TECHNICAL UNIVERSITY OF CRETE
ELECTRONIC & COMPUTER ENGINEERING DEPARTMENT
MICROELECTRONICS DIVISION
2009-2010

Investigation of Non-linearity in MOS transistors

By:
Andreas A. Bompos
abobos@isc.tuc.gr

Submitted on November 11th, 2010 in partial fulfillment of the requirements for the Electronic and Computer Engineering diploma degree.

Advisor: Assistant Professor Matthias Bucher
Co-advisor: Associate Professor Konstantinos Ballas
Co-advisor: Professor Konstantinos Kalaitzakis

Abstract

Since its introduction in the 1960's MOS transistor technology has been subject to a ceaseless progress in performance and decrease in overall transistor dimensions becoming the key component of today's integrated circuits (ICs). Meanwhile, the boost of wireless applications and communication technology in general, has created mass consumer markets for Radio-Frequency (RF) electronics. Applications such as WLAN, Bluetooth, 3G/4G networks and portable communication devices are fueling the demand for RF CMOS.

However, as the CMOS dimension is scaled to the nanometer regime ($< 100nm$) many difficulties arise. For example, in RF circuits phenomena which reduce the linearity are of great importance. Contemporary MOSFET models usually struggle to predict the non-linearities mainly because of an inaccurate description of the DC behavior. Nonetheless, since an accurate description of the DC behavior is a prerequisite for a good high-frequency behavior, this thesis focuses on the DC behavior of MOS transistors. Since MOSFETs non-linear behavior can be determined by the high-order terms of the drain current with respect to any MOSFET terminal, our analysis is focused mainly on gate, source and drain induced non-linearities of a MOSFET operating in saturation. The normalized transconductances and high-order derivatives of the drain current are calculated directly from on-wafer measurements of a CMOS 110nm technology and are presented as a function of the Inversion Coefficient for a more detailed analysis. Furthermore, we calculate the Intrinsic Voltage Gain (AV_i) and Voltage Third Intercept Point (VIP_3); a linearity Figure of Merit (FOM) which can be calculated directly from the DC measurements. Additionally, we present and analyze a simple Linearization technique, the so-called Source Degeneration. Finally, we present the evaluation results of the EKV3 MOSFET model; a compact MOSFET model designed for computational efficiency and ease of parameter extraction.

Acknowledgements

I would like to take some time here and thank people who made the completion of this work possible.

First of all, my parents, who supported me morally and financially throughout all my undergraduate studies. Additionally, I am deeply grateful to my advisor Assistant Professor Matthias Bucher, who has oriented and supported me with promptness and have always been patient and encouraging in times of difficulties.

I am also indebted to the PhD Candidate, Antonopoulos Aggelos and postgraduate students, Mavredakis Nikolaos and Chalkiadaki Marianna with whom I had the pleasure to work with. Their support has been invaluable and also helped me learn how to tackle new problems and develop techniques to solve them.

List of symbols

In the following table we present a list of symbols that were used in this thesis. However, symbols which were only used locally, are not listed. In the following subscripts i and j denote one of the MOSFET terminals: drain(D), gate (G), source (S) or bulk (B).

Table 1: Symbols

Symbol	Definition	Unit
AV_i	Intrinsic Voltage Gain	V/V
C_{ox}	Gate oxide capacitance per unit area, $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$	F/m^2
D_2, D_3	2^{nd} and 3^{rd}-order harmonic distortion	-
E_x	Transversal electric field at surface	V/m
G_x	Normalized transconductances, $G_x = \frac{g_{mx}U_T}{I_D}$, where x= G, S, D, B	-
g_m	Gate transconductance, $g_m = \frac{\partial I_D}{\partial V_G}\|V_S, V_D$	A/V
g_{ds}	Output conductance, $g_{ds} = \frac{\partial I_D}{\partial V_D}\|V_S, V_G$	A/V
g_{ms}	Source transconductance, $g_{ms} = \frac{\partial I_D}{\partial V_S}\|V_B, V_G$	A/V
I_D	Drain current	A
I_S	Specific current	A
i_f, i_r	Normalized forward and reverse currents	A
IC	Inversion Coefficient	-
L	Effective channel length	m
L_{mask}	Mask channel length	m
n	Slope factor	-
R_S	Source Resistance	Ω
THD	Total Harmonic Distortion	-
t_{ox}	Gate oxide thickness	m
U_T	Thermal voltage	V
V_{TH}	Threshold voltage	V
V_P	Pinch-off voltage	V
VIP_3	Voltage Third Intercept Point	V
W	Effective channel width	m
W_{mask}	Mask channel width	m
γ	Substrate factor, $\gamma = \sqrt{2q\epsilon_{si}N_{sub}} \cdot C'_{ox}$	$V^{\frac{1}{2}}$
ϵ_{si},ϵ_{ox}	Dielectric permittivity of Silicon, Silicon-oxide	F/m
μ	Carrier mobility	m^2/V_S
Ψ_0	Bulk Fermi potential	V

Contents

List of Figures

Chapter 1

Introduction

1.1 General Introduction

Since its introduction in the 1960s MOS transistor technology has been subject to a ceaseless progress in performance and decrease in overall transistor dimensions becoming the key component of today's integrated circuits (ICs). The continuous scaling of CMOS transistors has led to IC's with higher speed, lower power dissipation and higher packing density. The technological advances justify that the CMOS evolution is governed by the Moore's law, which states that the transistor density on integrated circuits doubles every couple of years.

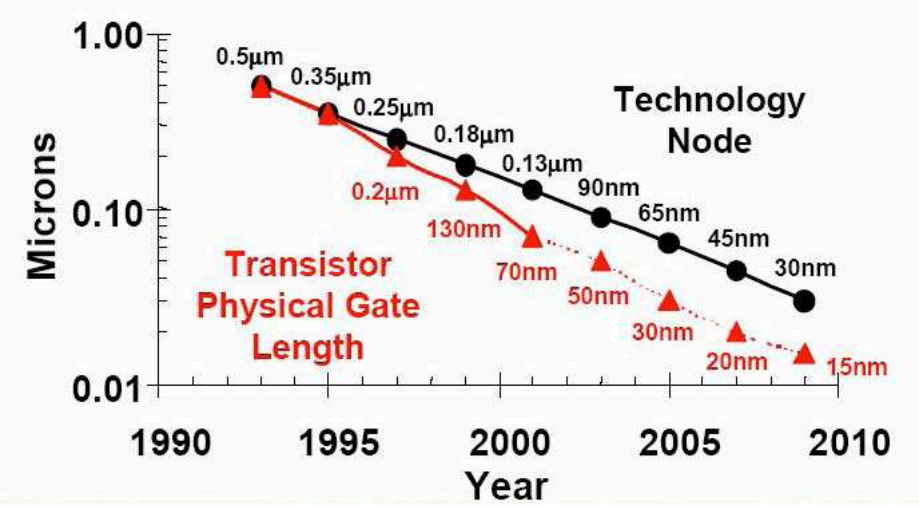


Figure 1.1.1: Trend of transistor gate length throughout the years.

Meanwhile, the boost of wireless applications increased the interest for the use of CMOS in Radio Frequency (RF) circuits. RF CMOS can be a cost-efficient solution for integration of digital logic circuits with analog and RF functions. However, as the CMOS dimension is scaled to the nanometer regime ($< 100nm$) (Fig. 1.1.1) many difficulties arise.

The production of MOSFETs with channel lengths much smaller than a micrometer is a challenge, and the difficulties of semiconductor device fabrication are always a limiting factor in advancing IC technology. Furthermore, in recent years, the small size of the MOSFET, below a few tens of nanometers, has also created critical operational problems (i.e increased gate-oxide

leakage, lower transconductance, heat production etc.). Taking into account that the transistors are the key components of today's computers and communication systems, we are obliged to study more carefully phenomena which decrease their overall performance and may cause further problems depending on the application. For example, in RF circuits phenomena which reduce the linearity are of great importance. A circuit designer needs to fully understand what design parameters affect the linearity and how, in order to be able to design high performance RF building blocks.

1.2 Thesis Outline

As becomes clear from the previous section, the investigation of the non-linear behavior in modern CMOS technologies, is of great importance since MOSFETs are the key components in most communication applications where linearity is a key factor. Contemporary MOSFETs models usually struggle to predict the non-linearities mainly because of an inaccurate description of the DC behavior. Nonetheless, since an accurate description of the DC behavior is a prerequisite for a good high-frequency behavior, this thesis will focus on the DC behavior of MOS transistors. The investigation and modeling of the non-linear sources of the MOSFET's AC behavior is reserved for future work.

In order to be able to investigate the sources which are mainly responsible for the non-linear behavior, an experimental setup has been developed and is described in Appendix. The experimental results of this setup are used throughout this thesis.

The following Chapter gives a general review of the MOSFET device and operation principles. Chapter 3 is an overview of the Linearity theory mainly described by signal and systems relations. A general distortion analysis is attempted by defining the terms of harmonic distortion and intermodulation. Furthermore, two representative measures of linearity, i.e the total harmonic distortion (THD) and the third intercept point (IP_3) are defined. We also express some considerations about the output-conductance contribution to the non-linear behavior of the MOSFET. Finally, we present indicative linearization techniques used in many applications and guarantee high levels of linearity. In Chapter 4, the theoretical model on which our calculations and assumptions were based is described, followed by the measurement results. A critical discussion is attempted based on the second-order effects and by a circuit design perspective. Chapter 5 is an overview of the main characteristics and advantages of the EKV MOSFET model over other compact MOSFET models, followed by the evaluation results of our measured data and a discussion concerning the second-order effects. Finally, Chapter 6 summarizes the results that were presented throughout this thesis followed by a number of future work ideas.

Chapter 2

The MOSFET device

As we mentioned in the previous chapter the evolution of the Integrated Circuits (ICs) is major. The basic elements of these circuits are the transistors and not passive elements such as resistors, capacitors etc. The major difference between these categories is that the transistor's current and voltage characteristics depend on the voltage (or current) on a control terminal.

The transistors can be distinguished into two subcategories based on the different physical principles that obey; the bipolar transistors and the filed effect transistors (FET). One type of bipolar transistor exists and specifically the bipolar junction transistor (BJT) while on the other hand we have two types of FETs - the junction field effect transistor (JFET) and the metal oxide semiconductor field effect transistor (MOSFET). Each type of device has its specific field of application and usefulness. For example, the BJT remain a device that excels in some applications such as discrete circuit design, due to the very wide selection on BJT types available and because of its high transconductance and output resistance compared to MOSFETs. The BJT is also a choice for demanding analog circuits, especially for very-high-frequency applications, such as radio-frequency circuits for wireless systems. On the other hand, MOSFETs appear in more than 90% of ICs and they have many advantages on digital circuitry such as reduced power consumption and large input impedance while these advantages do not translate into supremacy in all analog circuits. The modern trend is to combine Bipolar transistors with MOSFETs in an integrated circuit by using a BiCMOS process of wafer fabrication to create circuits that take advantage of the application strengths of both types of transistor.

In the following section we provide a detailed representation of the MOSFET device, we present the commonly used symbols and finally we explain analytically the way an NMOS is constructed.

2.1 Physical Structure of MOSFET

MOSFETs have three signal terminals: gate (G), source (S) and drain (D), plus a bulk terminal (B) to which the gate, source and drain voltages are referenced. As the name metal-oxide-semiconductor (MOS) suggests, the MOS transistor consists of a semiconductor on which a thin layer of insulating oxide (SiO_2) of thickness t_{ox} is grown. The **gate** electrode which is a conducting layer is deposited on top of the oxide. The p-type doped silicon region, commonly referred to as the **bulk** or the substrate, is contacted via the bulk contact. Furthermore, on either side of the gate, two heavily doped regions of depth X_j, called the **source** and the **drain**, are formed in the substrate on either side of the gate. From Figure 2.1.1 we can observe that the source and drain regions overlap slightly with the gate. The region between the drain and

source is called the channel region, which has a length L (in the y-direction) and a width (in the z-direction). Due to the manufacturing tolerances the mask length/width differ slightly from the final gate polysilicon length/width, furthermore the lateral under diffusion of the source and drain junctions also has its effect on the actual channel length/width. As a result both L and W may differ from the actual mask lengths, L_{mask} and W_{mask}. A variety of symbols are

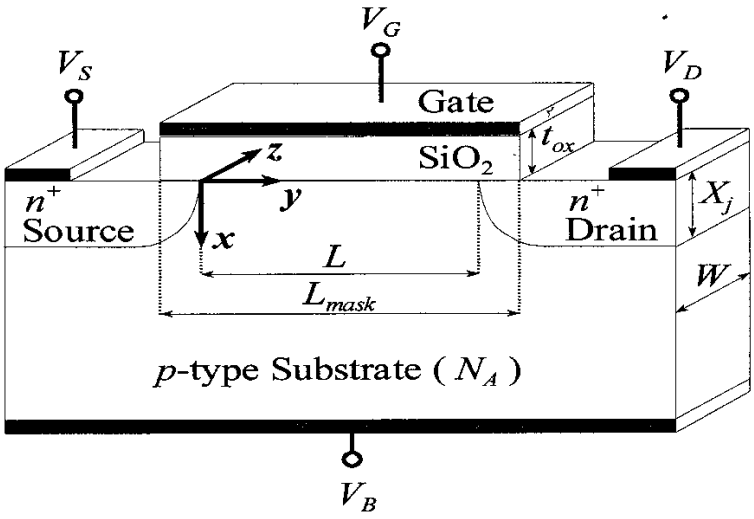


Figure 2.1.1: The basic structure of an n-type MOSFET [1].

used for the MOSFET. The basic design is generally a line for the channel with the source and drain leaving it at right angles and the bending back at right angles into the same direction. Another line is drawn parallel to the channel for the gate. The bulk connection, if shown, is connected to the back of the channel with an arrow indicating NMOS or PMOS. Arrows always point from P to N, so an NMOS (N-channel in P-substrate) has the arrow pointing from the bulk to the channel. Figure 2.1.2 depicts the commonly used symbols for MOSFETs where the bulk symbol is either labeled or implied.

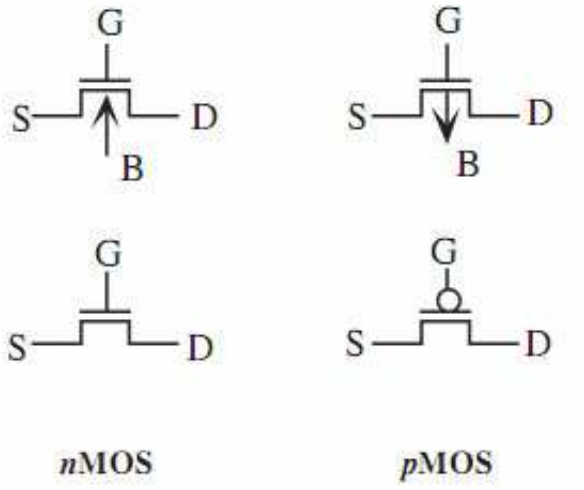


Figure 2.1.2: Symbols used at the circuit level.

2.2 MOSFET's Operation

Transistor terminals must have proper voltage polarity so as the device can operate correctly. The bulk or substrate of an NMOS transistor (PMOS) must always be connected to the lower (higher) voltage.

As shown in Fig. 2.2.1 , it is assumed that the bulk and source terminals are connected to simplify the description.

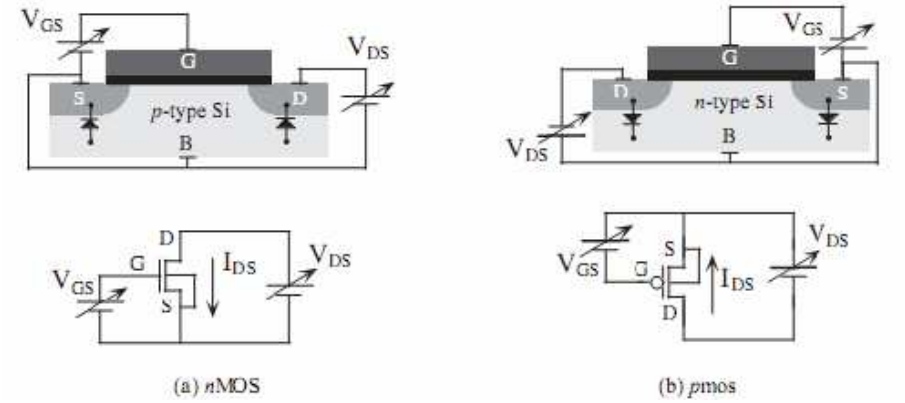


Figure 2.2.1: Normal transistor biasing (a) NMOS (b) PMOS[2].

The positive convention current in an NMOS (PMOS) device is from the drain (source) to the source (drain), and is referred to as I_{DS} or just I_D since drain and source current are equal. When a positive (negative) voltage is applied to the drain terminal, the drain current depends on the voltage applied to the gate control terminal (for PMOS transistors, V_{GS}, V_{DS} and I_{DS} are negative). If V_{GS} is zero, then an applied drain voltage reverse-biases the drain-bulk diode (Fig. 2.2.1), and there are no free charges between the drain and the source. As a result, there is no current when $V_{GS} = 0$ for NMOS devices (the same hold for PMOS devices). This is the nonconducting state of the transistor.

First of all, we will analyze the transistor's operation assuming that the source and substrate are at the same voltage. When the voltage at the gate terminal of a NMOS (PMOS) transistor is slightly increased (decreased), a vertical field starts to exist between the gate and the substrate across the oxide. In n-type (p-type) transistors, the electrons (holes) of the NMOS (PMOS) substrate close to the silicon-oxide interface initially are affected by this electrical field and move away from the interface.

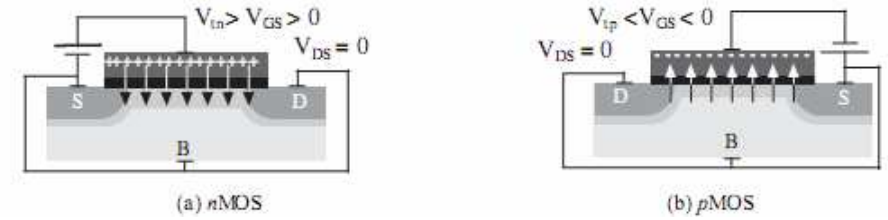


Figure 2.2.2: Depleting the NMOS (PMOS) with small positive (negative) values of V_{GS}[2].

As a result, a depletion region forms beneath the oxide interface for this small gate voltage (Fig. 2.2.2). The depletion region contains no mobile carriers, so the application of a drain voltage provides no drain current, since free carriers still do not exist in the channel.

If the gate voltage of the NMOS (PMOS) device is further increased (decreased), then the vertical electric field is strong enough to attract minority carriers (electrons in the NMOS device and holes in the PMOS device) from the bulk toward the gate. These carriers are attracted to the gate, but the silicon dioxide insulator stops them, and the electrons (holes) accumulate at the silicon-oxide interface. They form a conducting plate of mobile carriers (electrons in the p-type bulk of the NMOS device, and holes in the n-type bulk of the PMOS device). These carriers form the inversion region or conducting channel, which can be viewed as "short" circuit to the drain/source-bulk diodes. This connection is illustrated in Figure 2.2.3.

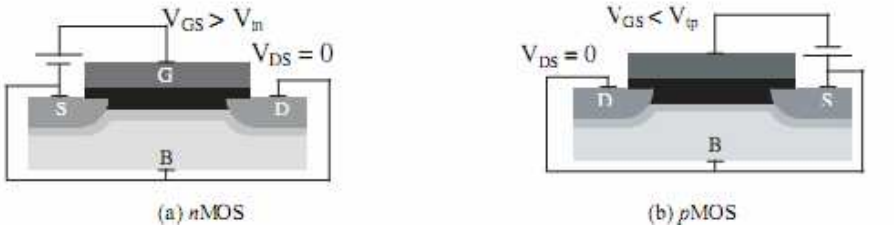


Figure 2.2.3: Creation of a conducting channel for (a) NMOS and (b) PMOS[2].

Since the drain (D) and source (S) are at the same voltage, the channel carrier distribution is uniform along the device. The gate voltage for which the conducting channels respond is an intrinsic parameter of the transistor called the threshold voltage, V_T. As an approximation, V_T can be considered constant for a given technology. The threshold voltage of a n-type transistor is positive, while for a p-type transistor is negative. Since NMOS and PMOS transistors have different threshold voltages, V_{thn} refers to the NMOS transistor and V_{thp} to the PMOS transistor.

An NMOS (PMOS) transistor has a conducting channel when the gate-source voltage is greater than (less than) the threshold voltage, i.e $V_{GS} > V_{thn}$, $V_{GS} < V_{thp}$. When the channel forms in the NMOS (PMOS) transistor, a positive (negative) drain voltage with respect to the source created a horizontal electric field, moving the channel carriers towards the drain and forming a positive (negative) drain current. If the horizontal electric field is of the same order or smaller than the vertical one, the inversion channel remains almost uniform along the device length. This happens when:

$$\begin{aligned} V_{DS} &< (V_{GS} - V_{thn}), \text{NMOS transistor} \\ V_{DS} &> (V_{GS} - V_{thn}), \text{PMOS transistor} \end{aligned} \tag{2.1}$$

The above conditions state that the vertical electric field dominates the horizontal one. The transistor is in its linear region, also called the ohmic region or non-saturated region. The MOSFET operates like a resistor, controlled by the gate voltage relative to both the source and drain voltages. If the drain voltage increases beyond the limit of Eq. 2.1, the horizontal electric field becomes stronger than the electric field at the drain end, creating an asymmetry of the channel carrier inversion distribution. The drain electric field is strong enough so that carrier inversion is not supported in this local drain region. The conducting channel retracts from

the drain, and no longer "touches" this terminal. When this happens, the inversion channel is said to be "pinched-off" and the device is in the saturation region. The pinch-off point is the location that separates the channel inversion region from the drain depletion region. It varies with changes in bias voltages. The channel distribution in this bias is shown in Figure 2.2.4 Although there are no inversion charges at the drain end of the channel, the drain region is still

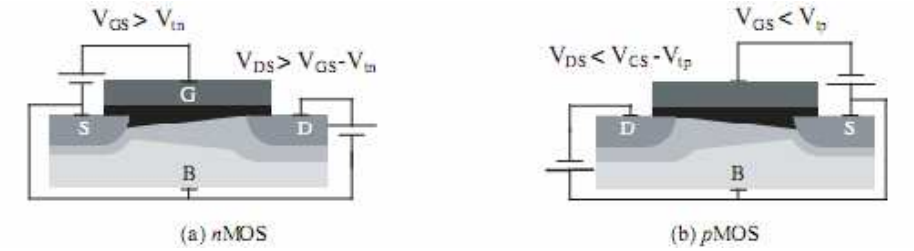


Figure 2.2.4: Channel pinch-off for (a) NMOS and (b) PMOS[2].

electrically active. Carriers depart from the source and move under the effect of the horizontal field. Once they arrive at the pinch-off point of the channel, they travel from that point to the drain, driven by the high electric field of the depletion region. CMOS ICs use all three states describe here: off-state, saturated state and the linear state.

2.3 RF CMOS

The boost of wireless applications and communication technology in general, has created mass consumer markets for Radio-Frequency (RF) electronics. Taking mobile communication as an example, we can observe that it is one of the fastest growing areas with huge market potentials. Applications such as WLAN, Bluetooth, 3G networks and portable communication devices are fueling the demand for RF CMOS.

The RF electronics were once considered the exclusive domain of III-V or silicon bipolar technologies. However, RF CMOS is capable of producing a smaller die size because of the smaller geometries used in advanced CMOS processing. The advantage in total die size is a function of the ratio of the digital-to-analog areas of the design. Designs with a large digital-to-analog ratio typically benefits from the lower geometries of RF CMOS.

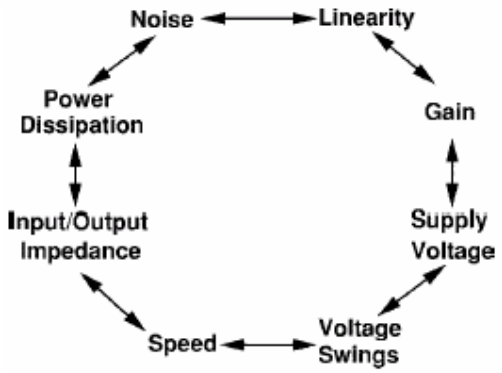


Figure 2.3.1: Analog design octagon[3].

Contrast to digital CMOS design, which is optimized and characterized for primarily one trade-off between speed and power dissipation, analog/RF circuits entail a multidimensional design space. This is illustrated in Fig. 2.3.1, where almost every two parameters trade with each other. We can observe from the figure above, **Linearity** is one of the RF figures of merit and has implications for gain and noise.

Chapter 3

MOSFET Linearity

Although many analog and RF circuits can be approximated with a linear model to obtain their response to small signals, non-linearities often lead to interesting and important phenomena[3]. In most analog building blocks, such as low noise amplifier and power amplifier, nonlinearity results in harmful effects which degrade the system performance. As a result, non-linearities need to be minimized to achieve high performance. The investigation of the circuit parameters and the correct modeling of MOSFET device physics which lead to non-linear effects is crucial.

3.1 General Distortion Theory

The transfer function of a low-frequency, continuous two-port system can be written by the following equation:

$$\begin{aligned} S_{out} &= a_0 + a_1 S_{in} + a_2 S_{in}^2 + a_3 S_{in}^3 + ... \\ &= \sum_{i=0}^{\infty} a_i S_{in}^i \end{aligned} \tag{3.1}$$

where S_{in} and S_{out} are the input and the output signal respectively. The coefficients are given by:

$$a_i = \frac{1}{i!} \frac{d^i S_{out}}{dS_{in}^i} \tag{3.2}$$

In an ideal world we would expect that the higher order terms of Eq. 3.1 would be zero, i.e the output would be the same as the input. However, this is not a plausible scenario. In reality, the higher-order terms are unavoidable causing various nonlinear phenomena [3]. The deviation from the linear behavior introduced by these higher-order terms is referred to as *distortion*.

3.1.1 Harmonics

Now that we have the full expansion of the equation which describes the output of the system we can consider the MOSFET's non-linear behavior being a memoryless time-variant system. A common method of describing the non-linearity of such a system is the specification of *harmonic distortion*[1].

Assuming v_{in} the input voltage and i_d the output-drain current, we can express the output as follows[4]:

$$i_d = a_0 + a_1 v_{in} + a_2 v_{in}^2 + a_3 v_{in}^3 + ... \tag{3.3}$$

Where a_j is a general function of time. If a sinusoid is applied to a nonlinear system, the output generally exhibits frequency components that are integer multiples of the input frequency. If $v_{in} = \hat{A}sin(\omega_0 t)$, then

$$\begin{aligned} i_d &= a_0 + a_1\hat{A}sin(\omega_0 t) + a_2\hat{A}^2 sin^2(\omega_0 t) + a_3\hat{A}^3 sin^3(\omega_0 t) \\ &= \sum_{i=0}^{\infty} a_i \hat{A}^i sin^i \omega_0 t \end{aligned} \tag{3.4}$$

In Eq. 3.4, the term with the input frequency is called the "fundamental" and the higher-order terms the "harmonics".

