and validate the equivalence of drawn dimension and physical dimension, i.e.

$$W_{eff} \cong W, \quad L_g \cong L$$

Once we have obtained accurate values of the effective mobility μ_{eff}, we will be at a position for analyzing the possible mechanisms responsible for the mobility enhancement or degradation in MOSFETs. However, the $\mathbf{C_{gg(DUT)}}$ extracted after an open deembedding contains not only inversion channel but also fringing capacitances from gate sidewall and finger ends, which cannot be removed even after an open deembedding to M1 (openM1). These gate related fringing capacitances are not scalable with the scaling of device dimensions, such as gate length and width. As a result, the fringing capacitances occupy a significant rate in $\mathbf{C_{gg(DUT)}}$ and may dominate that from intrinsic channel region in very short

and/or very narrow MOSFETs. Under this condition, the inversion carrier calculated directly by $\mathbf{C_{gg(DUT)}}$ will be overestimated and then the extracted μ_{eff} is underestimated. The increase of W_{eff} from STI TCR, namely ΔW and an accurate extraction of ΔW from I-V and C-V methods will be described in sections 3.4 and **3.6**. Through an appropriate correction on W_{eff}, which is indispensable for very narrow MOSFETs, the μ_{eff} can be accurately extracted for narrow-OD and multi-OD MOSFETs to clarify STI stress effect or STI TCR induced ΔW effect. The methods developed in this thesis can facilitate a rigorous investigation on the mechanisms responsible for mobility enhancement or degradation in miniaturized MOSFETs.

The STI stresses introduced in MOSFETs with standard and narrow-OD layouts (Fig. 3.2) are illustrated in **Fig. 3.5** to assist an analysis and understanding of layout effect on STI stress and the electrical characteristics. Note that STI stress is classified as longitudinal stress, denoted as $\sigma_{//}$, which is in parallel with the channel length, and transverse stress, namely $\sigma_{\perp}$, which is transverse to the channel length. In this work, the longitudinal stress $\sigma_{//}$ is considered to be similar for all of devices with various layouts, due to fixed gate length and poly gate edge to OD edge distance **[31]**. Referring to Table 3.1, the stress favorable for electron mobility in NMOS is tensile stress in both longitudinal direction ($\sigma_{//}$) and transverse direction ($\sigma_{\perp}$). However, the stress favorable for hole mobility in PMOS becomes compressive stress in longitudinal orientation ($\sigma_{//}$) but keeps tensile stress in transverse direction ($\sigma_{\perp}$) **[29]**.

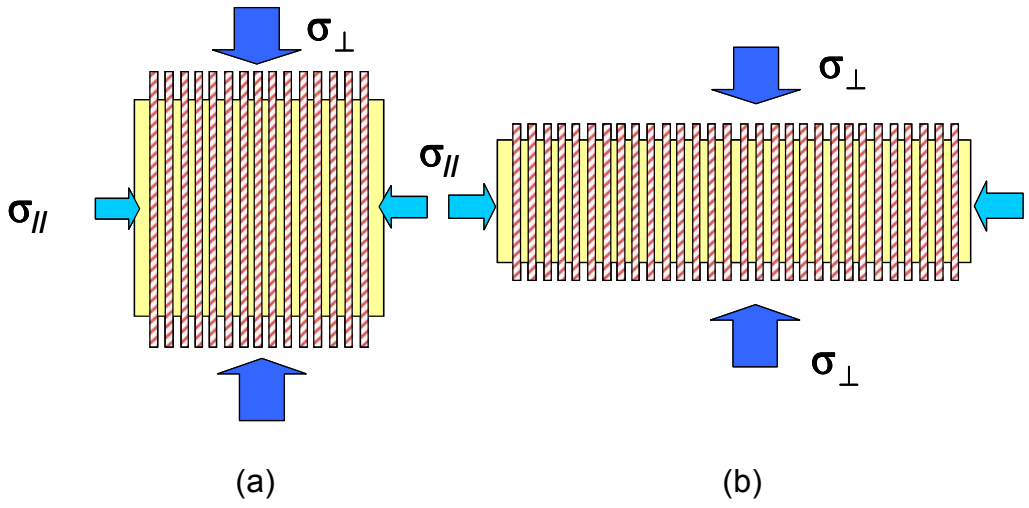


Fig.3.5 Schematics of STI stresses along the longitudinal and transverse directions, defined as $\sigma_{//}$ and $\sigma_{\perp}$ in MOSFETs with different layouts (a) standard multi-finger MOSFET (b) narrow-OD MOSFET

It has been known that STI process generally leads to V_T lowering with channel width scaling and it is the so called inverse narrow width effect (INWE) **[32]**. As shown in **Fig. 3.6**, the V_T vs. W_{OD} for narrow-OD NMOS present an obvious INWE. Note that INWE is determined by collective effects from doping profile, 2D charge sharing effect, trench corner field crowding, STI stress, and STI TCR (ΔW), etc. Considering V_T variations from the mentioned effects, $V_{GT}=V_{GS}-V_T$ is used for electrical characterization and analysis. **Fig. 3.7** presents the drain current (I_{DS}) and transconductance (G_m) under varying V_{GT}, measured from narrow-OD NMOS with two splits, such as W1N32 (1μm×32) and W05N64 (0.5μm×16) and the standard one W2N16 for a comparison. The results indicate that the smaller W_{OD} ($=W_F$) leads to lower G_m and also the maximum G_m ($G_{m,max}$). It is found that $G_{m,max}$ of W1N32 ($W_{OD}=W_F=1\mu m$) is degraded by around 2% but that of W05N64 ($W_{OD}=W_F=0.5\mu m$) is degraded by as large as 8%, compared with the standard NMOS, i.e. W2N16 ($W_{OD}=W_F=2\mu m$). The monotonic degradation of G_m with W_{OD} scaling in narrow-OD devices suggests that the increase of STI compressive $\sigma_{\perp}$ is the dominant factor responsible for μ_{eff} degradation and the resulted G_m degradation.

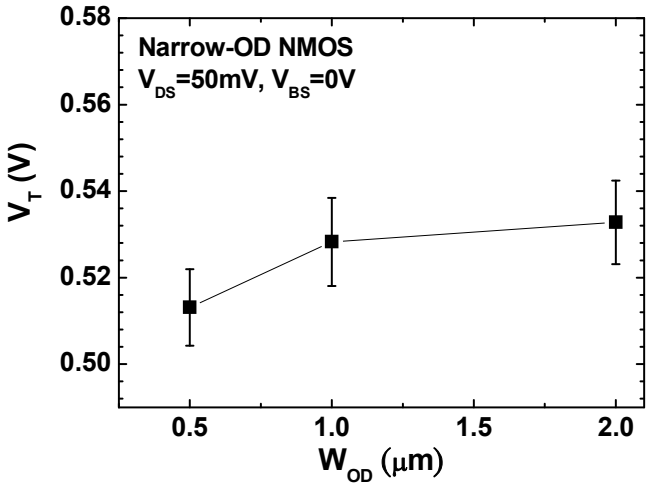


Fig.3.6 Linear V_T versus W_{OD} for narrow-OD NMOS with various layouts like W2N16

(W_{OD}=2μm, N_F=16), W1N32 (W_{OD}=1μm, N_F=16), W05N64 (W_{OD}=0.5μm, N_F=16) under the biases of V_{DS}=50mV and V_{BS}=0V.

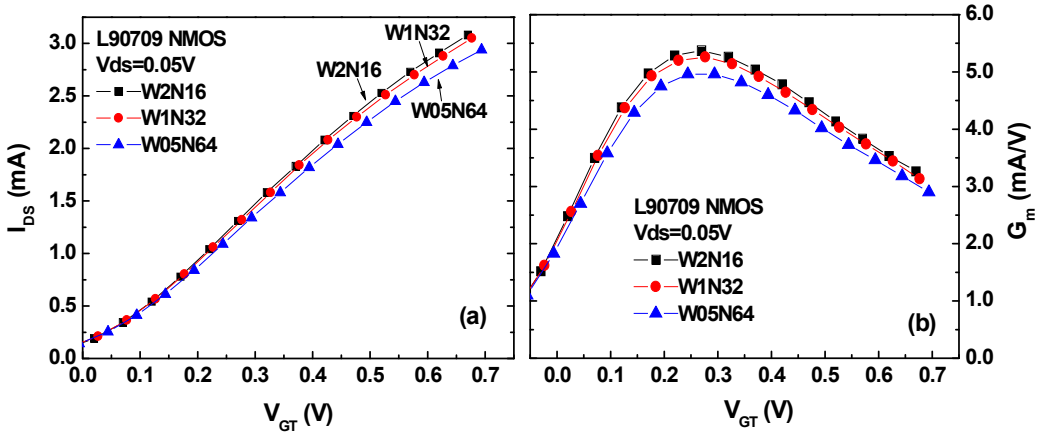
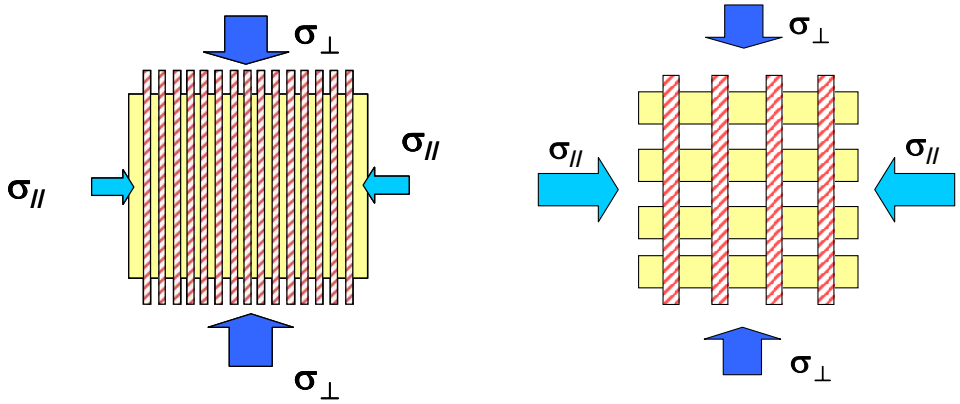


Fig.3.7 (a) The drain current I_{DS} vs. V_{GT} (b) transconductance G_m vs. V_{GT} measured from narrow-OD NMOS W1N32 (W_F=1μm, N_F=32) and W05N64 (W_F=0.5μm, N_F=64), and standard multi-finger NMOS W2N16 (W_F=2μm, N_F=16). All of the devices have the same total finger width, $W_F \times N_F$=32μm

Fig.3.8 illustrates the STI stresses introduced in MOSFETs with standard and multi-OD layouts (**Fig.3.3**) to facilitate an analysis of STI stress from an aggressive OD width scaling (W_{OD}=125nm for OD16) in the form of multiple OD fingers, and its impact on electrical characteristics. Again, STI stress is classified as longitudinal stress, denoted as $\sigma_{//}$, which is in parallel with the channel length, and transverse stress, namely $\sigma_{\perp}$, which is transverse to the channel length.



(a) (b)

Fig.3.8 Schematics of STI stresses along the longitudinal and transverse directions, defined as

$\sigma_{//}$ and $\sigma_{\perp}$ in MOSFETs with different layouts (a) standard multi-finger MOSFET (b) multi-OD MOSFET with multiple OD finger and multiple gate finger.

For multi-OD NMOS, the V_T vs. W_{OD} shown in **Fig. 3.9** indicates a strong INWE with V_T fall-off to 0.42V for OD16 with the minimal W_{OD} to 0.125μm. Again, INWE is determined by combined effects from doping profile, 2D charge sharing effect, corner field crowding, STI stress, and STI TCR (ΔW), etc. Note that the W_{OD} dependence of V_T for narrow-OD and multi-OD NMOS actually follow a universal curve in V_T vs. W_{OD}. Again, **Fig. 3.10** presents the I_{DS} and G_m under varying V_{GT} measured from multi-OD NMOS with two splits, such as OD8 (N_{OD}=8, W_{OD} =0.25μm) and OD16 (N_{OD}=16, W_{OD} =0.125μm) and their comparison with the standard one, i.e OD1 (N_{OD}=1, W_{OD} =2μm). The results indicate that the $G_{m,max}$ of OD8 is degraded by around 20% as compared to OD1 but the continuous scaling of W_{OD} to 0.125μm in OD16 leads to an increase of G_m compared to OD8 and the $G_{m,max}$ degradation compared to OD1 is shrunk to 11%. The result looks very interesting and cannot be explained by STI compressive stress alone.

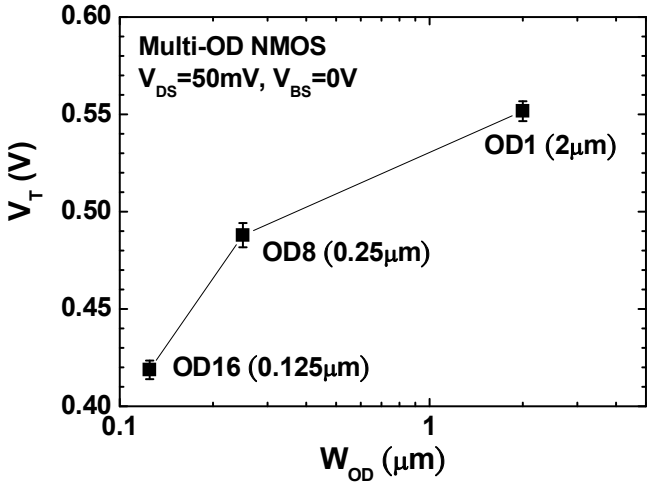


Fig.3.9 Linear V_T versus W_{OD} for multi-OD NMOS with various layout dimensions : OD1 (N_{OD}=1, W_{OD}=2μm), OD8 (N_{OD}=8, W_{OD}=0.25μm), and OD16 (N_{OD}=16, W_{OD}=0.125μm). Bias condition : V_{DS}=50mV and V_{BS}=0V.

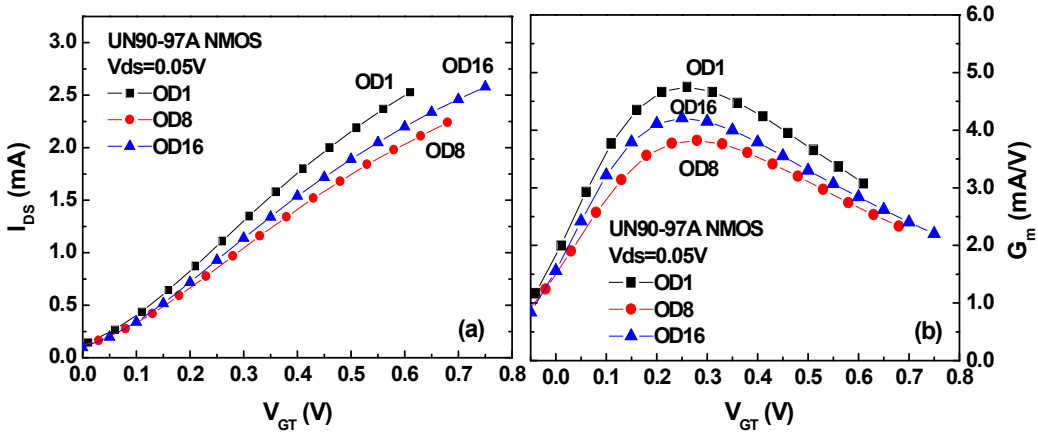


Fig.3.10 (a) The drain current I_{DS} vs. V_{GT} (b) transconductance G_m versus V_{GT} measured from multi-OD NMOS OD8 (N_{OD}=8, W_{OD}=0.25μm) and OD16 (N_{OD}=16, W_{OD}=0.125μm), and standard NMOS OD1 (N_{OD}=1, W_{OD}=2μm). All of the devices have the same finger number, N_F=16.

As for donut MOSFET with the layouts shown in **Fig. 3.4**, the STI stresses introduced in this donut device is illustrated in Fig. **3.11 (b).** Through a comparison with that of standard multi-finger MOSFET in Fig. **3.11(a),** the major difference is that donut layout can keep the MOSFET free from STI transverse stress $\sigma_{\perp}$, along the channel width and eliminate the impact from σ_{∞} on mobility. Another difference is that the channel current in donut MOSFETs is contributed from two directions in perpendicular to each other, which may have certain difference in the carrier mobility (μ_{eff}).

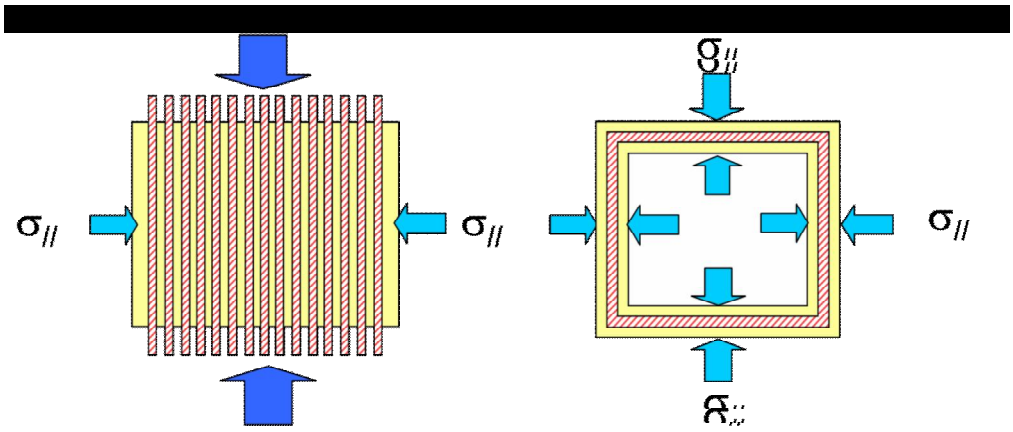


Fig.3.11 Schematics of STI stresses along the longitudinal and transverse directions, defined as $\sigma_{//}$ and $\sigma_{\perp}$ in MOSFETs with different layouts (a) standard multi-finger MOSFET (b) donut MOSFET with 4-side polygons for gate and OD.

Fig. 3.12 presents the drain current (I_{DS}) and transconductance (G_m) measured from donut **NMOS** and standard multi-finger NMOS (W2N32) with the same total width. As compared to the standard NMOS, it is found that $G_{m,max}$ of D1S1 is degraded by around 9.7% whereas that of D10S10 is enhanced by 7.5%. The experimental suggests that the compressive $\sigma_{//}$ from STI, which is maximized in D1S1 due to the minimum gate to STI edge space is the primary factor responsible for $G_{m,max}$ degradation. As for D10S10, much lower $\sigma_{//}$ due to 10 times larger space and eliminated $\sigma_{\perp}$ for donut layout contributes to $G_{m,max}$ improvement.

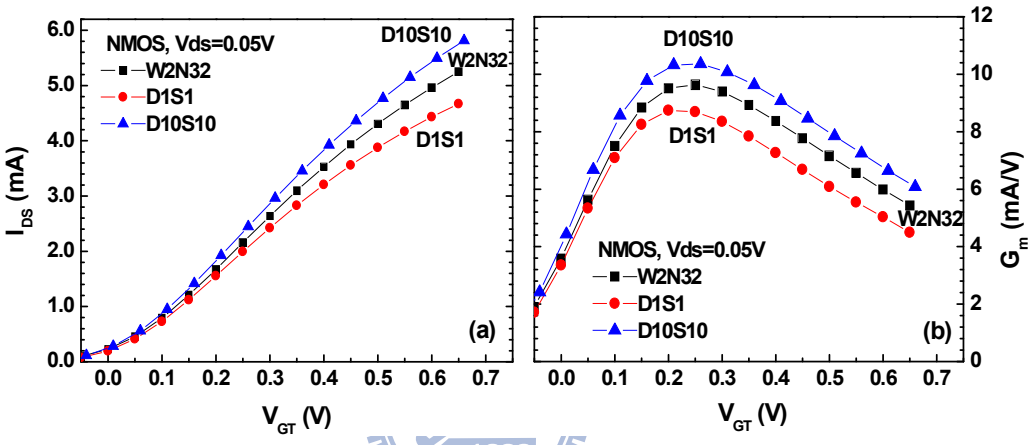


Fig.3.12 (a) The drain current I_{DS} vs. V_{GT} (V_{DS}=0.05V) (b) transconductance G_m vs. V_{GT} measured from donut NMOS D1S1 and D10S10 (16μmx4) and standard multi-finger NMOS W2N32 (W_F=2μm, N_F=32).

Again, **Fig. 3.13** shows I_{DS} and G_m measured from donut **PMOS** and standard multi-finger PMOS (W2N32) with the same total width. The experimental indicates that the donut PMOS, D1S1 and D10S10 demonstrate 12.2% and 7.6% higher $G_{m,max}$ than the standard PMOS, shown in Fig. 3 13(b). It is believed that D1S1 with the minimal gate to STI edge distance, resulting the highest compressive $\sigma_{//}$ and minimized $\sigma_{\perp}$ can benefit the most in hole mobility. The standard PMOS with relieved $\sigma_{//}$ in multi-finger structure and largest $\sigma_{\perp}$ due to the narrow width will suffer the worst hole mobility.

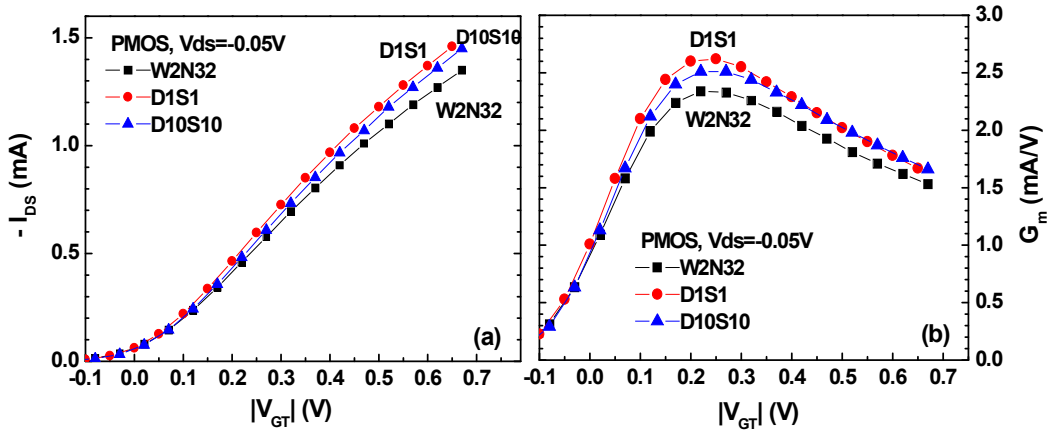


Fig.3.13 (a) The drain current I_{DS} vs. $|V_{GT}|(V_{DS}$=-0.05V) (b) transconductance G_m vs. $|V_{GT}|$ measured from donut PMOS D1S1 and D10S10 (16μmx4) and standard multi-finger PMOS W2N32 (W_F=2μm, N_F=32).

3.4 STI Top Corner Rounding (TCR) Effect on W_{eff} and G_m

To explain the extraordinary result measured from multi-OD devices with extremely narrow OD width (OD16) shown in **Fig. 3.10**, the increase of W_{eff} due to STI top corner rounding (TCR) is proposed as the primary mechanism trading with mobility degradation from STI compressive stress to determine the channel current (I_{DS}) and transconductance G_m. The details of STI TCR process with the profile and resulted ΔW contributing to W_{eff} will be described in section 3.5. Note that G_m is the key parameter responsible for analog and RF circuit performance.

It is proposed that G_m should decrease with the reduction of OD width (W_{OD}) since the mobility decreases as the transverse compressive stress $\sigma_{\perp}$ increases with the OD width scaling. However, in this work, we found that OD width scaling from W_{OD}=0.25μm in OD8 to W_{OD}=0.125μm in OD16 led to an increase of $G_{m,max}$ shown in **Fig. 3.10(b)** rather than degradation predicted by stress model. This experimental result suggests the transverse compressive stress $\sigma_{\perp}$ from STI, which is maximized in OD16 due to the minimum W_{OD} cannot fully explain the largest $G_{m,max}$ degradation in OD8 instead of OD16. The extraordinary results suggest that the variation of G_m with OD width scaling is determined not

only by STI stress effect on mobility (μ_{eff}) but also by STI TCR effect on the effective width (W_{eff}), namely ΔW effect **[33,34]**. For both OD8 and OD16, the large compressive stress $\sigma_{\perp}$ from STI spreads into the active region and degrades the μ_{eff}. However, for OD16 devices, the ΔW effect dominates and causes the increase of G_m. To simulate the proposed STI stress and TCR effects on G_m in miniaturized MOSFET, the semi-empirical formulas are derived as follows **[35]**.

$$I_{DS} = W_{eff} C_{ox} (V_{GS} - V_T - \alpha V_{DS}) \mu_{eff} \frac{V_{DS}}{L_{eff}} \tag{3.8}$$

$$G_m = \frac{\partial I_{DS}}{\partial V_{GS}} = W_{eff} C_{ox} \mu_{eff} \frac{V_{DS}}{L_{eff}} \tag{3.9}$$

let

$$\beta = C_{ox} \frac{V_{DS}}{L_{eff}} \tag{3.10}$$

$$\mu_{eff}(W_{OD}) = \mu_0 + \Delta\mu(W_{OD}) = \mu_0 \left[1 - k \cdot \ell og \left(\frac{W_{ref}}{W_{OD}} \right) \right] \tag{3.11}$$

$$W_{eff} = (W_{OD} + \Delta W) \times N_{OD} \times N_F \tag{3.12}$$

then

$$G_m = \beta \mu_0 \left[1 - k \cdot \ell og \left(\frac{W_{ref}}{W_{OD}} \right) \right] (W_{OD} + \Delta W) N_{OD} N_F \tag{3.13}$$

where

ΔW is the increase of OD finger width due to STI TCR

In this work, the standard device with W_F= W_{OD}=2 μm is specified as the reference, that is W_{ref}=2 μm to minimize the transverse stress $\sigma_{\perp}$ and its impact on mobility. The derived model given by (3.11) and (3.13) can accurately predict μ_{eff} and G_m measured from multi-finger MOSFETs with various OD layouts, such as standard, narrow-OD, and multi-OD. The simultaneous best fitting to μ_{eff} and G_m for multi-OD NMOS can be realized under the condition of ΔW=43 nm and k=0.2888. Note that ΔW is dependent on the OD layout and varies between the standard and multi-OD devices. For standard device with more gradient

trench profile, ΔW =24 nm appears to be smaller than that of multi-OD devices. For narrow-OD devices even with the minimum OD width in W05N64, its OD width (W_{OD}=W_F=0.5μm) is four times larger than that of OD16 (W_{OD}=0.125 μm) and ΔW effect is not strong enough to overcome μ_{eff} degradation from STI stress. It can be understood that the ΔW ratio is around 10.6% for narrow-OD device W05N64 with W_{OD}=W_F=0.5μm whereas that of multi-OD device OD16 with W_{OD}=0.125μm is significantly increased to 32.1%, which is around three times larger than that of narrow-OD device W05N64.