3.1.2 Intermodulation

While the analysis of harmonic distortion is of prime importance for the integration of analog and mixed integrated circuits using CMOS technology, certain cases require other measures of nonlinear behavior.

In bandpass filters, where higher-order harmonics are attenuated by the filter transfer function, the so-called *intermodulation* is often used to characterize filter non-linearity. When signals with

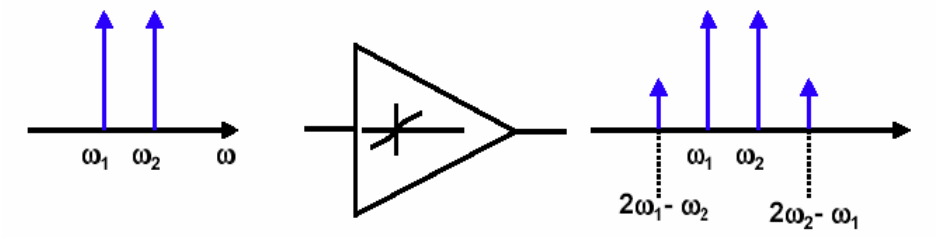


Figure 3.1.1: Intermodulation of a nonlinear system[3].

different frequencies are passing through a nonlinear system, intermodulation(IM) products are created and we get spectral components at the output that are not present in the input (Fig. 3.1.1). This phenomenon arises from multiplication of the signals when their sum is raised to a power greater than unity.

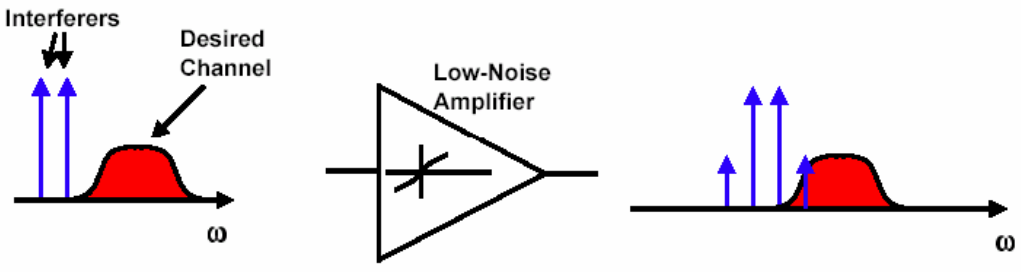


Figure 3.1.2: Corruption of a signal due to intermodulation between two interferers[3].

For example, if we have two strong interferers outside, but close to, the wanted signal band, they may create undesired frequency components in the signal band (Fig. 3.1.2). In terms of electronics, if we have the following input voltage:

$$v_{in} = \hat{A}(sin\omega_1 t + sin\omega_2 t) \tag{3.5}$$

and the fundamental frequencies ω_1 and ω_2 are not too far apart, the resulting higher-order harmonics $|\omega_1 - \omega_2|$, $|2\omega_1 - \omega_2|$ and $|2\omega_2 - \omega_1|$ may pass through the filter, which is often undesirable.

3.2 Calculation of Linearity

Mixed-signal integrated circuits have become increasingly prevalent lately, as both analog and digital circuits are being integrated in a single System-on-Chip (SoC) for increased functionality and reduced cost. However, the mixed ICs need large input signals which lead to non-linearities, two of which were mentioned in the previous section. Any type of non-linearity could limit the overall performance of an application. As we previously mentioned, MOS transistors are used in various fields such as telecommunication applications where non-linear effects should be predicted and if possible canceled.

The need of a circuit designer to have knowledge about how a MOS transistor behaves on various conditions such as temperature, different bias conditions and a wide range of frequency, made essential the development of methods, the so-called *linearity figure of merits* to quantify and predict the level of distortion. Depending on the type of non-linearity different methods for calculation and prediction has been developed. For example, the harmonic distortion is usually predicted using the fast Fourier transform, the five-point method[5], and the calculation of the higher-order derivatives of the input-output characteristics[6]. On the other hand intermodulation distortion is usually calculated using a Taylor expansion series[7] and Volterra Series[8]-[9].

3.2.1 Calculation of harmonic distortion

Expanding Eq. 3.4 in a Fourier series we obtain:

$$i_d = c_0 + \sum_{i=1}^{\infty}[c_{2i} \cdot cos2i\omega_0 t + c_{2i-1} \cdot sin(2i-1)\omega_0 t] \tag{3.6}$$

where the coefficients c_i are given by[1]:

$$\begin{aligned} c_0 &= a_0 + \sum_{n=1}^{\infty} \frac{a_{2n}}{2^{2n-2}} \binom{2n}{n} \hat{A}^{2n} \\ c_{2i} &= \sum_{n=i}^{\infty} (-1)^i \frac{a_{2n}}{2^{2n-1}} \binom{2n}{n+i} \hat{A}^{2n} \\ c_{2i-1} &= -\sum_{n=i}^{\infty} (-1)^i \frac{a_{2n-1}}{2^{2n-2}} \binom{2n-1}{n-i} \hat{A}^{2n-1} \end{aligned} \tag{3.7}$$

A convenient measure is given by the ratio $D_i = \frac{c_i}{c_1}$(for $i \geq 2$), which represents the correlation of the i^{th}-order harmonic amplitude to that of the fundamental one. Usually we are only concerned

for the second-order and third-order harmonic because even if we apply a low amplitude signal, the dependence of these harmonics on the signal amplitude $\hat{A}$ is still strong. In practice, the 2^{nd} and 3^{rd}-order harmonics are given by:

$$|D_2| \approx \frac{1}{2}\frac{a_2}{a_1}\hat{A} \tag{3.8}$$

$$|D_3| \approx \frac{1}{4}\frac{a_3}{a_1}\hat{A}^2 \tag{3.9}$$

In many applications which use transistors, it would be useful to have a convenient measure of the distortion caused by all powers of the input amplitude; the so-called *Total Harmonic Distortion* or *THD* has been defined[1] and can be expressed as follows:

$$THD = \sqrt{\sum_{i=2}^{\infty} D_i^2} \tag{3.10}$$

The calculation method of the higher-order harmonics which was just analyzed is not binding. Depending on the CMOS application new expressions for calculating harmonic distortion coefficients have been derived [10]. Furthermore, a new method based on specific integration of the DC current-voltage characteristics of the device has been proposed [11]-[12]. This method provides analytical expressions for the amplitudes of the second and third-order harmonics as functions of the device parameters. However, the dependence of the harmonic performance on the device parameters is not obvious and several iterations must be performed before a reliable conclusion can be reached.

3.2.2 Calculation of intermodulation distortion

In practice signals are complex and are usually formed of unequal amplitude multi-sinusoids. It is, therefore, argued that the THD figures are irrelevant and as a consequence intermodulation distortion (IMD) figures were proposed which are far more relevant and reflect how a nonlinear device performs.

By studying Eq. 3.3 we note that when the input voltage v_{in} is increased, the fundamental increases proportional to v_{in}, but the 3^{rd} order IM product increases in proportion of v_{in}^3.

At a certain input level, the 3^{rd} IM product will have the same magnitude as the fundamental. This level is defined as the **IP$_3$**, i.e the intersection of the two lines (Fig. 3.2.1(a))[3]. IP$_3$ can be referred to as the power level at the input(IIP$_3$) or at the output (OIP$_3$). If we know a_1 and a_3, we can, according to [2], calculate the amplitude at the output at IP3 as:

$$A_{IP_3} = \sqrt{\frac{4}{3}\left|\frac{a_1}{a_3}\right|} \tag{3.11}$$

This linearity figure is extremely important because the higher the value of the third-order intercept point, the more linear the device.

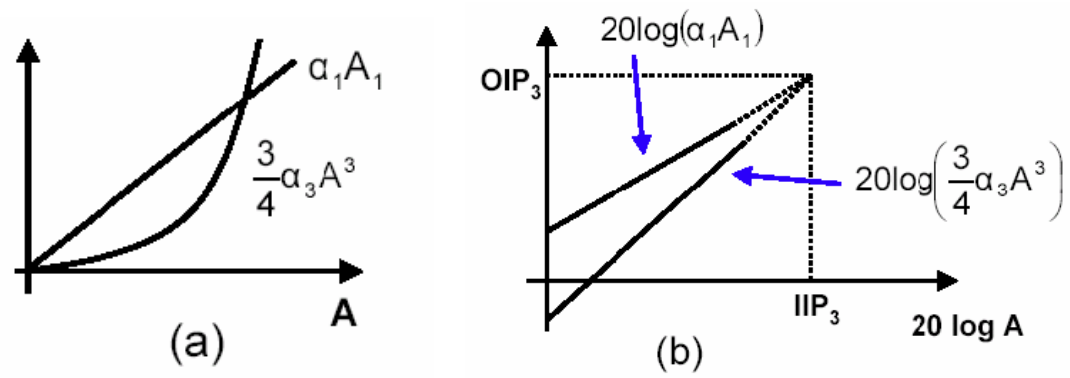


Figure 3.2.1: Growth of output components in an intermodulation test [2].

In practice, a Taylor series expansion of the drain current is usually used[7] which suggests that the drain current can be expressed as follows:

$$\begin{aligned} i_{ds}(v_{GS}, v_{DS}) &= I_{DS}(V_{GS}, V_{DS}) + G_m v_{gs} + G_d v_{ds} \\ &\quad + G_{m2} v_{gs}^2 + G_{ms} v_{gs} v_{ds} + G_{d2} v_{ds}^2 \\ &\quad + G_{m3} v_{gs}^3 + G_{m2d} v_{gs}^2 v_{ds} \\ &\quad + G_{md2} v_{gs} v_{ds}^2 + G_{d3} v_{ds}^3 + \dots . \end{aligned}$$

If we assume that the drain is shorted at a signal frequency, then all output-conductance terms (i.e G_d, G_{d2}, G_{d3},...) and cross modulation terms (G_{ms}, G_{md2}, G_{md3},...) are vanished and only transconductance terms (G_m, G_{m2}, G_{m3}, ...) remain. In this case, the IP_3 is given as follows:

$$IP_3 = \sqrt{\frac{4}{3}\frac{G_m}{G_{m3}}} \tag{3.12}$$

This formula is not accurate enough, because of the fact that it does not consider output-conductance non-linearity which is of equal importance and will be analyzed in the following section.

3.3 Output Conductance

Until now, we have only accounted for nonlinearities due to the MOSFET transconductance. The higher-order derivatives of the drain current with respect to the gate terminal are not the only causes of non-linear behavior.

The output conductance g_{ds} also has a dependence on the operation point as well as frequency, hence introduce further nonlinearities. Thus, the correct modeling of g_{ds} and g_{ds2} is essential in MOS analysis. Since in analog circuit design, a MOSFET is often driven by the drain terminal, we come to the conclusion that the output conductance could have a great impact on the linearity performance of a device. For example, the load device in an amplifier structure is often a drain terminal MOSFET operating in saturation region. Taking into consideration that in most applications the design of an amplifier asks for a trade-off between signal-to-noise ratio and linearity, it is more than obvious the need of a compact MOSFET model that predicts accurately the higher-order derivatives of the drain current with respect to the drain terminal.

However, we must mention that measurements of especially g_{ds2} and g_{ds3} are very sensitive to noise and as a result their correct modeling is a challenging task.

In order to have an insight on the behavior of the nonlinear sources of a MOS transistor you can refer to Chapter 5 where the measured and simulated results of the EKV3.0 model are presented.

3.4 Linearization techniques

Linearity is an important performance parameter when evaluating a power amplifier for use in certain communications applications, especially systems with advanced modulation formats. Certain applications use amplifiers with global feedback (i.e a feedback between the input and the output) to achieve high linearity performance. However, stability and settling issues of feedback arise, limiting their performance in high-speed applications. As a result, new methods for linearization, the so-called *Linearization Techniques*, were introduced. The basic principle behind linearization is to reduce the dependence of gain of the circuit upon the input level.

Source degeneration by means of a linear resistor is the simplest linearization technique[13]. For a common-source stage, degeneration reduces the signal swing applied between the gate and the source of the transistor, thereby making the input/output characteristic more linear. We can observe from Fig. 3.4.1 that the addition of a source resistance R_S in the source circuit

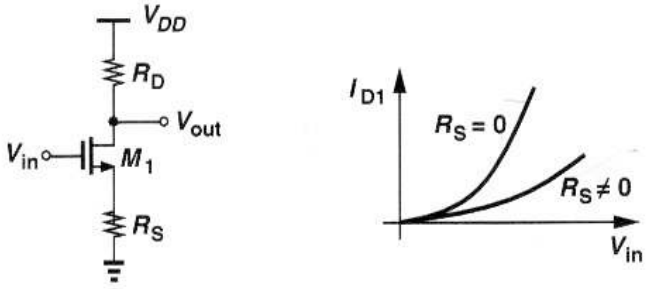


Figure 3.4.1: Common-source stage with resistive degeneration[13].

provides a more linear relation between the drain current of M1 and the input voltage. If we neglect the body effect, the overall transconductance of the stage is given by:

$$G_m = \frac{g_m}{1 + g_m R_S} \tag{3.13}$$

which for large values of $g_m R_S$ approaches $\frac{1}{R_S}$, an input-independent value. Furthermore, we must note that the amount of linearization depends on $g_m R_S$ rather on R_S alone. With a relatively constant G_m, the voltage gain, $G_m R_D$, is also independent of the input and the amplifier is linearized.

The use of resistive degeneration has implications for linearity, power dissipation and noise. Furthermore, this method requires high-quality resistors, a feature unavailable in many of today's CMOS technologies. For this reason, a replacement of the resistor by a transistor operating in ohmic region is usually applied (Fig. 3.4.2). The drawback in this technique is

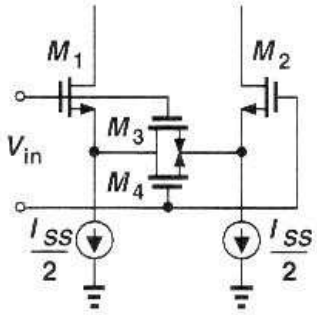


Figure 3.4.2: Differential pair degenerated by a MOSFET operating in deep triode region[13].

that when large input swings are applied, the transistor which serves the role of a resistor may not remain in deep triode region. Based on the same principle, another linearization technique was proposed[13], and guarantees the linear operation of the circuit even if one of the two degeneration devices goes into saturation region. This technique is illustrated in Fig. 3.4.2. By analyzing the above circuit we get that if $V_{in} = 0$ then:

$$V_{G1} = V_{G2} \Leftrightarrow V_{GS1} = V_{GS2} \tag{3.14}$$

Additionally, it holds that:

$$V_{G3} = V_{G1} \tag{3.15a}$$
$$V_{G4} = V_{G2} \tag{3.15b}$$
$$V_{D3} = V_{S1} \tag{3.15c}$$
$$V_{D4} = V_{D3} = V_{S1} \tag{3.15d}$$
$$V_{S4} = V_{S3} = V_{S2} \tag{3.15e}$$

As a result, if the input voltage is zero, we can replace Eq. 3.15a, 3.15b to Eq. 3.14 and get that:

$$V_{G3} = V_{G4} \tag{3.16}$$

Assuming that M3 operates in the ohmic region, the following equation should hold:

$$V_{D3} < V_{G3} - V_{TH}, \text{where } V_{TH} \text{ is the threshold voltage.} \tag{3.17}$$

By replacing Eq. 3.15c and Eq. 3.15a instead of V_{D3} and V_{G3} respectively, in Eq. 3.17, and solving the inequality we get that:

$$V_{GS1} > V_{TH} \tag{3.18}$$

which implies that M1 is "on". Subsequently, the primary assumption that M3 operates in ohmic region holds. Based on the same approach the same holds for the M4.

If the gate voltage of M1 becomes more positive than the gate voltage of M2, i.e $V_{G1} > V_{G2}$ then M3 continues to operate in the linear region based on the approach which was mentioned above. The drain voltage of M3 is given by

$$V_{D3} = V_{G3} - V_{GS1}. \tag{3.19}$$

Substituting Eq. 3.15a to Eq. 3.19 gives

$$V_{D3} = V_{G1} - V_{GS1}. \tag{3.20}$$

This implies that the increase in the drain voltage of M3 is less than the corresponding increase in the gate voltage of M3. This ensures that M3 remains in the linear region. On the other hand, the drain voltage increase of M4 together with the gate voltage decrease of M4 will drive the latter into the saturation region. Thus, the circuit remains relatively linear even if one degeneration device goes into saturation.

Chapter 4

Theoretical model and Experimental procedure

In this Chapter we will describe the small-signal model and the experimental procedure we followed to perform the measurements. Furthermore, we will describe the theoretical background in which our calculations and assumptions were based.

4.1 Small-Signal model

In order to model phenomena that appear when transistors operate at low frequencies, a common method is to replace the transistor with a structure of transconductances with real and zero imaginary parts. In Fig. 4.1.1 the small-signal model is depicted. We note that this analysis does not take into account the resistances formed in gate, source and drain terminals. These resistances are supposed to be zero with correct design layout.

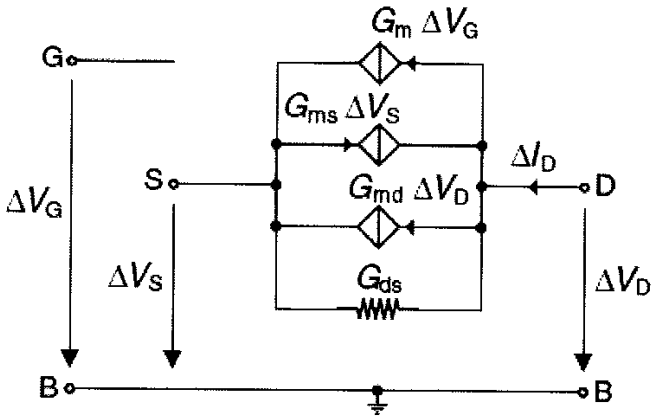


Figure 4.1.1: Small-Signal model.

The transoncudactances with respect to gate, source and drain terminal can be respectively expressed as

$$g_{mg} = \frac{\partial I_D}{\partial V_G} \mid V_{S,D}$$

$$g_{md} = \frac{\partial I_D}{\partial V_D} \mid V_{S,G} \tag{4.1}$$

$$g_{ms} = -\frac{\partial I_D}{\partial V_S} \mid V_{G,D}.$$

For models which refer to the source terminal, the substrate transconductance is written as

$$g_{mb} = -\frac{\partial I_D}{\partial V_B} \mid V_{G,D}. \tag{4.2}$$

Finally, the relationship between the transconductances can be expressed as

$$g_{ms} = g_{mg} + g_{md} + g_{mb}. \tag{4.3}$$

Another two important equations that include the slope factor n are the following

$$\begin{aligned} g_{mg} &= \frac{g_{ms} - g_{md}}{n} \\ g_{mb} &= \frac{n-1}{n}(g_{ms} - g_{md}), \end{aligned} \tag{4.4}$$

where the slope factor is given by[14][15]

$$n = \left[\frac{\partial V_p}{\partial V_g}\right]^{-1} = 1 + \frac{\gamma}{2\sqrt{\Psi_0 + V_p}}, \tag{4.5}$$

γ is the substrate factor and is expressed by $\gamma = \frac{\sqrt{2q\epsilon_{si}N_{sub}}}{C'_{ox}}$, and Ψ_0, i.e the Bulk Fermi potential, which is given by $\Psi_0 \simeq 2U_t \ln(\frac{N_{sub}}{n_i})$. The pinch off voltage, V_p, is used by the EKV MOSFET model as a reference for the channel voltage and is approximated[14] by the following equation

$$V_p \simeq \frac{(V_G - V_{TO})}{n}, \tag{4.6}$$

where V_{TO} is the threshold voltage. The slope factor (n) and the pinch off voltage (V_p) are depended on the channel length (L) due to the reverse short channel effect (RSCE), charge sharing and DIBL effect which affect the substrate factor (γ) and the threshold voltage (V_{TO})[16].

The inversion charges of the source and drain, q_s and q_d respectively, are given by[17]

$$\begin{aligned} \frac{V_P - V_S}{U_T} &= 2q_s + \ln q_s \\ \frac{V_P - V_D}{U_T} &= 2q_d + \ln q_d. \end{aligned} \tag{4.7}$$

The current that flows from drain to source is written as[17]

$$I_D = I_{spec}(i_f - i_r), \tag{4.8}$$

where I_{spec} is the specific current and i_f, i_r are the normalized forward and reverse currents respectively. These currents can be expressed as a function of source and drain charges as follows[17]

$$\begin{aligned} i_f &= q_s^2 + q_s \\ i_r &= q_d^2 + q_d. \end{aligned} \tag{4.9}$$

As a function of q_s and q_d the source and drain transconductances can be expressed[17]

$$\begin{aligned} g_{ms} &= q_s \frac{I_{spec}}{U_T} \\ g_{md} &= q_d \frac{I_{spec}}{U_T}. \end{aligned} \tag{4.10}$$

Another set of important equations which express g_m, g_{ms}, I_D and U_t as a function of q_s and i_f are the following[17]:

$$\begin{aligned} \frac{g_{ms}}{I_D} U_t &= \frac{1}{1+q_s} = \frac{1}{\frac{1}{2}+\sqrt{\frac{1}{4}+i_f}} \\ \frac{g_m}{I_D} U_t &= \frac{1}{n}\frac{1}{1+q_s} = \frac{1}{n}\frac{1}{\frac{1}{2}+\sqrt{\frac{1}{4}+i_f}} \end{aligned} \tag{4.11}$$

Finally, the slope factor (n) in the weak inversion (W.I) is given by:

$$n = \frac{g_{ms}}{g_m} = 1 + \frac{g_{mb}}{g_m} \tag{4.12}$$

In order to be able to compare the results for different channel lengths and for all the inversion regions , the drain current (I_D) and the transconductances are subjected to normalization. The normalization coefficient which is called specific current (I_{spec}) is given by

$$I_{spec} = I_0 \frac{W_{eff}}{L_{eff}}. \tag{4.13}$$

where $I_0 = 2n\mu {U_T}^2 C'_{ox}$, n is the slope factor, μ the carrier mobility, $U_T = \frac{kT}{q}$, the thermal voltage and finally $C'_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$ the gate oxide capacitance per unit area. The ratio of the drain current (I_D) to the specific current coefficient (I_{spec}) is called inversion coefficient (IC)[17] and is given by the following equation:

$$IC = \frac{I_D}{I_{spec}}. \tag{4.14}$$

The inversion coefficient (IC) distinguishes the level of inversion into three sub-areas:

- Weak inversion (W.I) : For IC$<$ 0.1
- Moderate inversion (M.I) : For 0.1 $<$ IC $>$ 10
- Strong inversion (S.I) : For IC$>$ 10

The level of inversion and also the channel length have a great impact on transconductances, the Early voltage and the intrinsic voltage gain. As a result, the inversion coefficient is a very important parameter for the circuit designers because they can observe the behavior of small-signal quantities and consciously select the desirable level of inversion. For this reason, we choose in this chapter to present the normalized source and gate transconductances, the higher-order terms of the drain current and the Intrinsic voltage gain as a function of IC.

The normalized transconductances are given by

$$G_x = \frac{g_{mx} U_T}{I_D}, \tag{4.15}$$

where x = g, s, b, d are the MOSFET terminals.
By observing Eq.(4.8) and Eq.(4.14) we can conclude that IC$\simeq i_f$ holds. As a result, in saturation region, Eq.4.11 could give the ideal relation of the normalized transconductance and the drain current and can be expressed as follows[16]:

$$G(IC) = \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + IC}}. \tag{4.16}$$

Depending on the value of the inversion coefficient we get that

$$\begin{aligned} IC &<< 1 : G(IC) \to 1 \\ IC &>> 1 : G(IC) \to \frac{1}{\sqrt{IC}} \end{aligned} \tag{4.17}$$

By combining the above with Eq.4.4, we come to the conclusion that the ideal relation between the drain current and the normalized transconductance is the following

$$\begin{aligned} G_s &\simeq G(IC) \\ G_g &\simeq \frac{G(IC)}{n} \\ G_b &\simeq \frac{n-1}{n} G(IC). \end{aligned} \tag{4.18}$$

As far as the drain terminal is concerned, the equivalent equation is given by[16]

$$G_d \simeq \frac{\partial V_p}{\partial V_d} G(IC) + \frac{U_T}{n} \frac{\partial n}{\partial V_d}. \tag{4.19}$$

However, this equation is quite complicated and is very challenging due to second-order effects which make this equation less accurate in small geometries.

Finally, we introduce the Intrinsic Voltage Gain A_{Vi} which is the maximum voltage gain for a specific transistor and is written as

$$A_{Vi} = \frac{g_m}{g_{ds}}. \tag{4.20}$$

4.2 Measurement Setup

The purpose of this thesis was to measure the drain current and its high-order derivatives with respect to the gate, source and drain terminal. The equipment we used and the experimental procedure that we followed is analytically described in Appendix.

The following bias conditions were defined on IC-CAP 2008:

- I_D measurement with respect to the gate (G) terminal.
  - V_G: From -300 mV to 1.2 V, with step size 30 mV.

  - V_S: From -200 mV to 600 mV, with step size 200 mV.
  - V_D: The drain voltage was set to 1 V so as the DUT to operate in saturation and 50 mV to operate in linear region.
  - V_B: The bulk voltage was set to 0 V.

- I_D measurement with respect to the source (S) terminal.
  - V_G: From 200 mV to 0.8 V, with step size 150 mV.
  - V_S: From -300 mV to 900 mV, with step size 30 mV.
  - V_D: The drain voltage was set to 1 V in order the DUT to operate in saturation.
  - V_B: The bulk voltage was set to 0 V.

- I_D measurement with respect to the drain (D) terminal.
  - V_G: From 200 mV to 1.2 V, with step size 200 mV.
  - V_S: The source voltage was set to 0 V
  - V_D: From 30 mV to 1.2 V with step size 30 mV.
  - V_B: The bulk voltage was set to 0 V.

The measurements were taken on room temperature, T = 27°C

4.3 Calculations and Assumptions

After successfully measuring I_D, we calculated the higher terms of the drain current through the IC-CAP using the integrated derivative function. These higher-order terms are expressed as

$$\frac{\partial^i I_D}{\partial V_j^i} \text{ , i = 1,...,3 and j = G, S, D.} \tag{4.21}$$

Furthermore, we calculated the normalized transconductances which are given as we already mentioned in Section 4.1 by

$$G_x = \frac{g_{mx} U_T}{I_D} \text{ , where x = G, S.}$$

In order to calculate the inversion coefficient (IC) which is given by Eq. 4.14, we firstly had to calculate I_0 which was defined in Section 4.1 and is given by

$$I_0 = 2n\mu {U_T}^2 C'_{ox},$$

where n is the slope factor, μ the carrier mobility, $U_T = \frac{kT}{q}$ the thermal voltage and $C'_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$ the gate oxide capacitance per unit area. Due to the dependence of the mobility and slope factor from the level of inversion, I_0 is not constant. However, we used a common practice[18], where n and μ have fixed values(n_0 and μ_0 respectively) independently of the level of inversion, so I_0 is supposed to be constant.

The fixed value of n (n_0) is calculated from the measured data using Eq. 4.12. As we mentioned

previously, Eq. 4.12 is an approximation for the weak inversion. In moderate inversion we notice 5% decrease and as we move to the deep strong inversion the percentage becomes noticeably bigger, mainly due to the velocity saturation effect.

Finally, we calculated $\mu_0 \cdot C'_{ox}$ according to the following procedure:
In saturation, the drain current is written as

$$I_D = \frac{n\beta}{2}\left[V_p - V_S\right], \tag{4.22}$$

where $\beta = \mu_0 \cdot C'_{ox} \cdot \frac{W}{L}$. By setting $V_S = 0$ and taking the square-root of I_D, we get that

$$\sqrt{I_D} = \sqrt{\frac{n\beta}{2}} V_p. \tag{4.23}$$

By calculating the derivative of Eq. 4.23, we find the maximum of this function and then we calculate β. More analytically, we get

$$\frac{\partial I_D}{\partial V_g} = \sqrt{\frac{\beta}{2n}} \frac{\partial V_p}{\partial V_g} \cdot n \tag{4.24}$$

Making use of Eq. 4.5, we write

$$\frac{\partial I_D}{\partial V_g} = \sqrt{\frac{\beta}{2n}}. \tag{4.25}$$

Finally, we calculate the value of β and subsequently the value of $\mu_0 \cdot C'_{ox}$ by finding the maximum value of Eq. 4.25 and replacing the fixed value n_0 instead of n.

Even though, the mobility μ is greatly affected in strong inversion by the velocity saturation and the Vertical Field Mobility Reduction (VFMR) effect, a fixed mobility value μ_0 is a good approximation for the inversion regions with the biggest interest, i.e weak and moderate inversion regions.

In conclusion, the inversion coefficient which we will use is associated with the fixed value of the current I_0

$$IC = \frac{I_D}{I_0 \frac{W}{L}} = \frac{I_D}{2n_0 \mu_0 {U_T}^2 C'_{ox} \frac{W}{L}} \tag{4.26}$$

4.4 Gate transconductance

The gate transconductance g_m is the ratio of the current change to the source-gate voltage (V_{GS}) change and is expressed as

$$g_m = \frac{\partial I_D}{\partial V_G} | V_S, V_D \tag{4.27}$$

We define the normalized form of g_m that we will use throughout this section

$$G_g = \frac{g_m U_T}{I_D}, \text{where } U_T \text{ is the thermal voltage.} \tag{4.28}$$

The MOSFET transconductance is a very important design parameter because it decides its gain and also affects the bandwidth, the thermal noise and the intrinsic voltage gain. The aim of the circuit designer is to keep the value of g_m as high as possible which leads to low

thermal noise, high voltage gain and bandwidth. As it will be shown in the following figures, the maximum value of the transconductance is achieved in weak inversion region due to the exponential relation of the drain current in that region. The drain current is expressed as

$$I_D(W.I) = 2n\mu C'_{ox} {U_T}^2 \frac{W}{L} e^{\frac{V_p - V_s}{U_T}} \tag{4.29}$$

In moderate inversion, g_m starts to reduce, reaching the lowest value in deep strong inversion region. Neglecting second-order effects, the expected behavior of transconductance is described, as we have already mentioned in Section 4.1, by the G(IC) function which is expressed as follows:

$$G_g = \frac{G(IC)}{n}, \text{where } G(IC) = \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + IC}} \tag{4.30}$$

4.4.1 Second-order effects

The gate transconductance is mainly affected by the following second-order effects:

- **Velocity saturation** : The transconductance is proportional to hole or electron mobility (depending on device type), at least for low drain voltages. As MOSFET size is reduced, the dopant impurity and field levels increase. Both charges reduce carrier mobility, and hence the transconductance. As channel lengths are reduced without proportional reduction of the drain voltage, raising the electric field in the channel, the result is velocity saturation of the carriers, limiting the current and the transconductance. The horizontal electric field that is formed between the drain and the source which reduces the carriers' mobility is given by:

  $$E_x = \frac{(V_{GS} - V_T)}{L} \tag{4.31}$$

  By observing Eq. 4.31 we come to the conclusion that the smaller the gate length (L), the bigger impact has the velocity saturation in short transistors.

- **VFMR (Vertical Field Mobility Reduction)** : When large voltages are applied to the gate of a MOSFET, electrons are strongly attracted to the top of the channel, which is the junction between the silicon and the gate oxide. This interface has dangling bonds and contaminants that slow down the electronics, reducing the effective electron velocity. The larger the gate voltage, the more the electrons are attracted to this junction.

- **Reverse Narrow Channel Effect** : In order to reduce the current leakage, MOSFETs are fabricated using a sidewall implantation technology(STI, LOCOS). Taking STI technique as an example, the key steps of this process involve etching a pattern of trenches in the silicon, depositing one or more dielectric materials to fill the trenches. The effect of the trench edge has given rise to what has recently been termed to the "reverse narrow channel effect".

4.4.2 Measurement Results and Discussion

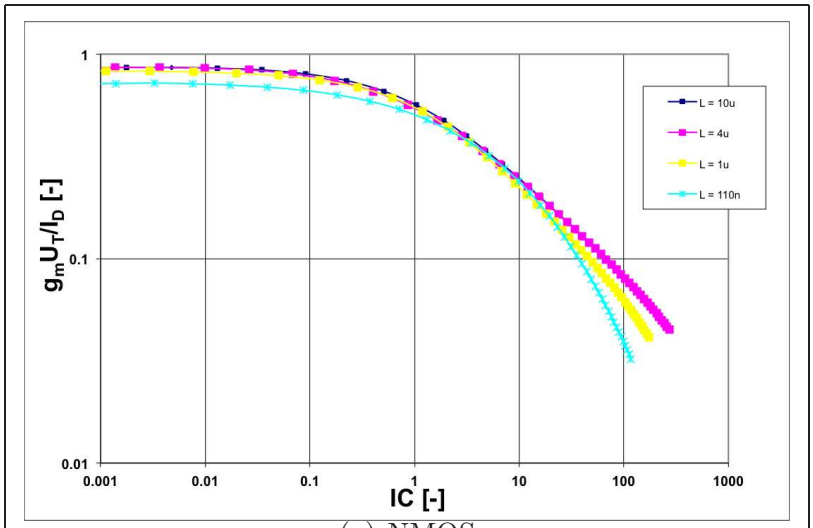
In this subsection we will present the results of the measurements that we performed in a CMOS 110nm technology. More specifically, we will present the normalized transconductance $\frac{g_m U_T}{I_D}$, the first, second and third differentials of the drain current, g_m, g_{m2} and g_{m3} respectively. All

these important quantities are plotted versus the inversion coefficient (IC), in order to have a more detailed view on all operation regions of the transistor and are grouped by the channel length (L) so as to be able to observe how they are affected by the channel length scaling. The Devices Under Test (DUTs) are presented in the following table:

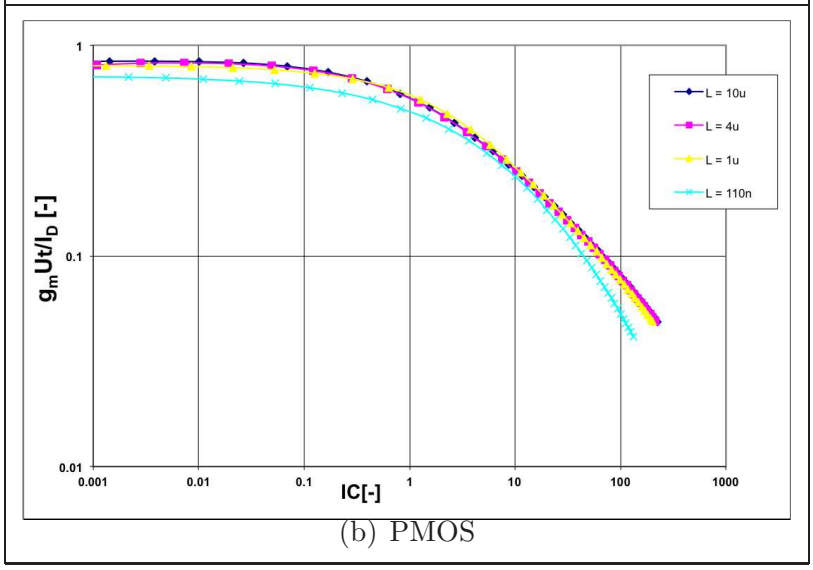
DUT	Width	Length	Number of Fingers
NMOS\PMOS	10u	10u	1
	10u	4u	1
	10u	1u	1
	10u	110n	1

Figure 4.4.1: Devices Under Test (DUTs)

- $\mathbf{\frac{g_m U_T}{I_D}}$ **vs. IC,** $\mathbf{V_D = 1V}$, $\mathbf{V_S = 0V}$, $\mathbf{V_B = 0V}$



(a) NMOS



(b) PMOS

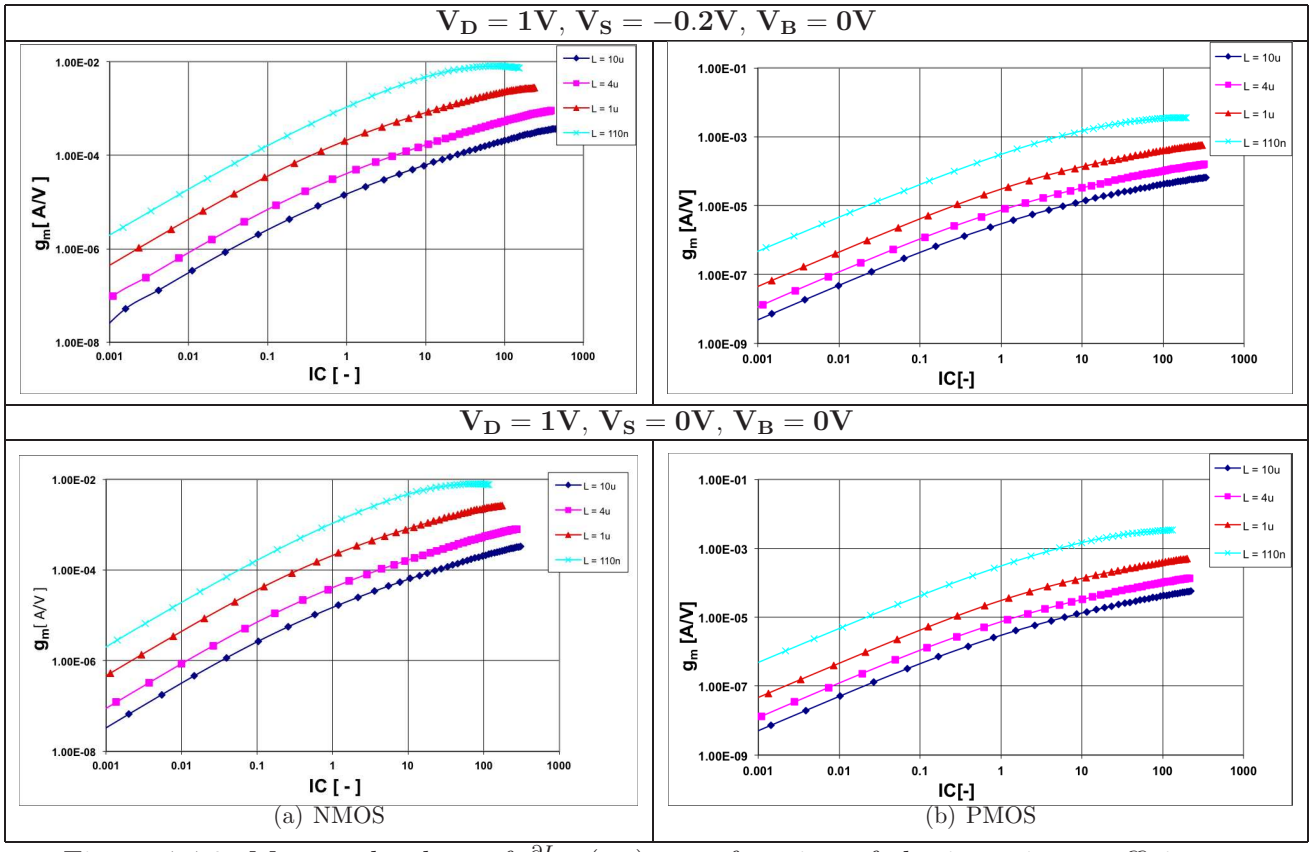
Figure 4.4.2: Normalized transconductance versus IC, $V_D = 1V, V_S = 0V, V_B = 0V$. (a)NMOS and (b)PMOS, $N_f = 1$.

As we previously mentioned the designer's aim is to keep the transconductance's value as high as possible. From Fig. 4.4.2 we can observe that the maximum value occurs in the week inversion region (IC< 0.1) while in the moderate inversion is starting to reduce reaching the lowest value in deep strong inversion region. This behavior illustrates why weak and moderate inversion regions are the most preferable. Usually, moderate inversion is the first choice of a designer because among the numerous advantages, provides high bandwidth too. The same results were obtained for different V_{DS} values. However, the relevant figures are not listed in this thesis.

By carefully examining the above figure, we notice the contribution of the second-order effects as the channel length becomes shorter. The starting value of the normalized transconductance in the weak inversion is smaller for the short transistors reaching a smaller value in the strong inversion region too. More specifically, we must note that the normalized transconductance of the short channel NMOS transistor experiences a decrease of about 12% contrary to the long channel transistor. In general, second-order effects seem to affect more the NMOS transistor than the PMOS.

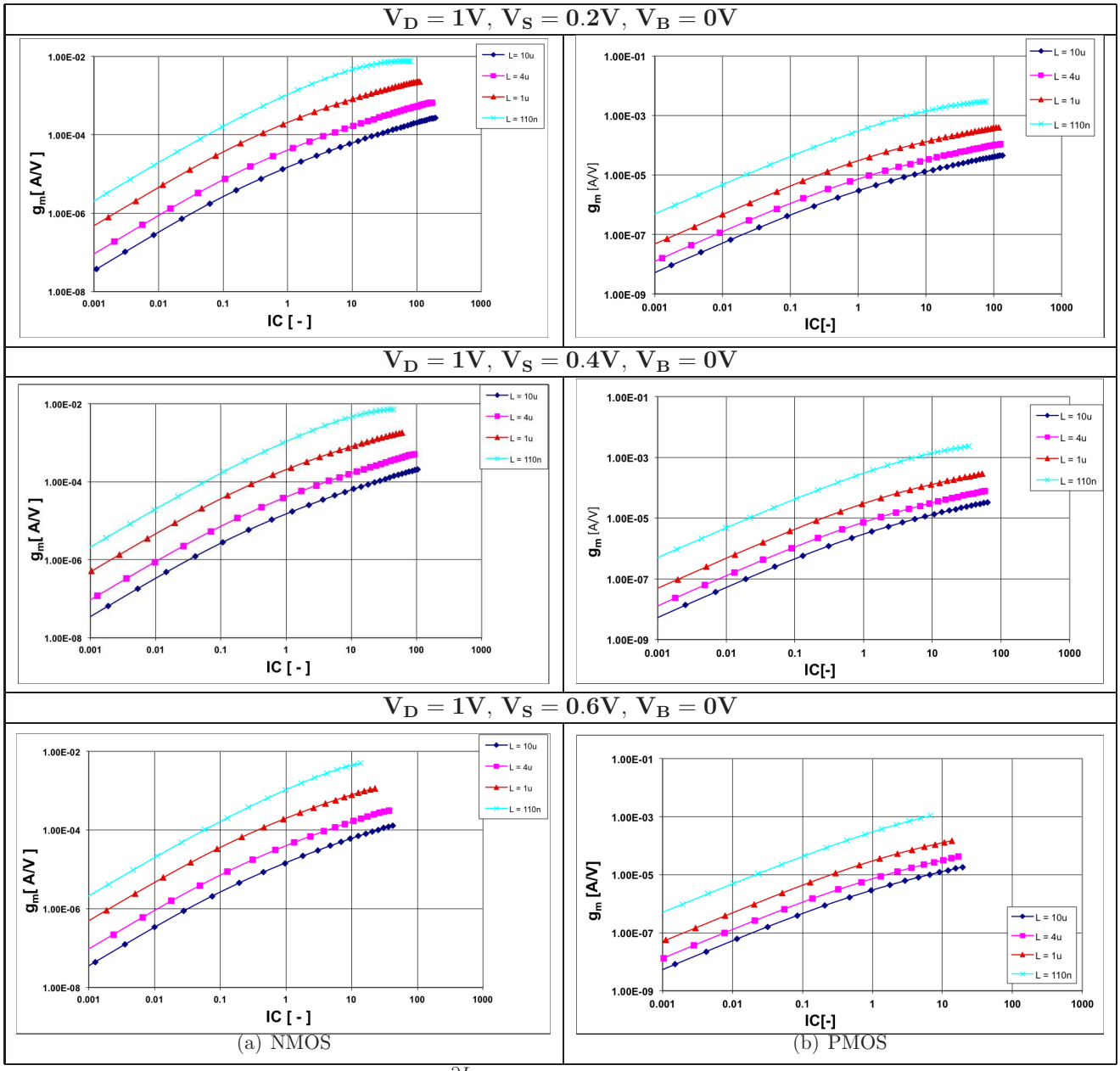
Additionally, we present some indicative figures concerning the non-normalized transconductance in order to be able to have an insight of how g_m is affected by the different bias conditions, the channel length scaling and the level of inversion.

- **g_m vs. IC**



(a) NMOS (b) PMOS

Figure 4.4.3: Measured values of $\frac{\partial I_D}{\partial V_G}$ (g_m) as a function of the inversion coefficient (IC) for various bias conditions. (a)NMOS (b)PMOS. (Cont.)



(a) NMOS (b) PMOS

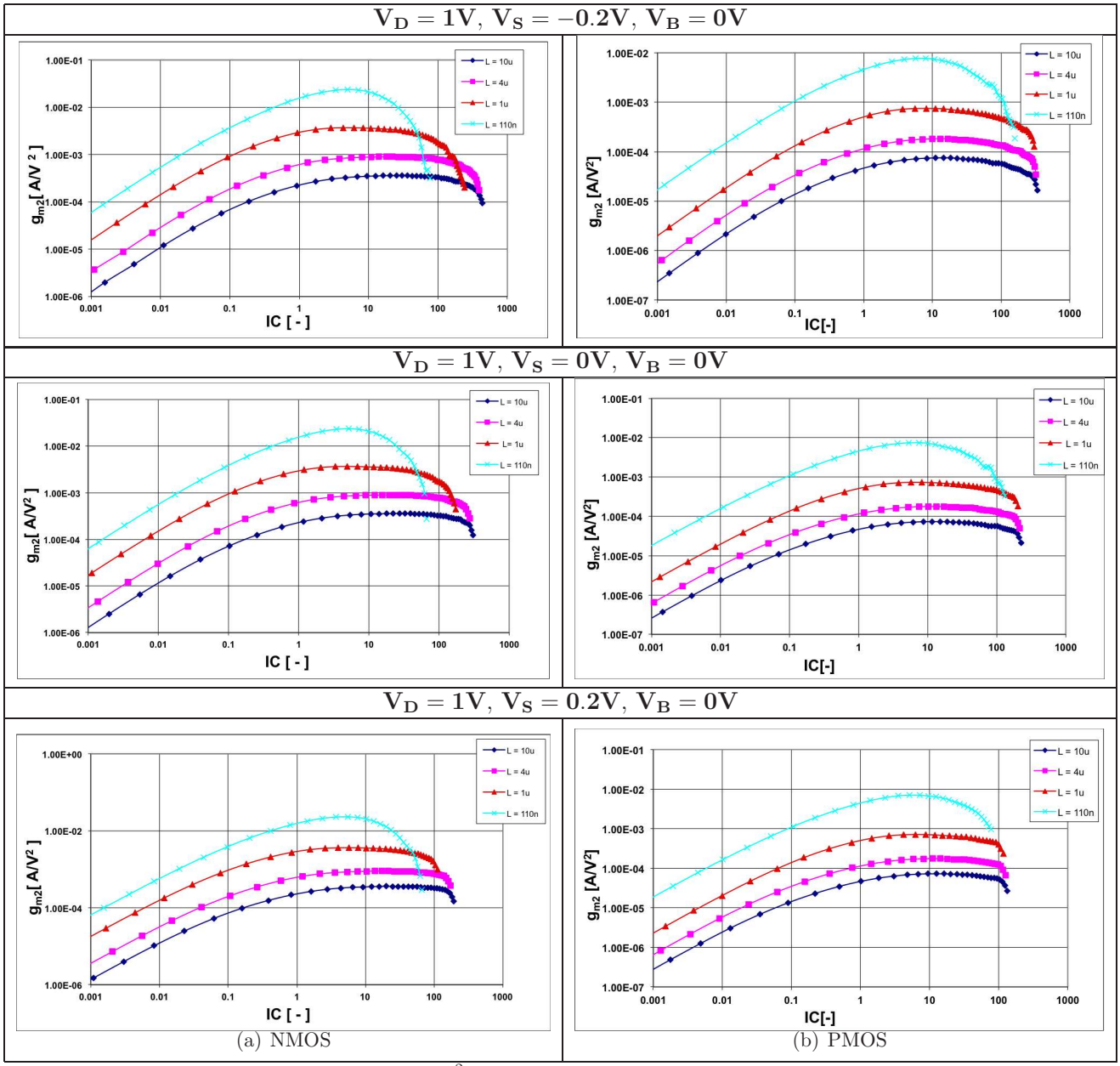
Figure 4.4.4: Measured values of $\frac{\partial I_D}{\partial V_G}$ (g_m) as a function of the inversion coefficient (IC) for various bias conditions. (a)NMOS (b)PMOS.

By observing Fig. 4.4.3 - 4.4.4, we can already draw some general conclusions about the behavior of $\frac{\partial I_D}{\partial V_G}$. More specifically, the shorter the gate length of the transistor the higher the value of g_m for all the levels of inversion, independently of the source voltage (V_S) value. The behavior is explained by the fact that the drain current is bigger in the short transistors so as the first derivative. For low source voltage values g_m increases smoothly reaching the highest value in deep strong inversion region. As V_S becomes more positive the transconductance reaches in a smaller level of inversion since the drain current decreases. Even though the PMOS transconductance is smaller than the NMOS, their overall behavior does not deviate at all.

The second and third-order derivative of the drain current with respect to the gate voltage is equally important to the gate transconductance when we investigate the non-linear behavior

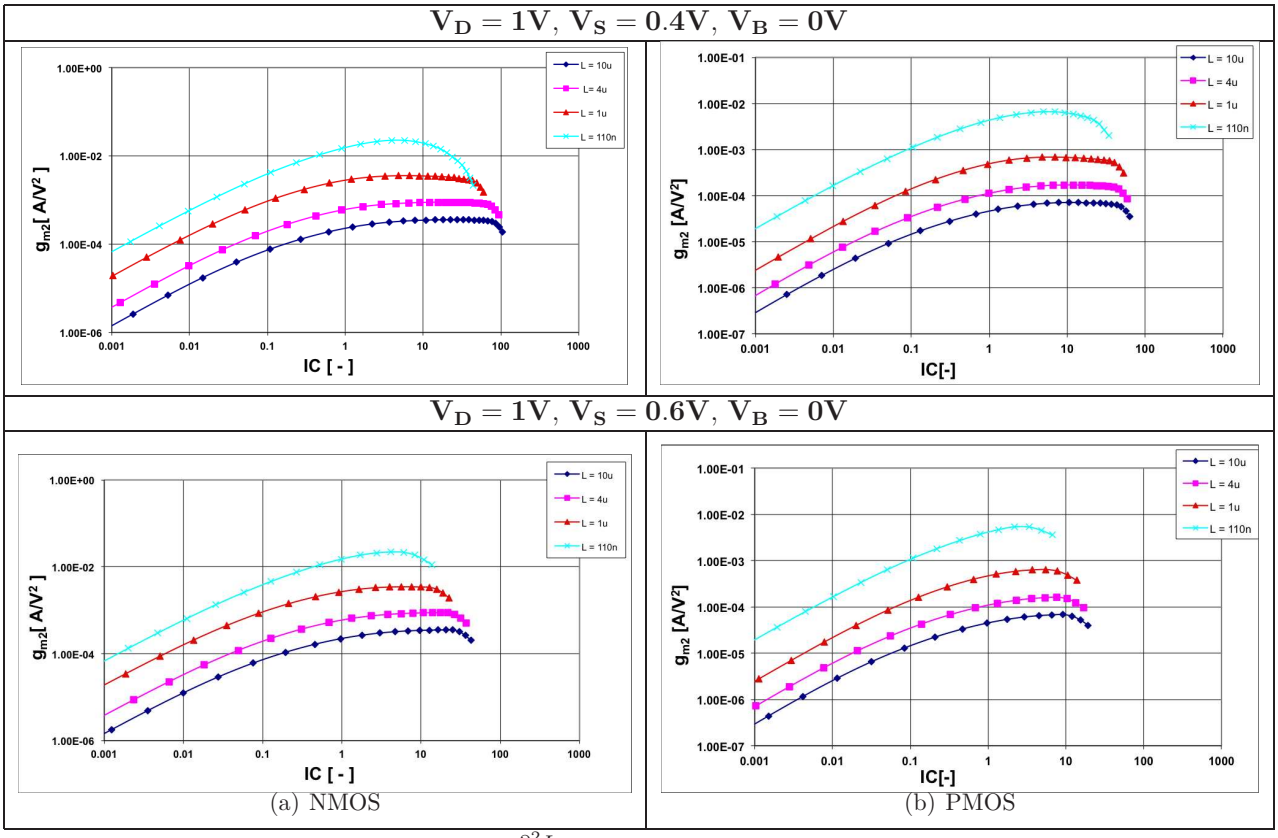
of a MOSFET device. The circuit designer should have a detailed knowledge of how these high-order terms behave as a function of the level of inversion so as to be able to choose the appropriate inversion region on which the device will operate. At this point we present the measured values of $\frac{\partial^2 \mathbf{I_D}}{\partial \mathbf{V_G^2}}$ ($\mathbf{g_{m2}}$) as a function of IC for different bias conditions. Finally, we will discuss how g_{m2} is affected by the channel length scaling, level of inversion and different bias conditions.

- $\mathbf{g_{m2}}$ **vs. IC**



(a) NMOS (b) PMOS

Figure 4.4.5: Measured values of $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) as a function of the inversion coefficient (IC) for various bias conditions. (a)NMOS (b)PMOS (Cont.).