3.5 STI Top Corner Rounding (TCR) Technique – Stress and ΔW

As mentioned previously, STI has been the standard isolation technology in advanced CMOS processes, beginning from 0.25μm node and keeping a successful enabler for CMOS technology scaling to nanoscale regime. STI becomes indispensable for achieving higher packing density and performance, attributed to tighter design rule and smaller sidewall junction capacitances **[36,37]**. However, trench corner engineering appeared as one of critical challenges in STI technology development to meet the requirements of a smooth top corner and prevention from gate wrap-around. The reason is that a sharp trench corner with gate wrap-around may lead to undesirable double-hump in the subthreshold I-V, inverse narrow width effect (INWE), poor gate oxide integrity, and enhanced gate tunneling leakage. To solve the mentioned problems, a number of STI top corner rounding (TCR) methods have been developed and published **[38-41]**. The conventional STI process starts with a thin pad oxide, a nitride cap layer, and a hard mask for active area patterning and trench etching. Following the trench formation and implantation, liner re-oxidation is performed as an annealing process to recover the damage caused by the etching and implantation, which is localized near the trench top/bottom corners. This step may produce a limited corner rounding but insufficient to fix the problems like subthreshold double-hump, gate tunneling leakage, and stress induced leakage

current (SILC) after high field injection **[38]**. Some more elaborated scheme like STI with LOCOS edge may enhance TCR effect but suffer high process complexity and difficulty in process control **[39]**. A relatively simple method, namely post CMP high temperature re-oxidation (HTR-STI) was proposed as a manufacturable TCR solution and demonstrated electrical characteristics free from INWE and subthreshold double-hump in 0.18μm CMOS process **[40]**. The effect of STI TCR can be verified by the improvement over the mentioned electrical characteristics and the results suggest that the larger corner radius can help reduce the stress and field crowding near the trench corner. Moreover, the larger corner radius from trench edge recess and TCR can lead to larger ΔW, as shown in **Fig. 3.14** and then an increase of effective channel width, i.e. W_{eff} . The increase of W_{eff} from TCR induced ΔW becomes significant in vary narrow devices and reveals its effect on channel current (I_{DS}) and transconductance (G_m), which can explain the extraordinary results measured from multi-OD devices with extremely narrow W_{OD}, shown in **Fig. 3.10**. The evidence of ΔW from STI TCR supports previous I-V model derivation that I_{DS} and G_m are determined by the trade-off between increase of W_{eff} from ΔW and mobility degradation from STI compressive stress $\sigma_{\perp}$. An accurate extraction of ΔW is indispensable to determine μ_{eff} in very narrow MOSFETs. It emerges as one of major research topics in this thesis.

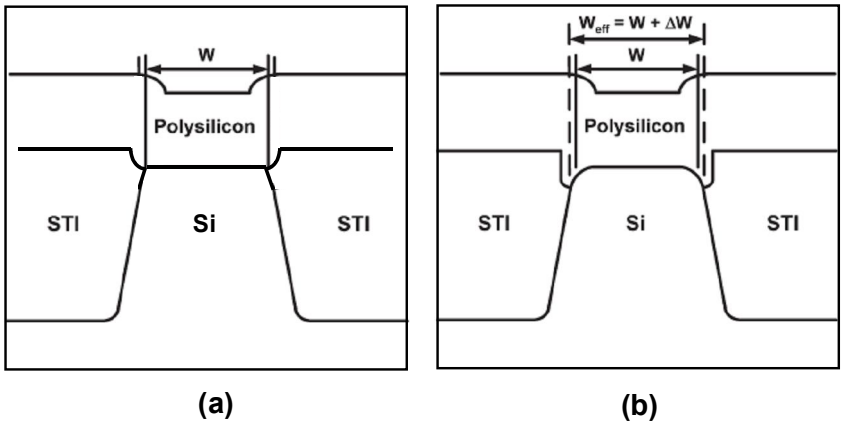


Fig.3.14 Schematically drawn cross-sectional view of a MOSFET in the channel width direction (a) the sharp STI corner without edge recess and top corner rounding (TCR) (b) the rounded STI corner contributes larger W_{eff} due to ΔW from edge recess and TCR

3.6 C-V Characterization and Fringing Capacitances Simulation for Accurate Mobility Extraction

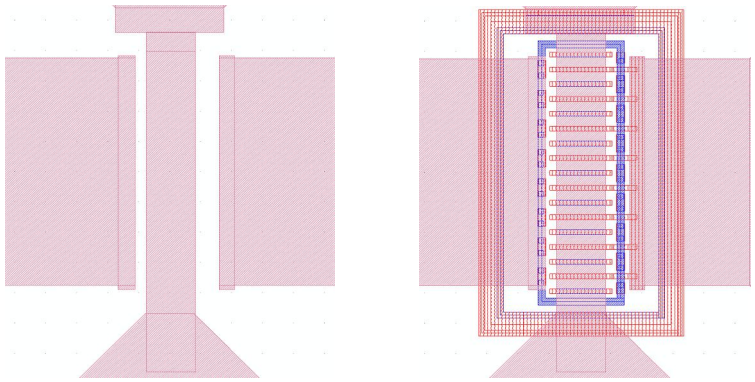
Referring to (3.1) ~ (3.4) for μ_{eff} extraction from linear I-V characteristics, Q_{inv}, W_{eff}, and L_{eff} appear as three key parameters determining the accuracy of extracted μ_{eff}. In previous sections, ΔW (from STI TCR) can be extracted from simultaneous best fitting to μ_{eff} and G_m in **(3.11)** and **(3.13)**. In this way, W_{eff} can be precisely determined, which is particularly critical for very narrow devices, such as narrow-OD and multi-OD MOSFETs. Unfortunately, one more challenge emerged from the determination of Q_{inv} and L_{eff} in miniaturized MOSFETs. According to **(3.4)**, Q_{inv} can be calculated from $C_{ox(inv)}$ and V_{GT} under strong inversion condition. However, the extraction of $C_{ox(inv)}$ become a tough job in miniaturized devices with short channel length and/or narrow width. Note that $C_{ox(inv)}$ represents the intrinsic gate capacitance density per unit area, which is corresponding to intrinsic channel region and excluding all other parasitic capacitances. However, the measured gate capacitance, namely $C_{gg,mea}$ before deembedding indeed incorporates all of the parasitic capacitances. In a standard flow for high frequency characterization, an open deembedding must be performed, trying to extract and remove all of the parasitic capacitances. However, if the open deembedding process can enable a truly clean deemebedding is critically dependent on the design of dummy open pads. A detailed discussion on this subject is described as follows.

3.6.1 Open deembedding Methods for Intrinsic C_{gg} Extraction

In this thesis, two dummy open test structures, namely openM3 (open deembedding to M3) and openM1 (open deembedding to M1) illustrated in **Fig. 3.15** are designed to investigate the differences and the final impact on μ_{eff} extraction. Note that the gate capacitances after openM3 and openM1 deembedding are defined as $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$ shown in (3.14) and (3.15).

$$C_{gg(DUT,OM3)} = C_{gg,mea} - C_{gg(OM3)} \quad : DUT's\ C_{gg}\ after\ openM3\ deembedding \qquad (3.14)$$

$$C_{gg(DUT,OM1)} = C_{gg,mea} - C_{gg(OM1)} \quad : DUT's\ C_{gg}\ after\ openM1\ deembedding \tag{3.15}$$



(a) (b)

Fig. 3.15 Layout of dummy open test structures for multi-finger MOSFETs (a) dummy openM3 for open deembedding to M3 (b) dummy openM1 for open deembedding to M1

Table 3.2 presents the gate capacitances measured from multi-finger MOSFET ($C_{gg,mea}$), dummy openM3 and dummy openM1 ($C_{gg(OM3)}$ and $C_{gg(OM1)}$), and those after openM3 and openM1 deembedding, denoted as $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$. Note that multi-finger MOSFETs with simultaneously varied W_F and N_F under fixed $W_{tot}= W_F \times N_F$, i.e. W2N16, W1N32, and W05N64 as shown in **Fig. 3.2** are adopted for this characterization and analysis. The bias condition is fixed at $|V_{DS}|$=0.05V and $|V_{GS}|$=1.2V for NMOS and PMOS to enable a linear I-V characteristics appropriate for the proposed model (3.1)~(3.4) for μ_{eff} extraction. The results indicate that $C_{gg(OM3)}$ measured from dummy openM3 is independent of the variation of N_F whereas $C_{gg(OM1)}$ measured from dummy openM1 increases with increasing N_F. Moreover, $C_{gg(OM1)}$ is apparently larger than $C_{gg(OM3)}$ and the difference increases with N_F. The difference between $C_{gg(OM1)}$ and $C_{gg(OM3)}$, denoted as $\Delta C_{gg(M1\sim M3)}$ represents the parasitic capacitance from M3 through M2 and down to M1 and the dependence of N_F comes from that associated with M2 and M1 following the multi-finger layout. Raphael simulation was performed to calculate the 3-D parasitic capacitances contributed from dummy openM3 and dummy openM1 as shown in **Fig. 3.16 (a)** and **(b),** respectively. Note that dummy openM1

incorporates M3/M2/M1 stack layers in **Fig. 3.16(b). Table 3.3** summarizes the difference $\Delta C_{gg(M1\sim M3)}$ between $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$, which are extracted from openM3 and openM1 deembedding on multi-finger NMOS with N_F=16 and 32 (W2N16, W1N32). The $\Delta C_{gg(M1\sim M3)}$ from measurement indicates a linear dependence on N_F, and the difference divided by N_F, namely $\Delta C_{gg0(M1\sim M3)}$ keeps nearly a constant of 0.32 ± 0.1 fF/finger. Furthermore, $C_{gg(M3)}$ and $C_{gg(M1\sim M3)}$ calculated by Raphael simulation for dummy openM3 and dummy openM1 (N_F=1) achieve the difference $\Delta C_{gg0(M1\sim M3)}$=0.32 fF/finger. The good match between measurement and simulation proves the accuracy and the improvement of openM1 over openM3 deembedding. It is assumed that $C_{gg(DUT,OM1)}$ extracted through openM1 deembedding should be the intrinsic gate capacitance, which is determined by $C_{ox(inv)}$ and total gate area $A_g=N_F \times W_F \times L_g$. For multi-finger MOSFETs with fixed $W_{tot}=W_F \times N_F$, the intrinsic gate capacitances should be independent of N_F and keep a constant under varying N_F. To verify this point, $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$ extracted from openM3 and openM1 deembedding are plotted versus N_F, as shown in **Fig. 3.17** (a) and (b). The results indicate that both $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$ reveal a linear dependence on N_F, but $C_{gg(DUT,OM1)}$ vs. N_F, presents much smaller slope than that of $C_{gg(DUT,OM3)}$. The experimental with proven accuracy highlights that some other components of parasitic capacitance cannot be removed, even using openM1 deembeeding, i.e. the best one of existing methods for a clean deembedding.

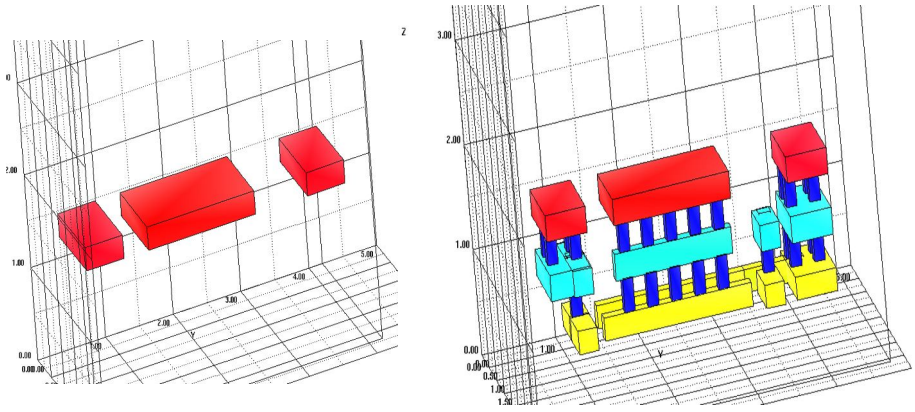
Table 3.2 Gate capacitances measured from multi-finger MOSFET ($C_{gg,mea}$), dummy open structures ($C_{gg(OM3)}$, $C_{gg(OM1)}$), $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$ after openM3 or openM1 deembedding

NMOS W_F=2um	2-port, open_M3 NMOS, V_{DS}=0.05V, V_{GS}=1.2V		
N_F	$C_{gg,mea}$ (fF)	$C_{gg(OM3)}$ (fF)	$C_{gg(DUT,OM3)}$ (fF)
16	65.89	**20.6**	**45.29**
32	72.11	**20.6**	**51.51**
64	85.54	**20.6**	**64.94**

NMOS W_F=2um	2-port, open_M1 NMOS, V_{DS}=0.05V, V_{GS}=1.2V		
N_F	$C_{gg,mea}$ (fF)	$C_{gg(OM1)}$ (fF)	$C_{gg(DUT,OM1)}$ (fF)
16	68.07	**28.12**	**39.95**
32	74.1	**32.57**	**41.53**

Table 3.3 $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$, extracted from openM3 and openM1 deembedding on multi-finger NMOS with N_F=16, 32. $C_{gg(M3)}$ and $C_{gg(M1\sim M3)}$ calculated by Raphael simulation for dummy openM3 and dummy openM1. A comparison of $\Delta C_{gg(M1\sim M3)}=C_{gg(DUT,OM3)}-C_{gg(DUT,OM1)}$ from measurement and $\Delta C_{gg(M1\sim M3)}=C_{gg(M1\sim M3)}-C_{gg(M3)}$ from simulation.

	Measurement		Simulation
NMOS finger numbers	N_F=16	N_F=32	N_F=1
C_{gg}	$C_{gg(DUT,OMx)}$ (fF)	$C_{gg(DUT,OMx)}$ (fF)	$C_{gg(OMx)}$ (fF)
openM3	45.29	51.51	0.126
openM1	39.95	41.53	0.446
$\Delta C_{gg(M1\sim M3)}$ (fF)	5.34	9.98	0.32
$\Delta C_{gg0(M1\sim M3)}$ (fF/finger)	**0.334**	**0.312**	**0.32**



(a) (b)

Fig. 3.16 3-D dummy open test structures for Raphael simulation (a) M3 only for openM3 (open deembedding to M3) (b) M3/M2/M1 stack for openM1 (open deembedding to M1)

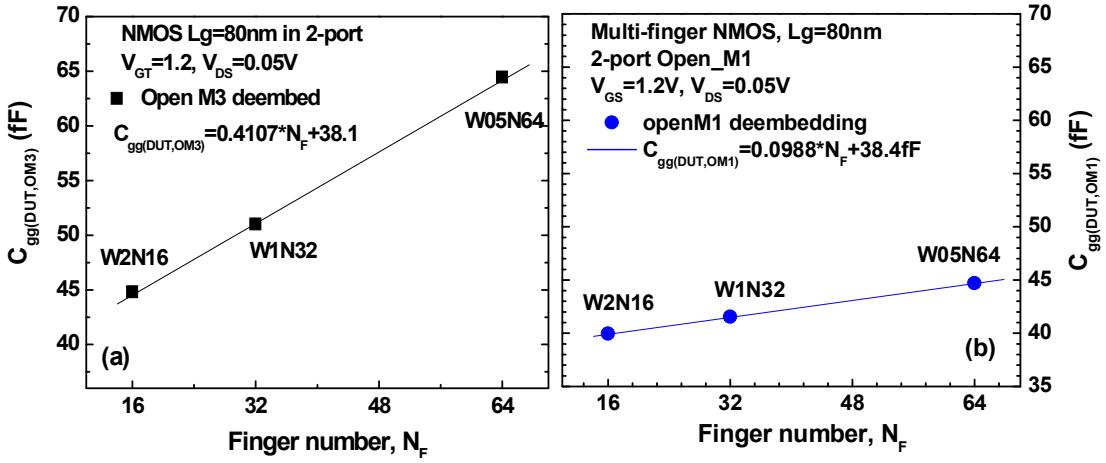


Fig. 3.17 Gate capacitances $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$ extracted through openM3 and openM1 deembedding on multi-finger NMOS (W2N16, W1N32, W05N64) with fixed W_FN_F (a) $C_{gg(DUT,OM3)}$ vs. N_F : linear function α'=0.4107 fF/finger, β'=38.1 fF (b) $C_{gg(DUT,OM1)}$ vs. N_F : α=0.0988 fF/fingerr, β=38.4 fF .

To explore the mechanism responsible for this new oberservation, a rigorous analysis was performed by using Raphael simulation and the results suggest that the parasitic capacitances lumped into $C_{gg,mea}$ can be classified into two categories : one is contributed from from pads, interconnection lines, and substrate, and the other from gate sidewall and finger ends, namely C_{of} (sidewall fringing capacitance) and $C_{f(poly\text{-}end)}$ (finger end fringing capacitance). The former one is a kind of extrinsic parasitic elements and can be removed through a dedicated open deembedding like openM1. As shown previously, openM1 can realize an open deembedding to the bottom metal, i.e. M1 to achieve a clean deembedding to the layers above poly gate. However, the latter one is actually a kind of intrinsic parasitic capacitance arising from the poly-gate fingers and the surrounding conductors like S/D diffusion regions and the contact plugs, and cannot be removed using any existing deembedding methods. To solve this problem, 3-dimensional capacitance simulation is performed using Raphael to calculate C_{of} and $C_{f(poly\text{-}end)}$ as follows.

3.6.2 3-D Fringing Capacitances Simulation and Analysis using Raphael [42]

In the following, 3-D Raphael simulation has been conducted to calculate the fringing capacitances from gate sidewall, i.e. $C_{of} = C_{g,Difff} + C_{g,CT}$ and another component from gate finger ends, namely $C_{f(poly\text{-}end)}$. **Fig.3.18** illustrates a 3-D MOSFET structure for Raphael simulation. The 3-D structure incorporates four conducting regions, such as poly gate, channel region, source/drain diffusion region, and contact to source/drain region. The physical profiles of the dielectric layers between every two conductors can be referred to technology file in foundry PDK. As individual electrode is specified for each conducting region as mentioned, the three components of coupling capacitances from the gate to other three regions can be calculated. Among the three components, gate to channel region is the intrinsic gate capacitance responsible for Q_{inv} and I_{DS}, and the other two components, i.e. gate to S/D diffusion regions ($C_{g,Difff}$) and gate to contact ($C_{g,CT}$) constitute the sidewall fringing capacitance, given by $C_{of} = C_{g,Difff} + C_{g,CT}$. **Fig. 3.19 (a)** and **(b)** depict the structure of multi-finger MOSFETs in planar view and cross-sectional view, which were built in Raphael for 3-D fringing capacitances simulation.

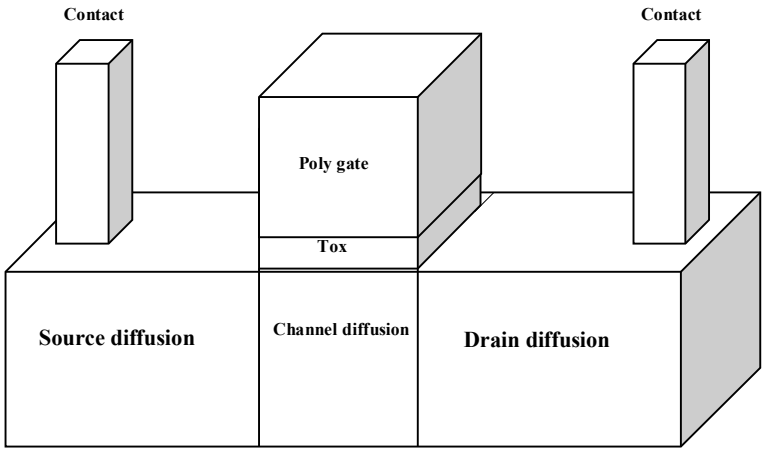
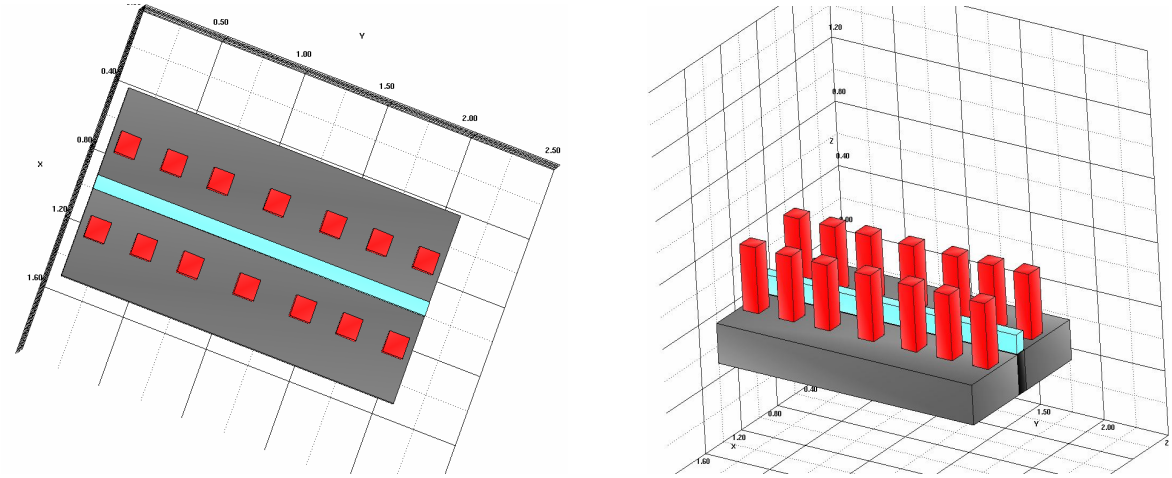


Fig.3.18 Schematical drawing of 3-D MOSFET structure for Raphael simulation. The 3-D structure incorporates four conducting regions, such as poly gate, channel region, source/drain diffusion region, and contact to source/drain



(a) (b)

Fig.3.19 Multi-finger MOSFET structure built in Raphael for 3-D capacitance simulation (a) planar view (b) cross sectional view

Fig. 3.20(a)-(d) present $C_{g,CT}$, $C_{g,Diff}$, and C_{of}, calculated for MOSFETs with square contact corresponding to various geometry parameters, such as L_g, L_{gct}, T_g, and T_{ox} over a wide range of variations. All of the parasitic capacitances indicate a weak dependence on L_g over a wide range of variation from 80 nm to 160 nm, shown in **Fig. 3.20(a)** and C_{of} keeps around 0.28fF/μm. The results suggest that the gate sidewall fringing capacitances $C_{of} = C_{g,CT}+C_{g,Diff}$ are not scalable with L_g and the impact on high frequency performance will go up with L_g scaling. For L_{gct} fixed at 140 nm on layout, according to 90 nm low leakage CMOS design rule, $C_{g,Diff}$ dominates near 80% of the total parasitic capacitance C_{of} and $C_{g,CT}$ contributes the remaining portion, i.e. around 20%. **Fig. 3.20(b)** presents the parasitic capacitances subject to varying L_{gct}. Interesting results are demonstrated with a significant dependence on L_{gct} that is in contrast with the weak dependence on L_g. Note that the increase of fringing capacitances with L_{gct} scaling highlights the impact from the inter-electrode space shrinkage in miniaturized devices on high frequency or high-speed circuit performance.

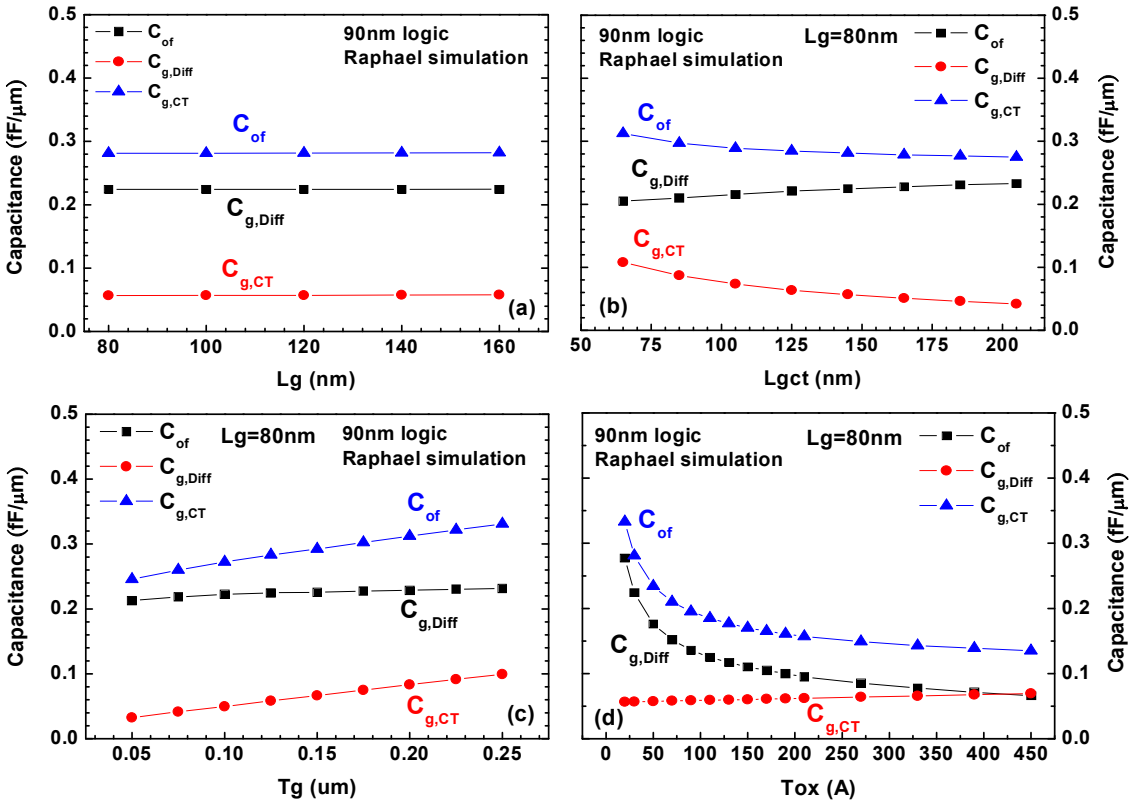


Fig.3.20 Parasitic capacitances $C_{g,CT}$, $C_{g,Diff}$, C_{of}, simulated by Raphael for MOSFETs with square contact versus device geometry parameters (a) gate length, L_g (b) gate to CT space, L_{gct} (c) gate electrode thickness, T_g (d) gate dielectric thickness, T_{ox} [42]

Besides the sidewall fringing capacitances C_{of}, finger-end fringing capacitance, namely $C_{f(poly\text{-}end)}$ is another element of the parasitic capacitances, which always exists in MOSFETs and cannot be removed by existing open deembedding methods. Again, Raphael simulation was employed to calculate $C_{f(poly\text{-}end)}$, which is required to enable an accurate extraction of truly intrinsic gate capacitance ($C_{gg,int}$) and determination of ΔW. Both C_{of} and $C_{f(poly\text{-}end)}$ are not scalable with device scaling and may dominate intrinsic gate capacitance in miniaturized MOSFETs. As a result, C_{of} and $C_{f(poly\text{-}end)}$ appear as key parameters to be known for an accurate extraction of μ_{eff} in multi-finger MOSFETs with various layout dimensions. **Fig. 3.21** illustrates the planar view of a multi-finger MOSFET in which three components of fringing capacitances, such as $C_{f(poly\text{-}end)}$, $C_{g,Diff}$, and $C_{g,CT}$ are depicted. This graphical analysis explains that the finger-end fringing capacitance is proportional to N_F but is independent of W_F and

W_{tot}. On the other hand, the sidewall fringing capacitance is determined by both W_F and N_F and in a linear proportional to W_{tot}. For narrow-OD MOSFETs (**Fig. 3.2**) with simultaneously varied N_F and W_F under a specified $W_{tot}= W_F \times N_F$, $C_{f(poly\text{-}end)}$ will increase with increasing N_F and its weighting factor increases dramatically in very narrow MOSFETs with large N_F and small W_F. According to 90 nm CMOS design rule, $C_{f(poly\text{-}end)}$ is around 0.064 fF/finger from Raphael simulation.