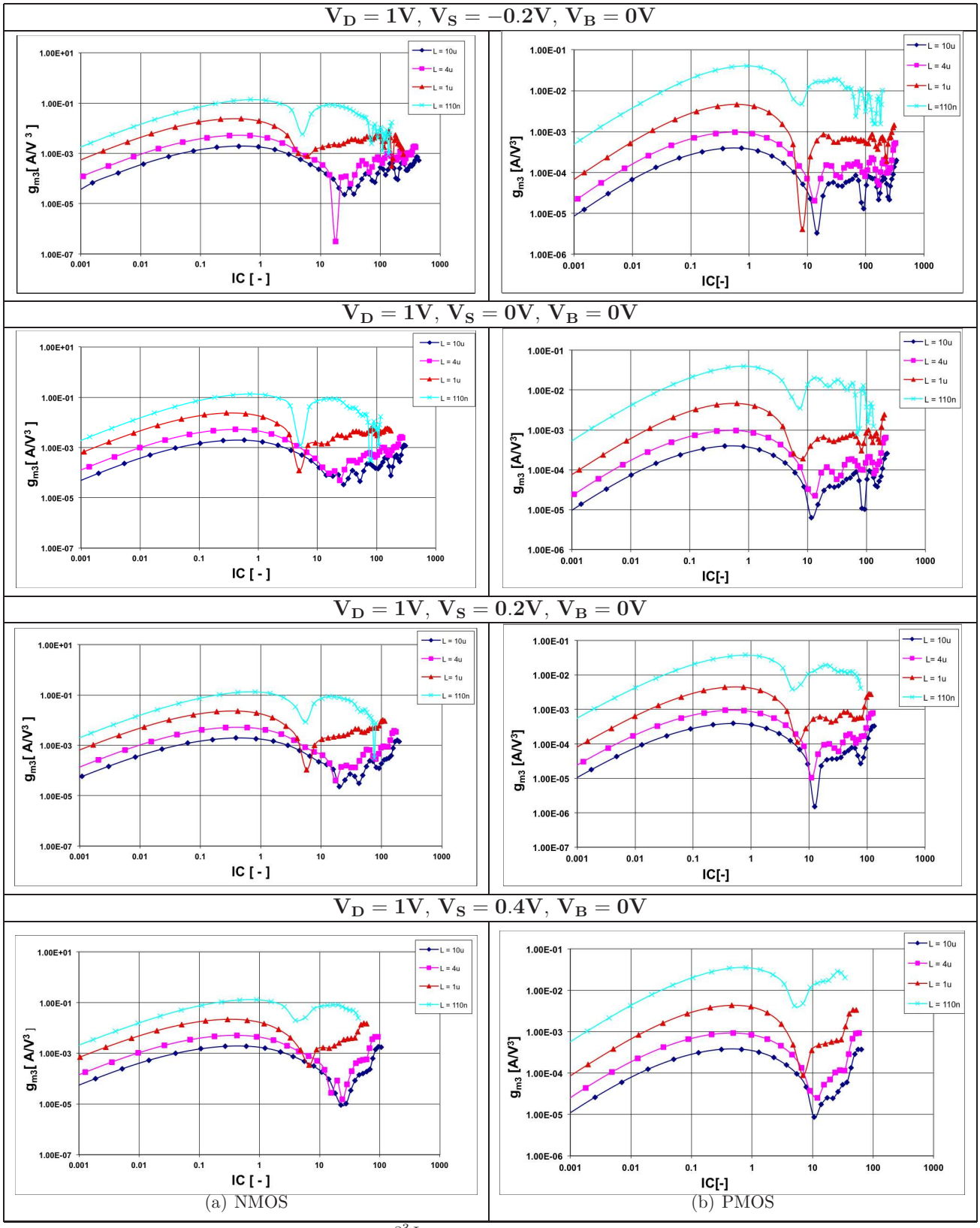
(a) NMOS (b) PMOS

Figure 4.4.6: Measured values of $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) as a function of the inversion coefficient (IC) for various bias conditions. (a)NMOS (b)PMOS.

By observing Figure 4.4.5 - 4.4.6, we can already draw some general conclusions about the behavior of $\frac{\partial^2 I_D}{\partial V_G^2}$. For low source voltage values the maximum value of g_{m2} occurs either in high moderate or low strong inversion region. Further increase of V_S does not change drastically the maximum point of the second-order derivative. Additionally, by comparing the results for different gate lengths we can observe that short channel transistors undergo a more sudden decrease when entering to the deep S.I compared to the other gate lengths. The reason for this behavior is the influence of the second-order effects to the short-transistor specifically as we mentioned previously.

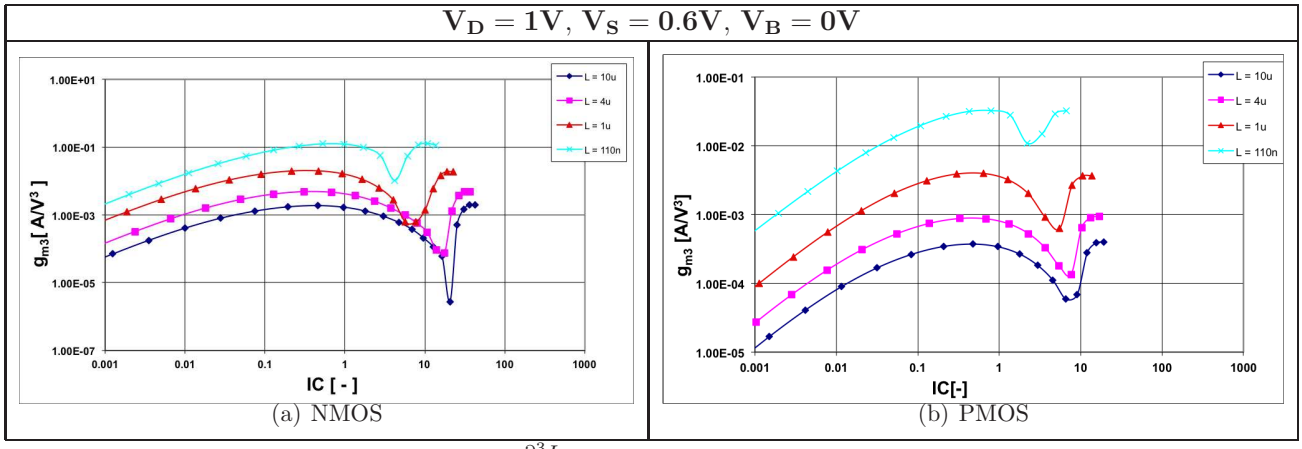
Finally, we will present the measured values of the $\frac{\partial^3 \mathbf{I_D}}{\partial \mathbf{V_G^3}}$ ($\mathbf{g_{m3}}$) as a function of the inversion coefficient, for different bias conditions and at the end we will discuss how g_{m3} is affected by channel length scaling, the inversion coefficient and the various bias conditions.

- **g_{m3} vs. IC**



(a) NMOS (b) PMOS

Figure 4.4.7: Measured values of $\frac{\partial^3 I_D}{\partial V_G^3}$ (g_{m3}) as a function of the inversion coefficient (IC) for various bias conditions. (a)NMOS (b)PMOS. (Cont.)



(a) NMOS (b) PMOS

Figure 4.4.8: Measured values of $\frac{\partial^3 I_D}{\partial V_G^3}$ (g_{m3}) as a function of the inversion coefficient (IC) for various bias conditions. (a)NMOS (b)PMOS.

The third-order term of the drain current, g_{m3}, is especially troublesome for RF systems, since it may lead to intermodulation, i.e distortion of the fundamental amplitude via signals in the adjacent bands. However, for proper DC-bias conditions can be made equal to zero. For this reason, the above experimental results are extremely important. For low source voltage values, the minima of g_{m3} appears to be between high M.I and low S.I region depending on the channel length. However, we notice substantial deviations of g_{m3} function which lead to more local minimums. As V_S becomes more positive, one local minimum exists for each channel length and occurs in the above mentioned boundaries. As a result, we come to the conclusion that for low distortion operation, the optimal inversion region is between high moderate and low strong inversion region. Finally, we must note that short-channel transistors suffer more from non-linearities since g_{m3} is greater than on other transistors.

4.5 Source Transconductance

The source transconductance g_{ms} is the ratio of the current change to the source voltage (V_S) change and is given by:

$$g_{ms} = \frac{\partial I_D}{\partial V_S}|V_{B,G} \tag{4.32}$$

Source transconductance can also be expressed as function of the other transconductances as follows:

$$g_{ms} = g_m + g_{mb} + g_{ds} \simeq g_m + g_{mb} \tag{4.33}$$

In various occasions g_{ms} is expressed by the ratio of g_m to g_{ms}, ($\frac{g_m}{g_{ms}}$). The normalized form of g_{ms} which we will use throughout this section is defined by:

$$G_S = \frac{g_{ms} U_T}{I_D}, \text{where } U_T \text{ is the thermal voltage.} \tag{4.34}$$

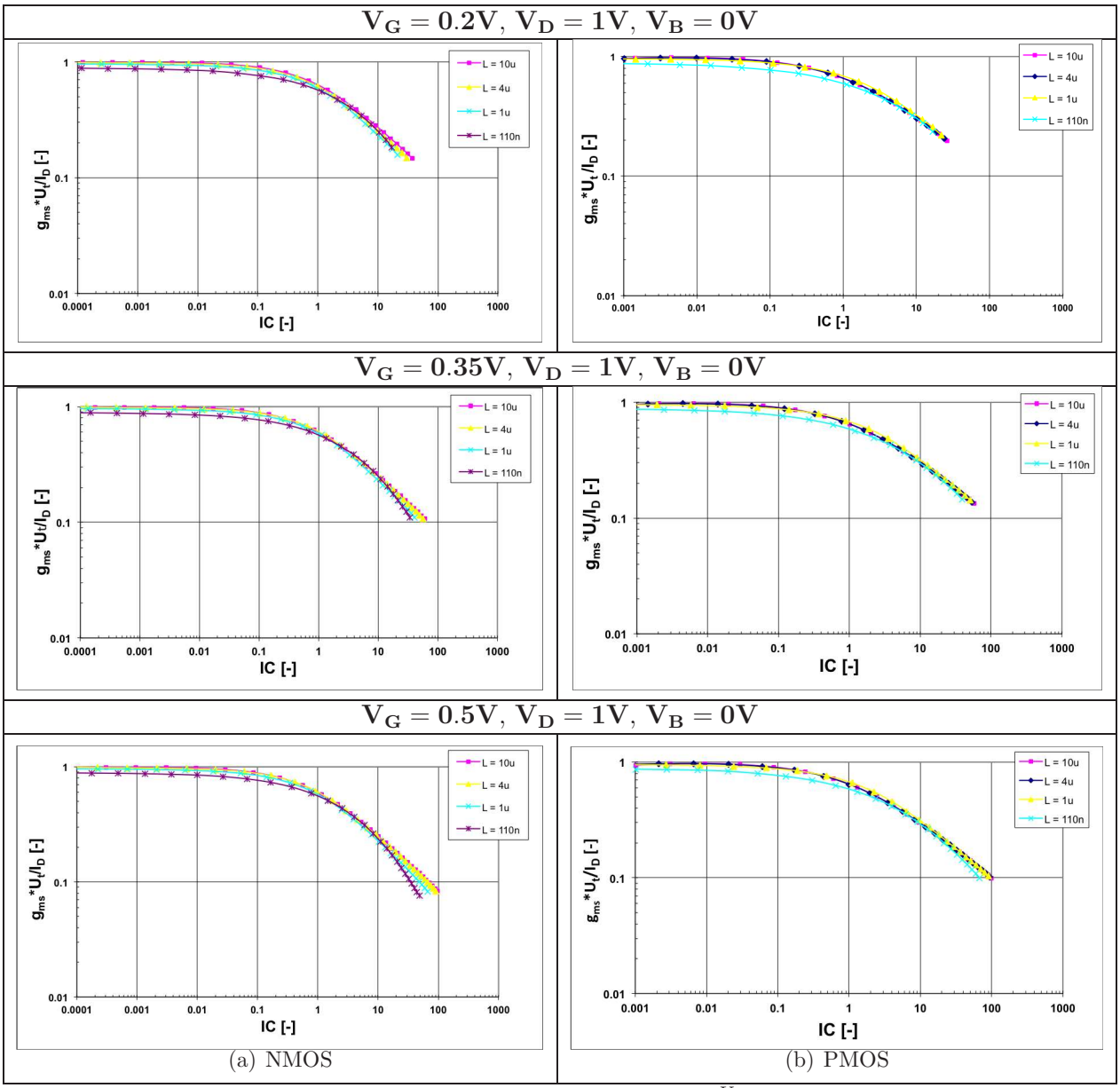
By observing the Eq. 4.5.3, we come to the conclusion that g_{ms} is affected by the same second-order effects which affect g_m and g_{mb}. The circuit designer's aim is to keep the value of g_{ms} as high as possible in order to achieve low input resistance.

Throughout the next subsection we will present the measurement results concerning the normalized source transconductance, the 1^{st}, 2^{nd} and 3^{rd} order derivatives of the drain current as a function of the inversion coefficient (IC).

4.5.1 Measurement results and Discussion

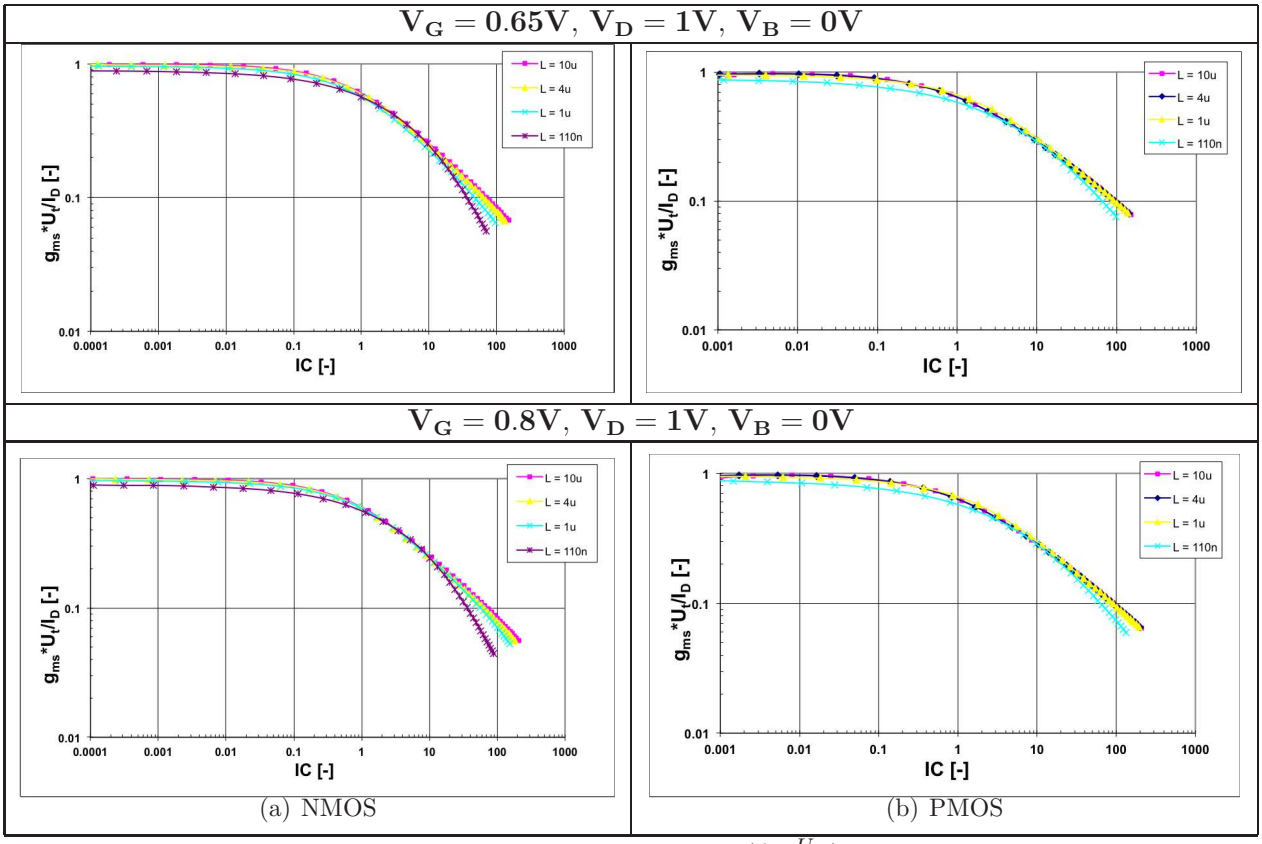
Having calculated the function G(IC) which shows the expected behavior of the normalized source transconductance, we were able to calculate $\frac{g_{ms}U_T}{I_D}$ for all the DUTs which are shown in Fig. 4.4.1. In the following figures we present the calculation results for various bias conditions illustrating the influence of second-order effects.

- $\frac{\mathbf{g_{ms}U_T}}{\mathbf{I_D}}$ **vs. IC**



(a) NMOS (b) PMOS

Figure 4.5.1: Normalized source transconductance ($\frac{g_{ms}U_T}{I_D}$) as a function of the inversion coefficient (IC) for various bias conditions.(a)NMOS (b)PMOS. (Cont.)



(a) NMOS (b) PMOS

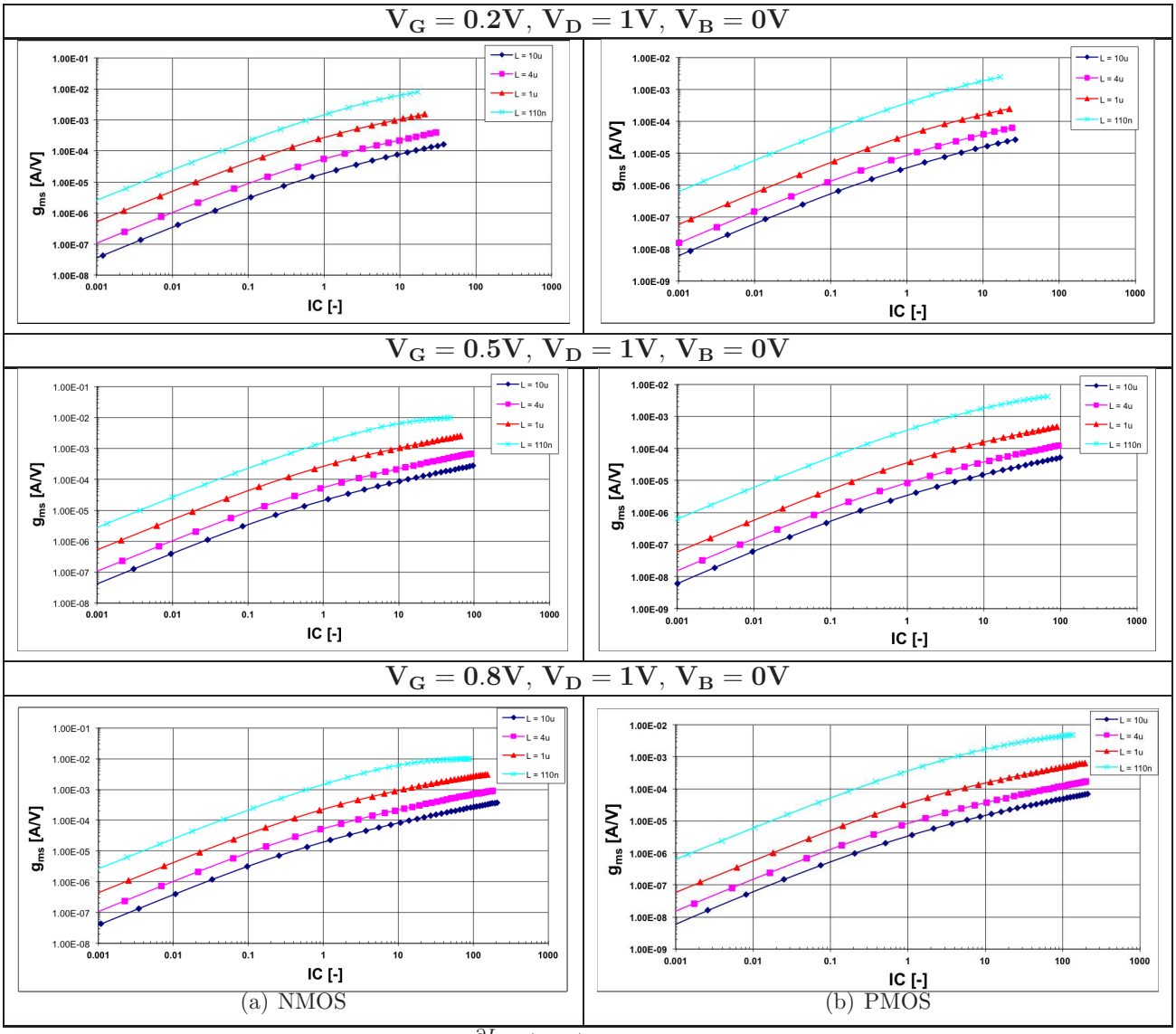
Figure 4.5.2: Normalized source transconductance ($\frac{g_{ms}U_T}{I_D}$) as a function of the inversion coefficient (IC) for various bias conditions.(a)NMOS (b)PMOS.

By carefully examining Figure 4.5.1- 4.5.2 we can draw very important conclusions about the behavior of the normalized source transconductance. First of all, the maximum value occurs in the week inversion region (IC< 0.1) while in the moderate inversion is starting to reduce reaching the lowest value in the deep strong inversion region. This is the reason why the weak and moderate inversion regions are the most preferable. By increasing the gate voltage we notice that lowest value of the normalized g_{ms} is shifted more and more in the deep strong inversion region. Moreover, the normalized g_{ms} of the NMOS short transistor seems to suffer a more sudden decrease when enters in strong inversion.

The behavior of the normalized source transconductance is determined by the second-order effects which were described in Subsection 4.4.1. More specifically, from weak to low moderate inversion region, the STI effect has a great impact especially on the short transistors leading to a lower g_{ms} value than the expected, i.e $\frac{g_{ms}U_T}{I_d} \simeq 1$. By entering the strong inversion, velocity saturation and VFMR are the main effects which lead to a substantial deviation of the optimal value that we calculated with the use of the G(IC) function, especially for NMOS short transistors.

At this point we present the calculation results of $\frac{\partial \mathbf{I_D}}{\partial \mathbf{V_S}}$ ($\mathbf{g_{ms}}$) as a function of the inversion coefficient, for different bias conditions and at the end we discuss how g_{ms2} is affected by the channel length scaling, second-order effects, level of inversion and the different bias conditions.

- $\mathbf{g_{ms}}$ **vs. IC**



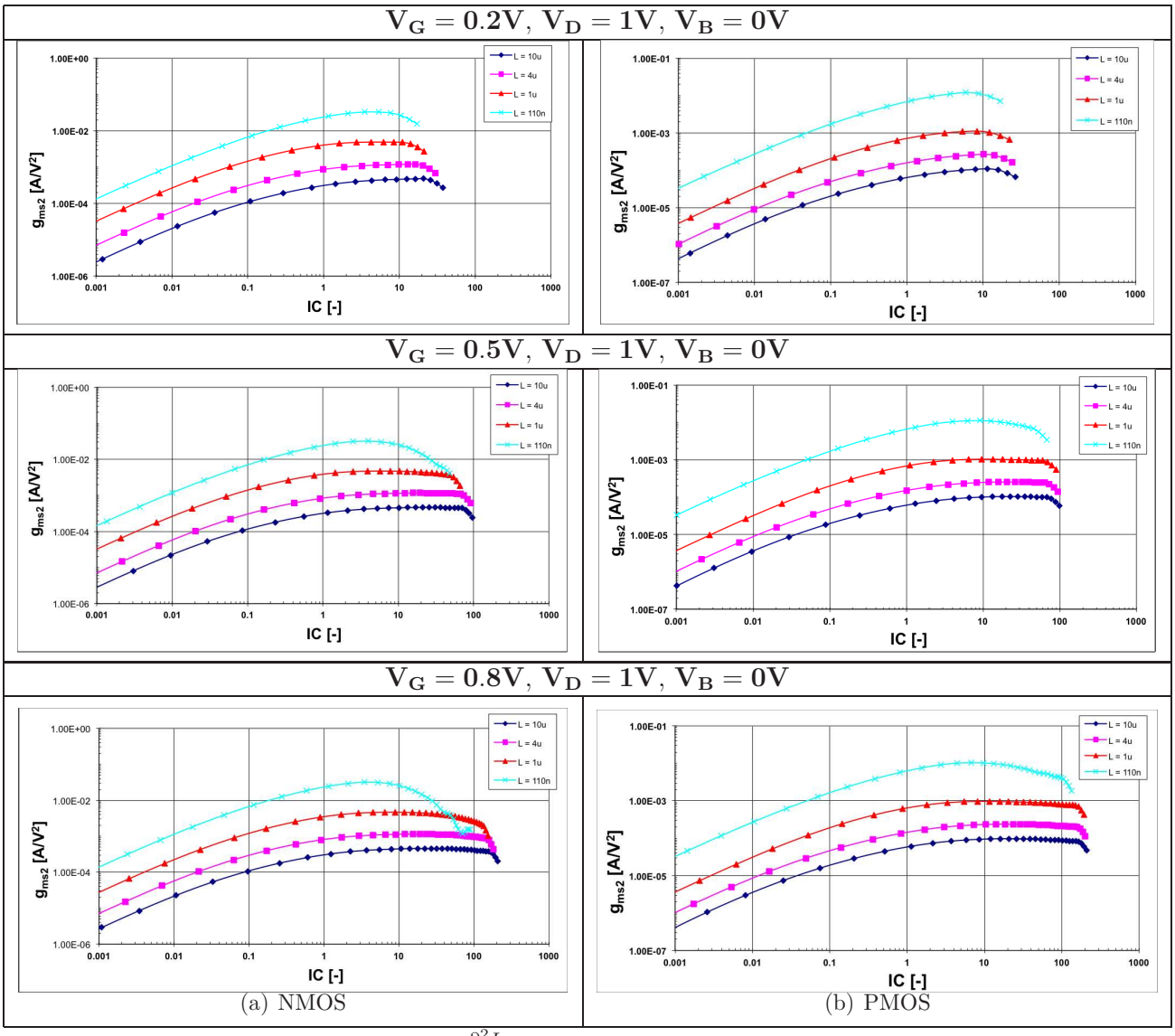
(a) NMOS (b) PMOS

Figure 4.5.3: Measured values of $\frac{\partial I_D}{\partial V_S}$ (g_{ms}) as a function of the inversion coefficient (IC) for various bias conditions.(a)NMOS (b)PMOS

By examining Fig. 4.5.3, we can observe that the maximum value of g_{ms} occurs between low and deep S.I region depending on the channel length and gate voltage value. Additionally, by comparing the results for different gate lengths we can observe that NMOS short channel transistors undergo a more sudden decrease when entering into the S.I compared to the other transistors. The reason for this behavior is the substantial influence of the second-order effects on the short transistor and NMOS transistors in general.

Furthermore, we will present the calculation results of the $\frac{\partial^2 \mathbf{I_D}}{\partial \mathbf{V_S^2}}$ $(\mathbf{g_{ms2}})$ as a function of the inversion coefficient, for different bias conditions and at the end we will discuss how g_{ms2} is affected by channel length scaling, second-order effects, level of inversion and different bias conditions.