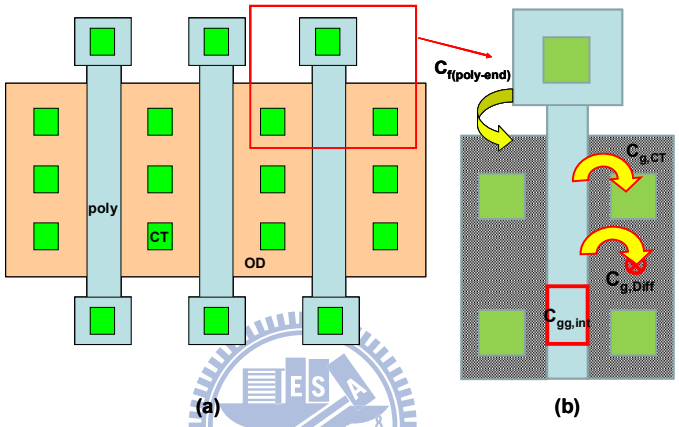


Fig. 3.21 (a) Schematical illustration of multi-finger MOSFET layout (b) three components of fringing capacitances associated with each gate finger : $C_{f(poly\text{-}end)}$ represents finger-end fringing capacitance, $C_{g,Diff}$, and $C_{g,CT}$ are gate sidewall fringing capacitances.

Based on an extensive simulation and analysis on two categories of parasitic capacitances, the intrinsic gate capacitance extraction flow and analysis can be derived as follows. At first, $C_{gg(DUT,OM1)}$ can be re-written as two parts, one of which is proportional to N_F and another is independent of N_F, given by (3.16)

$$C_{gg(DUT,OM1)} = N_F\left[(\Delta W \cdot L_g)C_{ox(inv)} + C_{f(poly-end)}\right] + (C_{ox(inv)}L_g + C_{of})W_{tot} \tag{3.16}$$

where

$$C_{gg,\text{int}} = C_{ox(inv)}L_g \tag{3.17}$$

$$C_{ox(inv)} = \varepsilon_{ox}\varepsilon_0 / T_{ox(inv)} \tag{3.18}$$

$$W_F N_F = W_{tot} \tag{3.19}$$

then $C_{gg(DUT,OM1)}$ of multi-finger MOSFETs with various N_F but fixed $W_F N_F = W_{tot}$ can be expressed as a linear function of N_F with the slope and intercept defined as α and β as follows

$$C_{gg(DUT,OM1)} = \alpha N_F + \beta \tag{3.20}$$

$$\alpha = (\Delta W \cdot L_g) C_{ox(inv)} + C_{f(poly-end)} \tag{3.21}$$

$$\beta = (C_{ox(inv)} L_g + C_{of}) W_{tot} \tag{3.22}$$

$$C_{of} = \frac{\beta}{W_{tot}} - C_{ox(inv)} L_g \tag{3.23}$$

$$L_g = \frac{\beta}{W_{tot} C_{ox(inv)}} - \frac{C_{of}}{C_{ox(inv)}} \tag{3.24}$$

Note that the first term of the intercept β in (3.22) is the intrinsic gate capacitance ($C_{gg,int} W_{tot}$), which contributes the inversion carriers $Q_{inv} W_{tot}$ responsible for channel current I_{DS} in (3.1). With C_{of} known from Raphael simulation, $C_{gg,int}$ can be extracted from β in (3.22). In fact, C_{of} and L_g extracted from (3.23) and (3.24) should be self-consistent to validate the accuracy and it has been proven that C_{of} =0.28 fF/μm and L_g =80 nm are the right ones to meet the target $C_{ox(inv)}$. Also, with C_{of} determined by Raphael simulation and L_g calculated by (3.24), ΔW can be extracted from (3.21). The accurate extraction of C_{of}, $C_{ox(inv)}$, L_g, and ΔW is pre-requisite to accurate determination of intrinsic gate capacitance $C_{gg,int}$, inversion carrier density Q_{inv}, and effective mobility μ_{eff} in intrinsic channel region.

Taking $\Delta C_{gg(M1\sim M3)}$, which is proportional to N_F and referring to (3.20)~(3.22) for that with openM1 deembedding, we can derive the expressions for that with openM3 deembedding as follows.

$$C_{gg(DUT,OM3)} = N_F \left[(\Delta W \cdot L_g) C_{ox(inv)} + C_{f(poly-end)} \right] + (C_{ox(inv)} L_g + C_{of}) W_{tot} + \Delta C_{gg(M1\sim M3)} \tag{3.25}$$

$$\Delta C_{gg(M1\sim M3)} = \Delta C_{gg0(M1\sim M3)} \cdot N_F$$

$$C_{gg(DUT,OM3)} = \alpha' N_F + \beta' \tag{3.26}$$

$$\alpha' = (\Delta W \cdot L_g) C_{ox(inv)} + C_{f(poly-end)} + \Delta C_{gg0(M1\sim M3)} \tag{3.27}$$

$$\beta' = (C_{ox(inv)} L_g + C_{of}) W_{tot} \tag{3.28}$$

The results derived for openM3 deembedding indicate that the intercept β' for openM3 in (3.28) keeps the same as β for openM1 in (3.22) under a fixed W_{tot}; however, the slope α' for openM3 in (3.27) appears larger than the slope α for openM1 in (3.21). In the following, α, β, and ΔW determined from multi-finger MOSFETs with openM1 deembedding, and (α', β') from those with openM3 deembedding will be presented for a comparison and verification on the proposed model.

Referring to **Fig. 3.17**(a) and (b) for $C_{gg(DUT,OM3)}$ and $C_{gg(DUT,OM1)}$ extracted by openM3 and openM1 deembedding. The results indeed demonstrate a linear relationship for both openM3 and openM1 deembedding, with nearly the same intercept, β=38.1~38.4 fF but significant difference in the slope. The result validates the model for β given by (3.22)=(3.28), and provides a simple solution that the intrinsic gate capacitance $C_{gg,int}$ defined as $C_{ox(inv)}L_g$ in (3.17) can be extracted from the slope β no matter which kind of dummy open structures was used. The slope for openM1, given as α=0.0988 fF/finger appears much smaller than that of openM3, α'=0.4107 fF/finger. The experimental results just match the prediction from our proposed model and the difference between α and α' comes from $\Delta C_{gg(M1\sim M3)}$, given by (3.21) and (3.27).

3.6.3 Effective Mobility Extraction for MOSFETs with Various Layouts

Following the characterization flow, openM1 deembedding and Raphael simulation can facilitate a thorough and precise extraction of 3-D parasitic capacitances and accurate determination of the intrinsic gate capacitance $C_{gg,int}$. Taking the developed characterization process with proven accuracy, the fundamental device parameters like $C_{ox(inv)}$, L_g, ΔW, C_{of}, $C_{f(poly\text{-}end)}$ can be determined with reliable accuracy for nanoscale MOSFETs. **Table 3.4** summarizes the key parameters for multi-finger NMOS.

Table 3.4 Multi-finger MOSFET device parameters extracted from DC I-V and RF C-V characterization flow

Device types		**NMOS_W2N16_openM1**
Bias condition		**Vds=0.05V, Vg=1.2V**
N_F		**16**
W_F	μm	**2**
L_g	μm	**0.08**
$T_{ox(inv)}$	A	**30**
$C_{ox(inv)}$	$fF/\mu m^2$	**11.51**
α	fF/finger	**0.0988**
β	fF	**38.37**
$C_{of,sim}$	fF/μm	**0.2812**
$C_{f(polyend)}$	fF/finger	**0.0640**
$C_{ox(inv)} =(\beta/W_F N_F - C_{of})/L_g$	$fF/\mu m^2$	**11.47**
$T_{ox(inv)}=\varepsilon_0\varepsilon_{ox}/C_{ox(inv)}$	A	**30.10**
$\Delta W=(\alpha - C_{f(polyend)})/(C_{ox(inv)}*L_g)$	μm	**0.0379**

In the following, the effective mobility μ_{eff} can be extracted from linear I-V model (3.3)~(3.4) in which W_{eff} and Q_{inv} have been determined with proven accuracy for miniaturized MOSFETs with various layouts. In this approach, the layout dependent STI stress effect on μ_{eff} can be determined with high precision due to the fact that all of the parasitic effects, such as ΔW from STI TCR, gate related 3-D fringing capacitances, and pads/interconnection lines introduced parasitic capacitance have been taken into the

characterization flow. **Fig. 3.22** presents μ_{eff} extracted from the group of narrow-OD NMOS, i.e. multi-finger NMOS with simultaneously varied W_F and N_F under fixed $W_{tot}=W_F \times N_F$, (W2N16, W1N32, and W05N64), shown in **Fig. 3.2**. Note that V_{GT} is taken to compensate V_T variation from INWE. The results indicate a monotonical degradation trend in μ_{eff} with finger width ($W_F=W_{OD}$) scaling from 2μm to 0.5μm and proves the impact from STI compressive σ_∞ on electron mobility. Taking W2N16 as the reference, μ_{eff} degradation in W1N32 is as small as 1.3% and that of W05N64 increases to around 6.8%. Referring to G_m versus V_{GT} (Fig. 3.7(b)), G_m degradation in W1N32 and W05N64 just follows the same trend and suggests that μ_{eff} degradation is the dominant factor. As for multi-OD NMOS with an aggressive channel width scaling, μ_{eff} versus V_{GT} shown in **Fig. 3.23** reveals substantially large degradation. For OD8 with W_{OD}=0.25μm, the μ_{eff} appears 33.6% lower than OD1 and the μ_{eff} degradation further increases to 37.3% for OD16 with W_{OD}=0.125 μm.

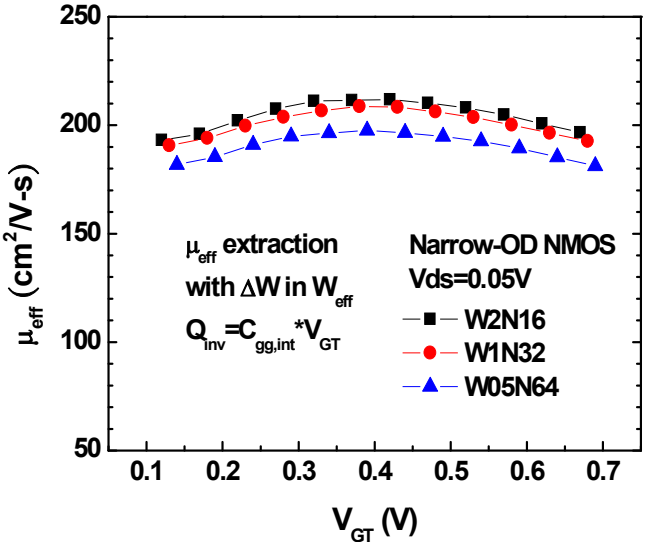


Fig. 3.22 Effective mobility μ_{eff} versus V_{GT} for narrow-OD NMOS determined with consideration of ΔW in W_{eff} and $C_{gg,int}$ for Q_{int}. Narrow-OD NMOS : W1N32, W05N54, standard : W2N16. Linear operation condition : V_{DS}=0.05V, $V_{GT}=V_{GS}-V_T$.

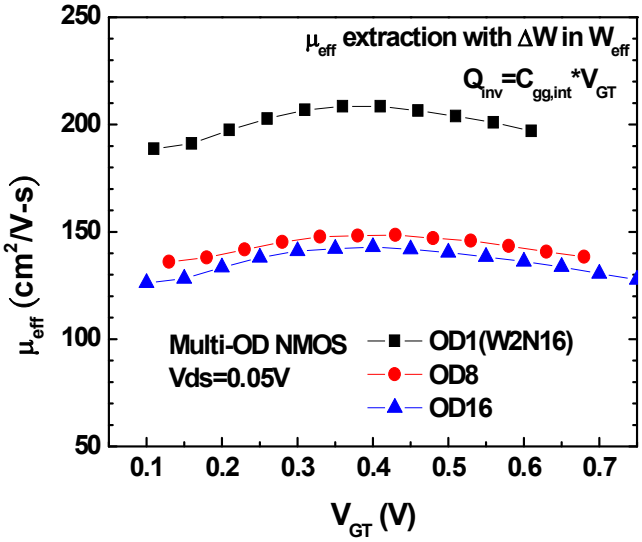


Fig. 3.23 Effective mobility μ_{eff} versus V_{GT} for multi-OD NMOS determined with consideration of ΔW in W_{eff} and $C_{gg,int}$ for Q_{int}. Multi-OD NMOS : OD8 (W_{OD}=0.25μm), OD16 (W_{OD}=0.125μm), standard : OD1 (W_{OD}=2μm)W2N16. Linear operation condition : V_{DS}=0.05V, V_{GT}=V_{GS} - V_T.

In comparison with narrow-OD and multi-OD MOSFET with enhanced σ_{∞}, donut MOSFETs were designed to verify another extreme case, that is the elimination of σ_{∞}, along the direction of channel width. For NMOS shown in **Fig. 3.24** (a), D1S1 suffer around 7.13% degradation while D10S10 gain 10.03% enhancement in μ_{eff} compared to the standard W2N16. The μ_{eff} enhancement in D10S10 NMOS just match the original expectation that the elimination of compressive σ_{∞}, can benefit electron mobility. As for μ_{eff} degradation in D1S1 NMOS compared to W2N32, the increase of $\sigma_{//}$ along the channel length, due to the minimal gate to STI edge space is considered the major cause responsible for electron mobility degradation. Regarding PMOS shown in **Fig. 3.24** (b), both donut devices, i.e. D1S1 and D10S10 gain μ_{eff} improvement over the reference W2N32. Interestingly, D1S1 PMOS can achieve 11.8% μ_{eff} enhancement, which is obviously higher than 6.1% realized by D10S10 PMOS. The opposite trend w.r.t. NMOS in D1S1 layout can be explained by the mechanism that compressive stress along the channel length, i.e. $\sigma_{//}$ can improve hole mobility but degrade electron mobility, as shown in Table 3.1. The experimental result proves that D1S1

PMOS with the minimal gate to STI edge distance, resulting the highest compressive $\sigma_{//}$ and minimized $\sigma_{\perp}$ can benefit the most in hole mobility. Again, the layout dependence of μ_{eff} is exactly the same as that of G_m shown in **Fig. 3.13**.

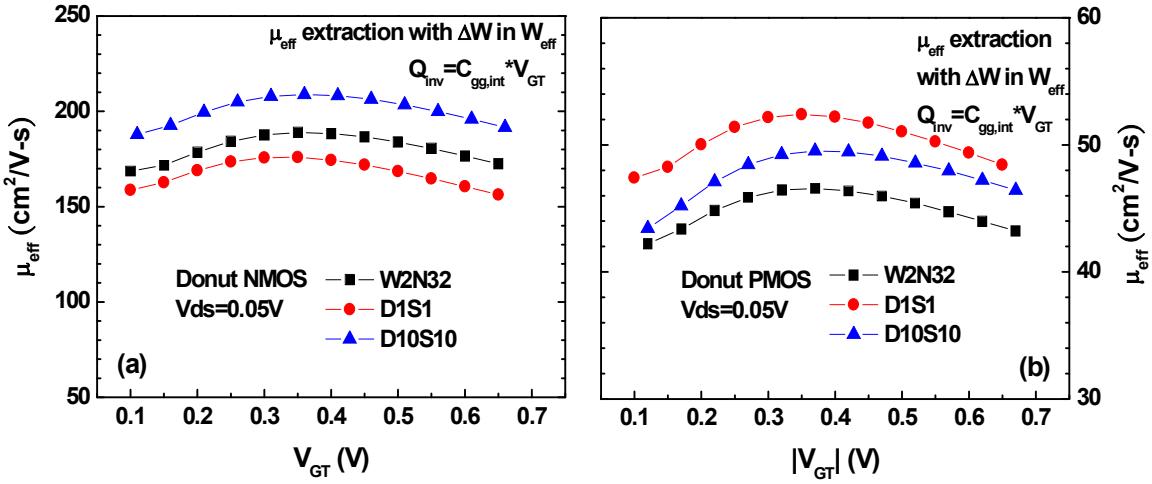


Fig. 3.24 Effective mobility μ_{eff} versus V_{GT} for donut MOSFET determined with consideration of ΔW in W_{eff} and $C_{gg,int}$ for Q_{int}. (a) NMOS (b) PMOS, donut : D1S1, D10S10, standard : W2N32. Linear operation condition : V_{DS}=0.05V, $V_{GT}=V_{GS}-V_T$.

To verify the scattering mechanisms responsible the mobility, the gate bias or gate overdrive V_{GT} is converted to an effective normal field E_{eff} according to the formulas (3.29) and (3.30) for NMOS and PMOS, respectively.

$$E_{eff} = \frac{1}{6}\frac{(V_{GT} + 2V_T)}{T_{ox(inv)}} \tag{3.29}$$

$$E_{eff} = \frac{1}{9}\frac{(V_{GT} + 3V_T)}{T_{ox(inv)}} \tag{3.30}$$

In this way, μ_{eff} vs. V_{GT} for all of the devices are transformed to μ_{eff} vs. E_{eff} as shown in Fig. 3.25 and Fig. 3.26 for narrow-OD NMOS and multi-OD NMOS, respectively. The universal mobility theory predicts that MOSFETs with various channel lengths, oxide thickness, and doping concentration should follow a universal curve when the effective mobility is expressed in the form of μ_{eff} vs. E_{eff} [43]. It can be explained by the mechanism that variations in the mentioned device parameters are actually incorporated in the body

charge and inversion carriers density (Q_b and Q_{inv}) and transformed into E_{eff}. However, the layout dependent stress is not the kind of parameter and may not impose its effect on μ_{eff} via E_{eff}. The mentioned argument explains why the μ_{eff} vs. E_{eff} for both narrow-OD and multi-OD NMOS with various W_{OD} would not follow a universal curve. Similar results are achieved for donut NMOS and PMOS shown in **Fig. 3.27**. A comparison between donut MOSFETs and multi-finger MOSFETs with the same W_{tot} is illustrated in Fig. 3.28(a) and (b) for NMOS and PMOS, respectively. However, μ_{eff} vs. E_{eff} indeed demonstrates three regions corresponding to three scattering mechanisms, such as coulomb scattering in low field, phonon scattering in medium field, and surface roughness scattering in high field, shown in **Fig. 3.29**. The relationship can be described by Matthiessen's rule in (3.31) and field dependence of each component in (3.32)~ (3.34),

$$\frac{1}{\mu_{eff}} = \frac{1}{\mu_{coul}} + \frac{1}{\mu_{ph}} + \frac{1}{\mu_{sr}} \tag{3.31}$$

$$\mu_{coul} = AE_{eff}^{\alpha 1} T^{\alpha 2} \tag{3.32}$$

$$\mu_{ph} = AE_{eff}^{-\beta 1} T^{-\beta 2} \tag{3.33}$$

$$\mu_{sr} = AE_{eff}^{-\gamma} \tag{3.34}$$

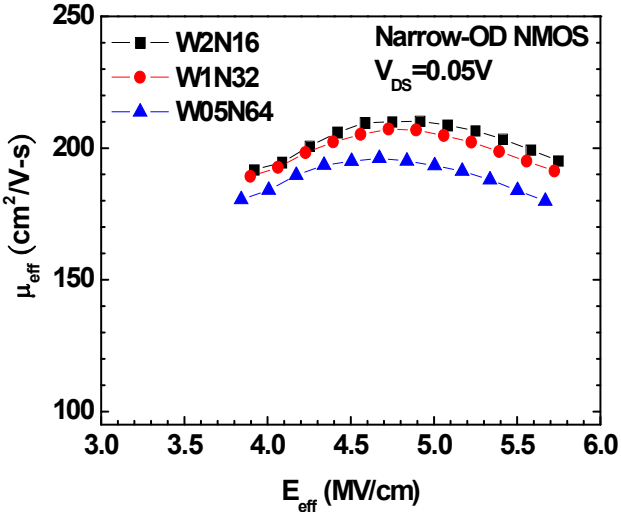


Fig. 3.25 Effective mobility μ_{eff} versus E_{eff} for narrow-OD NMOS determined with consideration of ΔW in W_{eff} and $C_{gg,int}$ for Q_{int}. Narrow-OD NMOS : W1N32, W05N54, standard : W2N16. V_{DS}=0.05V, $E_{eff} = (V_{GT} + 2V_T)/6T_{ox(inv)}$

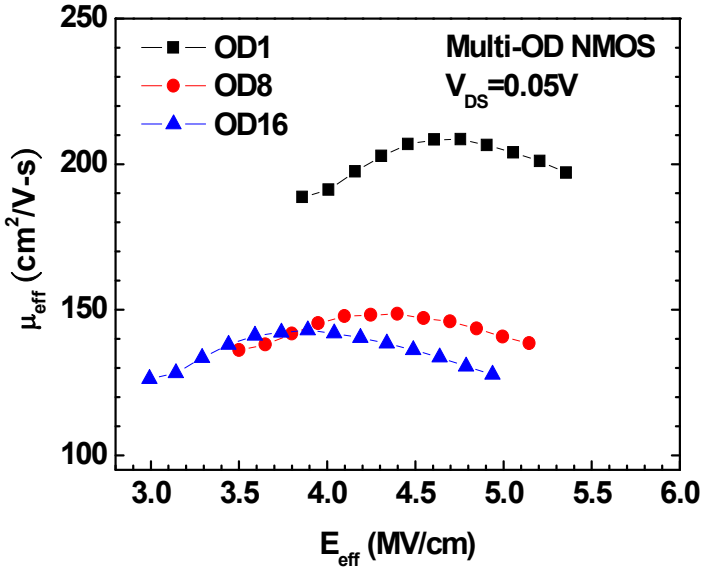


Fig. 3.26 Effective mobility μ_{eff} versus E_{eff} for multi-OD NMOS determined with consideration of ΔW in W_{eff} and $C_{gg,int}$ for Q_{int}. Multi-OD NMOS : OD8 (W_{OD}=0.25μm), OD16(W_{OD}=0.125μm), OD1 (W_{OD}=2μm). V_{DS}=0.05V, $E_{eff} = (V_{GT} + 2V_T)/6T_{ox(inv)}$

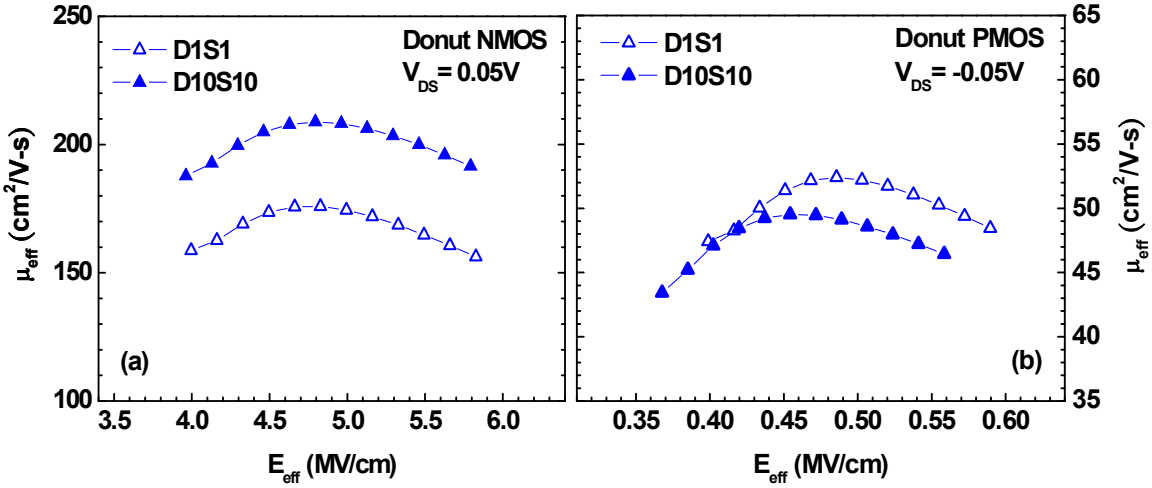


Fig. 3.27 Effective mobility μ_{eff} versus E_{eff} for donut MOSFET (D1S1, D10S10) determined with consideration of ΔW in W_{eff} and $C_{gg,int}$ for Q_{int} (a) NMOS : $E_{eff} = (V_{GT} + 2V_T)/6T_{ox(inv)}$ (b) PMOS $E_{eff} = (V_{GT} + 3V_T)/9T_{ox(inv)}$.

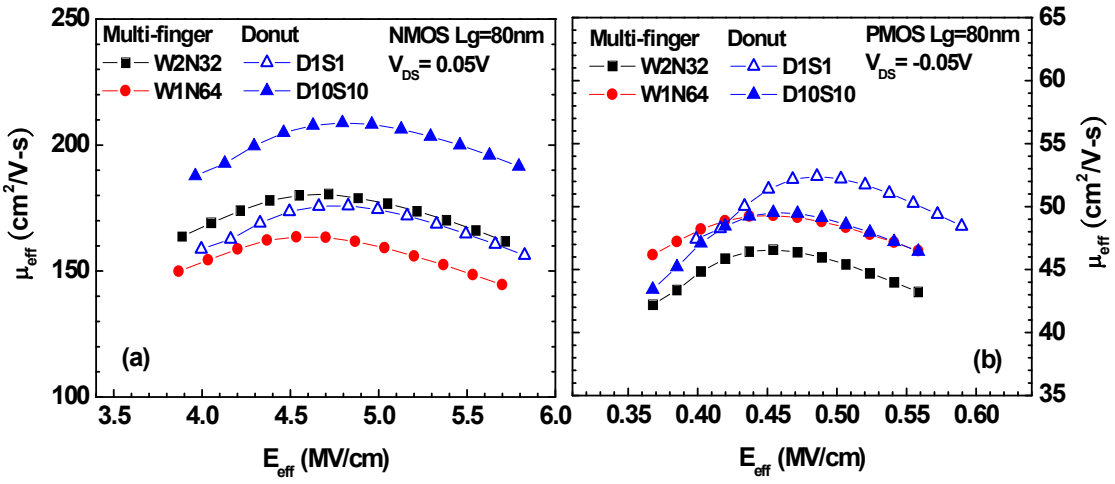


Fig. 3.28 Effective mobility μ_{eff} versus E_{eff} for donut MOSFET (D1S1, D10S10) and a comparison with multi-finger MOSFET (W2N32, W1N64) with the same W_{tot} (a) NMOS : $E_{eff} = (V_{GT} + 2V_T)/6T_{ox(inv)}$ (b) PMOS $E_{eff} = (V_{GT} + 3V_T)/9T_{ox(inv)}$.

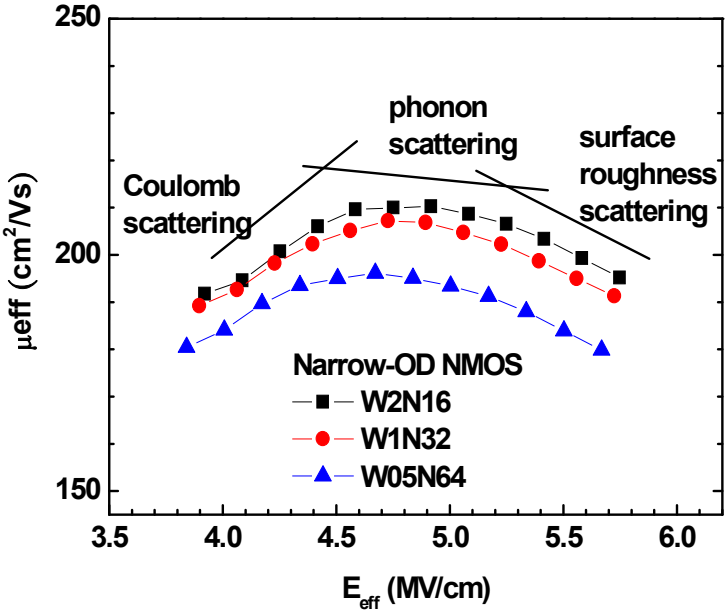


Fig. 3.29 Effective mobility μ_{eff} versus E_{eff} for multi-finger NMOS (W2N16, W1N32, W05N64). $E_{eff} = (V_{GT} + 2V_T)/6T_{ox(inv)}$. Three regions corresponding to three scattering mechanisms : coulomb scattering in low field, phonon scattering in medium field, and surface roughness scattering in high field.