- **g_{ms2} vs. IC**



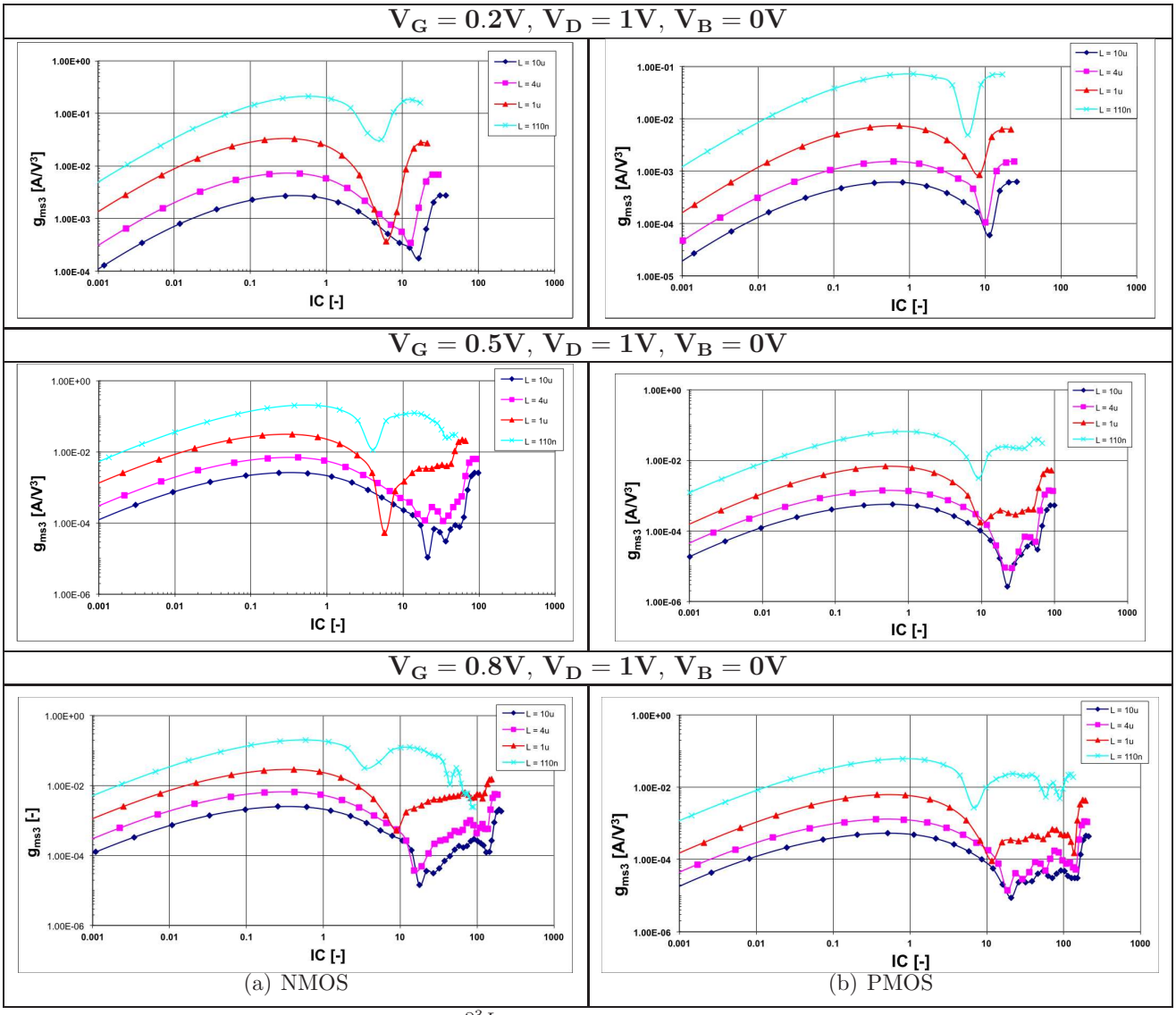
(a) NMOS (b) PMOS

Figure 4.5.4: Measured values of $\frac{\partial^2 I_D}{\partial V_S^2}$ (g_{ms2}) as a function of the inversion coefficient (IC) for various bias conditions.(a)NMOS (b)PMOS

By examining Fig. 4.5.4, we can observe that the maximum value of g_{ms2} occurs between the high moderate inversion and low strong inversion region depending on the channel length and gate voltage value. Additionally, by comparing the results for different gate lengths we can observe that NMOS short channel transistors undergo a more sudden decrease when entering into the deep S.I compared to the other transistors. The reason for this behavior is the influence of the second-order effects on the short-transistor as these were analyzed previously.

Finally, we will present the measured values of $\frac{\partial^3 \mathbf{I_D}}{\partial \mathbf{V_S^3}}$ ($\mathbf{g_{ms3}}$) as a function of the inversion coefficient, for different bias conditions and at the end we will discuss how g_{ms3} is affected by channel length scaling, the level of inversion and the bias conditions.

- **g_{ms3} vs. IC**



(a) NMOS (b) PMOS

Figure 4.5.5: Measured values of $\frac{\partial^3 I_D}{\partial V_S^3}$ (g_{ms3}) as a function of the inversion coefficient (IC) for various bias conditions.(a)NMOS (b)PMOS

By examining Fig. 4.5.5, we can observe that the maximum value of g_{ms3} occurs between the high moderate inversion or low strong inversion region depending on the channel length and gate voltage value. For low gate voltage values, one local minimum exists. However, the increase of V_G has a substantial effect on g_{m3} as it seems to lead in more than one local minima. Additionally, we must note that short-channel transistors suffer more from non-linearities as the g_{ms3} is greater than on other transistors.

4.6 Calculation of Figure of Merits

Depending on the application, different quality criteria must be fulfilled. For example, the quality of an amplifier can be characterized by a number of specifications, usually referred as Figures of Merit (FOMs), a number of which are listed below :

- **Gain**

- Efficiency
- **Linearity**
- Noise
- Output dynamic range
- Overshoot
- Stability

For each quality factor, various Figure of Merits have been proposed [ ]. In this thesis we consider Linearity and Intrinsic Voltage Gain as the FoMs' under investigation.

4.6.1 Intrinsic Voltage Gain

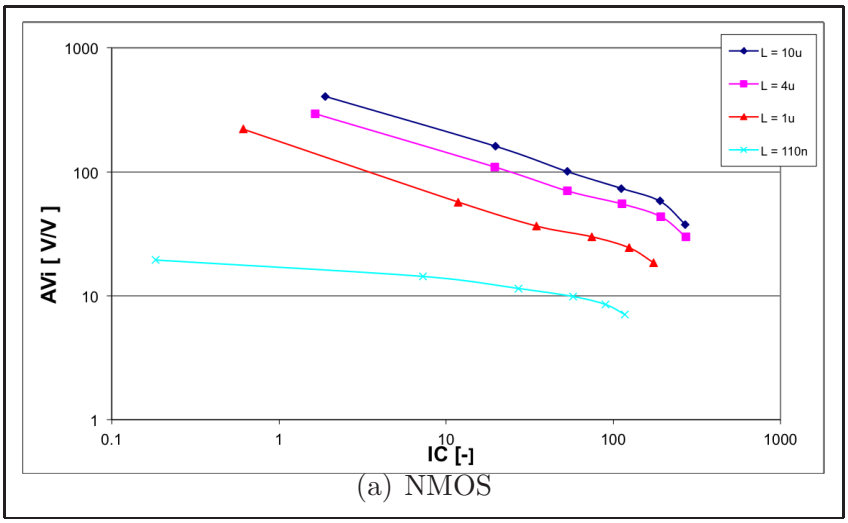
Gain is a measure of the ability of a circuit (often an amplifier) to increase the power or the amplitude of a signal from the input to the output. Taking into consideration that the transistors are the main components of an amplifier we can conclude that a designer would definitely desire the maximum voltage gain of each transistor. For this purpose, the Intrinsic Voltage Gain is written as

$$A_{Vi} = \frac{g_m}{g_{ds}} \tag{4.35}$$

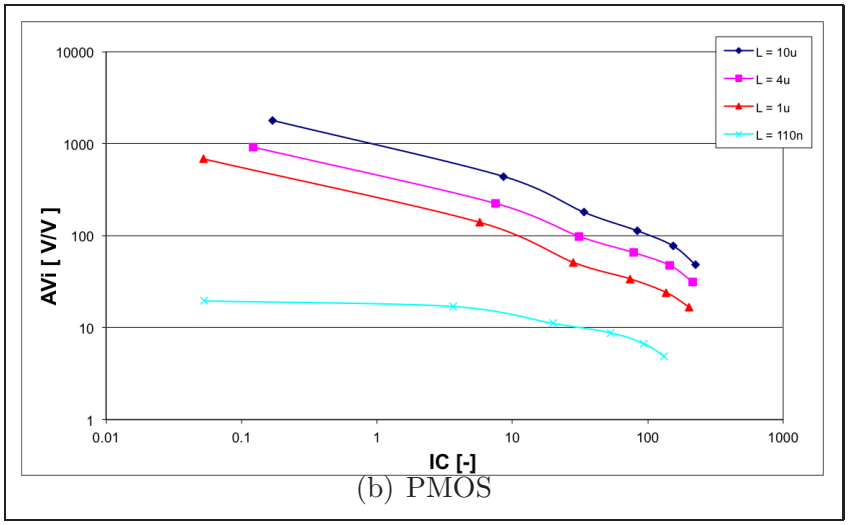
which describes the transistor's maximum voltage gain and as a consequence is a very important FOM in analog circuit design. By observing Eq. 4.35 we come to the conclusion that A_{Vi} would be affected by the same second-order effects that affect gate and output transconductance and are strongly related to the inversion coefficient (IC), the bias conditions and the channel length scaling.

In the following figure we present the calculation results of the Intrinsic Voltage Gain for both NMOS and PMOS devices operating in saturation.

- **A_{Vi} vs. IC**



(a) NMOS



(b) PMOS

Figure 4.6.1: Intrinsic Voltage Gain as a function of the inversion coefficient (IC), $V_G = 200mV, ..., 1.2V$ with step size 200mV, $V_D = 1V$, $V_S = 0V$, $V_B = 0V$.(a)NMOS (b)PMOS

By carefully examining Figure 4.6.1 important conclusions can be drawn. First of all, we notice the dependence of the intrinsic voltage gain on the channel length scaling and level of inversion. More specifically, AV_i undergoes a significant decrease as we move from M.I to S.I region. Second-order effects such as VFMR and velocity saturation are the main causes for this behavior. As a result, a circuit designer should consciously choose moderate inversion region where the gain is maximum. Finally, we must notice that PMOS transistors up to 1um channel length provide significantly bigger intrinsic voltage gain when compared to the NMOS.

4.6.2 Linearity Figure of Merits

As we already noted in Chapter 3, depending on the application and the calculation method, different linearity FOMs have been derived. We referred to the IP_3 is a reliable and simple metric for a direct evaluation of RF linearity. However, a variation of the IP_3 called **Voltage Third Intercept Point (VIP_3)** has been proposed[19] and can be calculated directly from the DC measurements.

In the following subsection, we define the VIP_3 and we present the calculation results for both NMOS and PMOS devices. At the end, we discuss how the channel length L influences the linearity based on the VIP_3 calculation results.

4.6.2.1 Voltage Third Intercept Point

The Voltage Third Intercept Point is defined as:

$$VIP_3 = \sqrt{\frac{24 \cdot g_m}{g_{m3}}} \tag{4.36}$$

Here g_m is the gate transconductance and g_{m3} is the third-order derivative of the drain current with respect to the gate terminal, i.e $\frac{\partial^3 I_d}{\partial V_G^3}$; both calculated using the integrated derivative function of IC-CAP 2008. VIP_3 is the extrapolated input gate bias at which 1^{st} and 3^{rd}-order amplitudes of the drain current are equal. We must note that **for low-distortion operation**

VIP_3 should be as high as possible.

In the following figure we present the calculated results of the VIP_3 as a function of the Inversion Coefficient:

- **VIP_3 vs. IC**

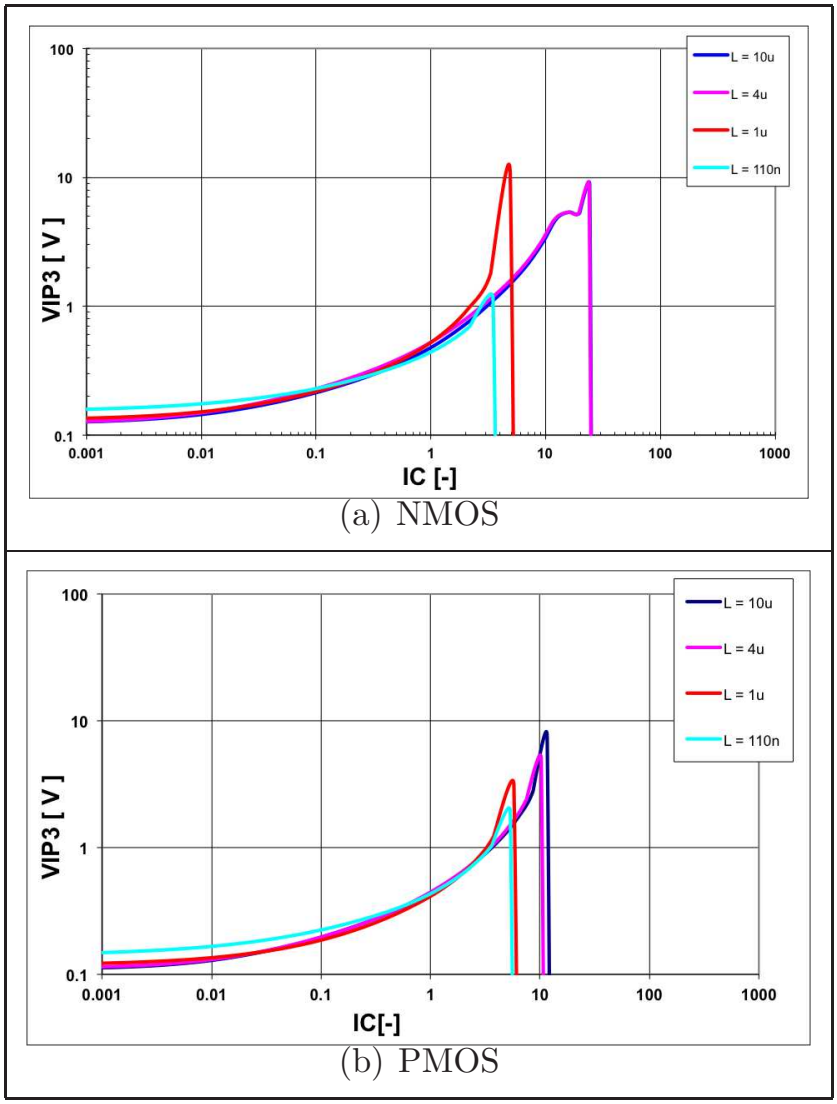


Figure 4.6.2: Voltage Third Intercept Point (VIP_3) as a function of the inversion coefficient (IC), $V_D = 1V$, $V_S = 0V$, $V_B = 0V$. (a)NMOS (b)PMOS

In general, with increasing gate bias, VIP_3 increases to a maximum value (i.e $g_{m3} = 0$) followed by a local minimum. With decreasing channel length the maximum in VIP_3 shifts to lower V_{GS} and its local minimum decreases. However, by observing Fig. 4.6.2, we notice that at the point where VIP_3 reaches the maximum value, it becomes zero. The reason for this, is that g_{m3} in modern CMOS technologies seems to become zero after a certain voltage gate value. As a result, modifications should be made to the VIP_3 calculation formula, if possible. However, even though we cannot have a detailed view on all inversion regions, the above mentioned inaccuracy does not make VIP_3 an irrelevant figure of merit. On the contrary, we can observe that the highest possible linearity is achieved between high moderate and low strong inversion

region depending on the channel length, which could be a major advantage considering that with the channel length scaling, gate bias conditions around threshold voltage will become more common place, implying that MOS transistors will be used principally in the moderate inversion. Finally, as far as the NMOS results are concerned, we can observe that by decreasing the channel length, the local minimum of the transistor with channel length 1um increases instead of decreasing. This inaccuracy does not cancel the importance of VIP_3 as a linearity figure of merit.

Chapter 5

Simulation Results and Discussion

The growth of circuit complexity has made indispensable the use of computer-aided simulation tools, the so called *circuit simulators*. These simulators use mathematical models for the quantitive description of the terminal behavior of the circuit elements allowing the circuit designer to predict and optimize the circuit behavior before the circuit is realized in silicon. The above mentioned models are often referred to as *compact models*. The success of a circuit simulator depends not only on the accuracy of the compact models used, but also on the efficiency, i.e the simulation time.

We can categorize the type of compact mosfet models based on the nature of the equations they use to simulate the behavior of the transistor. First of all, we have the *physics-based* models which use analytical expressions that have been derived directly from the device physics. Another category are the *empirical* models of which the relations are of a curve fitting nature and finally, the *table look-up* models where the characteristics are reconstructed via the measured data. The majority of the compact models used in circuit simulators are physics-based models because they make use of important physical parameters which obey well-defined geometrical and temperature scaling rules and allow statistical modeling.

More specifically, a good MOSFET model should satisfy the following criteria:

- An accurate prediction of the I-V characteristics and a fulfillment of the charge conservation principle.
- A good description of non quasi-static operation.
- An accurate prediction of both white and 1/f noise.
- An accurate description of all small-signal quantities: transconductance $g_m = \frac{\partial I_D}{\partial VGS}$, conductance $g_d = \frac{\partial I_D}{\partial VDS}$, substrate transconductance $g_b = \frac{\partial I_D}{\partial VSB}$ and all capacitances.
- Fulfillment of the above criteria over a large range of bias conditions, device geometries and temperature values.
- The model uses one parameter set which contains as few parameters as possible, is completely scalable and can be obtained using an efficient parameter extraction method.
- An accurate description of distortion behavior.

5.1 EKV3 Model

The EKV3 model is a compact mosfet model designed in order to have ease of parameter extraction, computational efficiency and an insight look into the device behavior. More specifically, EKV3 is a charged-based model which first calculates the dependency of the density Q_i of induced mobile charge on the voltages applied to the transistor. Then, it relies on Q_i, and on its particular values Q_{i_S} and Q_{i_D} at the source and drain ends of the channel, to calculate the drain current and to model all aspects of the device behavior.

The natural effects that occur inside the MOSFET are determined using a number of parameters, related to the technology used for the construction and the materials used. The total of the parameters can be divided into two subtotals: those that refer to the instance parameters of the transistor (gate width and length, number of fingers in a multi-finger transistor etc.), defined by the circuit designer, and those that are independent of specific dimensions, not defined by the designer and extracted through a wide range of measurements.

Although EKV3 model, has the fewest parameters to extract, comparatively to other models, lots of phenomena that influence parameters' optimal value exist making this procedure a quite hard task. However, an analytical methodology which summarizes and simplifies the extraction procedure with respect to the MOSFET physics has been proposed [20].

5.2 Evaluation of the EKV3 MOSFET model

The purpose of a good model is to predict accurately a transistor's behavior of a specific technology under all conditions and independently of structural characteristics of the transistor such as width, length and number of fingers. Our thesis is focused on the investigation of the nonlinearities, so a good MOSFET model should not only predict the I-V characteristics, but also predict the higher-order derivatives of the drain current which are the dominant sources of the non-linear behavior.

In the following table, we present the Devices Under Test (DUT) for the evaluation of the EKV3 MOSFET model.

DUT	Width	Length	Number of Fingers
NMOS\PMOS	10u	10u	1
	10u	4u	1
	10u	1u	1
	10u	110n	1

Figure 5.2.1: DUTs for evaluation of EKV3 MOSFET model

5.2.1 Linear Region

Even though our measurements were focused mainly in saturation region, we present above some indicative figures associated with the linear operation of the MOSFET.

- I_D **vs.** V_G

  - **NMOS**(Linear scale)

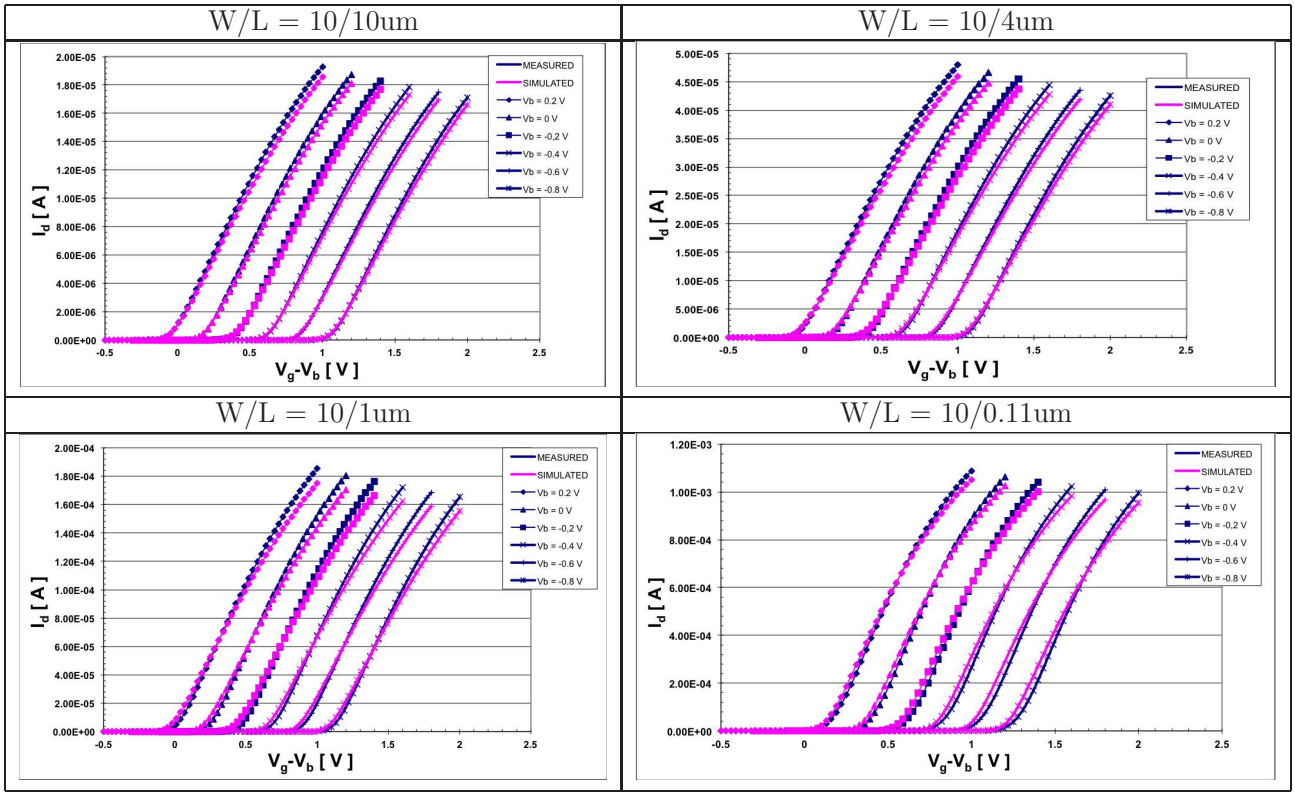


Figure 5.2.2: Measurement and simulation with EKV3.0 for I_D versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. NMOS transistor, $N_f = 1$.

  - **PMOS**(Linear scale)

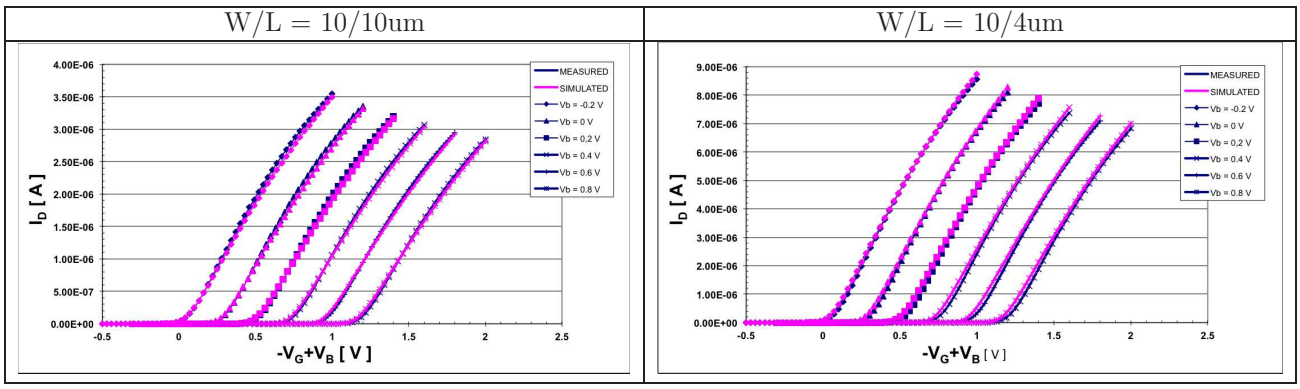


Figure 5.2.3: Measurement and simulation with EKV3.0 for I_D versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, -V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. PMOS transistor, $N_f = 1$. (Cont.)

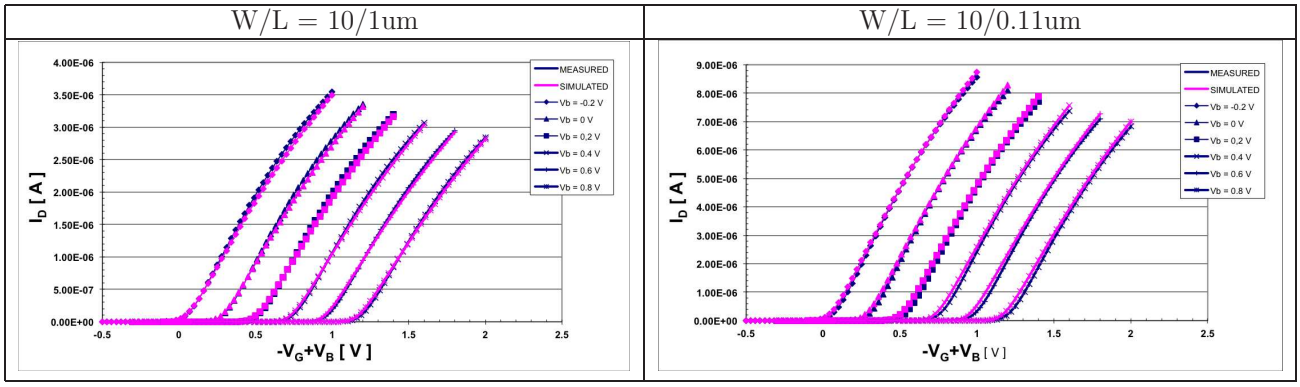


Figure 5.2.4: Measurement and simulation with EKV3.0 for I_D versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, -V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. PMOS transistor, $N_f = 1$.