3.7 CP Current Measurement for Interface States Extraction and Analysis of STI Stress Effect

Charge pumping technique has been widely used to characterize interface state densities in MOSFETs **[44,45]**. In thin gate films, leakage current is relatively high due to quantum mechanical tunneling of carriers through the gate. As a result, the traditional technique of extracting interface traps density. Collecting simultaneous re high frequency C-V measurement data and comparing the difference can't be used because the C-V is very hard to achieve at the leakage current level.

However, charge-pumping measurements can still be used to extract interface trap density, and the effect of gate leakage can be compensated for by measuring charge-pumping current at lower frequency and subtracting it from measurement results at higher frequencies.

Stress effect on interface traps is one of major concerns for deployment of stress engineering in the state-of-the-art process. In the form of localized interface-state generation, is a key reliability issue. In order to improve the understanding of the physical mechanism resulting in STI stress degradation, there is a great need for experimental results of variations in Si-SiO_2 interface state density and distribution under the gate of the device.

In this work, the use of charge pumping to accurately measure Si-SiO_2 interface trap density in four-terminal RF CMOS is described.

The charge pumping measurements are operated by Agilent 81110A signal generator, Keithley708A switch, and HP 4156 DC supply. **Fig. 3.30** illustrates the experimental facilities of CP measurement for a cross-section MOSFET in 4T structure. The basic charge-pumping technique involves measuring the substrate current while applying voltage pulses of fixed amplitude, rise time, fall time, and frequency to the gate of the transistor, with the source, drain, and body tied to ground. The pulse also can be applied with fixed amplitude, voltage base sweep or a fixed base, variable amplitude sweep. **Fig. 3.31** illustrates the pulse waveform

employed for the Elliot method in which the base level (V_{base}) is swept from accumulation to inversion while keeping the pulse amplitude (V_a) constant. Note that $I_{CP,max}$ can occur under the condition of $V_{base} < V_{FB}$ and $V_{base}+V_a = V_h > V_T$ and the accuracy is justified with a clear plateau for $I_{CP,max}$ and a precise linear dependence on amplitude (V_a) and frequency (not shown).

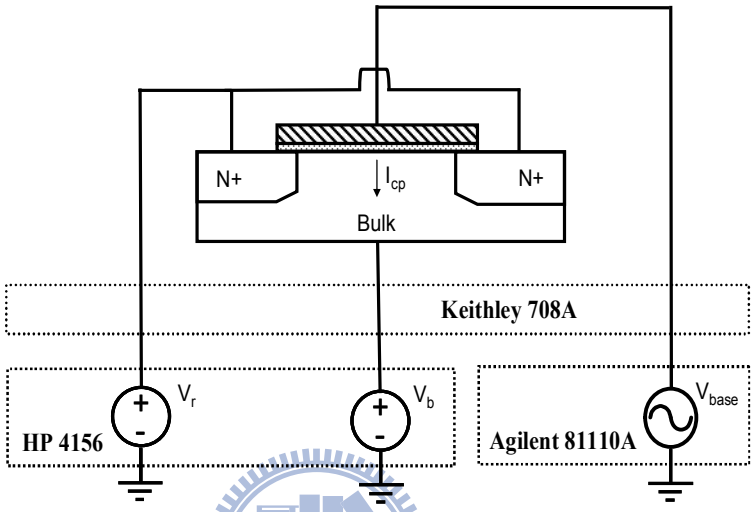


Fig.3.30 Schematic of experimental set-up of the charge pumping measurement applied to MOSFET with four terminals

Fixed amplitude and varying base level are obtained from an Agilent 81110A, which is connected to the gate terminal and varying the pulse base level from accumulation to inversion while keeping the amplitude of the pulse constant. Electrons captured in the traps in the inversion mode then recombined with holes in the accumulation mode. This electron-hole recombination process gave rise to a current flow from the substrate to the channel. Recombination of charge through interface traps in the substrate results in the flow of charge pumping current in the bulk.

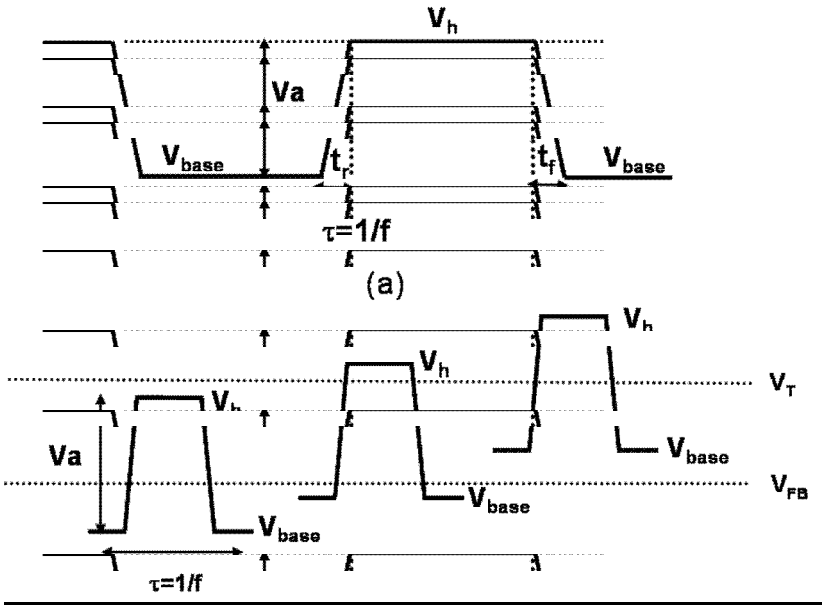


Fig.3.31 CP pulse waveform : t_r and t_f are the rising and falling times, V_{base} and V_h are swept from accumulation (V_{base} <V_{FB}) to inversion (V_h >V_T), under fixed pulse amplitude V_a=V_h-V_{base},

Chapter 4

RF MOSFET Layout effect on Low Frequency Noise

4.1 Low Frequency Noise Analysis in Two-port 3T MOSFET

MOSFET layout effect on I-V, G_m, μ_{eff}, and gate capacitance for Q_{inv}, have been presented in chapter 3. The experimental results can be explained by the collective effects from layout dependent STI stress and ΔW from STI TCR. Due to the fact that LFN is one of critical issues considered in RF and analog circuits design, MOSFET layout effect on LFN will be investigated and described in this chapter.

The multi-finger layout structure has been widely used for RF application to get better f_{MAX} by minimizing the parasitic gate resistance (R_g) [46], but the inconsistent STI stress effect on device characteristics need to be more discussed, and the multi-finger device type is insufficient for analyzing. In this chapter, four kinds of layout, namely multi-finger, narrow-OD, multi-OD, and donut MOSFETs with the same total width are designed to investigate channel wdith (W_{OD}) scaling effect on LFN. The STI stress might be highly considered, but on the other hand, the ΔW due to STI TCR is increasingly important [47]. Thus it is essential to distinguish the ΔW effect from the STI stress.

4.2 Narrow-OD MOSFET and Comparison with Standard MOSFET

The multi-finger structure with finger number N_F=16, 32, and 64, finger width W_F= 2, 1, and 0.5μm respectively. The total width and channel length are fixed at 32um and 0.09um. The he measured S_{ID}/I_{DS}^2 at frequency 50Hz are plotted versus I_{DS} for narrow-OD devices under various V_{GT} (0.1~ 0.7V).

The measurement result of narrow-OD for the STI stress mechanism analysis as shown in **Fig. 4.1** The normalized drain current noise S_{ID}/I_{DS}^2 of the W2N16 and W1N32 devices are

about 2 times higher than W05N64. The difference between W2N16 and W1N32 seems to be ignorable. The phenomena can be explained by the delta width effect overcome the STI stress on the performance for narrow-OD devices. The delta width effect dominated in the W05N64 LFN characteristic, but the delta width effect and STI stress effect are comparable in W1N32 LFN characteristic. From the above discussion, the compressive stress σ// effect is not obvious for NMOS narrow-OD device. But the influence of these two mechanisms (STI stress effect and delta width effect) is different in NMOS and PMOS, and this will be a further discussion in the following section.

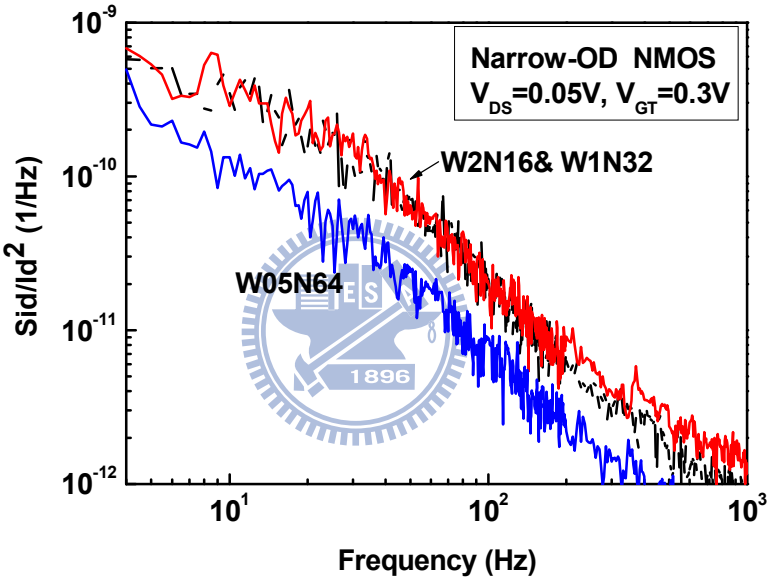


Fig.4.1 The low frequency noise S_{ID}/I_{DS}^2 measured for the multi-finger NMOS W2N16, W1N32, and W05N64.

The measured S_{ID}/I_{DS}^2 follows a function proportional to $1/I_{DS}^2$ over the whole range of bias conditions, which indicates that number fluctuation model given by (4.1) is the dominant mechanism appears in NMOS LFN.

$$\frac{S_{ID}}{I_{DS}^{\ 2}} = \frac{q^2 k_B T \lambda N_{it}}{W_{eff} L C_{ox}^2} \frac{1}{f} \frac{g_m^2}{I_{DS}^2} \qquad (4.1)$$

N_t : *the density of traps at quasi - Fermi level*

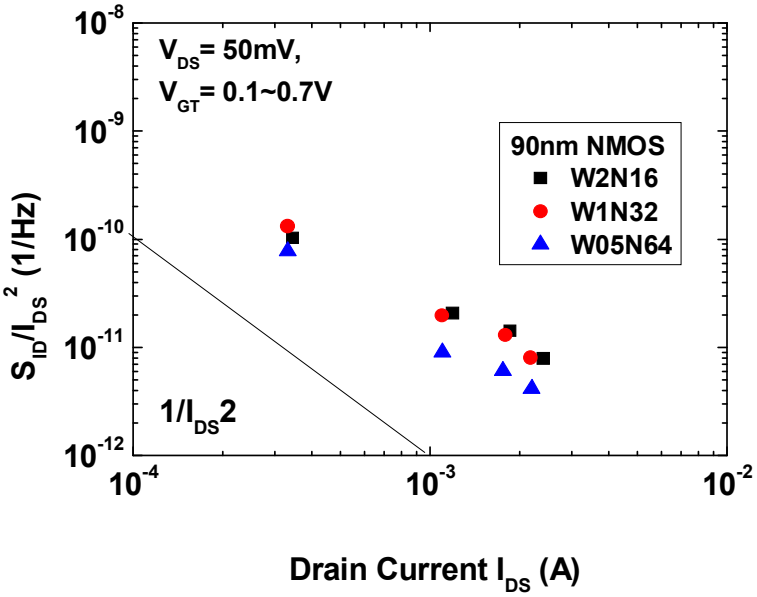


Fig.4.2 S_{ID}/I_{DS}^2 vs. I_{DS} under varying $|V_{GT}|$ (0.1~0.7V) for NMOS narrow-OD devices

4.3 Multi-OD MOSFET and Comparison with Standard MOSFET

The LFN in terms of S_{ID}/I_{DS}^2 for OD1 and OD16 NMOS in the frequency domain is shown in **Fig. 4.3** The noise spectrum follows 1/f characteristics over a wide frequency domain from 4 to 10K Hz. It means that the measured LFN is a typical flicker noise. It is interesting to show that the OD16 gets lower S_{ID}/I_{DS}^2 than OD1 over a wide range of frequencies, although the interface trap density N_{it} of OD16 is higher than that of OD1.

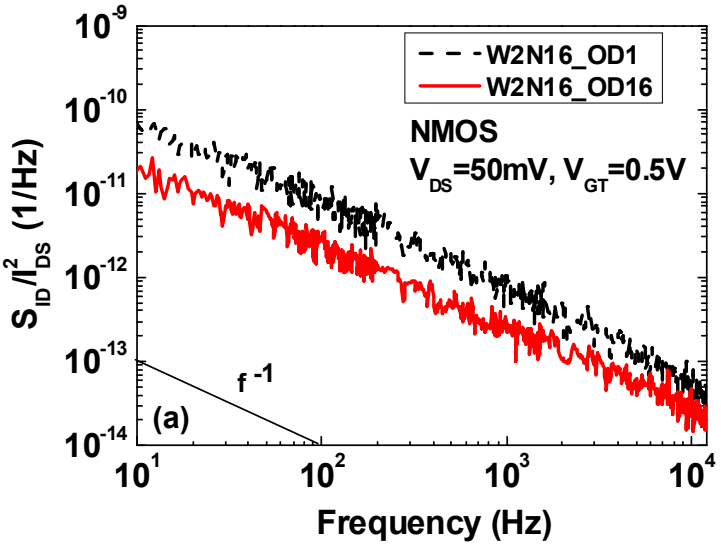


Fig.4.3 S_{ID}/I_{DS}^2 vs. I_{DS} under varying V_{GT} (0.1~0.7V) for OD1 and OD16 devices

To further analyze the mechanism responsible for LFN, the measured S_{ID}/I_{DS}^2 at frequency 50Hz are plotted versus I_{DS} for both OD1 and OD16 NMOS under various V_{GT} (0.1~ 0.7V) as shown in **Fig. 4.4.**

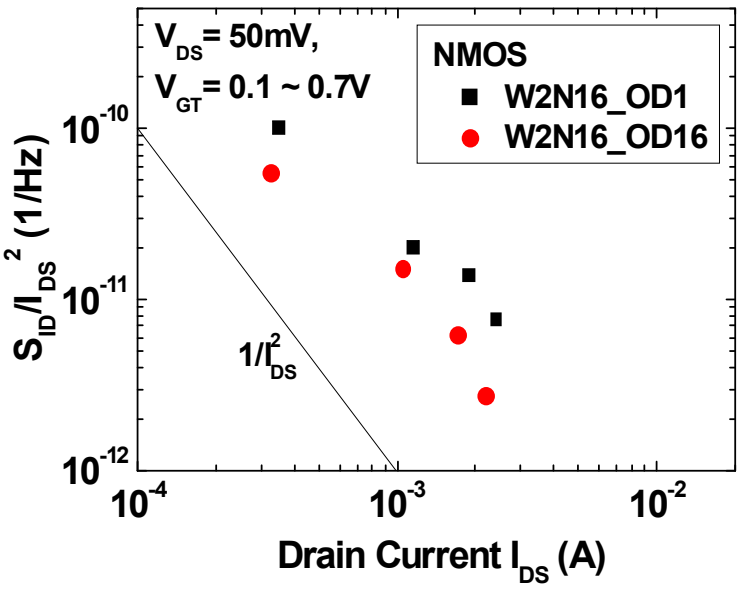


Fig.4.4 S_{ID}/I_{DS}^2 vs. I_{DS} under varying $|V_{GT}|$ (0.1~0.7V) for NMOS multi-OD devices

Herein, the measured NMOS multi-OD S_{ID}/I_{DS}^2 also follows a function proportional to $1/I_{DS}^2$ over the whole range of bias conditions, which indicates that number fluctuation model given by (4.1) is the dominant mechanism appears in NMOS LFN

Therefore, S_{ID}/I_{DS}^2 of NMOS is proportional to N_{it}/W_{eff} and that predicts the decrease of LFN with increasing the effective width W_{eff}. It is believed that the OD16 device suffers higher transverse compressive stress as well as interface traps N_{it}, which may aggravate the scattering effects and increase the flicker noise [50]. However, the larger W_{eff} can eliminate these effects. The mentioned mechanism can explain why the OD16 devices can have the lower LFN as compared to OD1.

As for the LFN in terms of S_{ID}/I_{DS}^2 for PMOS OD1, OD8, and OD16 in the frequency domain is shown in **Fig. 4.5** The devices are all with gate finger number at 16 and the poly width is 2um thus the total width is 32μm.

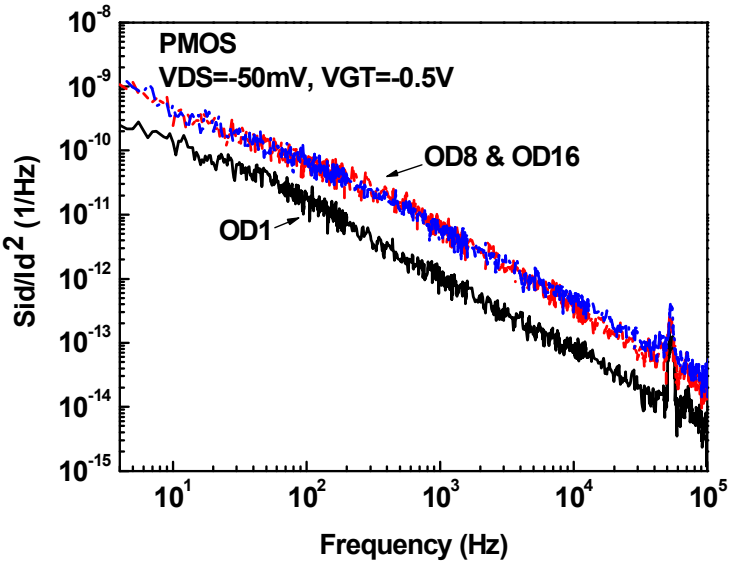


Fig. 4.5 The low frequency noise S_{ID}/I_D^2 measured for the multi-OD devices, OD1, OD8, and OD16.

It is easy to show that the OD8 and OD16 have higher S_{ID}/I_{DS}^2 than OD1 over a wide range of frequencies, and the OD8 is nearly equal to OD16. S_{ID}/I_{DS}^2 of PMOS is tells that the OD16 device suffers the most critical mobility degradation. However, the larger W_{eff} can eliminate these effects and thus the LFN of OD8 is very close to OD16. The mentioned W_{eff} mechanism is exist in PMOS but the effect W_{eff} cannot overcome the mobility degradation effect for PMOS multi-OD devices. The resulting STI stress and mobility degradation effect dominates over the delta width effect.

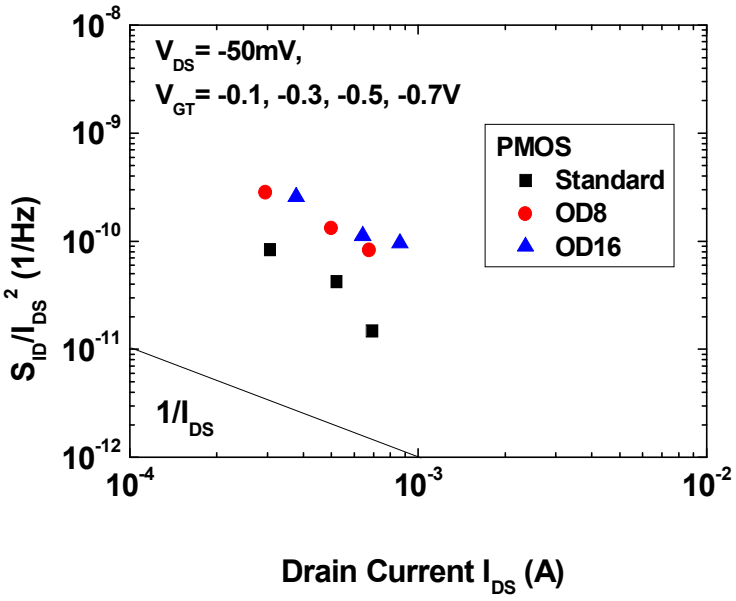


Fig.4.6 S_{ID}/I_{DS}^2 vs. I_{DS} under varying $|V_{GT}|$ (0.1~0.7V) for PMOS multi-OD devices

4.4 Donut MOSFET and Comparison with Standard MOSFET

Fig. 4.7 (a) and (b) make a comparison of LFN in terms of S_{ID}/I_D^2 between the standard and donut devices for NMOS and PMOS, respectively. The noise spectrum follows 1/f characteristics over a wide frequency domain from 4 to 10K Hz. It means that the measured LFN is a typical flicker noise. The standard device reveals near twice larger S_{ID}/I_D^2 as compared to donut devices for both NMOS and PMOS, under a specified gate overdrive voltage, $|V_{GT}|$= 0.7V. In contrast, the donut device D10S10 with the most extended gate to STI-edge distance indicates the lowest S_{ID}/I_D^2.

The results can be consistently explained by the fact that D10S10 can keep free from $\sigma\perp$ as well as interface traps near STI edge, and the smallest σ// due to 10 times larger space away from the STI edge compared to D1S1.

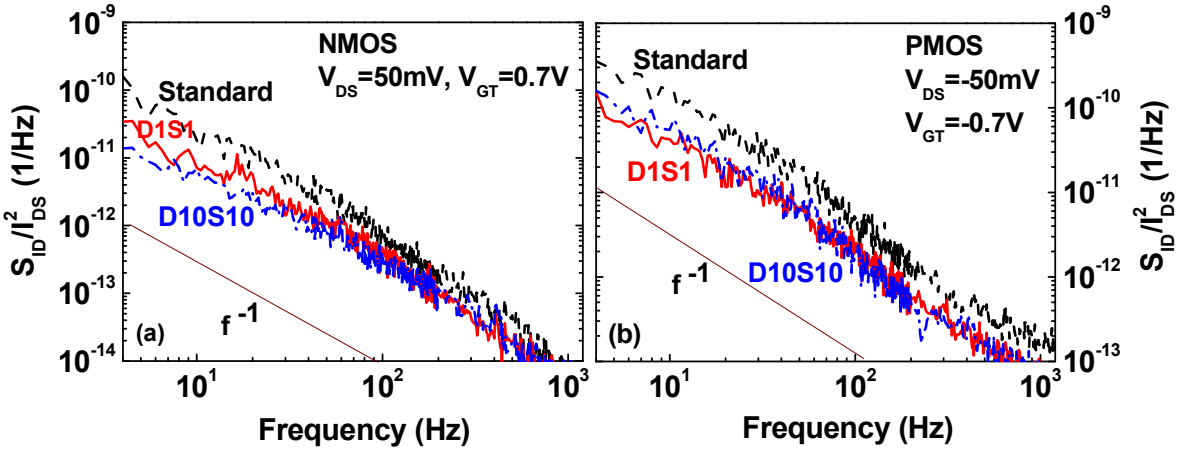


Fig.4.7 The low frequency noise S_{ID}/I_D^2 measured for the standard and donut devices (a) NMOS (b) PMOS: Standard (multi-finger W2N32), Donut D1S1, and D10S10.

To further explore the mechanism responsible for LFN, the measured S_{ID}/I_{DS}^2 at frequency 50Hz are plotted versus I_{DS} for three different devices, under various $|V_{GT}|$ (0.1~ 0.7V) shown in **Fig. 4.8** (a) and (b) for NMOS and PMOS, respectively. For NMOS devices, the measured LFN characteristic is dominated by number fluctuation model given by (4.1) in which S_{ID}/I_{DS}^2 is proportional to N_{it}/I_{DS}^2 and that predicts the increase of LFN with increasing the traps density N_{it}.

It is believed that the gate to STI-edge overlap region will suffer the most severe compressive strain as well as interface traps N_{it} [51], and the donut devices can eliminate these effects along the gate width, i.e. in the transverse direction. According to previous study, the stress generated traps may aggravate the scattering effect and increase the flicker noise. The mentioned mechanism can explain why the donut devices free from gate to STI-edge overlap region can have the lowest LFN.

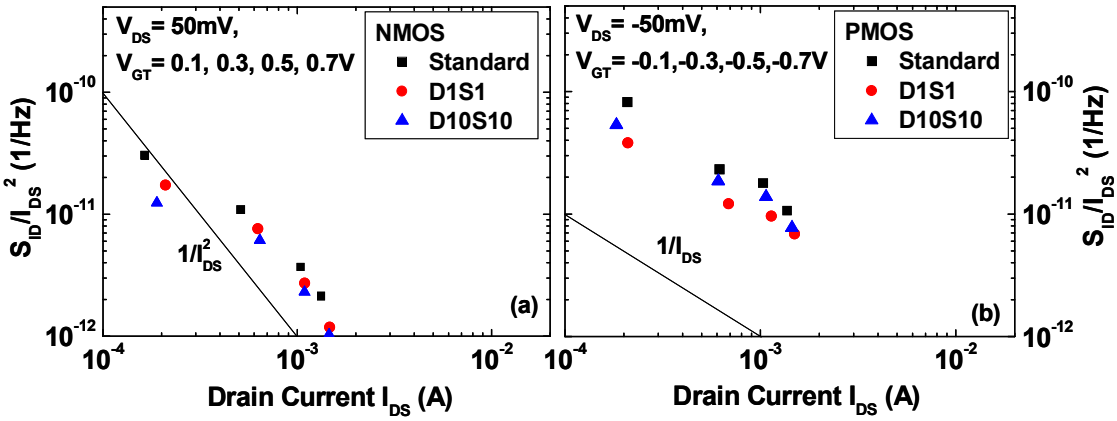


Fig. 4.8 S_{ID}/I_{DS}^2 vs. I_{DS} under varying $|V_{GT}|$ (0.1~0.7V) for standard and donut devices (a) NMOS (b) PMOS Standard : multi-finger W2N32, Donut, D1S1 and D10S10.

4.5 Comparison of LFN between NMOS and PMOS

LFN measurement was carried out using Celestry 9812B noise analyzer and Agilent 35670A dynamic signal analyzer. Noise data was collected under varying V_{GS} in linear region ($|V_{DS}|$=0.05V). According to LFN theory described in chapter 2, the normalized noise power spectral density (S_{ID}/I_{DS}^2) is taken to be an appropriate parameter for a comparison of LFN in various devices with different drain current I_{DS}. Note that the comparison was made under fixed V_{GT} (V_{GS}-V_T) to compensate V_T variations, which may come from INWE, STI TCR, and other factors.

Fig. 4.9 (a) and (b) present S_{ID}/I_{DS}^2 measured from multi-finger MOSFETs with W2N32 and W1N64, respectively. For each layout, NMOS and PMOS were put together for a comparison and the bias condition is specified at the same $|V_{GT}|$ and $|V_{DS}|$. As can be seen clearly that S_{ID}/I_{DS}^2 of PMOS is higher than NMOS by around one order at very low frequency (<10 Hz) and the difference decreases at higher frequency. The results suggest that conventional design using PMOS in VCO for low phase noise is no longer valid. Two potential causes are proposed responsible for the higher LFN revealed in PMOS : one is

thicker $T_{ox(inv)}$ due to worse P^+ poly gate depletion [48] and the other is lower μ_{eff} for holes in PMOS than electrons in NMOS [49]. The first cause can be identified from the measurement that $C_{gg,DUT}$ or $C_{gg,int}$ of PMOS appears smaller than that of NMOS and $T_{ox(inv)}$ extracted from PMOS is around 3.2 nm, i.e. 0.2 nm thicker than that of NMOS. Referring to LFN models, described in chapter 2, the thicker $T_{ox(inv)}$, i.e. the lower $C_{ox(inv)}$ will lead to higher S_{ID}/I_{DS}^2. As for the second one, the lower μ_{eff} may bring worse mobility fluctuation and higher LFN.

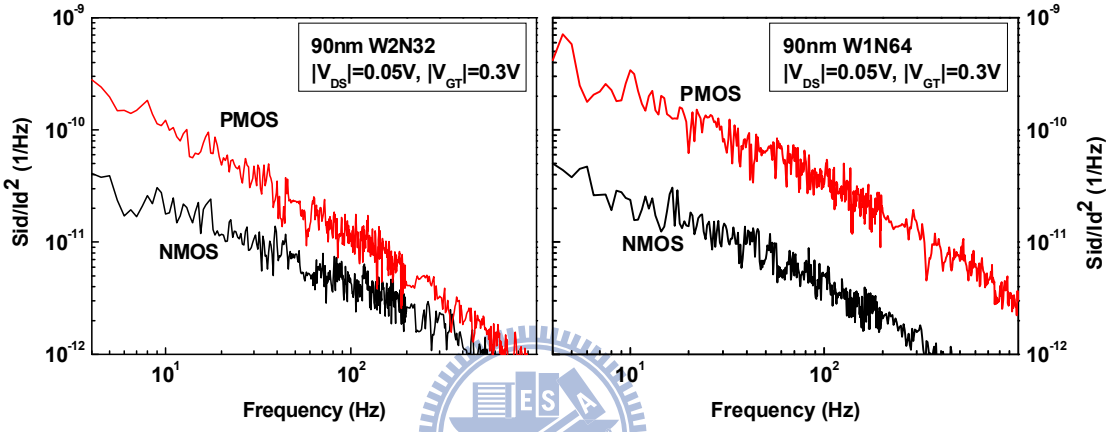


Fig. 4.9 Normalized PSD of drain current noise S_{ID}/I_{DS}^2 vs. frequency, measured from (a) W2N32 NMOS and PMOS (b) W1N64 NMOS and PMOS, under $|V_{GT}|$=0.3V and $|V_{DS}|$=0.05V.

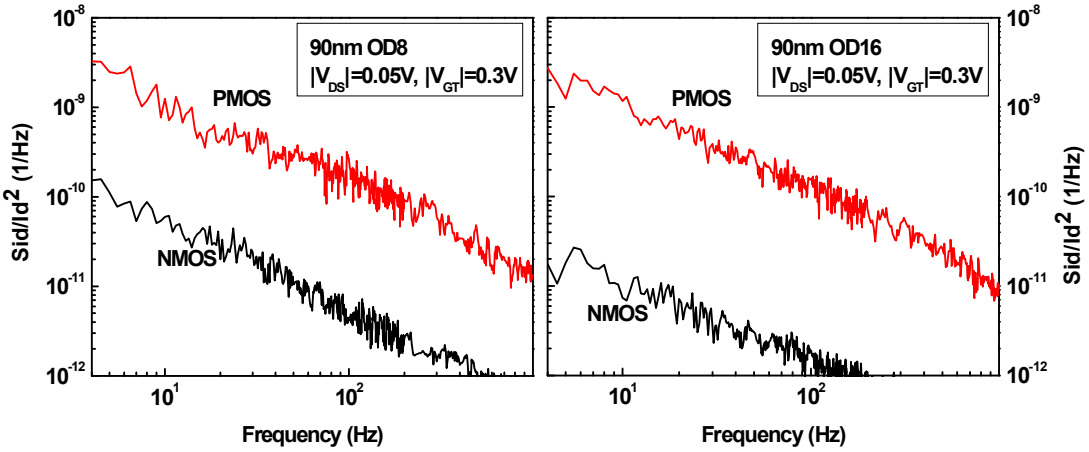


Fig. 4.10 Normalized PSD of drain current noise S_{ID}/I_{DS}^2 vs. frequency, measured from multi-OD MOSFET (a) OD8 ($W_{OD}xN_{OD}$=0.25μmx8) : NMOS and PMOS (b) OD16

($W_{OD}xN_{OD}$ =0.125μmx16) : NMOS and PMOS, under $|V_{GT}|$=0.3V and $|V_{DS}|$=0.05V. multi-OD MOSFET : N_F=16, $W_{OD}xN_{OD}$ =2 μm

Fig. 4.10 (a) and (b) indicate S_{ID}/I_{DS}^2 measured from multi-OD MOSFETs OD8 and OD16, respectively. Again, NMOS and PMOS with the same multi-OD layout and under the same bias condition ($|V_{GT}|$ and $|V_{DS}|$) are put in the plot to make a comparison. As can be seen, the noise level increases with increasing OD number for PMOS devices, over the entire drain current range. For the NMOS devices, the noise level get lower with increasing OD number. This is due to the delta width effect is dominant in NMOS LFN characteristic. The difference between NMOS and PMOS become higher under this condition.

4.6 Layout Dependence of Interface States and Low Frequency Noise

The basic charge-pumping technique involves measuring the substrate current while applying voltage pulses of fixed amplitude, rise time, fall time, and frequency to the gate of the transistor, with the source, drain, and body tied to ground. The pulse can be applied with fixed amplitude, voltage base sweep or a fixed base, variable amplitude sweep. The rise time and fall time used in this research is 100nS. The gate is switched from inversion to accumulation and vice versa using an Agilent 81110A waveform generator. In inversion, some of the minority carriers pumped by the n+ terminal are trapped on the interface states. During the transition from inversion to accumulation, the mobile electrons of the inversion layer are collected by the n+ contact before the majority carriers, flowing from the p+ contact, reach the interface. Next, the recombination of the electrons trapped on the interface states with majority carriers gives rise to a net CP current I_{cp}, measured with an HP 4146. The experiment is computer-controlled by Sagi software which offers multiple options and real-time parameter extraction.

Charge-pumping current measured as a function of base level as shown in **Fig. 4.11** The

amplitude of the gate signal was changed from 1.5V to 1.7V. All experiments are very close indicates that no visible generation of maximum charge pumping current.

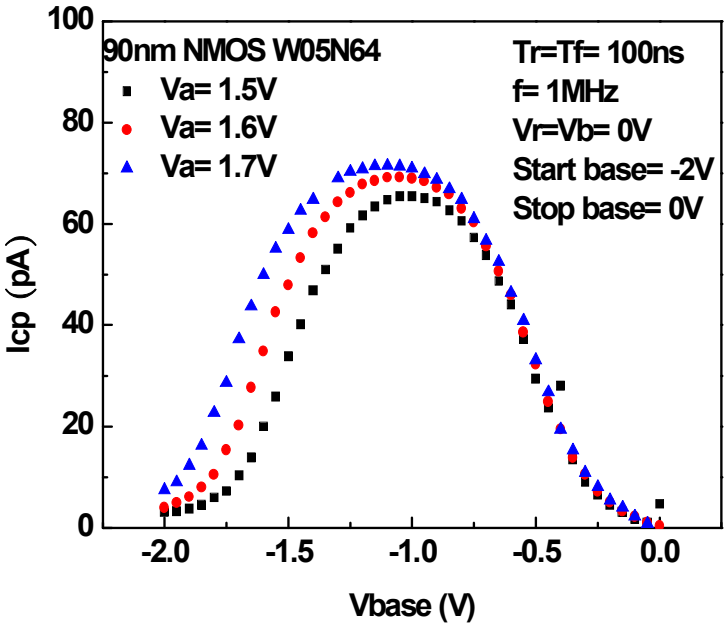


Fig.4.11 Comparison of measured charge pumping current with different pulse amplitude (Va)

Fig. 4.11 indicate the interface trap density N_{it} over the energy levels swept through the Fermi level, which is calculated from the base-level CP current for multi-OD devices as shown in (4.2).

$$I_{cp\_\max} = qfW_{eff}LN_{it} \tag{4.2}$$

Where f is the frequency of applied pulse, and W_{eff} L is the device area. According to (4.2), I_{cp_max} is proportional to Nit and Nit can be extracted from the I_{cp} under specified f and device dimension W_{eff} L.

Standard and Multi-OD MOSFET Nit Analysis

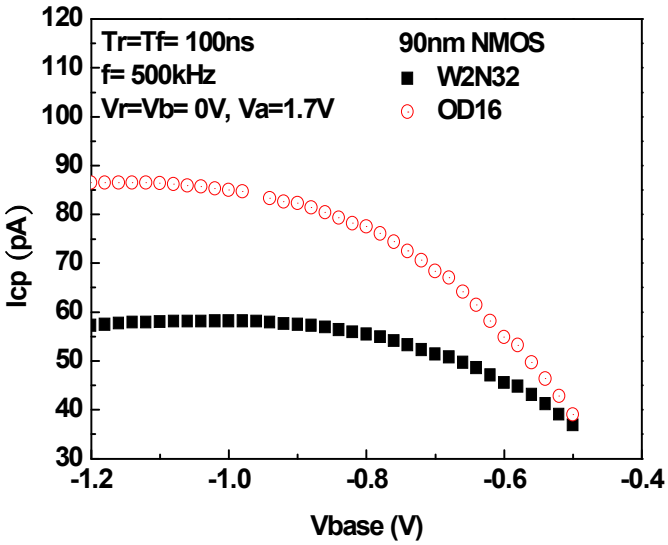


Fig. 4.12 Charge pumping current measurement of standard and multi OD device

Fig. 4.12 presents I_{CP} vs. V_{base} for W2N32 (standard multi-finger) and OD16 (multi-OD) NMOS. The measurement was repeated for different devices layouts and the extracted trap densities are summarized. A statistical drawing of the data in is shown in **Fig. 4.13**. It is noted that W_{total} is used to reflect the $\triangle W$ effect. In **Fig. 4.13**, it is found that the compressive stress may introduce additional interface traps in OD16 devices. However, the normalized difference of $\mathbf{N_{it}}$ between OD1 and OD16, denoted as $\Delta N_{it(OD16,OD1)})/ N_{it(OD1)}$ decreases from 55% to 10% due to the DW effect. This implies that for the multi-OD devices are affected by DW effect significantly, as well as low frequency noise.

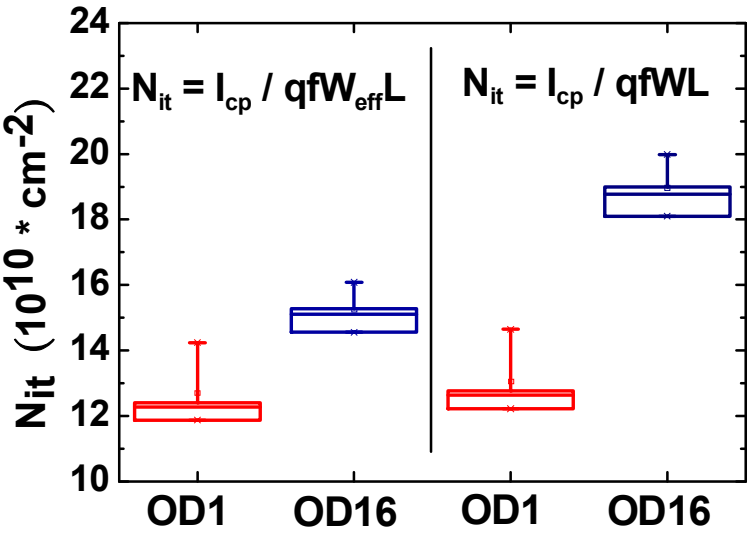


Fig.4.13 Interface trap density N_{it} of OD1 and OD16 devices extracted by using W_{eff} and W, measured at V_D=V_S= V_B= 0V and V_a = 1.7V.

Standard and Narrow-OD MOSFET Nit Analysis

As shown in **Fig. 4.14**, it is found that the charge pumping current is found to be the smallest for W05N64 compared to the W2N16. To confirm that the W05N64 has more interface traps than W2N16, stable and repeatable measurement should be taken statistically that only a slight fluctuation of data among devices is acceptable. However, the difference of each die (~37%) is much larger than device variation (~16%) which is shown in **Fig. 4.15**. The layout effect of narrow OD device on interface trap density remain in discussion here. This has to be devoted to explore the details in the trend for the LFN with narrow-OD devices.

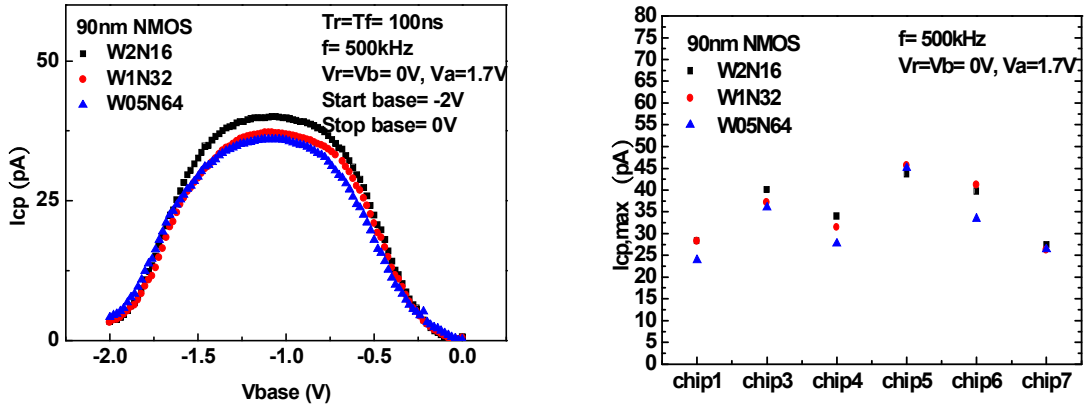


Fig. 4.14 Charge pumping current of three narrow OD devices, and maximum charge pumping current comparison among each measurement die

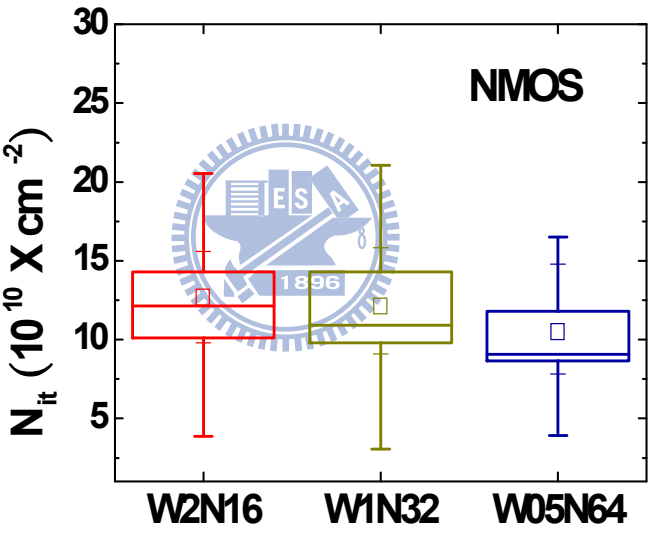


Fig. 4.15 Interface trap density N_{it} of narrow-OD devices extracted by using W_{eff} and W, measured at V_D=V_S= V_B= 0V and V_a = 1.7V.

4.7 Low Frequency Noise Model Parameter Extraction for Simulation

As for doughnut PMOS shown in **Fig.4.8** (b), the measured S_{ID}/I_{DS}^2 follows a simple power law of $1/I_{DS}$ and manifests itself governed by mobility fluctuation model, according to Hooge empirical formula expressed in. Note that the Hooge's parameter α_H is dimensionless

and may vary with biases and process technologies. The reduction of LFN measured from donut PMOS suggests the suppression of mobility fluctuation due to the eliminated compressive $\sigma_{\perp}$.

$$\frac{S_{ID}}{I_{DS}^{\ 2}} = \frac{1}{f}\frac{\alpha_H \mu_{eff}}{L^2}\frac{qV_{DS}}{I_{DS}} \qquad (4.3)$$

α_H : *the Hooge parameter*

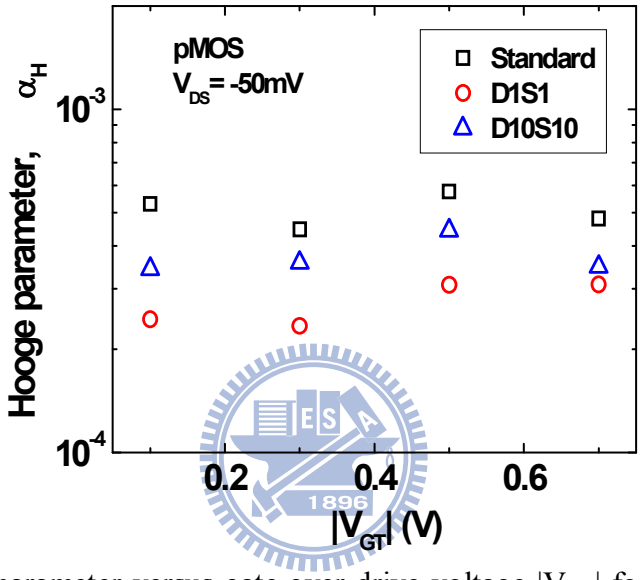


Fig. 4.16 Hooge's parameter versus gate-over-drive voltage $|V_{GT}|$ for standard and doughnut devices biased at $|V_{GT}|$= 0.1~0.7V

Hooge's parameter is a figure of merit for low frequency noise comparison between 10^{-7} and 10^{-3}. **Fig. 4.16** shows the extracted Hooge's parameter versus $|V_{GT}|$ from the mean value of several PMOS devices. The standard device shows larger αH than the doughnut devices. It is explained that the larger enhancement in the surface roughness mobility results in the more phonon scattering limited mobility for standard devices and thus the higher Hooge's parameter can be attributed to lower channel mobility.

Chapter 5

RF MOSFET Layout Effect on High Frequency Characteristics

5.1 Two-port S-parameters Measurement and Deembedding

2-port network parameters can be explained most easily by considering a network with only two ports [63], an input port and an output port, like the network shown in **Fig. 5.1**

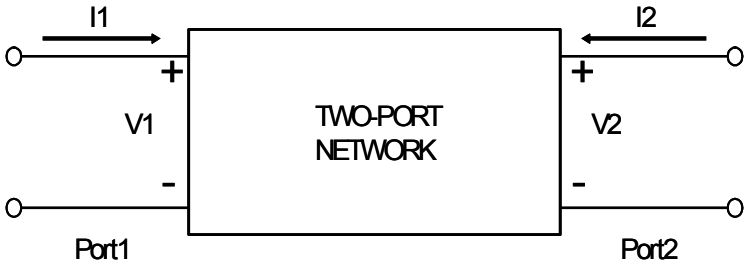


Fig. 5.1 General two-port network

To characterize the performance of such a network, any of several parameter sets can be used, each of which has certain advantages. Each parameter set is related to a set of four variables associated with the two-port model. Two of these variables represent the excitation of the network, and the remaining two represent the response of the network to the excitation. If the network of **Fig. 5.1** is excited by voltage sources V_1 and V_2, the network currents I_1 and I_2 will be related by the following equations (assuming the network behaves linearly):

$$I_1 = Y_{11}V_1 + Y_{12}V_2 \qquad (5.1)$$
$$I_2 = Y_{21}V_1 + Y_{22}V_2$$

If other independent and dependent variables had been chosen, the network would have been described, as before, by two linear equations similar to equations (5.1), except that the variables and the parameters describing their relationships would be different. However, all parameter sets contain the same information about a network, and it is always possible to calculate any set in terms of any other set.

Scattering parameters, which are commonly referred to as s-parameters, are a parameter

set that relates to the traveling waves that are scattered or reflected when a 2-port network is inserted into a transmission line. Another important advantage of s-parameters stems from the fact that traveling waves, unlike terminal voltages and currents, do not vary in magnitude at points along a lossless transmission line. This means that scattering parameters can be measured on a device located at some distance from the measurement transducers, provided that the measuring device and the transducers are connected by low-loss transmission lines.

Following the bias conditions and dc characterization, S-parameters were measured by using HP8510C vector network analyzer from 0.5GHz up to 40 GHz, and the DC supply is HP 4142. A valid calibration needs to be performed prior to the S-parameter measurement, which requires SOLT (Short, open, load, through) calibrations between the two ports of pad and SMA.

Open and short de-embedding are done on the measured two port S-parameters to extract the intrinsic ac characteristics in terms of Y-parameters for C-V model parameter extraction[64] and H-parameters for determination. Z-parameters and Y-parameter of the intrinsic MOSFETs were obtained from S-parameters after de-embedding and used to extract the resistance and such as R_g, R_s, R_d [64]. The intrinsic MOSFET incorporating the parasitic, as extracted was adopted by ADS simulation to do I-V and C-V model parameter extraction and optimization simultaneously. The accuracy of intrinsic MOSFET model has been extensively verified and validated by good match with the measurement.

5.2 Two-port 3T MOSFET with Multi-finger, Multi-OD, and Donut Structures

Test structures were fabricated using 90nm RFCMOS process technology, all devices are designed to 2-port test-key with GSG pads for High frequency measurement. Multi-finger structure with finger number N_F=16, 32, and 64, finger width W_F= 2, 1, and 0.5, the total width and channel length are fixed at 32um and 0.09um. Another set of multi-finger type with

N_F=32, 64, and 128, finger width= 2, 1, and 0.5, the total width and channel length are fixed at 64um and 0.09um, respectively. The multi-OD device with gate finger number N_F= 16, finger width W_F= 2, 0.25, 0.125um and OD number N_{OD}= 1, 8, 16 accordingly with fixed OD space 0.2μm. Doughnut devices with two layout dimensions D1S1 and D10S10 are also characterized here. D1S1 represents donut MOSFET in which the space from poly gate to STI edge follows the minimum rule, and D10S10 denotes donut MOSFET with 10 times larger space between poly gate and STI edge.

The parameters in equivalent circuit would be extracted and discussed in the following chapters. De-embedding method and parameter extraction are major subjects and small signal equivalent circuit model constructed by extracted parameters would be verified by using Agilent Advanced Design System (ADS) simulator to make the parameter extraction method and extracted parameters solider.

5.2.1 Intrinsic Device and Parasitic RLC Parameters Extraction and Analysis – Layout Effect and Bias dependence

The intrinsic gate resistance obtained by Y-method of standard and doughnut were extracted at V_{DS}= 1.2V and V_{GS}= 1V are plotted in **Fig. 5.2** From the extracted R_g at various frequencies, we can find easily it is frequency independent and does not appear non-quasi-static (NQS) effect because of sufficient high frequency and short channel.

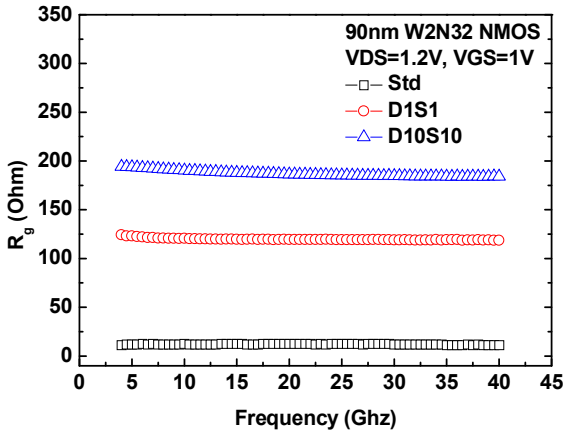


Fig. 5.2 The extracted R_g of standard and doughnut devices as a function of frequency with

V_{DS}= 1.2V, V_{GS}= 1V

Fig. 5.3 and **Fig. 5.4** shows the extracted R_g for NMOS and PMOS with standard and doughnut devices at saturation region. We can see R_g shows very weak gate bias dependence for short channel devices. Based on above results, we can see there is no different between Y-method and Z-method extracting results, and thus we can confirm the exact value of poly gate resistance for different test patterns.

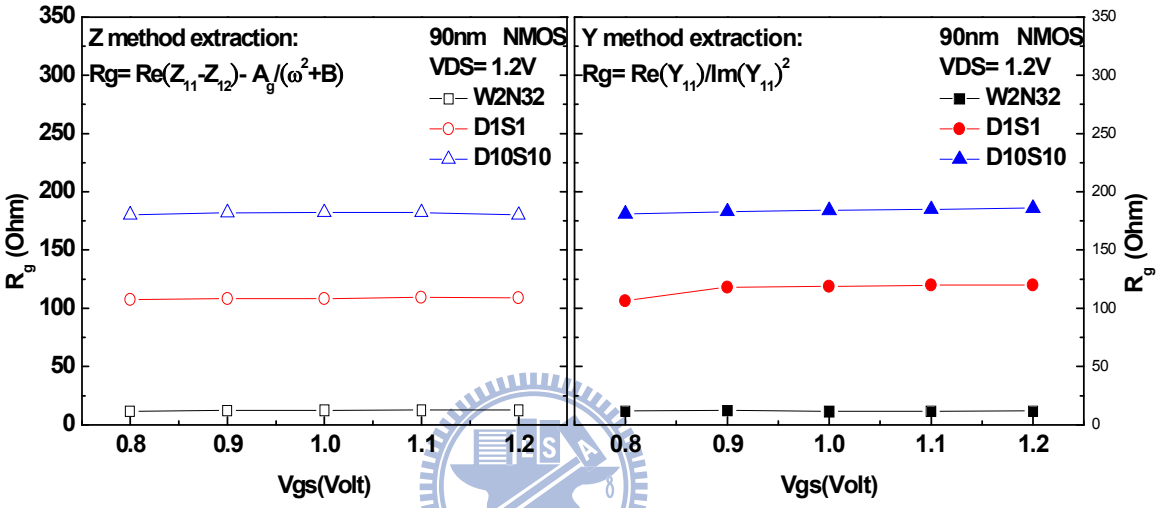


Fig. 5.3 Comparison of (a) Z-method and (b) Y-method R_g extraction for NMOS standard and doughnut devices

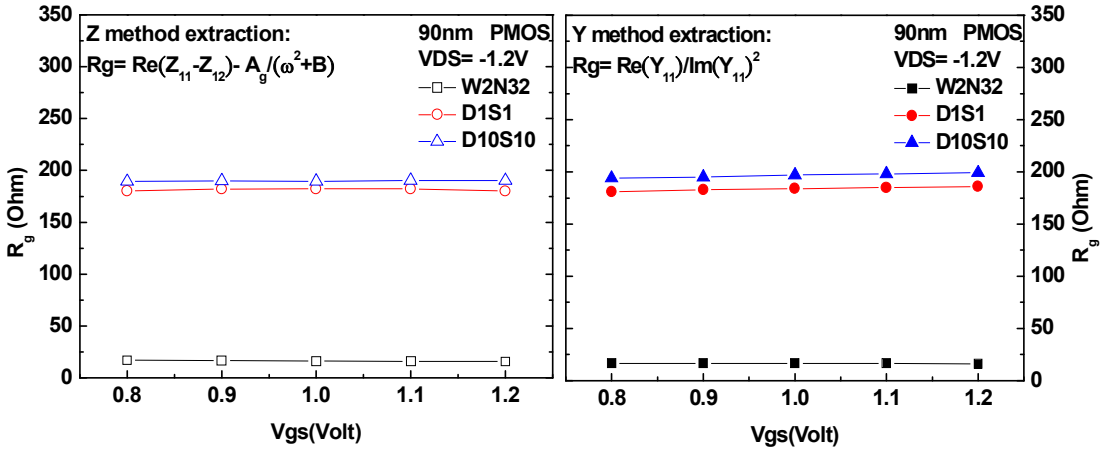


Fig. 5.4 Comparison of (a) Z method and (b) Y method R_g extraction for PMOS standard and doughnut device

Table. 5.1, **table. 5.2**, and **table. 5.3** is the parameter values which are for the gate resistance optimization of Z-method. The calculated A_g and B are mainly from the parameter g_{ds}, C_{gg}, C_{gd}, C_{gs}, g_m, and C_{ds}. The used devices are standard multi-finger W2N32, D1S1, and D10S10, including N type and P type MOSFETs.