- g_m **vs.** V_G

  - **NMOS**

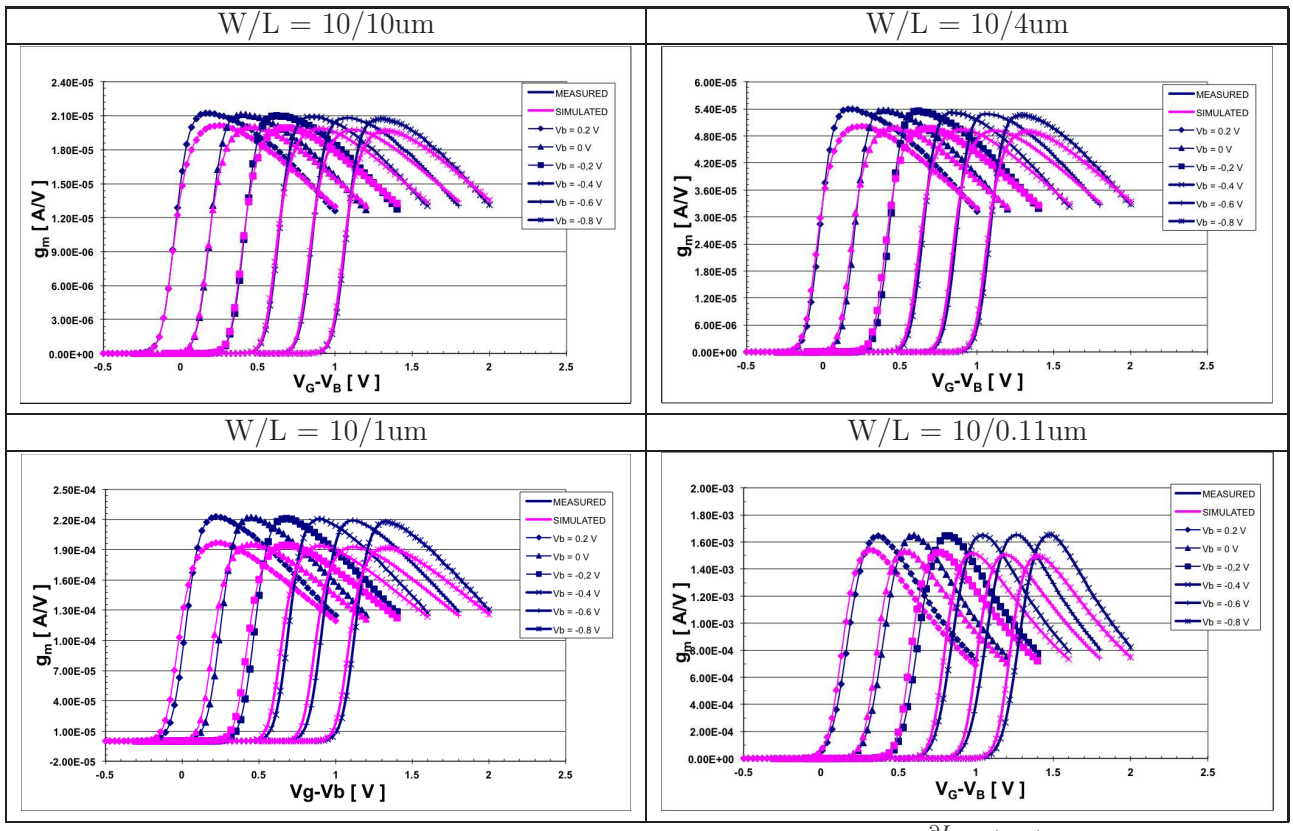


Figure 5.2.5: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_G}$ (g_m) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. NMOS transistor, $N_f = 1$.

– **PMOS**

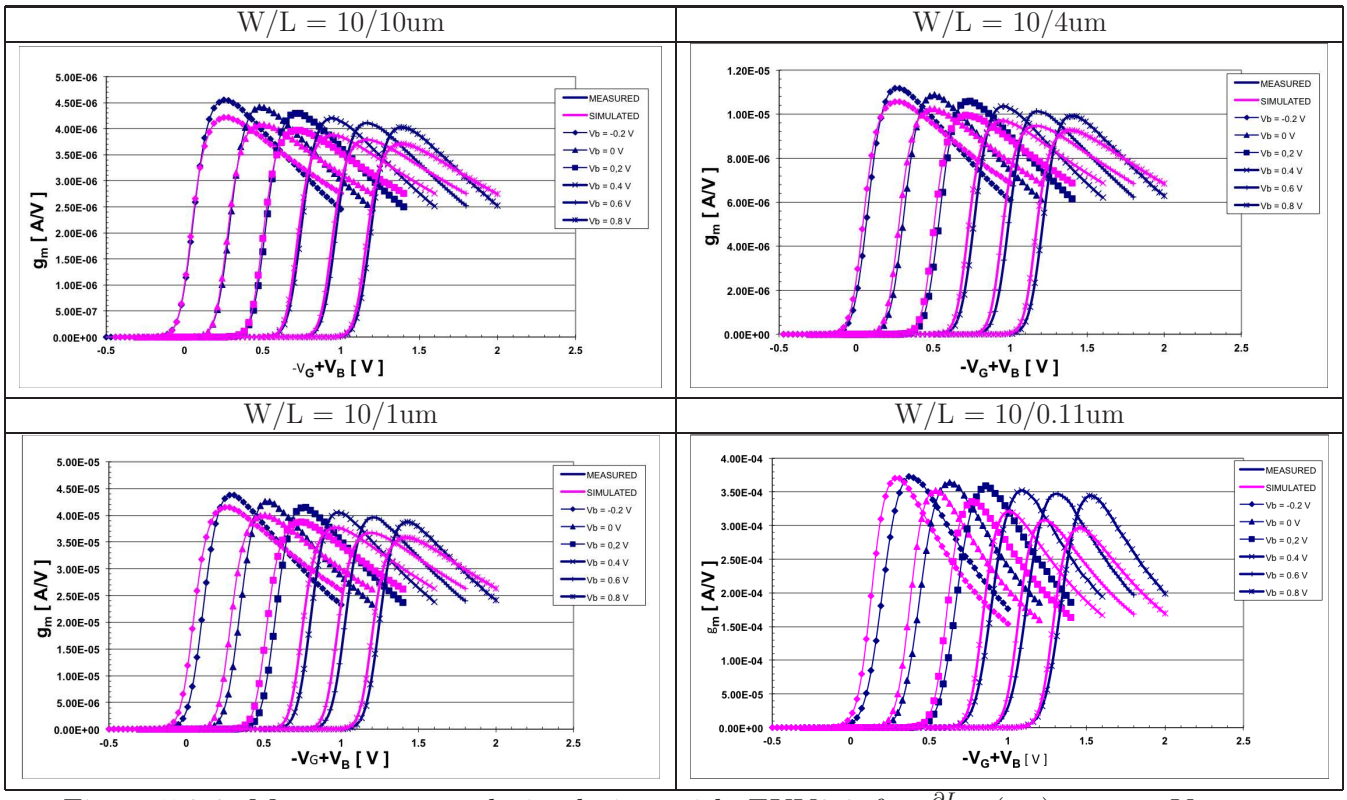


Figure 5.2.6: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_G}$ (g_m) versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, -V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. PMOS transistor, $N_f = 1$.

- g_{m2} **vs.** V_G

– **NMOS**

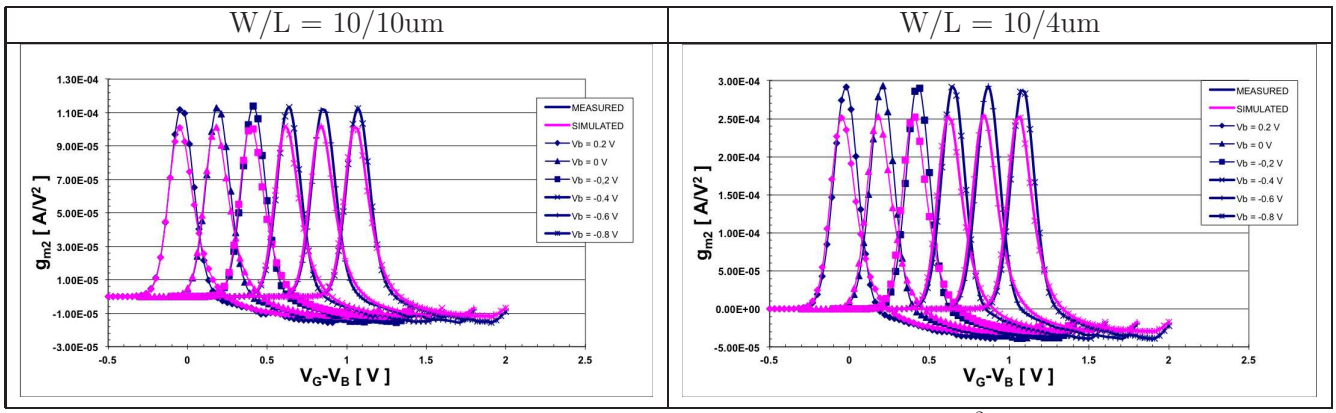


Figure 5.2.7: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. NMOS transistor, $N_f = 1$. (Cont.)

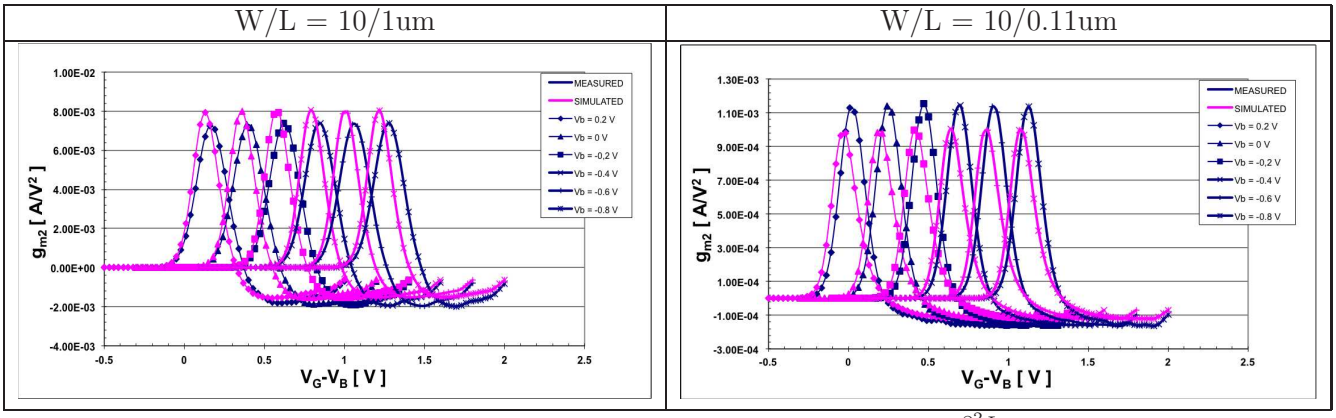


Figure 5.2.8: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. NMOS transistor, $N_f = 1$.

– **PMOS**

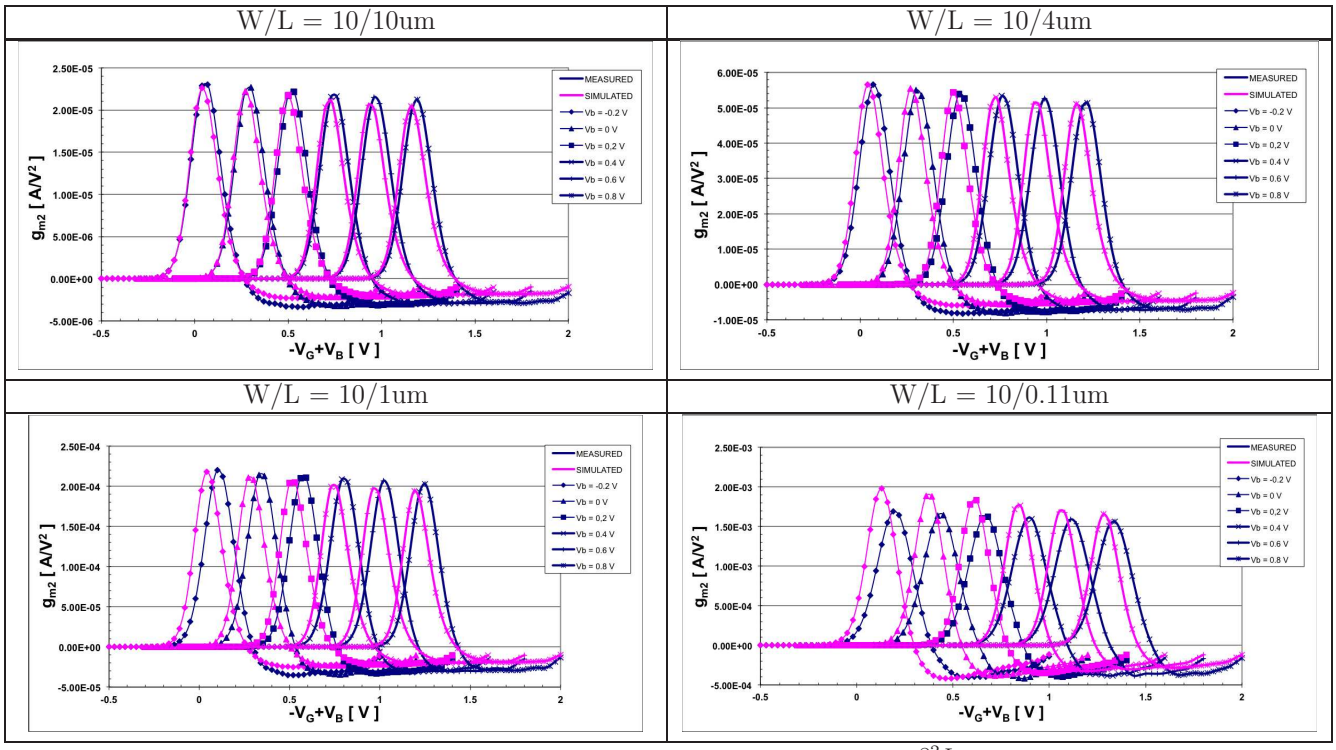


Figure 5.2.9: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, -V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. PMOS transistor, $N_f = 1$.

- g_{m3} **vs.** V_G

– **NMOS**

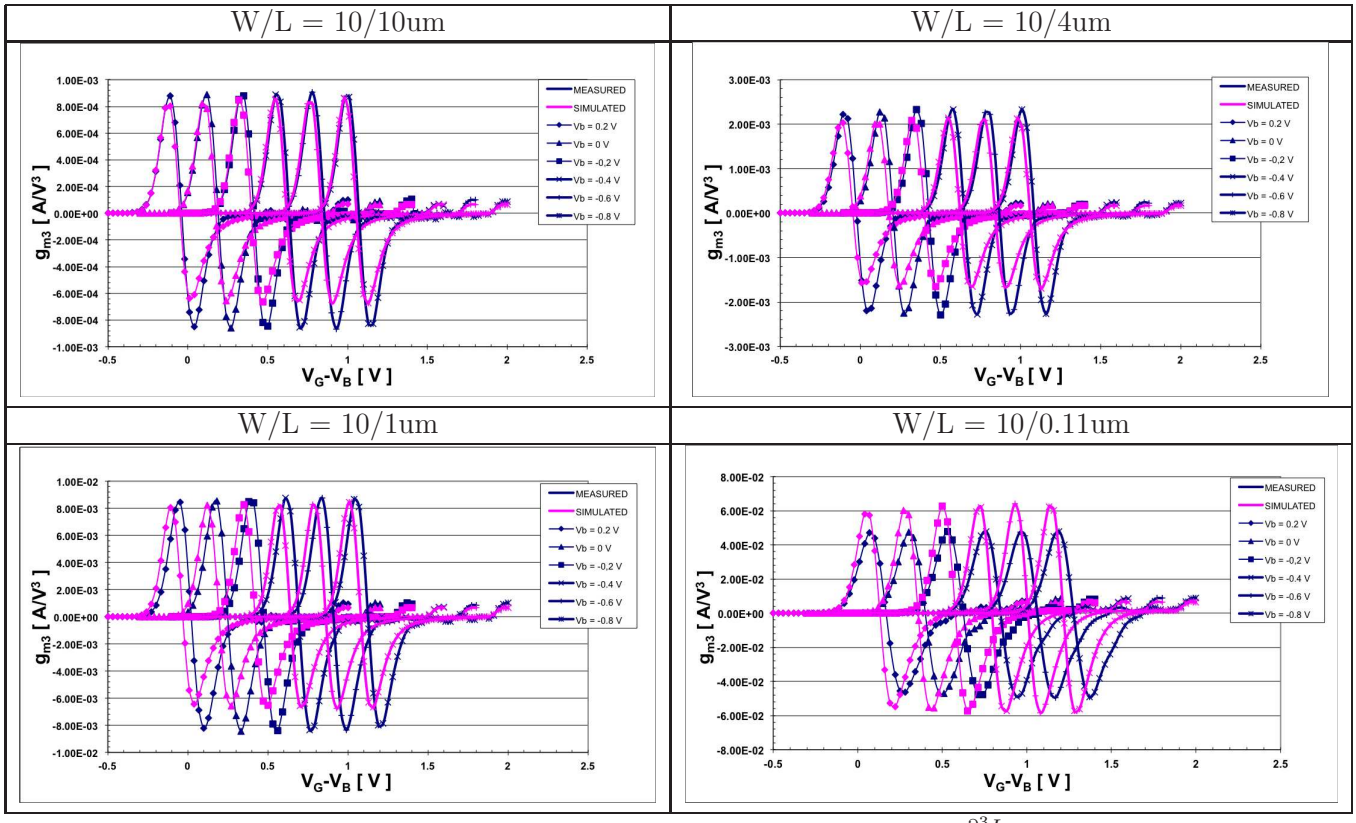


Figure 5.2.10: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_G^3}$ (g_{m3}) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. NMOS transistor, $N_f = 1$.

– **PMOS**

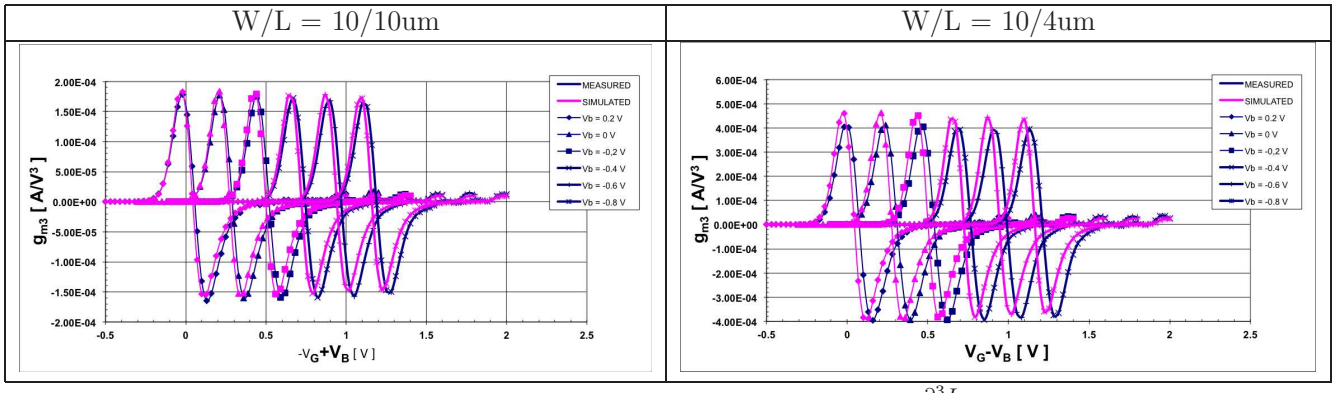


Figure 5.2.11: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_G^3}$ (g_{m3}) versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, -V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. PMOS transistor, $N_f = 1$. (Cont.)

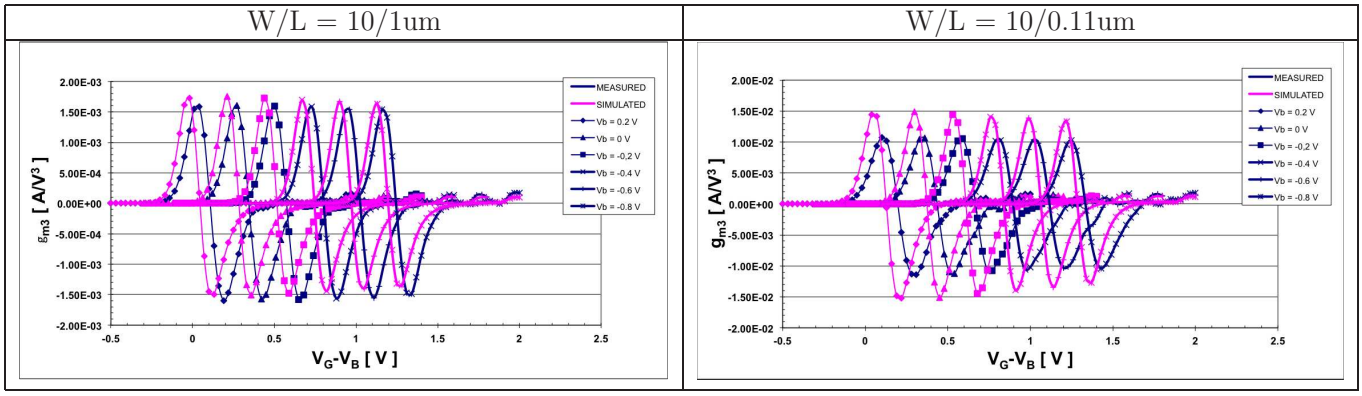


Figure 5.2.12: Measurement and simulation with EKV3.0 for g_{m3} versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, V_S = 0V, -V_D = 50mV, V_B = -80mV, -60mV, ..., 0.2V$. PMOS transistor, $N_f = 1$.

In this subsection we focused on the higher-order derivatives of the drain current (I_D) in the ohmic region, with respect to the gate voltage and we presented the results for both NMOS and PMOS devices for a number of representative gate lengths.

For the long channel n-type MOS transistor (L = 10um) where the series resistance effects are negligible and the derivatives are mainly determined by mobility degradation, we notice that EKV3.0 is very accurate. More specifically, our model predicts the peaks of g_m, g_{m2} and g_{m3}, nevertheless some adjustments should be made to have a perfect matching. As far as the PMOS transistor is concerned, we notice that we have better results than the n-type transistor and the model predicts more accurately the higher-order derivatives for all range of gate voltages.

As we move to semi-short and short channel devices, the effect of series resistance dominates the effect of mobility degradation. The higher-order terms of the drain current are increasing significantly, but the general behavior of these terms does not change very much with decreasing channel length L. As a result, the EKV3.0 model describes quite well their behavior but for the short channel device (L=0.11um) the model predicts higher values for these orders. As a result, improvements should be made to fit accurately all higher-order terms for all gate lengths and a wide range of gate voltages.

5.2.2 Saturation

The behavior of a single transistor is determined by the higher-order derivatives of the drain current I_D with respect to any terminal, so in this section we will present the results of the higher-order terms with respect to gate, source and drain terminals in saturation region. As we noted in Chapter 3, the output conductance also has a dependence in the operation point as well as frequency, hence introduce further nonlinearities. A good prediction of the g_{ds}, g_{ds2} and g_{ds3} is very important because often the load device in an amplifier is often a drain terminal driven MOSFET operating in saturation.

- I_D **vs.** V_G

– **NMOS**(Linear scale)

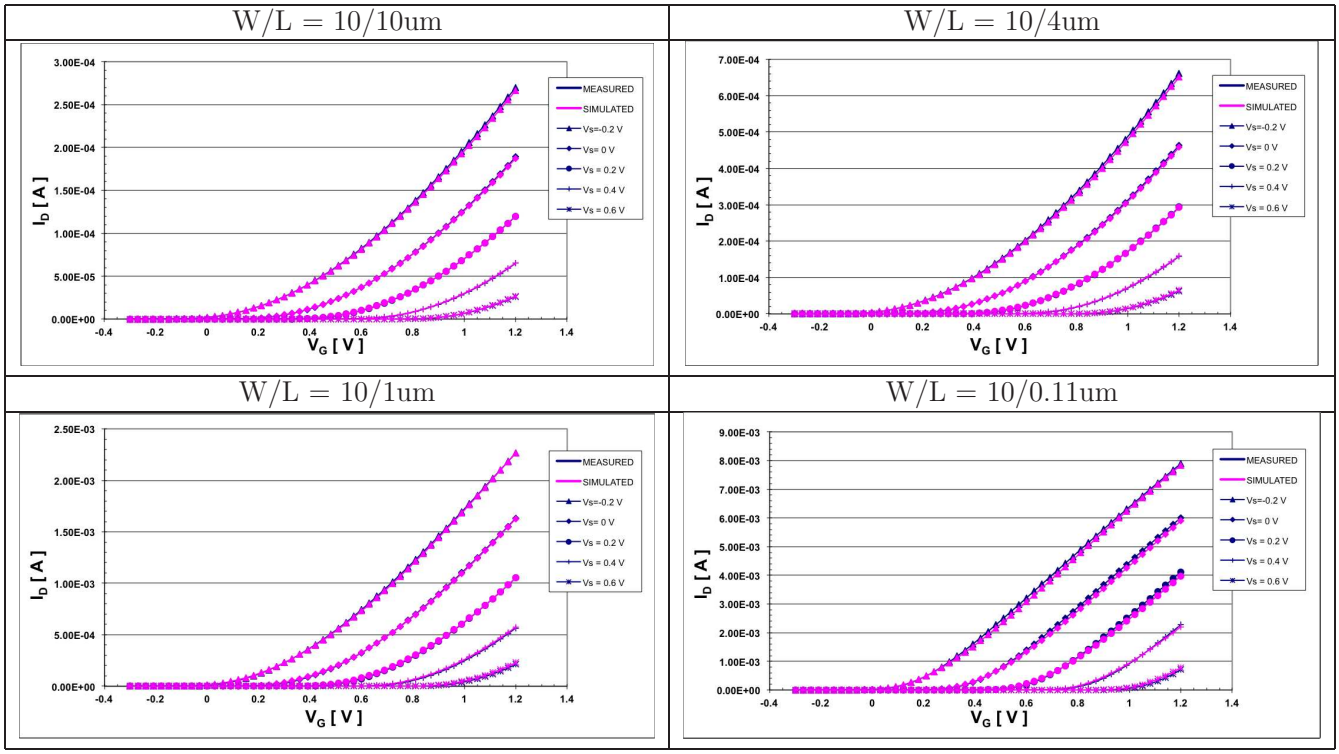


Figure 5.2.13: Measurement and simulation with EKV3.0 for I_D versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = -200mV, ..., 0.6V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**(Linear scale)

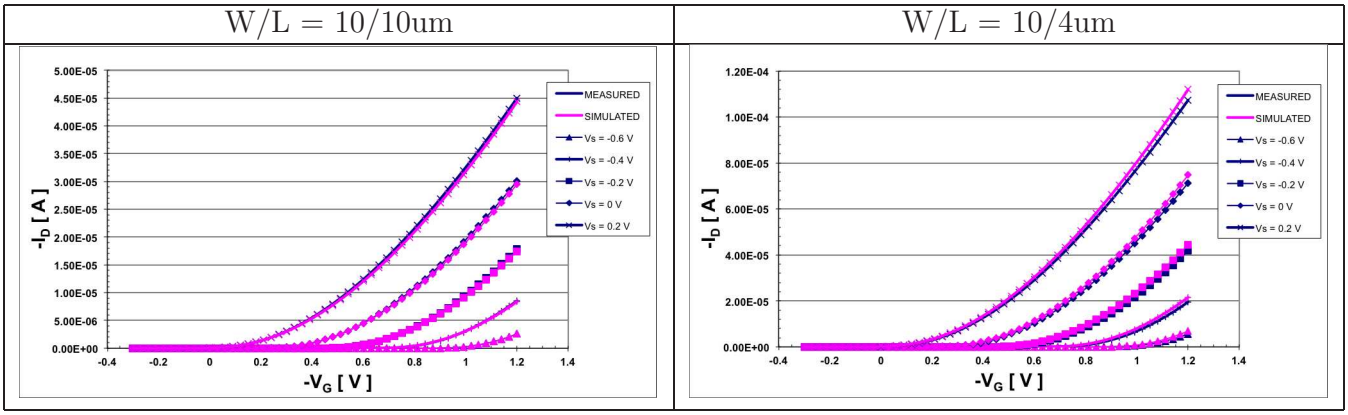


Figure 5.2.14: Measurement and simulation with EKV3.0 for I_D versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, -V_S = -200mV, ..., 0.6V, -V_D = 1V, -V_B = 0V$. PMOS transistor, $N_f = 1$. (Cont.)

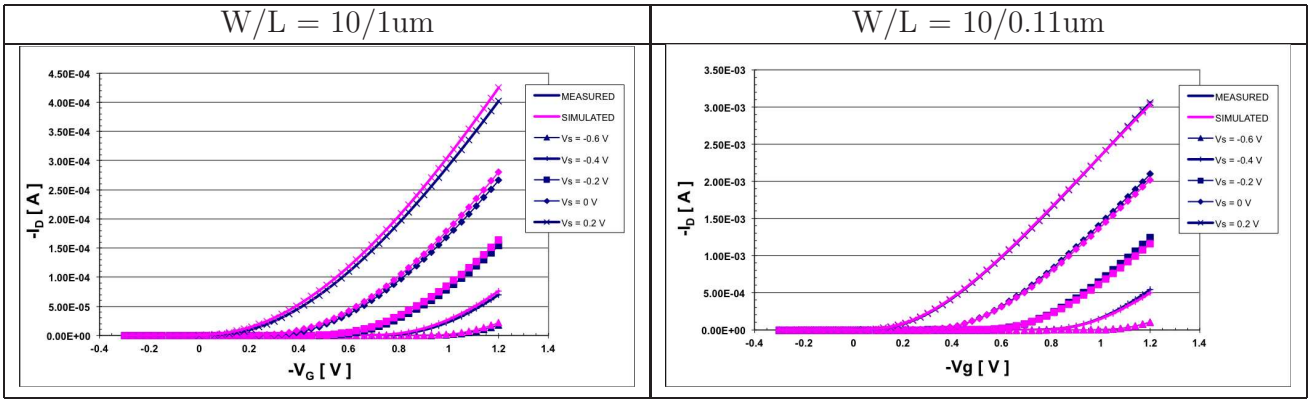


Figure 5.2.15: Measurement and simulation with EKV3.0 for I_D versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, -V_S = -200mV, ..., 0.6V, -V_D = 1V, -V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_m **vs.** V_G

  – **NMOS**

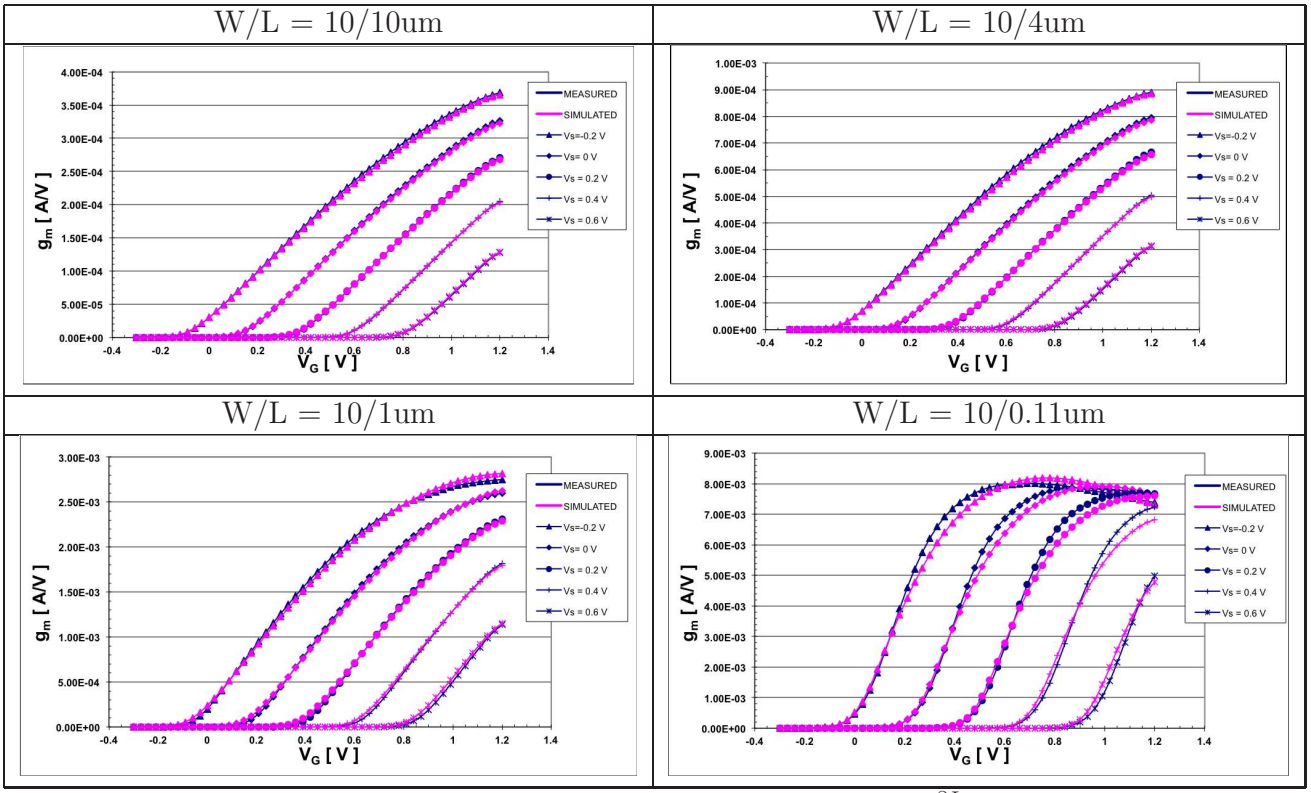


Figure 5.2.16: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_G}$ (g_m) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = -200mV, ..., 0.6V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**

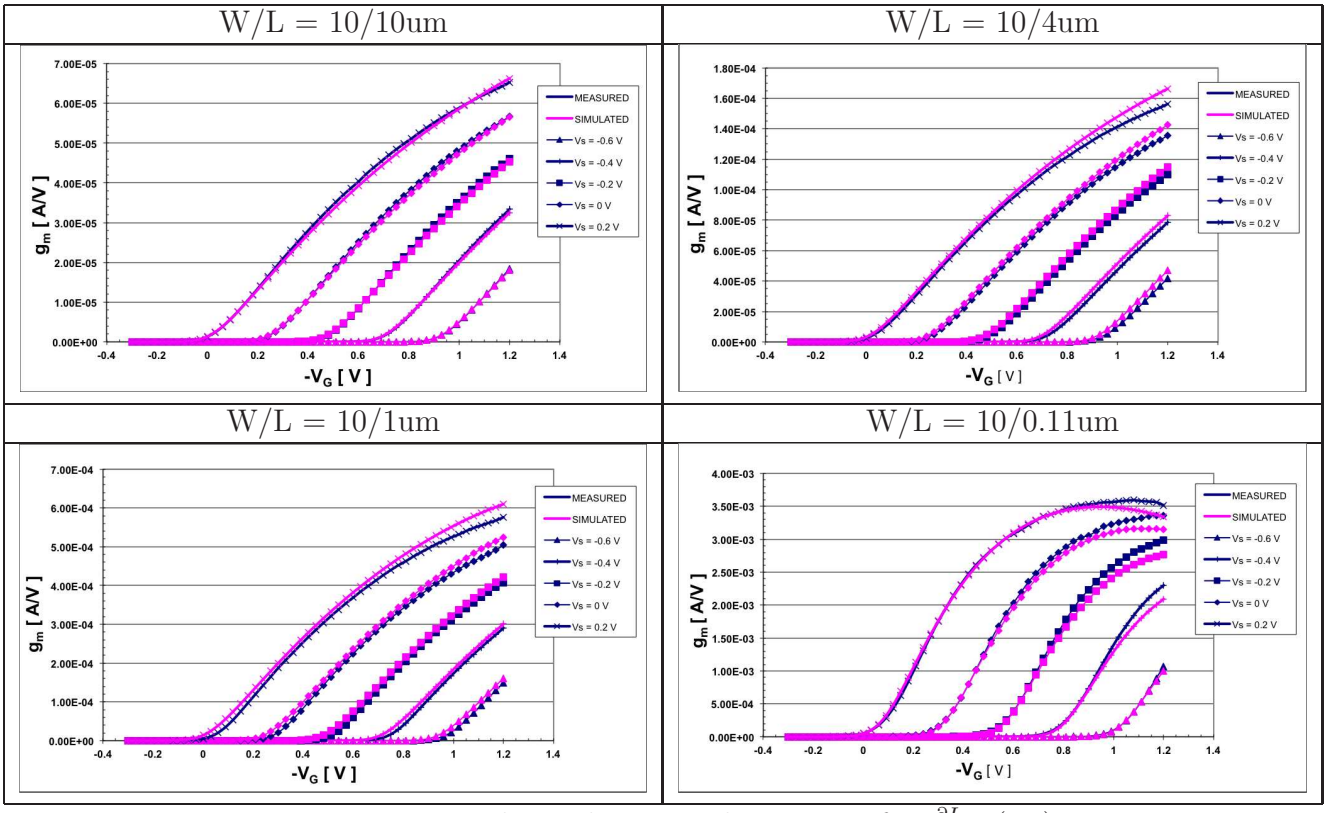


Figure 5.2.17: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_G}$ (g_m) versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, -V_S = -200mV, ..., 0.6V, -V_D = 1V, -V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_{m2} **vs.** V_G

  – **NMOS**

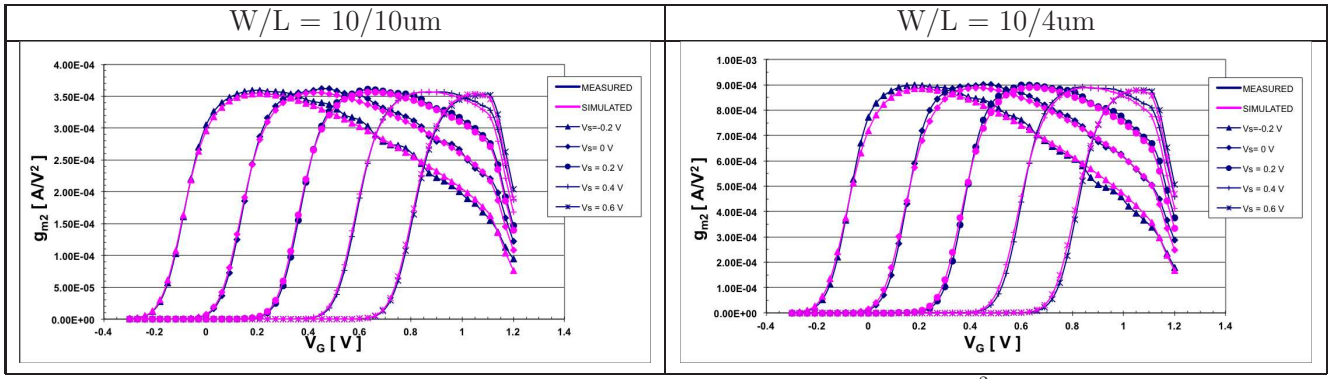


Figure 5.2.18: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = -200mV, ..., 0.6V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$. (Cont.)

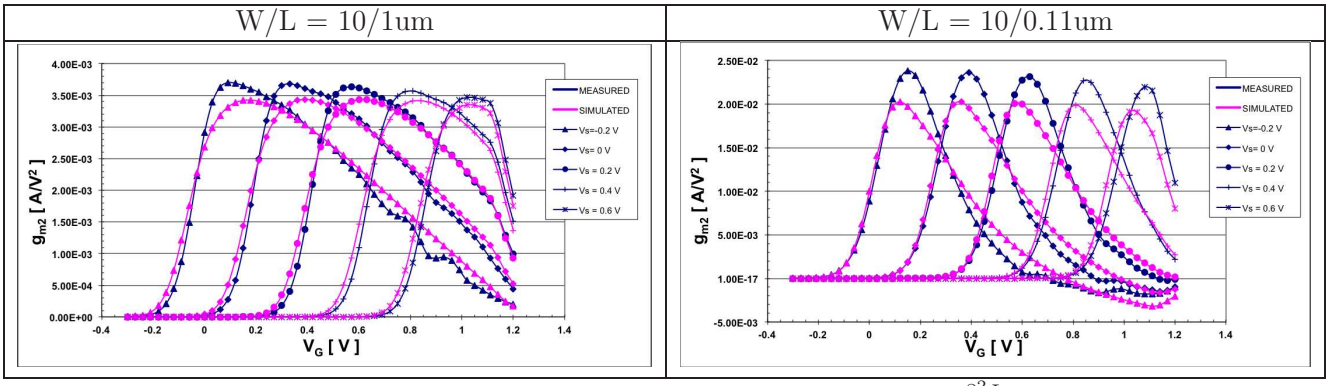


Figure 5.2.19: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = -200mV, ..., 0.6V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**

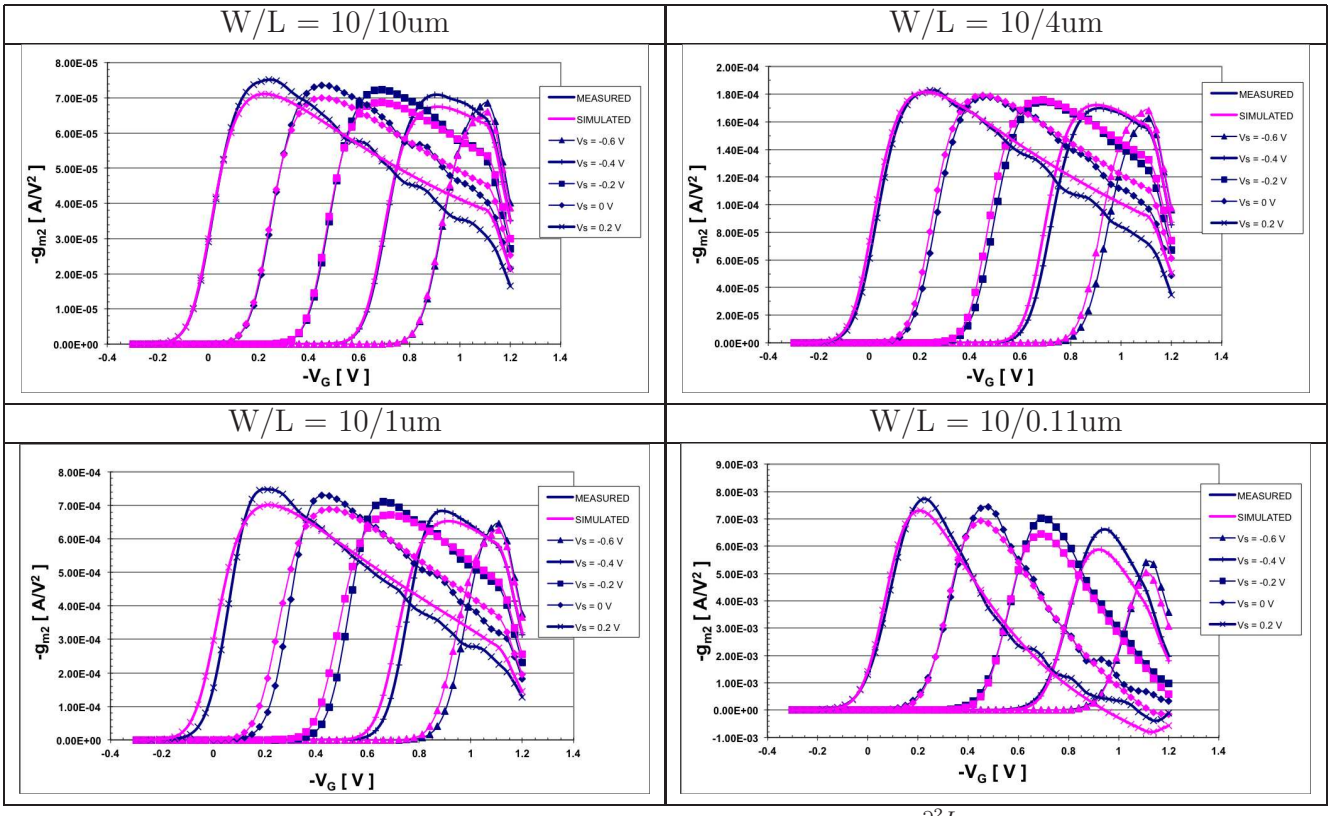


Figure 5.2.20: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_G^2}$ (g_{m2}) versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, -V_S = -200mV, ..., 0.6V, -V_D = 1V, -V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_{m3} **vs.** V_G

– **NMOS**

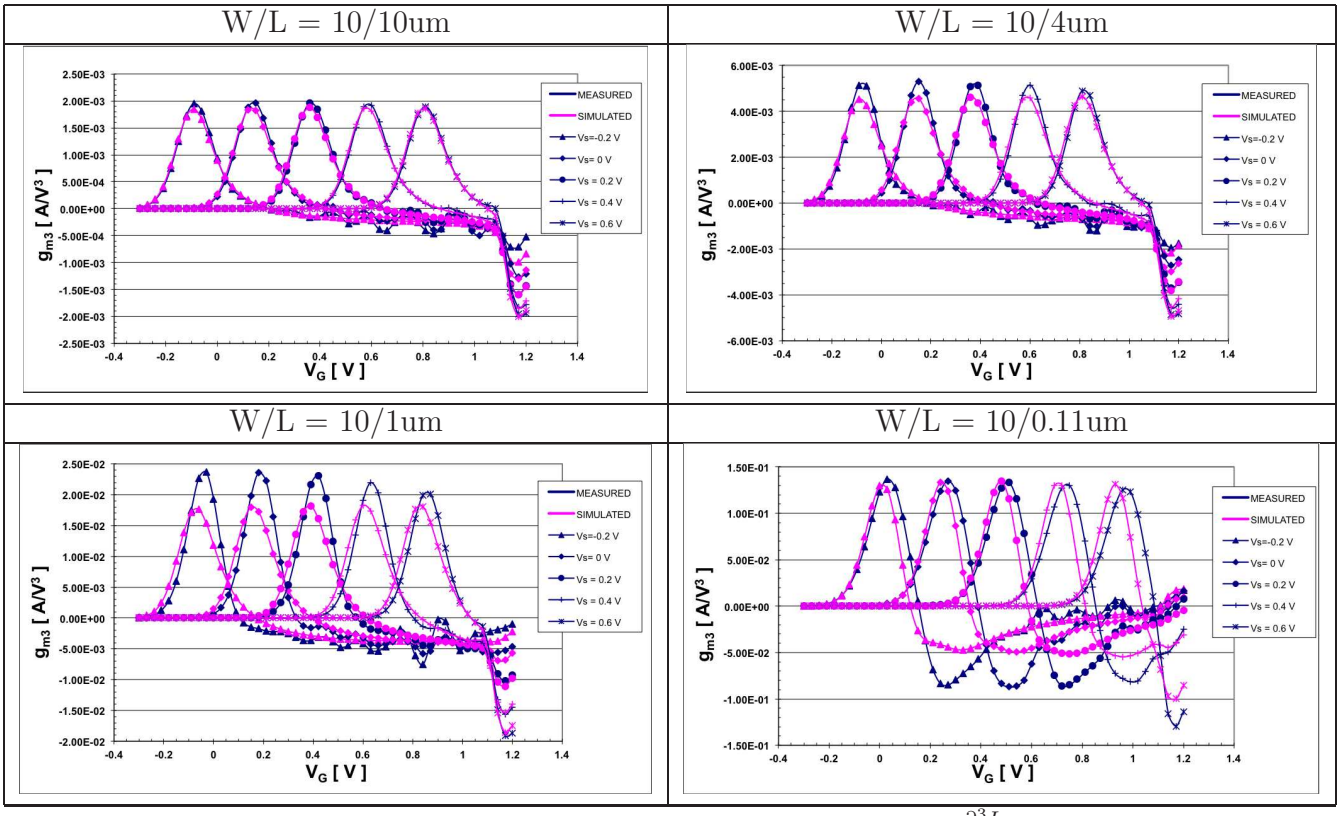


Figure 5.2.21: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_G^3}$ (g_{m3}) versus V_G, $V_G = -300mV, -270mV, ..., 1.2V, V_S = -200mV, ..., 0.6V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**

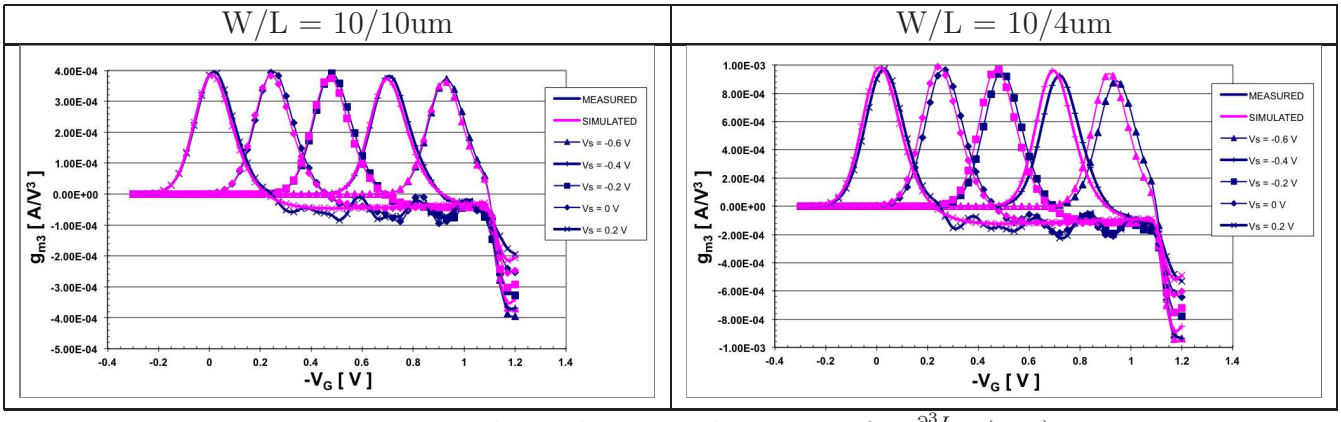


Figure 5.2.22: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_G^3}$ (g_{m3}) versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, -V_S = -200mV, ..., 0.6V, -V_D = 1V, -V_B = 0V$. PMOS transistor, $N_f = 1$. (Cont.)

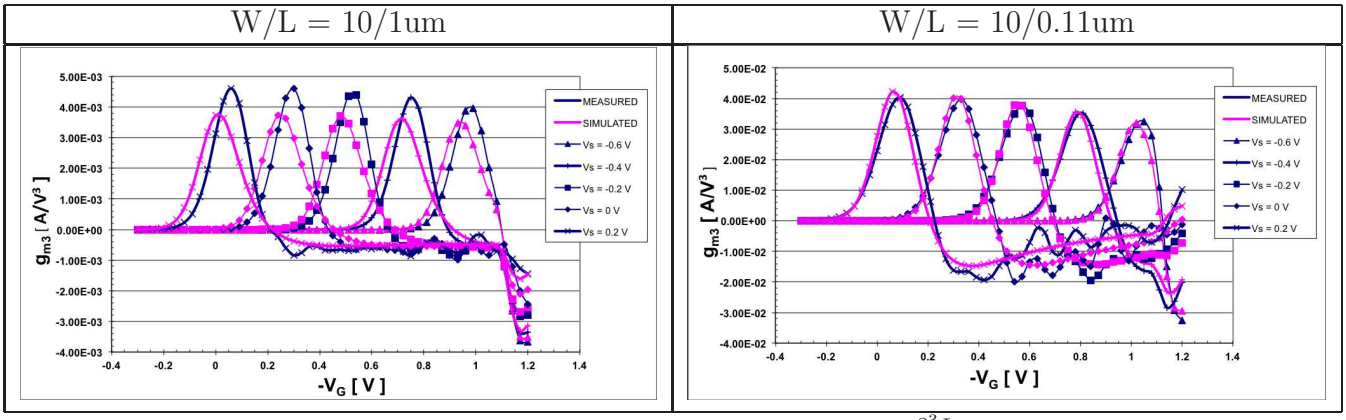


Figure 5.2.23: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_G^3}$ (g_{m3}) versus V_G, $-V_G = -300mV, -270mV, ..., 1.2V, -V_S = -200mV, ..., 0.6V, -V_D = 1V, -V_B = 0V$. PMOS transistor, $N_f = 1$.

Until now, we focused on the higher-order derivatives of the drain current (I_D) in saturation region, with respect to the gate voltage and we presented the results for both NMOS and PMOS devices for various representative gate lengths.

Second-order effects such as velocity saturation, Vertical Field Mobility Reduction (VFMR) and shallow Trench isolation (STI) usually affect the gate transconductance (g_m). These effects are mainly noticeable in strong inversion and usually affect short channel transistors. However, for the purposes of this thesis the simulation results that were presented above are completely satisfactory.

More specifically, for DUT's with gate length 10um and 4um, the EKV 3.0 model predicts accurately all higher-order derivatives of the drain current. Nevertheless, improvements should be made for better fitting of g_m, g_{m2} and g_{m3} for both NMOS and PMOS devices of gate length 4um. As the gate length becomes shorter, the second-order effects become obvious. The EKV3.0 model predicts the peaks of the transistor with channel length 1um, but the higher-order derivatives are not yet modeled very well. Finally, from the short channel transistor simulation results, i.e 0.11um, we can observe that our model has a very good performance, although improvements should be made to fit accurately g_{m2} and the negative values of g_{m3}.

Although, the harmonic distortion can be studied with an excitation on any of the MOSFET terminals, in practice setups with excitation on either gate (G) or drain (D) terminals are commonly used[21]. However, a good model should also fit correctly the drain current and its derivatives with respect to the **source (S)** terminal as well. Hence, we continue by presenting the measured and simulated results of the high-order terms of the drain current with respect to the source terminal where we illustrate the excellent capabilities of the EKV MOSFET model.