Table. 5.1 The optimized parameters (A_g and B) for Z method R_g extraction of standard W2N32 device

NMOS standard W2N32 bias dependent with vary Vgs (Vds=1.2V)									
Vgs	**gds (S)**	**Cgg (F)**	**Cgd (F)**	**Cgs (F)**	**gm (S)**	**Cds (F)**	**B**	**Ag**	**Ag/(w^2 +B), f=40GHz**
0.8	3.38E-03	7.68E-14	2.49E-14	5.19E-14	4.26E-02	1.89E-14	#######	#######	7.07
0.9	4.06E-03	7.87E-14	2.51E-14	5.36E-14	4.68E-02	1.79E-14	#######	#######	5.74
1	4.63E-03	8.00E-14	2.53E-14	5.47E-14	4.87E-02	1.72E-14	#######	#######	4.85
1.1	5.15E-03	8.10E-14	2.56E-14	5.54E-14	4.92E-02	1.67E-14	#######	#######	4.17
1.2	5.65E-03	8.17E-14	2.59E-14	5.58E-14	4.89E-02	1.63E-14	#######	#######	3.59
								Average	**5.08**

PMOS standard W2N32 bias dependent with vary Vgs (Vds= -1.2V)									
Vgs	gds (S)	Cgg (F)	Cgd (F)	Cgs (F)	gm (S)	Cds (F)	B	Ag	Ag/(w^2 +B), f=40GHz
-0.8	3.84E-03	7.16E-14	2.39E-14	4.78E-14	1.60E-02	2.77E-14	4.42E+22	6.34E+23	5.91
-0.9	4.37E-03	7.28E-14	2.43E-14	4.86E-14	1.84E-02	2.78E-14	5.70E+22	7.07E+23	5.88
-1	4.88E-03	7.39E-14	2.47E-14	4.92E-14	2.01E-02	2.81E-14	6.78E+22	7.40E+23	5.65
-1.1	5.45E-03	7.47E-14	2.52E-14	4.95E-14	2.13E-02	2.81E-14	7.96E+22	7.40E+23	5.19
-1.2	5.96E-03	7.56E-14	2.58E-14	4.98E-14	2.21E-02	2.81E-14	8.96E+22	7.20E+23	4.71
								Average	5.47

Table. 5.2 The optimized parameters (Ag and B) for Z method Rg extraction of doughnut D1S1 device

NMOS doughnut D1S1 bias dependent with vary Vgs (Vds=1.2V)									
Vgs	gds (S)	Cgg (F)	Cgd (F)	Cgs (F)	gm (S)	Cds (F)	B	Ag	Ag/(w^2 +B), f=40GHz
0.8	3.11E-03	7.73E-14	2.15E-14	5.57E-14	4.26E-02	2.58E-14	1.31E+23	1.96E+24	10.06
0.9	3.75E-03	7.93E-14	2.17E-14	5.76E-14	4.68E-02	2.48E-14	1.67E+23	1.98E+24	8.62
1	4.31E-03	8.05E-14	2.19E-14	5.85E-14	4.87E-02	2.42E-14	1.92E+23	1.95E+24	7.65
1.1	4.81E-03	8.15E-14	2.22E-14	5.93E-14	4.92E-02	2.36E-14	2.10E+23	1.85E+24	6.79
1.2	5.30E-03	8.22E-14	2.26E-14	5.96E-14	4.89E-02	2.33E-14	2.23E+23	1.75E+24	6.11
								Average	7.85

PMOS doughnut D1S1 bias dependent with vary Vgs (Vds= -1.2V)									
Vgs	gds (S)	Cgg (F)	Cgd (F)	Cgs (F)	gm (S)	Cds (F)	B	Ag	Ag/(w^2 +B), f=40GHz
-0.8	3.85E-03	7.06E-14	2.34E-14	4.72E-14	1.69E-02	2.67E-14	4.98E+22	7.06E+23	6.25
-0.9	4.27E-03	7.22E-14	2.38E-14	4.84E-14	1.97E-02	2.69E-14	6.31E+22	8.04E+23	6.37
-1	4.66E-03	7.34E-14	2.43E-14	4.92E-14	2.16E-02	2.72E-14	7.37E+22	8.53E+23	6.23
-1.1	5.05E-03	7.44E-14	2.48E-14	4.96E-14	2.30E-02	2.74E-14	8.38E+22	8.81E+23	6
-1.2	5.44E-03	7.51E-14	2.53E-14	4.98E-14	2.39E-02	2.75E-14	9.28E+22	8.84E+23	5.67
								Average	6.1

Table. 5.3 The optimized parameters (A_g and B) for Z method R_g extraction of doughnut D10S10 device

NMOS doughnut D10S10 bias dependent with vary Vgs (Vds=1.2V)									
Vgs	**gds (S)**	**Cgg (F)**	**Cgd (F)**	**Cgs (F)**	**gm (A/V)**	**Cds (F)**	**B**	**Ag**	**Ag/(w^2 +B), f=40GHz**
0.8	2.83E-03	7.82E-14	2.49E-14	5.33E-14	4.75E-02	3.33E-14	1.28E+23	2.31E+24	12.09
0.9	3.50E-03	7.99E-14	2.51E-14	5.48E-14	5.12E-02	3.27E-14	1.54E+23	2.34E+24	10.78
1	4.07E-03	8.10E-14	2.54E-14	5.56E-14	5.27E-02	3.21E-14	1.73E+23	2.31E+24	9.8
1.1	4.58E-03	8.18E-14	2.57E-14	5.61E-14	5.29E-02	3.14E-14	1.87E+23	2.24E+24	8.97
1.2	5.08E-03	8.25E-14	2.60E-14	5.65E-14	5.23E-02	3.07E-14	1.98E+23	2.14E+24	8.21
								Average	**9.97**

PMOS doughnut D10S10 bias dependent with vary Vgs (Vds= -1.2V)									
Vgs	**gds (S)**	**Cgg (F)**	**Cgd (F)**	**Cgs (F)**	**gm (A/V)**	**Cds (F)**	**B**	**Ag**	**Ag/(w^2 +B), f=40GHz**
-0.8	2.98E-03	7.12E-14	2.32E-14	4.80E-14	1.69E-02	3.22E-14	3.15E+22	8.02E+23	8.48
-0.9	3.65E-03	7.26E-14	2.36E-14	4.90E-14	1.94E-02	3.11E-14	4.48E+22	8.59E+23	7.96
-1	4.28E-03	7.36E-14	2.41E-14	4.95E-14	2.13E-02	3.01E-14	5.91E+22	8.91E+23	7.29
-1.1	4.90E-03	7.46E-14	2.47E-14	4.99E-14	2.26E-02	2.93E-14	7.30E+22	8.83E+23	6.49
-1.2	5.52E-03	7.54E-14	2.53E-14	5.01E-14	2.34E-02	2.87E-14	8.63E+22	8.49E+23	5.68
								Average	**7.18**

Narrow-OD R_g with Z-method and Y-method were extracted and plotted as a function gate bias, as shown in **Fig. 5.5** that were obtained by the two methods are compared. Similar values were obtained for W2N16, which has the minimum finger number here. But for the W1N32 and W05N64, a slight difference between Z-method and Y-method. The lowering of R_g by Y-method attribute to the term in the denominator of $Re(Y_{11})/Im(Y_{11})^2$, this capacitive components become higher due to the poly gate to metal coupling introduced by poly gate to metal coupling capacitance, which is not clean enough for open metal3 de-embedding result.

And this can be improved by applying metal1 de-embedding but will be at a significantly cost of chip area

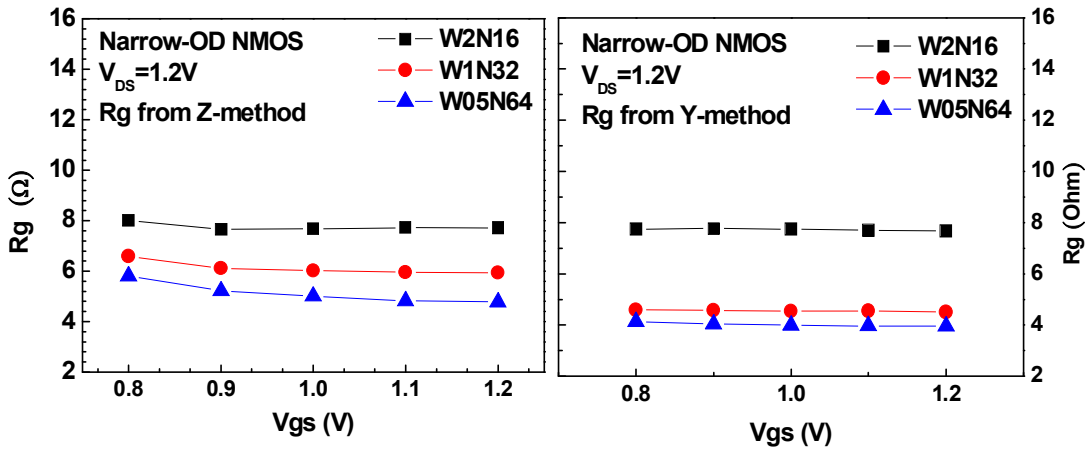


Fig. 5.5 Comparison of (a) Z method and (b) Y method R_g extraction for NMOS narrow-OD device

As discussed before, the R_g lowering effect is fine and acceptable for narrow-OD devices with Y-parameter. But for multi-OD with even larger extrinsic parasitic capacitance, the R_g is dramatically different between there two method. Multi-OD R_g with Z-method and Y-method were extracted and plotted as a function gate bias, as shown in **Fig. 5.6** that were obtained by the two methods are compared. Similar values were obtained for OD1 and OD8, which has nearly the same poly width here. But for OD16, a significant difference between Z-method and Y-method occurs and a totally different trends. The Y-method R_g of OD16 even lower than OD1. Same as narrow-OD devices, the lowering of R_g by Y-method attribute to the term in the denominator of $Re(Y_{11})/Im(Y_{11})^2$, this capacitive components become higher due to the longer poly gate to metal coupling introduced by poly gate to metal coupling capacitance, which is not clean enough for open metal3 de-embedding result. And this also can be improved by applying metal1 de-embedding for a more accurate result. We can judge from

the comparison result that the Z-parameter is suitable and more accurate than Y-parameter in most of the conditions and test patterns.

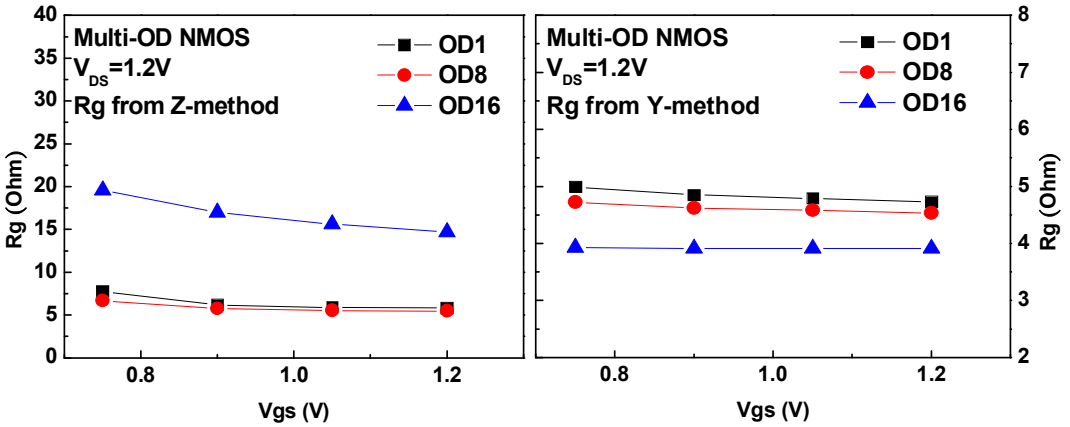


Fig. 5.6 Comparison of (a) Z method and (b) Y method R_g extraction for NMOS multi-OD device

5.2.2 High Frequency Performance Analysis (f_T, f_{MAX}, and NF_{min})

A model for the high frequency figures-of-merit can be derived based on the small signal equivalent circuit includes gate resistance (R_g), gate-source (C_{gs}), gate-drain (C_{gd}), gate-body (C_{gb}) capacitances, transconductance (G_m), body resistance (R_{bb}), and junction capacitance (C_{dsb}). Previous studies [67] have ignored the presence of C_{gb}, but it will affect the f_T and unilateral gain significantly, even with a small value that arises from fringing capacitance.

The cut-off frequency f_T and Maximum oscillation frequency f_{MAX} are the most important figures of merit for the frequency characteristics of RF transistors. They are often used to emphasize the superiority of newly developed semiconductors or technologies. The value of f_{MAX} can be determined by the unilateral power gain U_{gain} as defined by [68].

$$U_{gain} = \frac{\left|\frac{S_{21}}{S_{12}} - 1\right|^2}{2K\left|\frac{S_{21}}{S_{12}}\right| - 2\,\mathrm{Re}\left[\frac{S_{21}}{S_{12}}\right]} \tag{5.17}$$

Where K is Kurokawa's stability factor[] defined as

$$\text{where } K = \frac{1-|S_{11}|^2-|S_{22}|^2+|\Delta|^2}{2|S_{12}S_{21}|},\ \Delta = S_{11}S_{22}-S_{12}S_{21}\text{CC} \tag{5.18}$$

We can also expressed in the Y-parameter

$$U_{gain} = \frac{|Y_{21}-Y_{12}|^2}{4[\mathrm{Re}(Y_{11})\mathrm{Re}(Y_{22})-\mathrm{Re}(Y_{12})\mathrm{Re}(Y_{21})]} = \frac{|Y_{21}-Y_{12}|^2}{4[g_{11}g_{22}-g_{12}g_{21}]} \tag{5.19}$$

Therefore, f_{MAX} is the Maximum frequency at which the transistor still provides a power gain. An ideal oscillator would still be expected to operate at this frequency, hence the name Maximum oscillation frequency. Like the short circuit current gain H_{21}, Ugain drops with a slope of -20dB/dec.

f_{MAX} does not have to necessarily larger than f_T. Generally, transistors have useful power gains up to f_{MAX}, that above they cannot be used as power amplifiers any more. However, the importance of f_T and f_{MAX} depends on the specific application. Thus, there is no general answer whether f_{MAX} should be prioritized over f_T. Both figures should be as high as possible, and manufactures often strive for $f_T \sim f_{MAX}$ in order to enter many different application for their transistors.

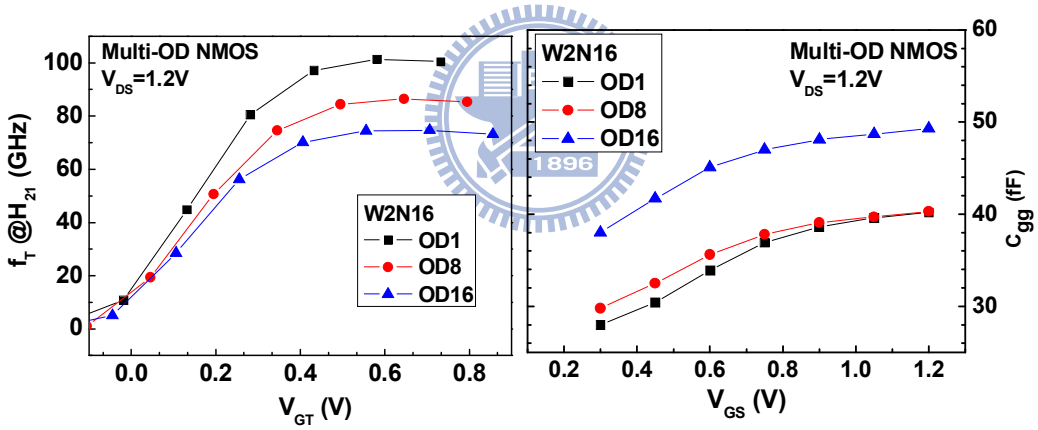
The impact from layout dependent STI stress on high frequency performance is crucial for RF MOSFETs and circuits design. Fig. illustrates the cutoff frequency f_T measured from NMOS with multi-OD layouts. (Noted that f_T is extracted from extrapolation of $|H_{21}|$ to unity gain).

The impact from layout dependent STI stress on high frequency performance is of special concern for RF MOSFETs and circuits design. **Fig. 5.7** illustrates the cutoff frequency f_T measured from NMOS with multi-OD layouts. (Note that f_T is extracted from extrapolation of short circuit current gain $|H_{21}|$ to unity gain). H_{21} basically characterizes the ratio between the

small-signal drain and gate current $|I_D/I_G|$. The cut-off frequency is normally extracted for various operating points.

It is found that OD1 gains the highest f_T while the OD16 reveals itself the worst one. An analytical model for calculating f_T, it is predicted that f_T is proportional to G_m and the enhancement of G_m can boost f_T under fixed gate capacitances (C_{gg} and C_{gd}). **Fig.5.13** (b) present C_{gg} measured from multi-OD layouts. The results indicate that OD16 gets larger C_{gg} and smaller G_m as compared to OD1. Thus, layout dependence of f_T just follows those of both C_{gg} and G_m.

Regarding other RF performance parameters, such as maximum oscillation frequency, f_{MAX} and noise figure, NF_{min} (not shown), the OD16 MOSFETs suffer significant degradation due to inherently larger gate resistances.



(a) (b)

Fig. 5.7 (a) The cut-off frequency f_T versus V_{GT} measured for multi-OD devices(b) C_{gg} versus V_{GS} extracted from Y-parameters for multi-OD devices

The second important RF figure of merit is the Maximum oscillation frequency f_{MAX}, which is related to the frequency at which the device power gain equals unity. The high frequency U_{gain} and f_{MAX}, on the other hand, are independent of C_{gb}, but are influenced by the pole formed by C_{db}. Since f_{MAX} are relatively insensitive to the change in body resistance, it makes sense that the power characteristics are also insensitive to body resistance. But the

impact of body resistance on unilateral gain is an important consideration for designers. The unilateral power gain, on the other hand, can depend on the body resistance. We used experimental data to verify the relationship between f_T and f_{MAX}.

$$f_{MAX} = \frac{f_T}{2\sqrt{g_{ds}(R_g + R_s) + 2\pi f_T C_{gd} R_g}} \tag{5.20}$$

Where R_g and R_s are the gate and source resistance respectively, and gds is output conductance. From the equation, it is noted that f_{MAX} is heavily dependent on the parasitic resistance and capacitance of the CMOS device. Hence, the layout of the device can significantly affect the value of f_{MAX}. **Fig. 5.8** and **Fig. 5.9** are the f_{MAX} for multi-OD and narrow-OD NMOS.

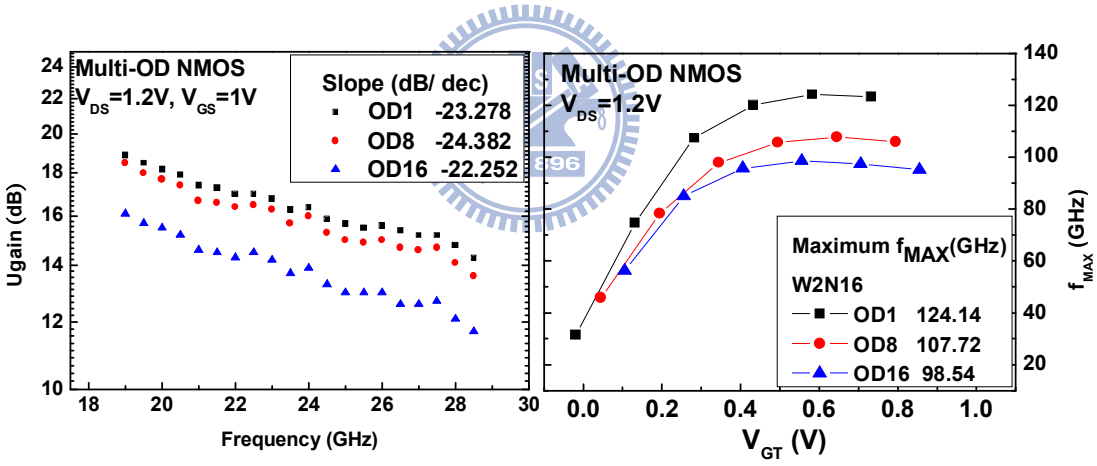


Fig. 5.8 (a) Ugain and (b) f_{MAX} extraction for multi-OD NMOS

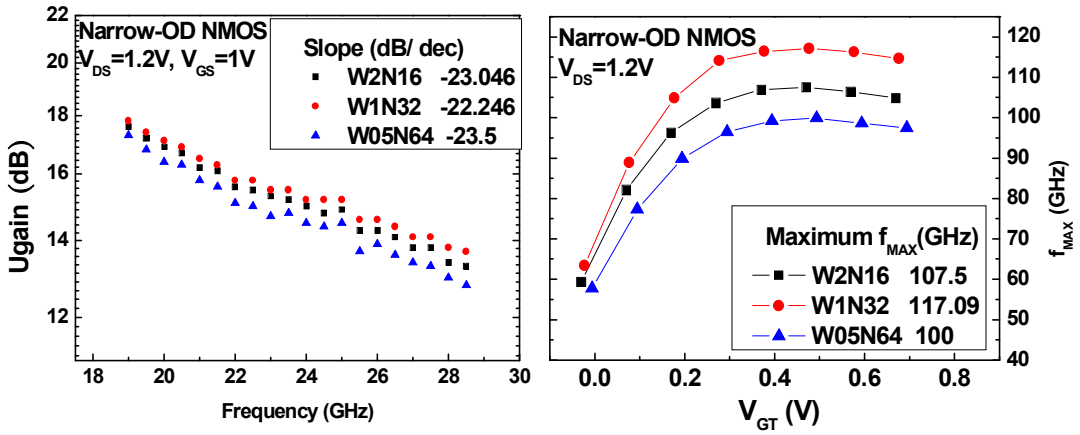


Fig. 5.9 (a) Ugain and (b) f_{MAX} extraction for narrow-OD NMOS

Standard and Doughnut device comparison

The impact from layout dependent STI stress on high frequency performance is of special concern for RF MOSFETs and circuits design. **Fig. 5.10** (a) and (b) illustrate the cutoff frequency f_T measured from NMOS and PMOS with donut and standard layouts. Note that f_T is extracted from the extrapolation of $|H_{21}|$ to unity gain. For NMOS in **Fig. 5.10** (a), D10S10 gains 5% improvement in the Maximum f_T compared to the standard and D1S1. The benefit from donut layout becomes particularly larger for PMOS. As shown in **Fig. 5.10** (b), D1S1 presents the best performance with the highest f_T and realizes 28% increase in the Maximum f_T than the standard device.

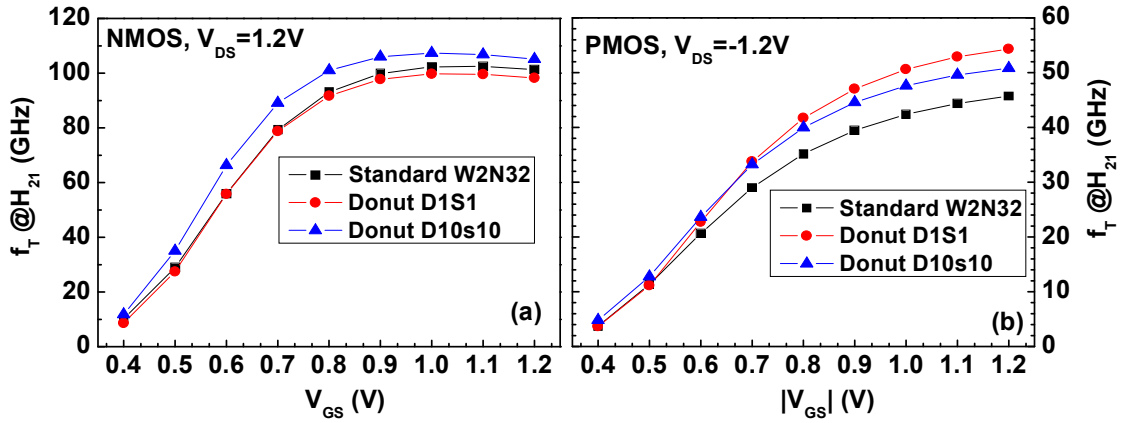


Fig. 5.10 The cut-off frequency f_T vs. V_{gs} measured for standard and donut devices (a) NMOS (b) PMOS. Standard : multi-finger W2N32. Donut : D1S1 and D10S10.

The resulted improvement on f_T in donut MOSFETs can be consistently explained by the enhancement of μ_{eff} and G_m. An analytical model for calculating f_T, it is predicted that f_T is proportional to Gm and the enhancement of Gm can boost f_T under fixed gate capacitances (C_{gg} and C_{gd}). **Fig. 5.11** (a) and (b) present C_{gg} measured from NMOS and PMOS with three different layouts. The results indicate much smaller difference in C_{gg} between donut and standard layouts, as compared with Gm. Thus, layout dependence of f_T just follows that of G_m.

$$f_T = \frac{G_m}{2\pi\sqrt{C_{gg}^2 - C_{gd}^2}} \tag{5.21}$$

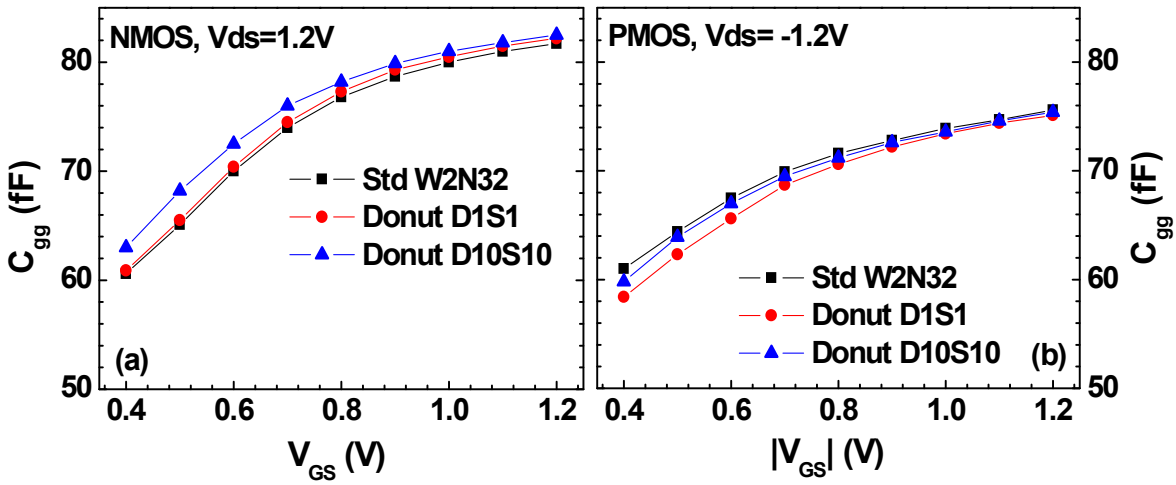


Fig. 5.11 C_{gg} vs. V_{gs} extracted from Y-parameters for standard and donut devices (a) NMOS (b) PMOS. Standard : multi-finger W2N32. Donut : D1S1 and D10S10.

Regarding other RF performance parameters, such as maximum oscillation frequency, f_{MAX} and noise figure, NF_{min}, the donut MOSFETs suffer significant degradation due to inherently larger gate resistances than the standard one with multiple gate fingers. The experimental suggests an innovative donut device layout is required to cover all of the RF and analog performance. Normally, noise parameters are measured using a relatively complicated system that measured both noise and S-parameters of the device under test (DUT) and these measurements are very time consuming.

An analytical method to extract the HF noise parameters of the MOSFET directly has been recently presented in [69]. Direct measurement of the noise parameters requires considerable amount of time since. The noise performance of any noisy two-port network can be represented by:

$$\begin{aligned} F_{\min} &= 1 + 2R_n \,\mathrm{Re}(Y_{sopt})[1 + R_n \,\mathrm{Re}(Y_{sopt})] \\ &= 1 + K_1 \cdot \frac{f}{f_T}\sqrt{gm(R_g + R_s)} \\ &= 1 + K_2 \cdot f \cdot C_{gs} \cdot \sqrt{\frac{(R_g + R_s)}{g_m}} \end{aligned} \tag{5.22}$$

$$NF_{\min} = 10 \log F_{\min} \tag{5.23}$$

$$g_{do} = \frac{W_{eff}}{L_{eff}} \mu_{eff} C_{ox} (V_{gs} - V_t) \tag{5.24}$$

Where F_{min} is the minimum noise factor (the minimum noise factor in dB is called minimum noise figure), which is defined as the signal-to-noise ratio at the input port divided by the signal-to-noise ratio at the output port of the noisy two-port. R_n is the equivalent noise resistance, Y_s (= G_s+ jB_s) is the source admittance and Y_{sopt} (= G_{sopt}+ jB_{sopt}) is the optimum source admittance which result in the F_{min}. The best noise figure in a circuit is achieved when the device is presented with optimum source impedance. The optimum input network to achieve this objective does not in general result in an excellent return loss match. Balanced amplifiers and isolators are sometimes used to achieve both the optimum noise figure and a good match.

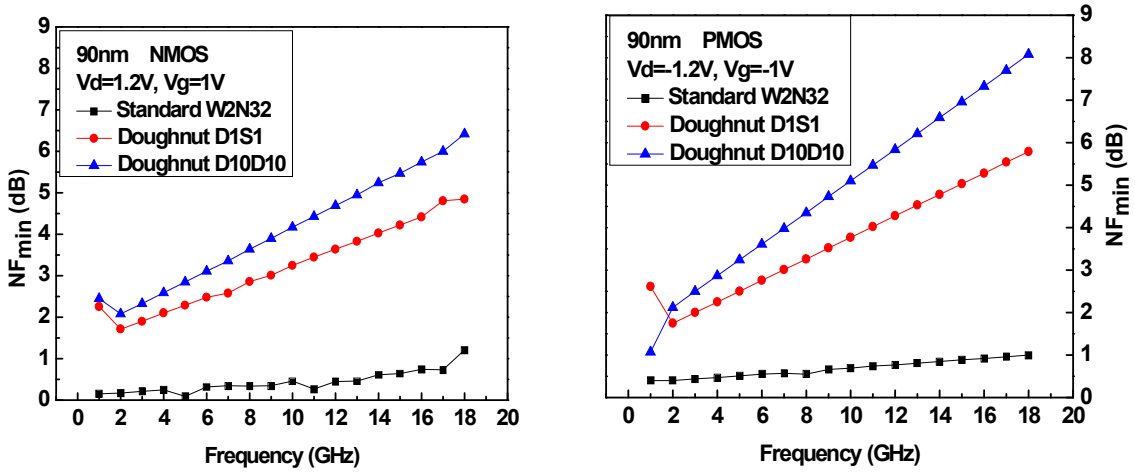


Fig. 5.12 Comparison of NF_{min} measurement result of standard and doughnut devices of NMOS (left) and PMOS (right).

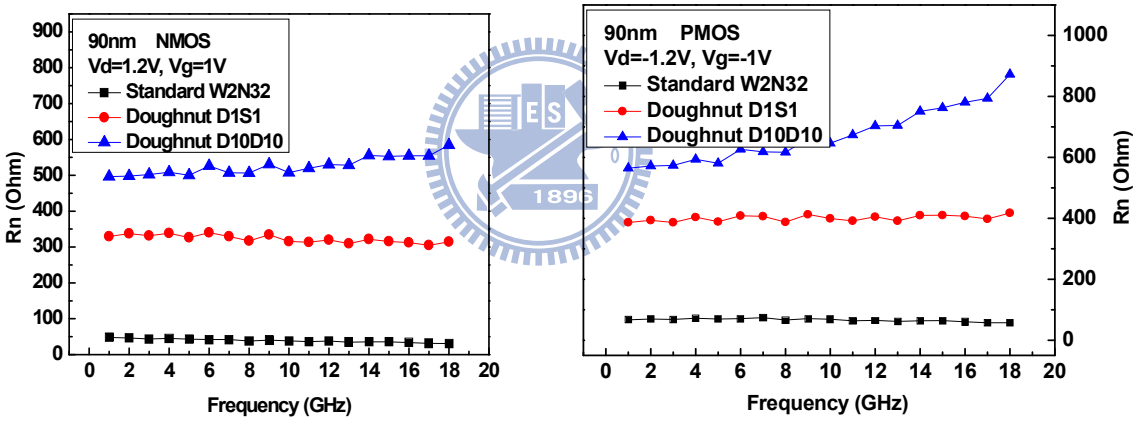


Fig. 5.13 Comparison of R_n measurement result of standard and doughnut devices of NMOS (left) and PMOS (right).

Fig. shows the measured values of all four noise parameters of the device versus frequency at V_{gs}= 1V and V_{ds}= 1.2V. It is observed that R_n does not vary much with frequency, but the doughnut devices have markedly difference R_n between standard multi-finger devices, which can be seen that the gate resistance have a great impact on noise resistance. The R_n of doughnut is D1S1 about 7 times larger than standard and we can verify it from the following relation formula:

$$R_n \approx \mathrm{Rg} + \gamma \frac{g_{do}}{g_m^{\,2}}, \quad \frac{2}{3} \le \gamma \le 1 \tag{5.25}$$

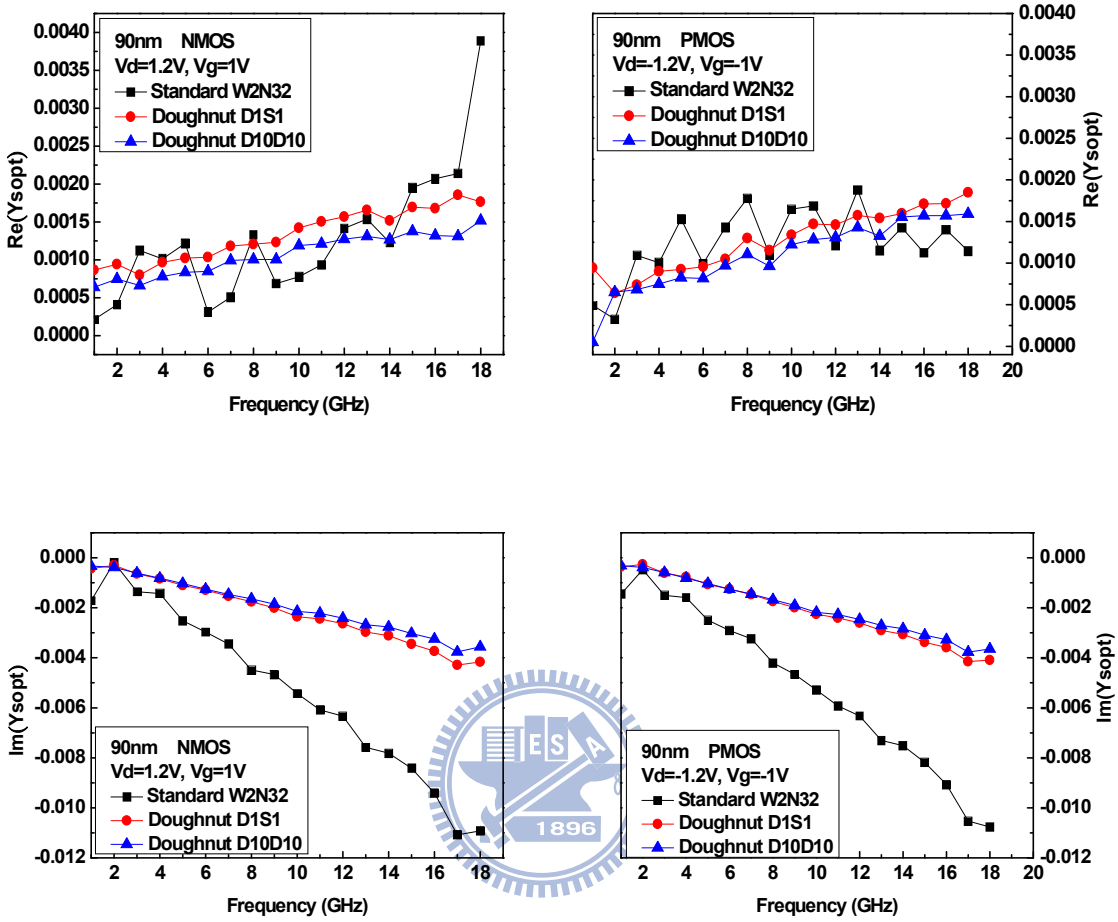


Fig. 5.14 Comparison of Y_{sopt} measurement result of standard and doughnut devices of NMOS (left) and PMOS (right).

Fig. 5.15 shows NF_{min} and R_n vs. frequency measured from the narrow-OD NMOS at V_{gs}= 0.8V and V_{ds}= 1.2V, which is the bias condition when G_m reach the Maximum value. It is observed that R_n shows no difference among three devices, NF_{min} is the worst for W05N64 but not that critical for a normal condition. It is believed that the multi-finger type transistors in 90nm technology are well-controlled at HF noise behavior.

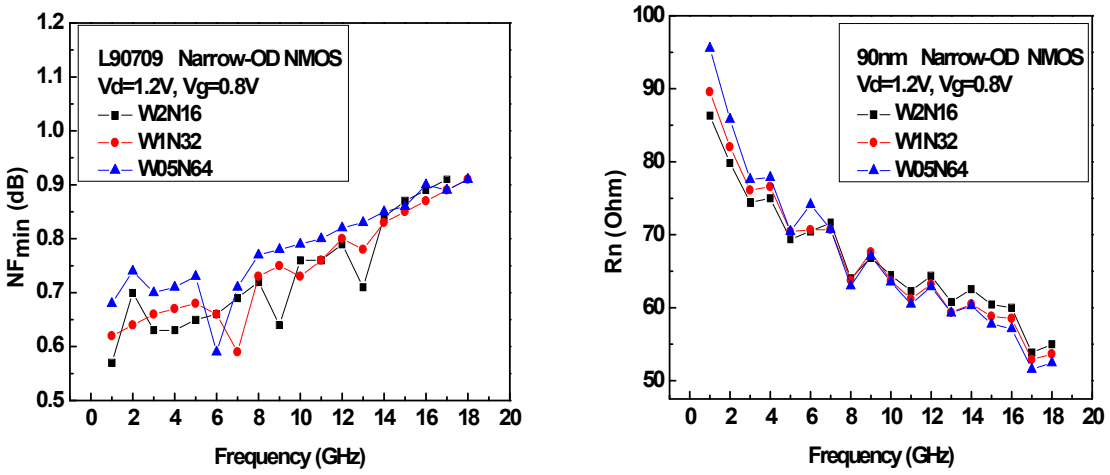


Fig. 5.15 Comparison of noise parameters measurement result of standard and narrow-OD devices, (a) NF_{min} (b) R_n

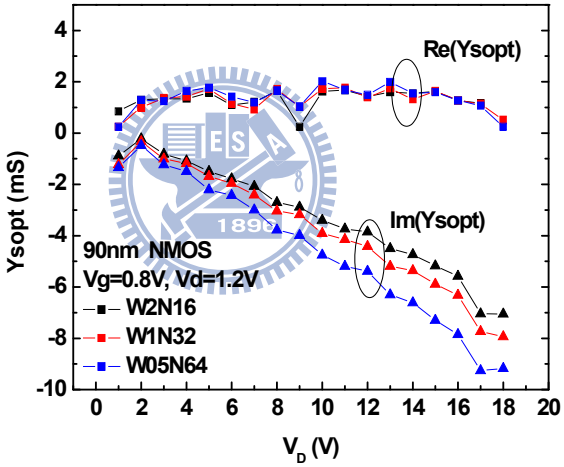


Fig. 5.16 Comparison of noise parameters $Re(Y_{sopt})$ and $Im(Y_{sopt})$ measurement result of standard and narrow-OD devices

The difference of RF noise parameters from layout variation in narrow-OD NMOS is not obvious, whereas the multi-OD devices have a totally different performance with respect to the narrow-OD ones, as shown in **Fig. 5.17**. The OD16 with 0.125um OD width has the worst NF_{min}, but the noise resistance of OD16 is not the largest one. The OD8 has a totally different slope of NF_{min} versus frequency compared to OD1 and OD16, and the noise resistance of OD8 is the largest one. For the frequency under 10Ghz, the NF_{min} of OD8 is higher than OD1,

when the frequency higher than 10Ghz, the NF_{min} of OD1 exceed the OD8. As a summury, OD16 with middle Rn but the worst parasitics effect (mainly from Cgs) make it NF_{min} worse than OD1 and OD8.

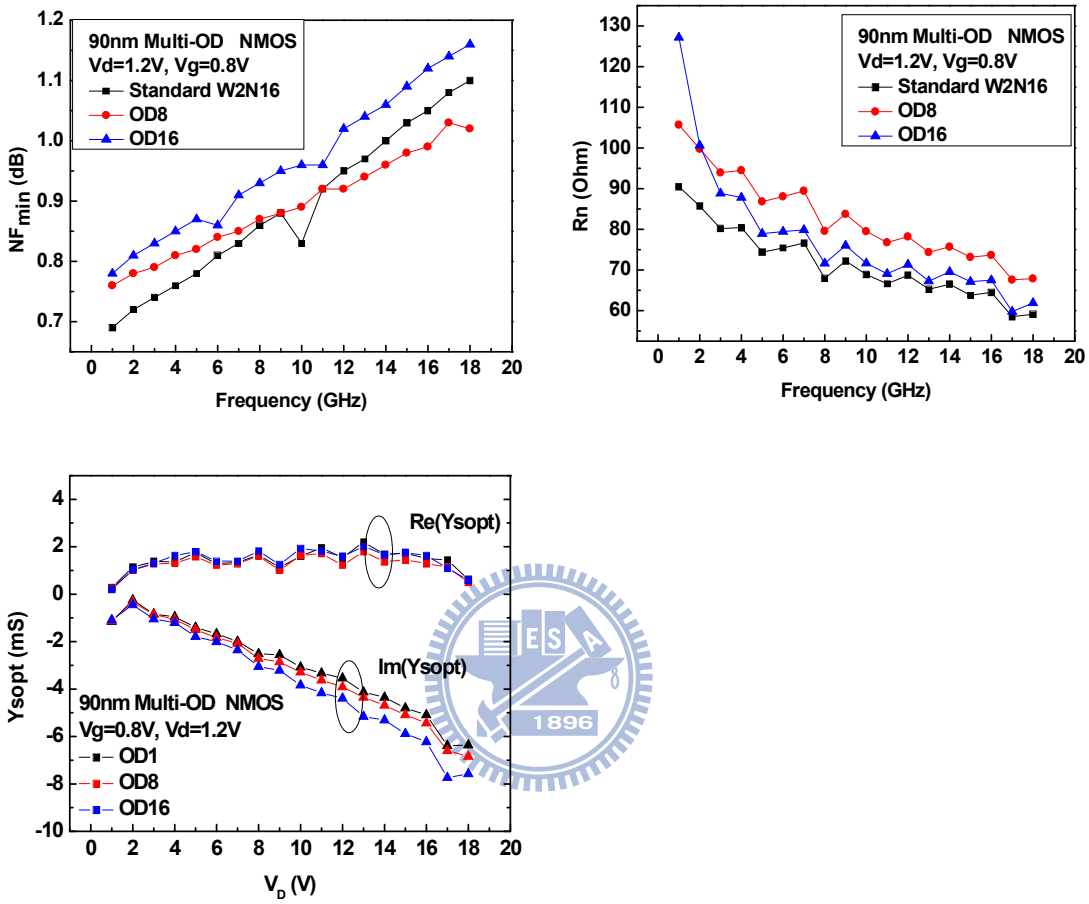


Fig. 5.17 Comparison of noise parameters measurement result of standard and multi OD devices, (a) NF_{min} (b) R_n (c) $Re(Y_{sopt})$ and $Im(Y_{sopt})$

5.3 Four-port 4T MOSFET with various Body Contact Layouts

Accurate extraction and modeling of the body network is of utmost importance in RF regime and application to CMOS RF circuit design [70-72]. However, in the conventional 2-port RF test structure used for extraction of elements of the body model, the source terminal is invariably shorted to body. Body network shunts the source junction making extraction of its capacitance difficult.

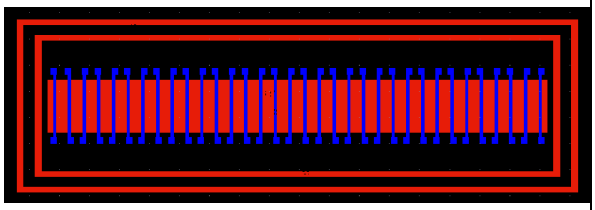
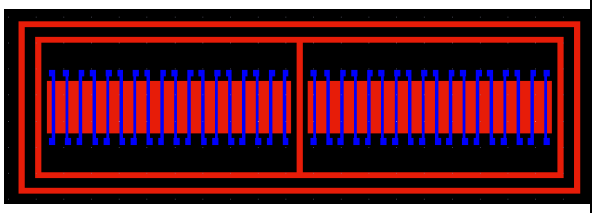
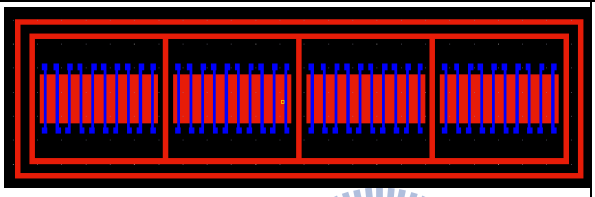
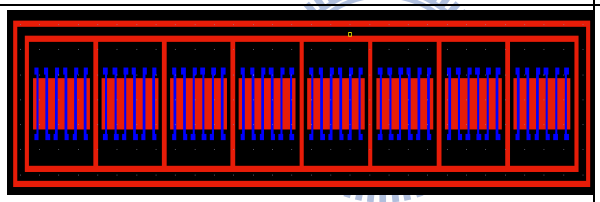
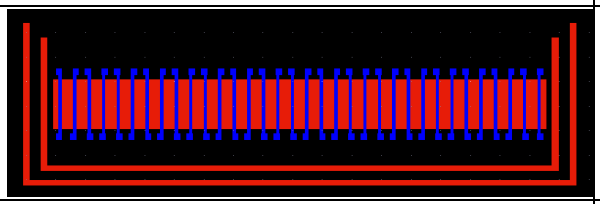
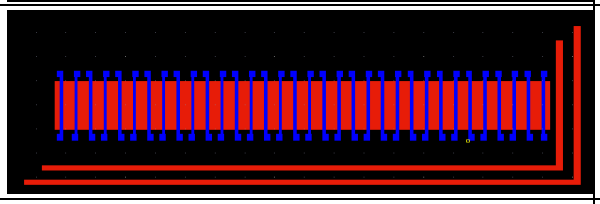
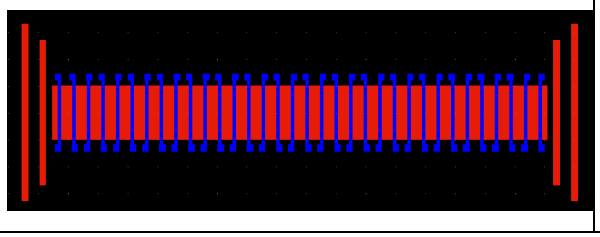
The four-port RF MOSFET test structure related to body pattern and device characterization were implemented in this section to alleviate these problems and help in accurately extracting the body networks. We proposed eight structures which include each type of body contact shape.

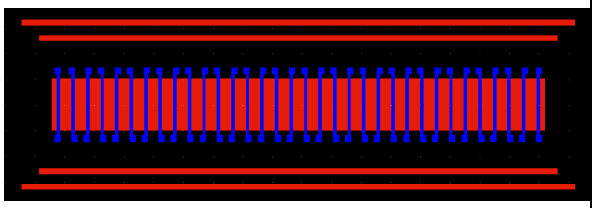
Due to the analytical expressions for body parameter modeling based on device geometry, we proposed each kind of layout with different body pickup. The conventional ring-shaped body contacts, we called it standard briefly. We usually approximate it to body coupling in vertical and horizontal directions, and it was the most common used structure in many works. The other structures include only side body contacts such as parallel and perpendicular, which means the body contacts are parallel or perpendicular to the gate fingers, and then the body coupling only in one side and less area of body contacts.

We also proposed some body shapes which are not in symmetry such as L-shape contacts and U-shape contacts which may reveal some interesting outcome of body network parameters compared with the above different types layout.

The third kind layout, we split the body contacts into multi-ring and the total width (W_{total}), finger number (N_f), and channel length (L_g) are the same as standard type(ring-shape) in this work. The detailed description and classification were shown in the following table.

Table. 5.4 The layout view and classification of each kind body contacts

Structure	Layout	Description
Standard		The ring-shape Body and N-well contact.
2-rings		Body contact split into 2 group and ring-shape N-well contact.
4-rings		Body contact split into 4 group and ring-shape N-well contact.
8-rings		Body contact split into 8 group and ring-shape N-well contact.
U shape		U-shape body and N-well contact.
L shape		L-shape body and N-well contact.
Parallel		Body and N-well contact parallel to poly-gate fingers.

Perpendicular		Body and N-well contact perpendicular to poly-gate fingers.

5.3.1 Four-port De-embedding Method

The 4-port here has the same analytical de-embedding method for 2-port. The following equation is the matrix de-embedding method for 4-port devices:

$$Y_{meas\_o} = Y_{meas} - Y_{open} \tag{5.26}$$

$$Y_{short\_o} = Y_{short} - Y_{open} \tag{5.27}$$

$$Z_{dut} = Z_{meas\_o} - Z_{short\_o} = ((Y_{meas} - Y_{open})^{-1} - (Y_{short} - Y_{open})^{-1})^{-1} \tag{5.28}$$

Thus, according to the above, the procedures of the two step de-embedding technique can be given as follows. First, we obtain the s-parameters (S_{meas}, S_{open}, and S_{short}) for DUT, open and short test structures and convert them to Y parameters (Y_{meas}, Y_{open}, and Y_{short}). Then perform the first step de-embedding by removing the parallel parasitics from both Y_{DUT} and Y_{short} according to the following equations

$$Y_{meas\_o} = Y_{meas} - Y_{open} \tag{5.29}$$

$$Y_{short\_o} = Y_{short} - Y_{open} \tag{5.30}$$

The last step of de-embedding is to perform the second de-embedding by removing the series parasitics Z_{short_o}, converting from Y_{short_o}, from Z_{meas_o}, converting from Y_{meas_o} according to the following equation

$$Z_{DUT} = Z_{meas\_o} - Z_{short\_o} \tag{5.31}$$

5.3.2 Intrinsic Device and Parasitic RLC Parameters Extraction and Analysis – Body Contact Layout Effect and Bias dependence

As shown in the table, A 2*32um device with standard body layout is selected as a reference device, and the C_{js} of multi-ring body devices have linearly increased with body ring number.

The extracted capacitances are plotted as a function in Fig, the C_{js} and C_{jd} have very strong frequency dependence at higher frequency. This is due to when the frequency increase, the term $\omega^2(C_{jd}+C_{js})^2R_{bb}^2$ in the denominator of Y_{43} and Y_{43} become significant. The high frequency effect of $Im(Y_{43})$ and $Im(Y_{34})$ starts to be dominant, and therefore the higher R_{bb} may cause the C_{jd} and C_{js} lowering. The Fig. shows the source and drain junction capacitance C_{js} and Cjd of standard, U shape, and L shape body contact layout. It is easily found that for the Cjs and Cjd without applied R_{bb} calibration, the capacitance at frequency higher than 500 Mhz dramatically drops because of the term $\omega^2(C_{jd}+C_{js})^2R_{bb}^2$ get higher. The phenomenon is more serious for the U shape and L shape body contact, which have less body contact number and hence larger R_{bb}.

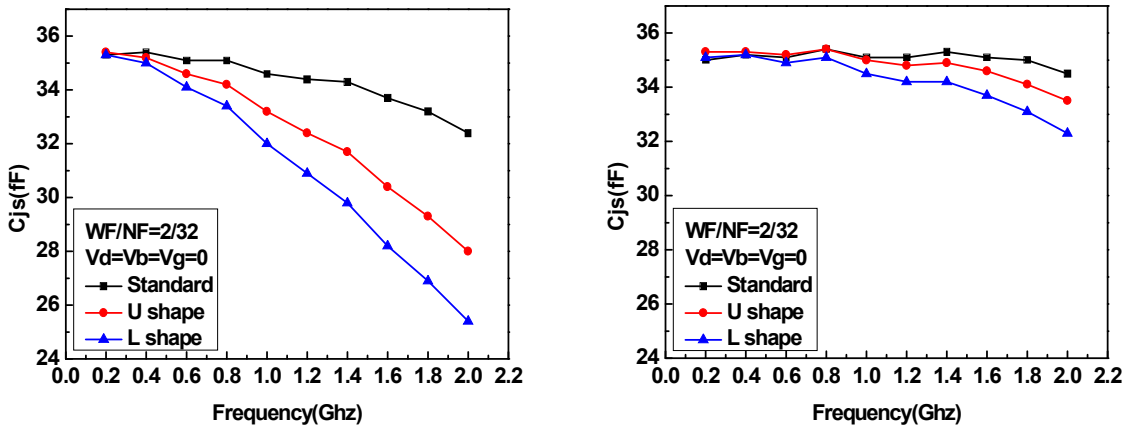


Fig. 5.18 C_{js} extraction versus frequency (a) without R_{bb} calibration and (b) after R_{bb}

calibration

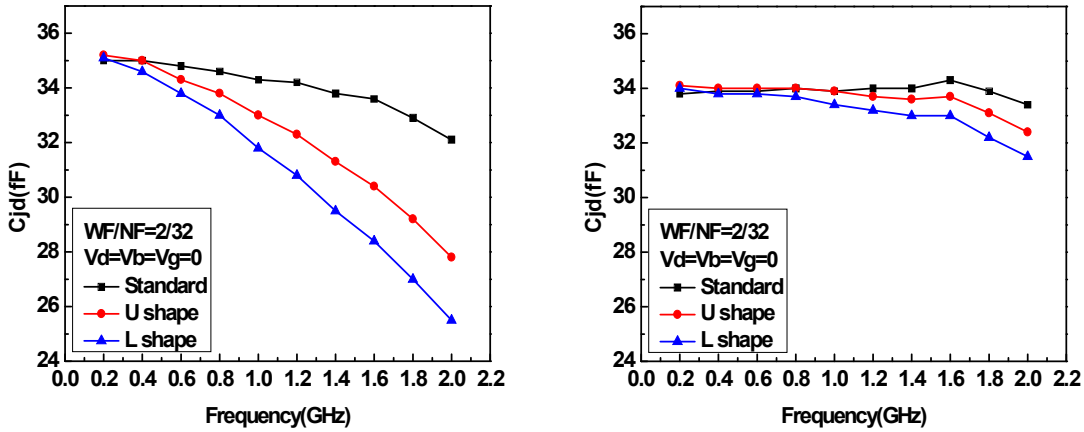


Fig. 5.19 C_{jd} extraction versus frequency (a) without R_{bb} calibration and (b) after R_{bb} calibration

Table. 5.5 C_{js} extraction versus V_{gb} extraction

Vbs(V)	Cjs(standard)	Cjs(2-rings)	Cjs(4-rings)	Cjs(8-rings)
-0.6	2.67E-14	2.88E-14	3.32E-14	4.21E-14
0	3.50E-14	3.80E-14	4.34E-14	5.48E-14
0.6	9.50E-14	1.05E-13	1.24E-13	1.63E-13

Table. 5.6 The increment of C_{js} with the body ring number increased

Vbs(V)	Capacitance of 2-rings increament(%)	Capacitance of 4-rings increament(%)	Capacitance of 8-rings increament(%)
-0.6	7.87	16.5	33.3
0	8.57	15.4	32.57
0.6	10.5	20	41.05
Equation	(Cjs_R2- Cjs_std)/ Cjs_std	(Cjs_R4- Cjs_R2)/ Cjs_std	(Cjs_R8- Cjs_R4)/ Cjs_std

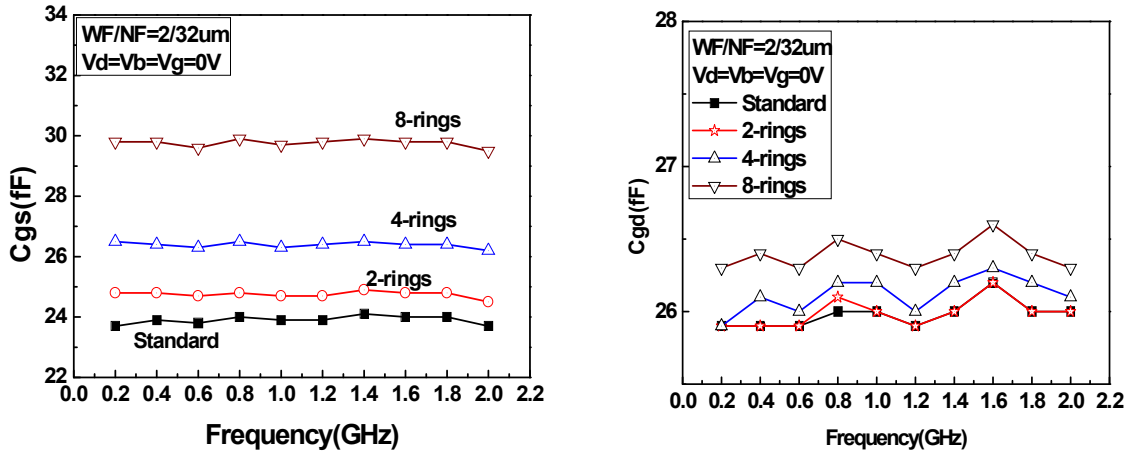


Fig. 5.20 C_{gs} and C_{gd} extraction versus frequency of ring type body contact layout

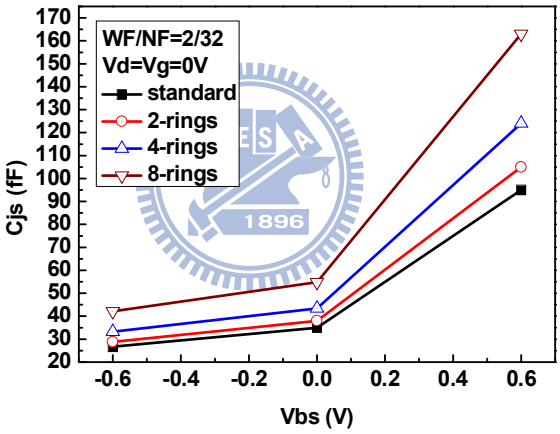


Fig. 5.21 C_{js} extraction versus frequency of ring type body contact layout

Fig. shows the measured C_{sdb} and C_{bsd} as a function of body bias V_{bs} for 1 ring and 8-Rings body contact layout devices. For a zero gate bias, the measured capacitance decreases with increasing junction reverse bias (V_{bs}< 0) due to the increase of the depletion widths of the individual junction components. It can also be seen that at V_{gb}= 0V (V_{gs}= V_{bs}= 0), corresponding to a depleted surface in the transistor channel. As the channel goes into accumulation (V_{gb}< 0V), the inner sidewall of source and drain junction comes into existence and the measured capacitance increases due to reduction of the depletion width of the

sidewall junction with increasing degree of accumulation.

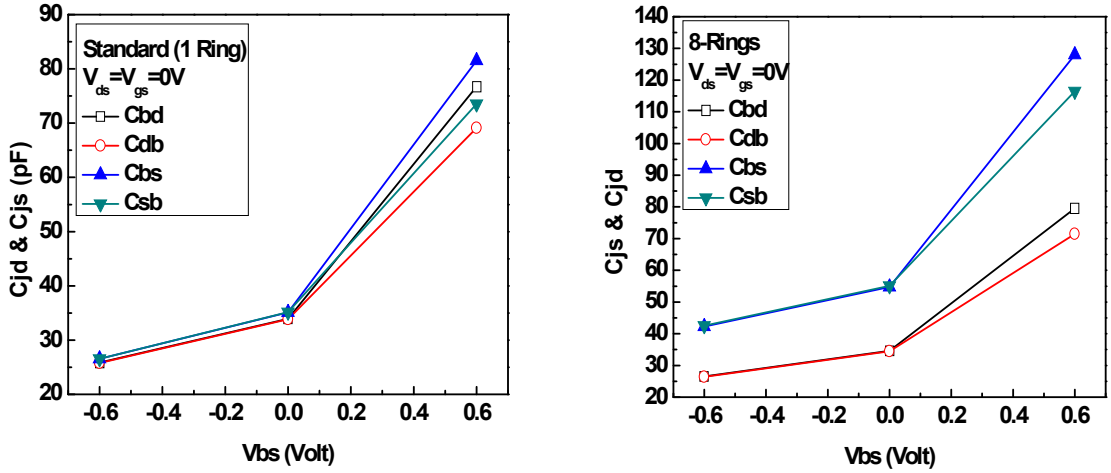


Fig. 5.22 C_{jd} (C_{bd} and C_{db}) and C_{jd} (C_{bs} and C_{sb}) extraction versus body bias (a) standard and (b) 8-Rings

It's obviously that C_{js} was larger than C_{jd}, this phenomenon especially occurs in multi-ring body contact structures (2-Rings, 4-Rings, and 8-Rings). It's due to the more body contact we split, the more source side transmission line distance that may impact the total capacitance of source side capacitances as shown in **Fig. 5.22**.

These are mainly inter connect coupling capacitances introduced by both metal-to-metal and metal-to-contact interconnect.

The body resistance shows weak bias dependence, but strongly dependence on device geometry such as number of fingers and length of body contacts. In these case, 1/ R_{bb} scales with body contact length which is due to the length of resistive path.

Table. 5.7 Body resistance and deep-N-well capacitance with different body shapes

Body layout	Rbb (Ohm)	Cdwn (fF)
Parallel	**1415.48**	**105.6**
L shape	**691.86**	**132.8**
U shape	**600.97**	**140.9**
Perpendicular	**409.15**	**151.8**
standard	**384.21**	**155.4**
2-Rings	**310.25**	**166.6**
4-Rings	**212.87**	**187.7**
8Rings	**181.27**	**231.2**

The extracted body resistance and capacitance values from Fig. can be viewed in Table. The body resistance and capacitance values have been obtained after applying a slight optimization so that the best match for the simulated and measured Y-parameters can be achieved. The body resistance and capacitance values have been obtained after applying a slight optimization so that the best match for the simulated and measured Y-parameters can be achieved.

Chapter 6

Conclusions

6.1 Summary

The potential impact from layout dependent STI stress on LFN and high frequency performance has been investigated on multi-finger MOSFETs with various layouts, such as narrow-OD and multi-OD. The monotonic decrease of G_m with finger width (W_F) scaling in narrow-OD NMOS proves μ_{eff} degradation from the compressive STI stress along transverse direction ($\sigma_\perp$). However, the multi-OD NMOS reveal an abnormal G_m increase for extremely narrow OD width to W_{OD}=0.125μm. The observed results suggest that STI stress is not the only mechanism governing the electrical property in miniaturized devices. STI TCR induced ΔW is identified as another key factor, which may overcome STI stress effect in determining channel current and G_m. Semi-empirical formulas have been derived to successfully predict W_{OD} scaling effect on μ_{eff} and G_m. Taking this method, ΔW can be precisely extracted based on a simultaneous best fitting to μ_{eff} and G_m and the resulted increase of effective width (W_{eff}) is dramatically large to around 34% for OD16 with W_{OD}=0.125μm. The larger W_{eff} becomes the major contributor to reducing LFN and overcome N_{it} effect in narrow-OD and multi-OD devices with sufficiently small W_{OD}. The reduction of LFN with OD width scaling is the other evidence reflecting STI TCR induced ΔW effect.

Unfortunately, the OD width scaling leads to a negative impact on high frequency performance like f_T and f_{MAX}, due to G_m degradation and undesired increase of C_{gg}. An improved open deembedding method can reduce the parasitic capacitances from inter-metal coupling but cannot eliminate gate related fringing capacitances. The multi-finger MOSFETs with miniaturized OD width cannot prevent from f_T degradation. The trade-off between LFN and high frequency performance identified from this research work provides an important layout guideline for analog and RF circuit design.

The proposed donut MOSFETs demonstrate the advantages over the standard multi-finger MOSFETs, such as the lowest S_{ID}/I_{DS}^2 in low frequency domain (1~ 10K Hz) and higher f_T in very high frequency region (100/50 GHz for N/P MOS). The elimination of STI stress and excess traps along the channel width is validated as the primary mechanism responsible for the enhancement of μ_{eff} as well as f_T, and reduction of LFN. The layout dependent stress mechanism can be applied to both NMOS and PMOS, even though their LFN are governed by different models. An innovative donut device layout for solving the potential degradation of f_{MAX} and NF_{min} emerges as an interesting and important topic in the future work for RF and analog applications.

The variations of body contact layout reveal significant effect on body resistance (R_{bb}), junction capacitances (C_{js} and C_{jd}), and gate capacitances (C_{gs}, C_{gd}, and C_{gg}). Multi-ring body contact layout offers the advantage of lower R_{bb}, but pays the penalty of larger junction capacitances and gate capacitance. The drawback of increased gate capacitances leads to degradation of high frequency performance like f_T. For 8-ring body contacts, the measured f_T is degraded by around 12.6%, compared to the standard multi-finger MOSFET with the same channel width (N_FxW_F). However, the fluence on f_{MAX} from multi-ring body contact becomes a benefit instead of penalty. f_{MAX} extracted from U-Gain method indicates around 9.84% and 8.41% improvement from 4-ring and 8-ring compared to the standard multi-finger MOSFET. The mechanism responsible for f_{MAX} improvement even under the condition of f_T degradation introduces another interesting topic in future work.

6.2 Future Work

The proposed donut MOSFETs can provide the advantages over the standard multi-finger MOSFETs, such as the lower LFN for VCO or down-conversion mixer design. Even though the improvement of μ_{eff} and G_m can gain another benefit such as higher f_T, the

significant increase of gate resistance R_g indeed leads to degradation of f_{MAX} and RF noise, which are key parameters determining LNA and PA performance. An improved donut MOSEFT layout for solving the potential degradation of f_{MAX} and NF_{min} emerges as an interesting and important topic in the future works for RF and analog applications.

The second interesting topic worthy of continuous research effort is to explore the mechanism responsible for f_{MAX} improvement from multi-ring body contacts layout even with the penalty of f_T degradation. The reduction of body resistance R_{bb} is considered one of key factors but the underlying mechanism is not truly understood. Furthermore, the impact from R_{bb} in single MOSFET on amplifiers like PA or LNA with cascade or cascade topology appears as another interesting subject in the future work.

The third interesting topic is the OD/STI density effect on STI stress and TCR profile, and their impact on μ_{eff}, G_m, I_{DS}, LFN, and high frequency performance (f_T, f_{MAX} , and NF_{min}). Extremely narrow MOSFET with single gate-finger and multi-OD and multiple gate-finger and single narrow OD will be designed to investigate the mechanism.

The last one interesting topic to remark in the future work is the gate/channel orientation dependence of mobility modulation from STI stress. Single gate-finger MOSFETs with x- and y- gate/channel orientations and different channel width (W_{OD}) will be implemented to explore the truth. Also, the results can facilitate analysis and modeling of donut MOSFETs with both x and y directions in the gate/channel orientation.

Bibliography

[1] K. Lee, I. Nam, et al. "The impact of semiconductor technology scaling on CMOS RF and digital circuits for wireless application," IEEE Transactions on Electron Devices, vol.52, no.7, pp. 1415-1422, 2005

[2] D. P. Tsarapkin, "Treatment of Flicker Noise in Oscillators and Methods of Its Suppression," in 2006 IEEE International Frequency Control Symposium and Exposition, 2006, pp. 452-456

[3] M. Miyamoto, H. Ohta, et al. "Impact of reducing STI-induced stress on layout dependence of MOSFET characteristics," IEEE Transactions on Electron Devices, vol.51, no.3, pp.440-443, 2004

[4] International Technology Roadmap for Semiconductors (ITRS), 2007 edition, http://www.itrs.net

[5] A. Hajimiri and T. H. Lee, "A general theory of phase noise in electrical oscillators," IEEE J. Solid-State Circuits, vol.33, pp.179-194, 1998

[6] T. H. Lee and A. Hajimiri, " Oscillator phase noise : a tutorial," IEEE J. Solid-State Circuits, vol.35, pp.326-336, 2000

[7] M. Sanden, F. Jonsson, O Marinov, M. J. Deen, and M. Ostling, " Up-conversion of 1/f noise to phase noise in voltage controlled oscillators," in Proc. 16th Int. Conf. Noise and Fluctuation (ICNF), 2001, pp.499-502

[8] F. Jonsson, M. von Maartman, M. Sanden, M. Ostling, " A voltage controlled oscillator with automatic amplitude control in SiGe technology," in Proc. 19th NORCHIP, 2001, pp.28-33

[9] S. Zhou and M.-C. F. Chung, " A CMOS passive mixer with low flicker noise for low power direct conversion receiver," IEEE J. Solid-State Circuits, vol.40, pp.1084-1093,

2005

[10] H. Darabi and A. A. Adibi, " Noise in RF-CMOS mixers : a simple physical model," IEEE J. Solid-State Circuits, vol.35, pp.15-25, 2000

[11] A. L. McWhorter, Semiconductor surface physics, University Pennsylvania, press, 1957

[12] E. Simoen and C. Claeys, " On the flicker noise in submicron Si MOSFETs," Solid-State Electronics, vol.43, pp.865-882, 1992

[13] G. Ghibaudo and T. Boutchacha, " Electrical noise and RTS fluctuations in advanced CMOS devices," Microelectron. Reliab., vol.42, pp.573-582, 2002

[14] J. Chang, A. A. Abidi, and C. R. Viswanthan, "Flicker noise in CMOS transistors from subthreshold to strong inversion at various temperatures," IEEE Trans. Electron Devices, vol. 41, pp.1965-1971, 1994

[15] L. K. J. Vanadamme, X. Li, and D. Rigaud, " 1/f noise CMOS devices, mobility or number fluctuations ? ," IEEE Trans. Electron Devices, vol. 41, pp.1936-1945, 1994

[16] B. Cretu, M. Fadlallah, G. Ghibaudo, J. Jomaah, F. Balestra, and G. Guegan, " Thorough characterization of deep submicron surface and buried channel PMOSFETs," Solid-State Electronics, vol.53, pp.971-975, 2002

[17] M. Fadlallah, G. Ghibaudo, J. Jomaah, M. Zoaeter, and G. Guegan, "Static and low frequency characterization of surface- and buried-mode 0.1μm P and NMOSFETs," CMOS devices," Microelectron. Reliab., vol.42, pp.41-46, 2002

[18] R. A. Wilcox, J. Chang, and C. R. Viswanathan, "Low-temperature characterization of buried-channel NMOST," IEEE Trans. Electron Devices, vol. 36, pp.1440-1447, 1989

[19] X. Li, C. Barros, E. P. Vandamme, and L. K. J. Vanadamme, "Parameter extraction and 1/f noise in a surface and a bulk-type p-channel LDD MOSFETs," Solid-State Electronics, vol.37, pp.1853-1862, 1994

[20] L. K. J. Vanadamme, " Bulk and surface 1/f noise," IEEE Trans. Electron Devices, vol.

36, pp.987-992, 1989

[21] A. van der Ziel, “ Unified presentation of 1/f noise in electronic devices : fundamental 1/f noise source,” IEEE Proc., vol.76, pp.233-258, 1988

[22] F. N. Hooge and L. K. J. Vanadamme, “ Lattice scattering causes 1/f noise,” Phys. Lett. A, vol.66, pp.315-316, 1978

[23] F. N. Hooge, “ 1/f noise is no surface effect,” Phys. Lett. A, vol.29a, pp.139-140, 1969

[24] A. Chatterjee et al., "A Shallow Trench Isolation Study for 0.25/0.18um CMOS Technologies and Beyond," in 1996 VLSI Tech. Symp. Digest, pp.156-157

[25] C. Chen et al., "A Novel 0.25um Shallow Trench Isolation Technology," in 1996 IEDM Tech. Digest, pp.837-840

[26] M. Nandakumar, A. Chatterjee, S. Sridhar, K. Joyner, M. Rodder, and. I.-C. Chen, “ Shallow Trench Isolation for advanced ULSI CMOS Technologies,” in 1998 IEDM Tech. Digest, pp. 133-136

[27] R. A. Bianchi, G. Bouche, and O. Roux-di-Buisson, "Accurate Modeling of Trench Isolation-Induced Mechanical Stress Effects on MOSFET Electrical Performance," in 2002 IEDM Tech. Digest, pp. 117-120.

[28] BSIM4.6.1 MOSFET Model User's Manual, www.device.eecs.berkeley.edu/~bsim3/bsim4.html

[29] Y. Luo and D. K. Nayak, “Enhancement of CMOS performance by process-induced stress” IEEE Trans. Semicond. Manuf., vol. 18, no. 1, pp. 63–68, Feb. 2005.

[30] P. L’opez, M. Oberst, H. Neubauer, D. Cabello and J. Hauer, “Performance analysis of high-speed MOS transistors with different layout styles”, IEEE Int. Symposium on Circuits and Systems, Kobe, Japan, 23-25 May, 2005.

[31] M. Miyamoto, H. Ohta, Y. Kumagai, Y. Sonobe, K. Ishibashi and Y. Tainaka, “Impact of reducing STI-induced stress on layout dependence of MOSFET characteristics,” *IEEE*

Trans. Electron Devices, vol. 51, 2004, pp.440-443.

[32] A. Ono, R. Ueno, and I. Sakai, "TED Control Technology for Suppression of Reverse Narrow Channel Effect in 0.1 um MOS Devices," in *IEDM Tech. Digest*, 1997, pp.227-230.

[33] P. B. Y. Tan, A. V. Kordesch, and O. Sidek, "Analysis of deep submicron CMOS transistor Vtlin and Idsat versus channel width," in *Proc. Asia-Pacific Microw. Conf.*, 2005, pp. 1569-1572.

[34] G. Scott, J. Lutze, M. Rubin, F. Nouri, and M. Manley, "NMOS drive current reduction caused by transistor layout and trench isolation induced stress," in *IEDM Tech. Digest*, 1999, pp.827-830.

[35] R. Li, L. Yu, H. Xin, Y. Dong, K. Tao and C. Wang, "A comprehensive study of reducing the STI mechanical stress effect on channel-width-dependent Idsat", *Semicond. Sci. Technol.,* vol. 22, pp.1292-1297, 2007.

[36] M. Nandakumar, A. Chatterjee, et al., "Shallow Trench Isolation for advanced ULSI CMOS Technologies," in 1998 IEDM Tech. Digest, pp. 133-136

[37] M. Nandakumar, S. Sridhar, et al., "A Shallow Trench Isolation for Sub-0.13um CMOS Technologies," in 1997 IEDM Tech. Digest, pp.657-660

[38] H. Watanabe, K. Shimizu, et al.,"Corner-Rounded Shallow Trench Isolation Technology to Reduce the Stress-Induced Tunnel Oxide Leakage Current for Highly Reliable Flash Memories," in 1996 IEDM Tech. Digest, pp.833-836

[39] A. Chatterjee, D. Rogers, et al., "A Shallow Trench Isolation using LOCOS Edge for Preventing Corner Effects for 0.25/0.18 μm CMOS Technologies and Beyond," in 1996 IEDM Tech. Digest, pp.829-832

[40] C. P. Chang, C. S. Pai, et al., "A Highly Manufacturable Corner Rounding Solution for 0.18 μm Shallow Trench Isolation," in 1997 IEDM Tech. Digest, pp.661-664

[41] S. Matsuda, T. Sato, et al., “Novel corner rounding process for shallow trench isolation utilizing,” MSTS (Micro-Structure Transformation of Silicon),” in 1998 IEDM Tech. Digest, pp. 137-140

[42] Jyh-Chyurn Guo and Chih-Ting Yeh “A New Three Dimensional Capacitor Model for Accurate Simulation of Parasitic Capacitance in Nanoscale MOSFETs,” *IEEE Trans. on Eelectron Devices*, vol.56, no.8, pp.1598-1607, Aug. 2009

[43] S.-I. 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 Concentration”, *IEEE Trans. Electron Devices*, vol. ED-41, no. 12, pp. 2357-2362, Dec. 1994.

[44] Gaitan, M., E. W. Enlow, et al. "Accuracy of the charge pumping technique for small geometry MOSFETs," *IEEE Trans. Nuclear Science*, vol.36, no.6, pp.1990-1997, 1989.

[45] D. Schmitt and G. Dorda, "Interface states in MOSFETs due to hot-electron injection determined by the charge pumping technique," *Electronics Letters*, vol. 17, no.20, pp.761-763, 1981

[46] H. Chien-Chihand J.-C Guo, et al., "0.13μm RF CMOS and varactors performance optimization by multiple gate layouts," Electron Devices, IEEE Transactions on, vol. 51, pp. 2181-2185, 2004.

[47] K. L. Wang, "A new method to electrically determine effective MOSFET channel width," Electron Devices, IEEE Transactions on, vol. 29, pp. 1825-1827, 1982.

[48] H. S. Momose, et al., "A study of flicker noise in n- and p-MOSFETs with ultra-thin gate oxide in the direct-tunneling regime," in Electron Devices Meeting, 1998. IEDM '98 Technical Digest., International, 1998, pp. 923-926.

[49] R. P. Jindal and A. van der Ziel, "Model for mobility fluctuation 1/f noise," Applied Physics Letters, vol. 38, pp. 290-291, 1981.

[50] S. Maeda, et al., "Impact of mechanical stress engineering on flicker noise

characteristics," in VLSI Technology, 2004. Digest of Technical Papers. 2004 Symposium on, 2004, pp. 102-103.

[51] L. Hyeokjae, et al., "Low-frequency noise degradation caused by STI interface effects in SOI-MOSFETs," Electron Device Letters, IEEE, vol. 22, pp. 449-451, 2001.

[52] P. Bai, *et al.*, "A 65nm Logic Technology Featuring 35nm Gate Lengths, Enhanced Channel Strain, 8 Cu Interconnect Layers, Low-k ILD and 0.57μm^2 SRAM Cells," in IEDM Tech. Digest, Dec. 2004, pp. 657-660

[53] S. Decoutere, et. al., "Technologies for sub-45nm Analog/RF CMOS – Circuit Design Opportunities and Challenges," in IEEE CiCC Proc., 2006, pp. 679-686

[54] H. Li, *et al.*, "Technology Scaling and Device Design for 350GHz RF Performance in 45nm Bulk CMOS Process," in VLSI Tech. Symp. Proc., June 2007, pp. 56-57

[55] T. Ohguro, *et al.*,"The Impact of Oxynitride Process, Deuterium Annealing and STI Stress to 1/f Noise of 0.11mm CMOS," in VLSI Tech. Symp. Proc., June 2003, pp.37-38

[56] S. Maeda, *et al.*, "Impact of Mechanical Stress Engineering on Flicker Noise Characteristics," in VLSI Tech. Symp. Proc., June 2004, pp.102-103

[57] T. Ueno, *et al.*, " Improved 1/f Noise Characteristics in Locally Strained Si CMOS using Hydrogen-Controlled Stress Liners and Embedded SiGe," in VLSI Tech. Symp. Proc., June 2006, pp.104-105

[58] F. Assaderaghi, *et al.*,"A Dynamic Threshold Voltage MOSFET (DTMOS) for Ultra-Low Voltage Operation," in IEDM Tech. Digest, Dec. 1994, pp. 809-812

[59] T. Sakurai**,** "Reducing power consumption of CMOS VLSI's through V_{DD} and VTH control," in ISQED Proc., Mar. 2000, pp. 417-423

[60] Jun-Wei Wu, *et al.*, "Pocket Implant Effect on Drain Current Flicker Noise in Analog n-MOSFET Devices," IEEE TED-51 , no.8, pp.1262 – 1266, Aug. 2004

[61] F. N. Hooge, et. al. "Lattice scattering causes 1/f noise," Phys. Lett. A, vol. 66, pp.

315-316, 1978.

[62] R. P. Jindal, et al. “Phonon fluctuation model for flicker noise in elemental semiconductors,” J. Appl. Phys. vol.52, no. 4, pp. 2884-2888, Apr. 1981.

[63] S. Ling, et al., "A method for on-wafer S-parameter measurement of a differential amplifier by using two-port network analyzer," in Microwave Conference Proceedings, 2005. APMC 2005. Asia-Pacific Conference Proceedings, 2005, p. 4 pp.

[64] G. T. Sasse, et al., "Gate-capacitance extraction from RF C-V measurements [MOS device applications]," in Solid-State Device Research conference, 2004. ESSDERC 2004. Proceeding of the 34th European, 2004, pp. 113-116.

[65] K. Myounggon, et al., "Separate Extraction of Gate Resistance Components in RF MOSFETs," Electron Devices, IEEE Transactions on, vol. 54, pp. 1459-1463, 2007.

[66] J-C Guo and Yi-Min Lin, "A new lossy substrate de-embedding method for sub-100 nm RF CMOS noise extraction and modeling," Electron Devices, IEEE Transactions on, vol. 53, pp. 339-347, 2006.

[67] Y. Cui, et al., "Modeling of anomalous frequency and bias dependences of effective gate resistance in RF CMOS," Electron Devices, IEEE Transactions on, vol. 53, pp. 2620-2626, 2006.

[68] M. Miyamoto, et al., "Impact of reducing STI-induced stress on layout dependence of MOSFET characteristics," Electron Devices, IEEE Transactions on, vol. 51, pp. 440-443, 2004.

[69] H. Fukui, "Optimal noise figure of microwave GaAs MESFET's," Electron Devices, IEEE Transactions on, vol. 26, pp. 1032-1037, 1979.

[70] N. Srirattana, et al., "A new analytical scalable substrate network model for RF MOSFETs," in Microwave Symposium Digest, 2004 IEEE MTT-S International, 2004, pp. 699-702 Vol.2.

[71] U. Mahalingam, et al., "Direct extraction of substrate network parameters for RF MOSFET modeling using a simple test structure," Electron Device Letters, IEEE, vol. 27, pp. 130-132, 2006.

[72] S. C. Rustagi, et al., "BSIM3 RF models for MOS transistors: a novel technique for substrate network extraction," in Microelectronic Test Structures, 2003. International Conference on, 2003, pp. 118-123.