- I_D **vs.** V_S

– **NMOS**

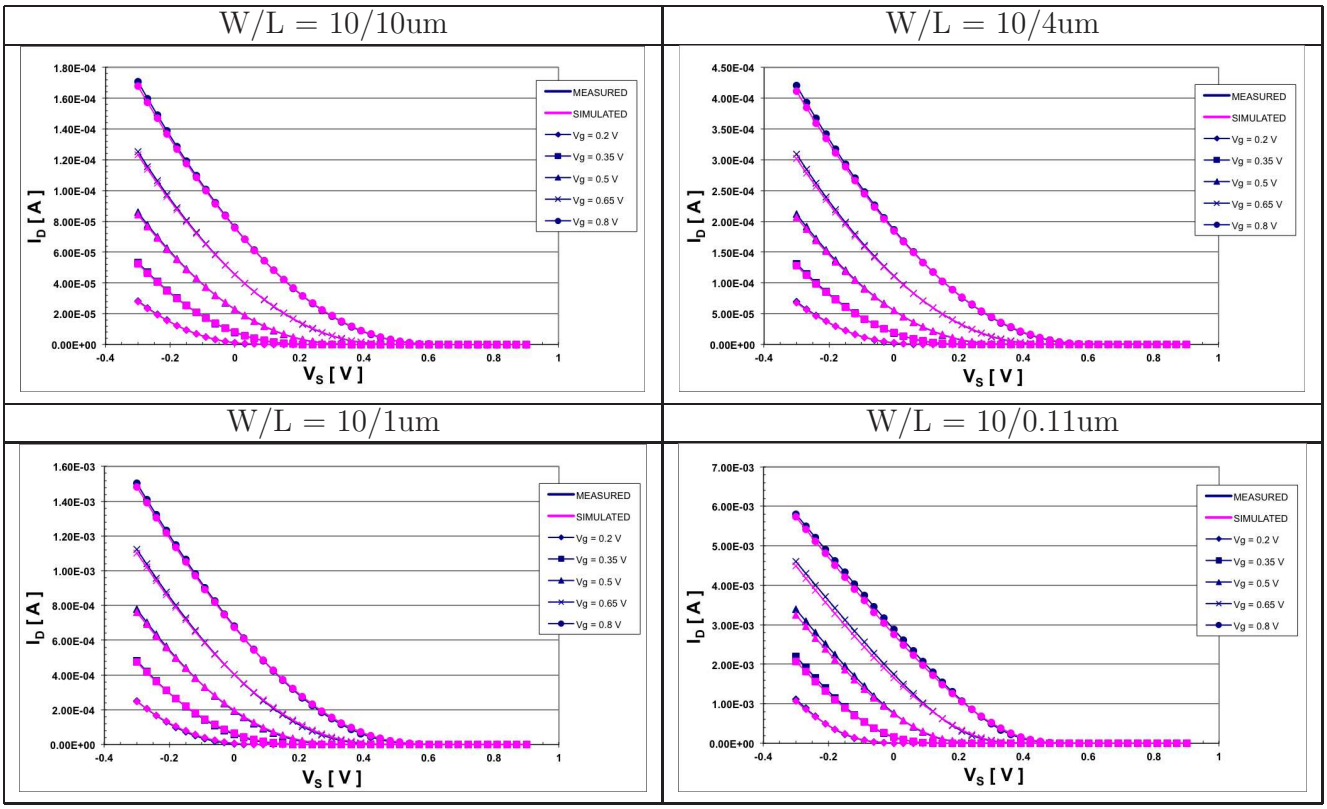


Figure 5.2.24: Measurement and simulation with EKV3.0 for I_D versus V_S, $V_S = -300mV, -270mV, ..., 0.9V, V_G = 200mV, 350mV, ..., 0.8V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**

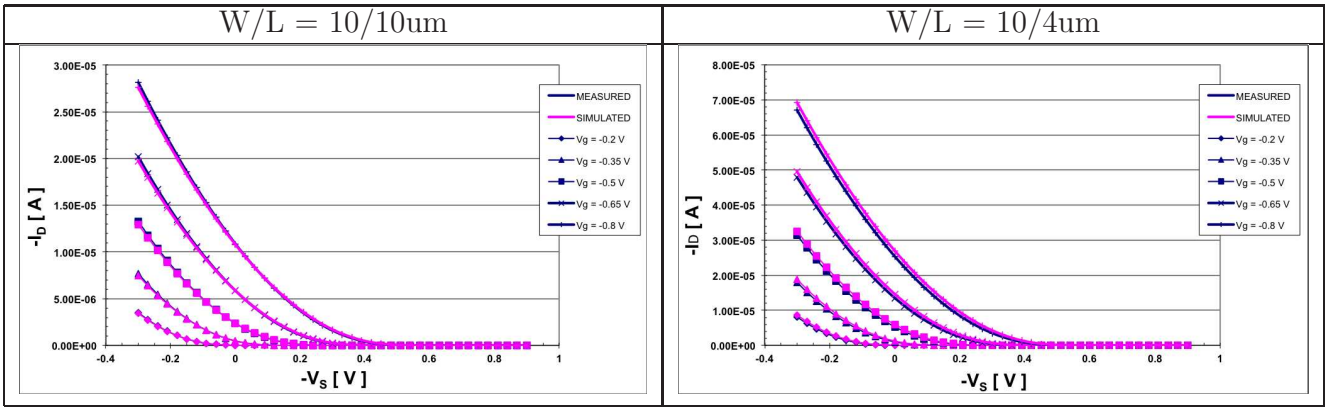


Figure 5.2.25: Measurement and simulation with EKV3.0 for I_D versus V_S, $V_S = -300mV, -270mV, ..., 0.9V, V_G = 200mV, 350mV, ..., 0.8V, V_D = 1V, V_B = 0V$. PMOS transistor, $N_f = 1$. (Cont.)

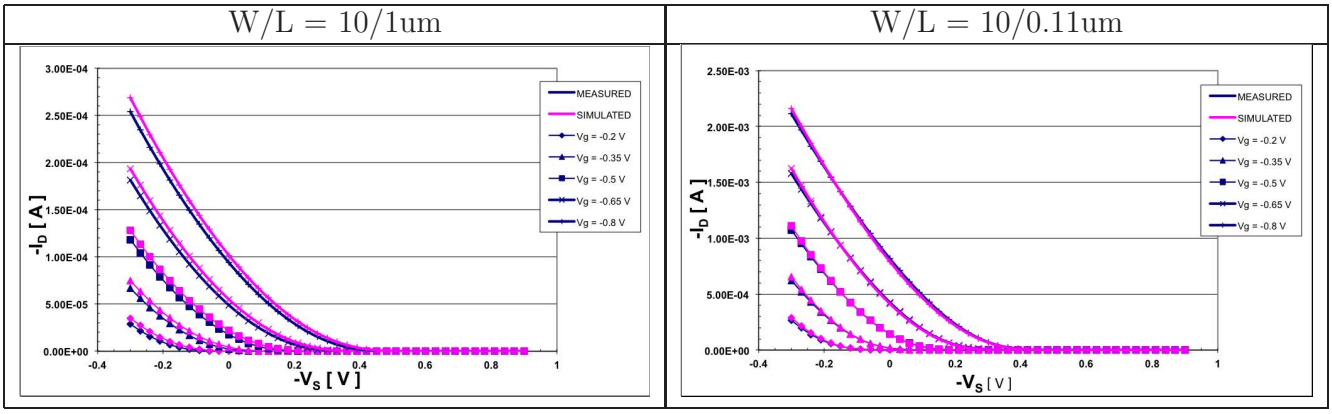


Figure 5.2.26: Measurement and simulation with EKV3.0 for I_D versus V_S, $-V_S = -300mV, -270mV, ..., 0.9V, -V_G = 200mV, 350mV, ..., 0.8V, -V_D = 1V, V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_{ms} **vs.** V_S
  - **NMOS**

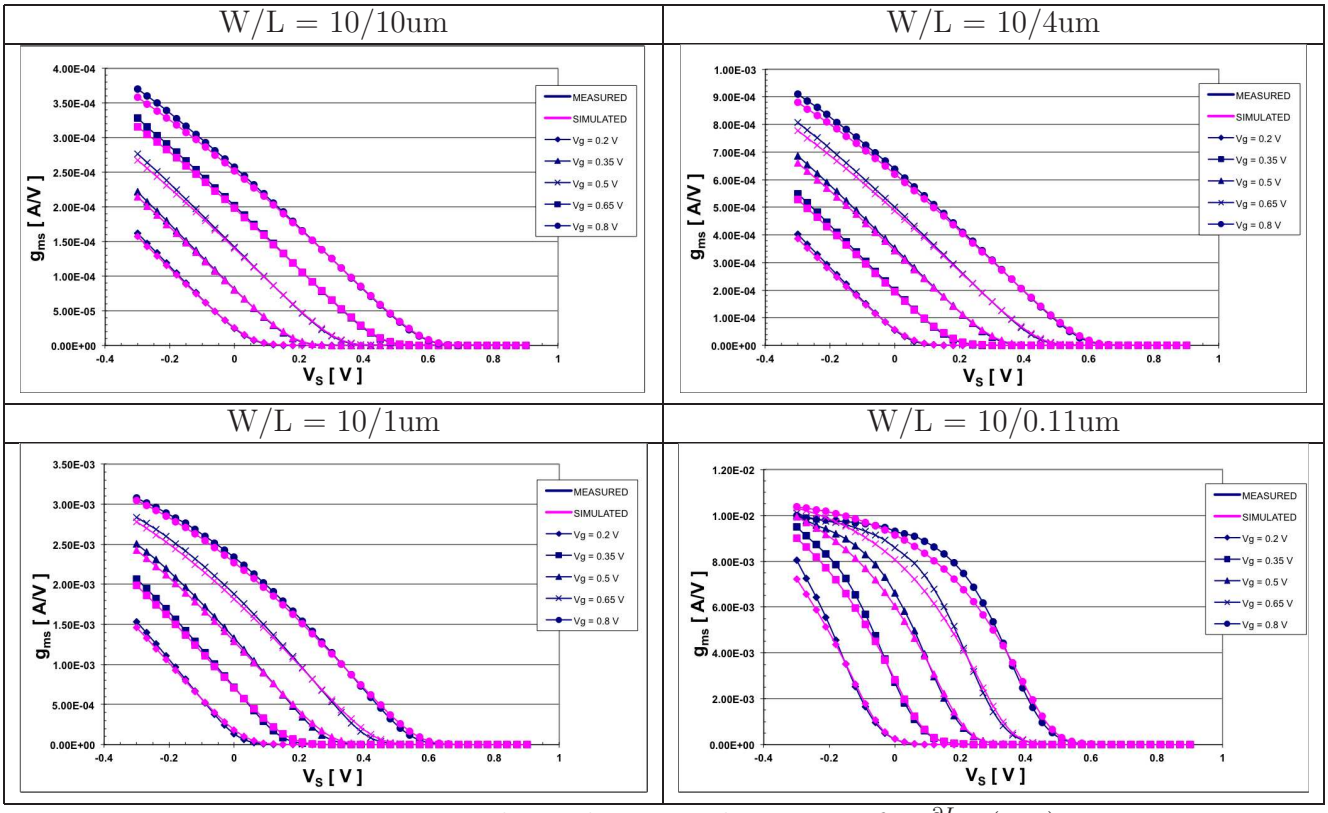


Figure 5.2.27: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_S}$ (g_{ms}) versus V_S, $V_S = -300mV, -270mV, ..., 0.9V, V_G = 200mV, 350mV, ..., 0.8V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**

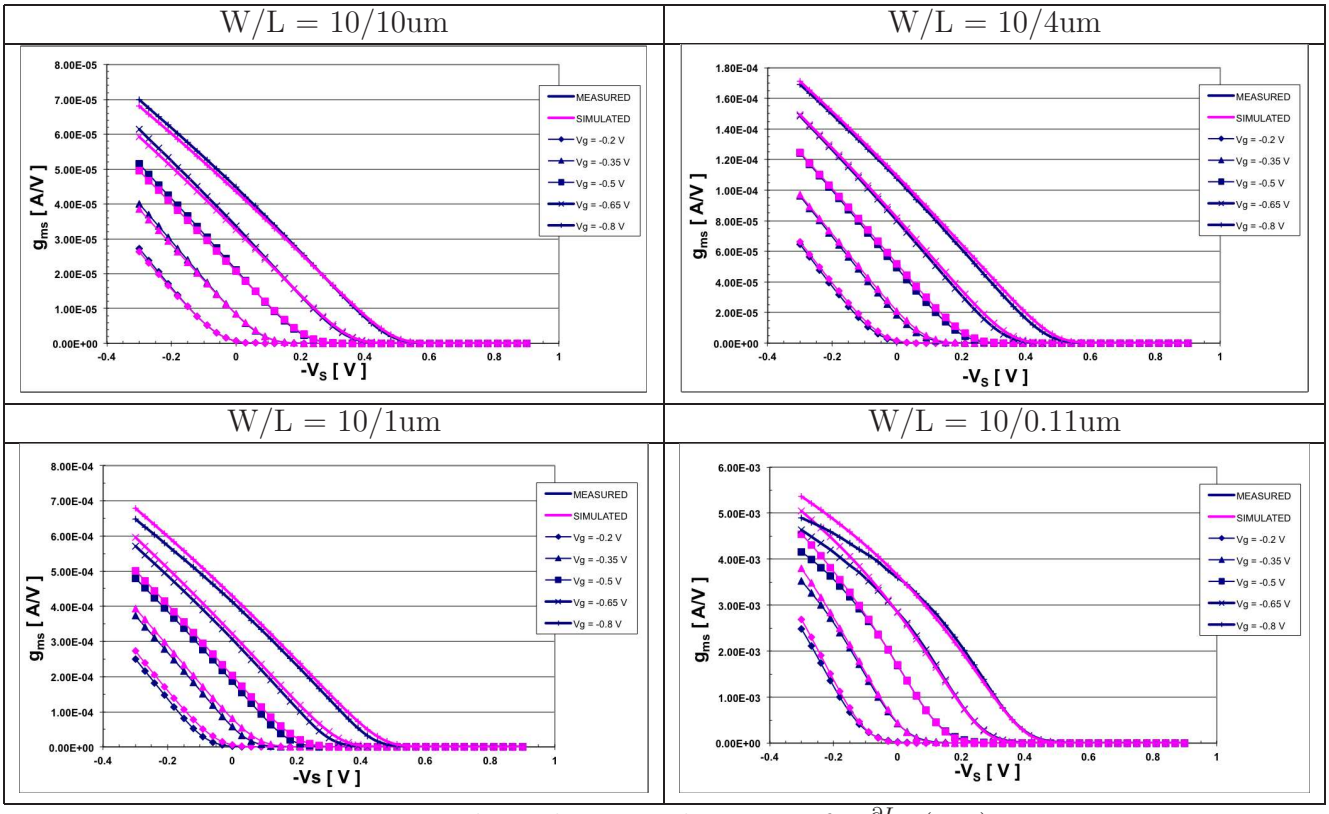


Figure 5.2.28: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_S}$ (g_{ms}) versus V_S, $-V_S = -300mV, -270mV, ..., 0.9V, -V_G = 200mV, 350mV, ..., 0.8V, -V_D = 1V, V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_{ms2} **vs.** V_S

– **NMOS**

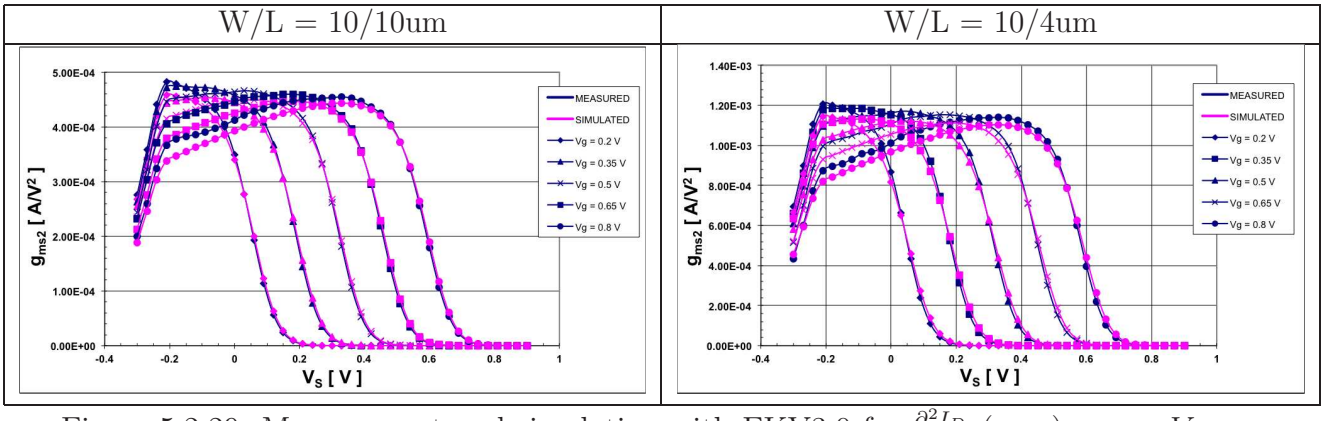


Figure 5.2.29: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_S^2}$ (g_{ms2}) versus V_S, $V_S = -300mV, -270mV, ..., 0.9V, V_G = 200mV, 350mV, ..., 0.8V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$. (Cont.)

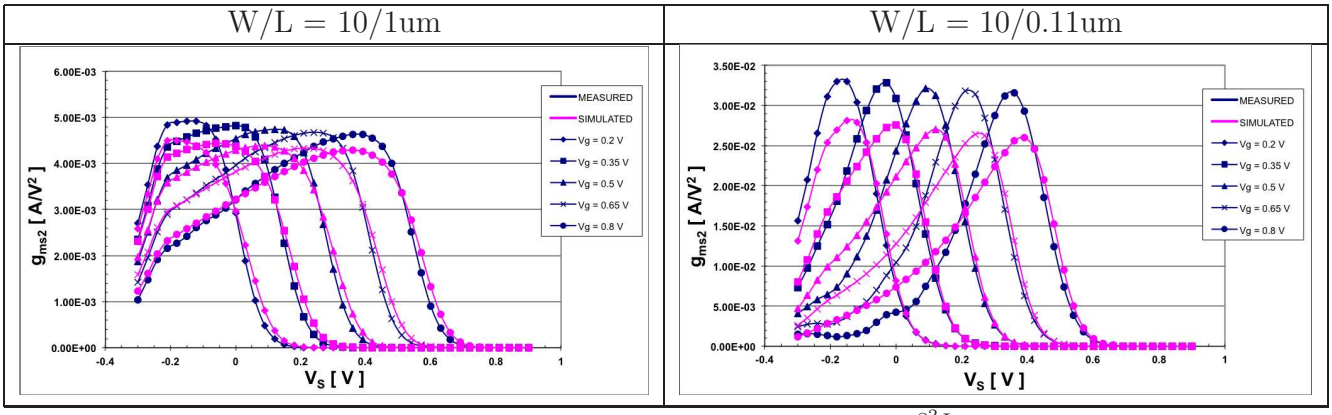


Figure 5.2.30: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_S^2}$ (g_{ms2}) versus V_S, $V_S = -300mV, -270mV, ..., 0.9V, V_G = 200mV, 350mV, ..., 0.8V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**

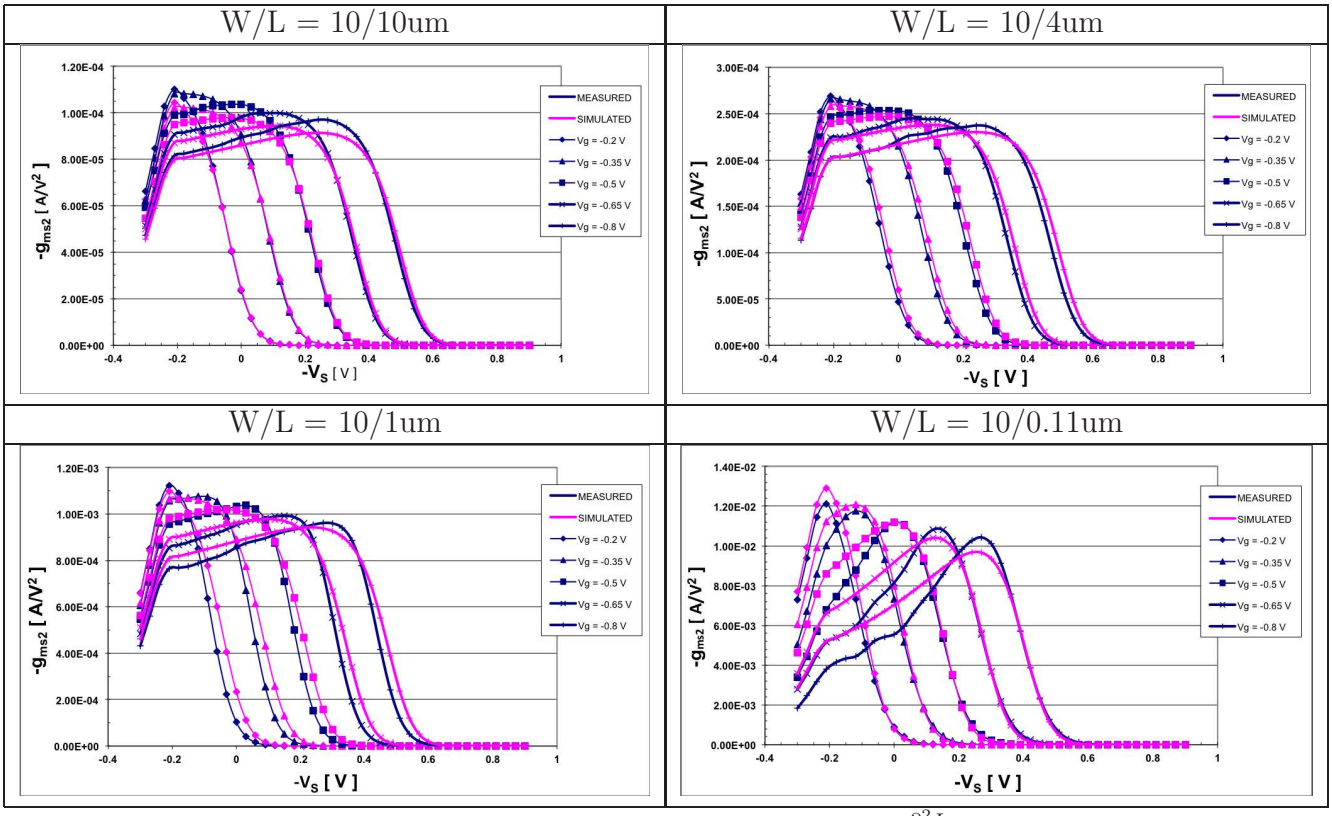


Figure 5.2.31: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_S^2}$ versus V_S, $-V_S = -300mV, -270mV, ..., 0.9V, -V_G = 200mV, 350mV, ..., 0.8V, -V_D = 1V, V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_{ms3} **vs.** V_S
  - **NMOS**

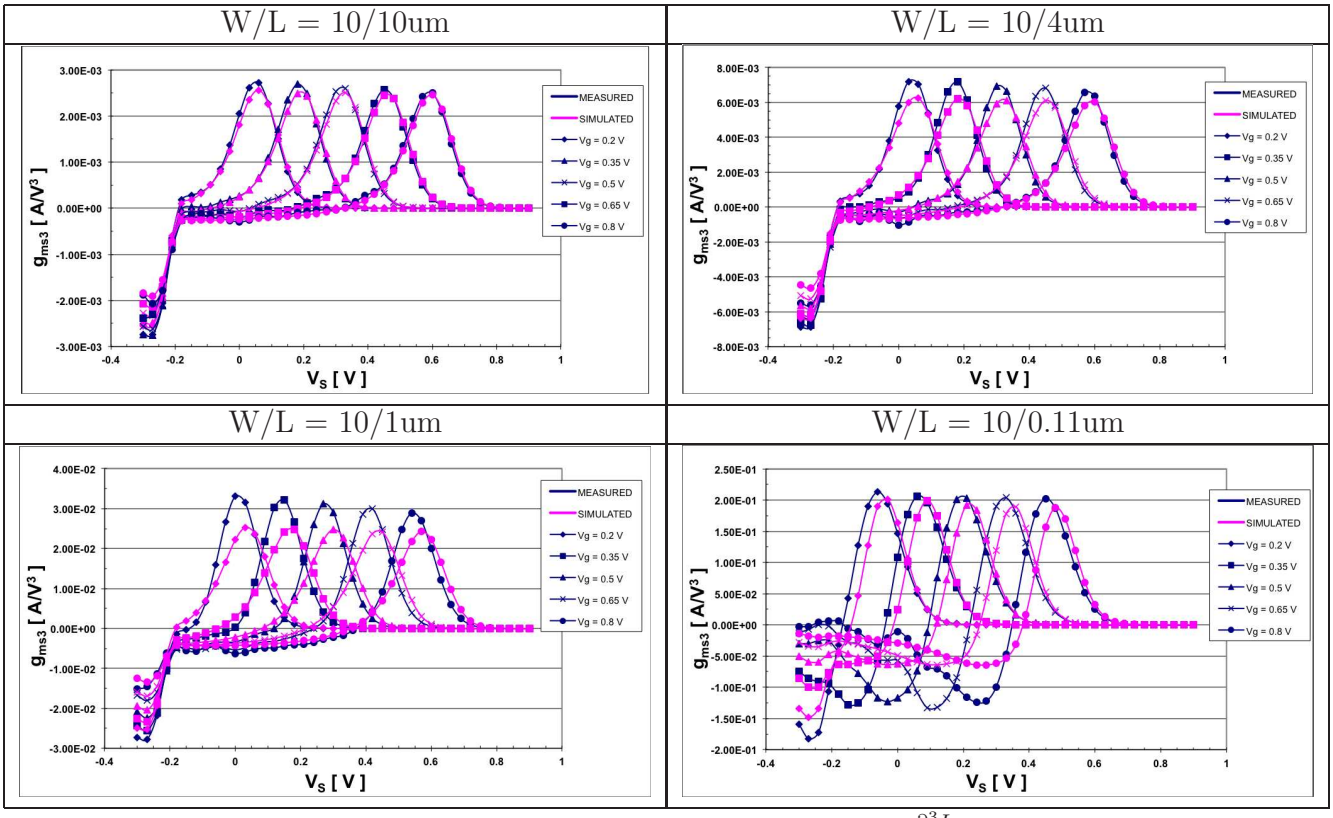


Figure 5.2.32: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_S^3}$ (g_{ms3}) versus V_S, $V_S = -300mV, -270mV, ..., 0.9V, V_G = 200mV, 350mV, ..., 0.8V, V_D = 1V, V_B = 0V$. NMOS transistor, $N_f = 1$.

  - **PMOS**

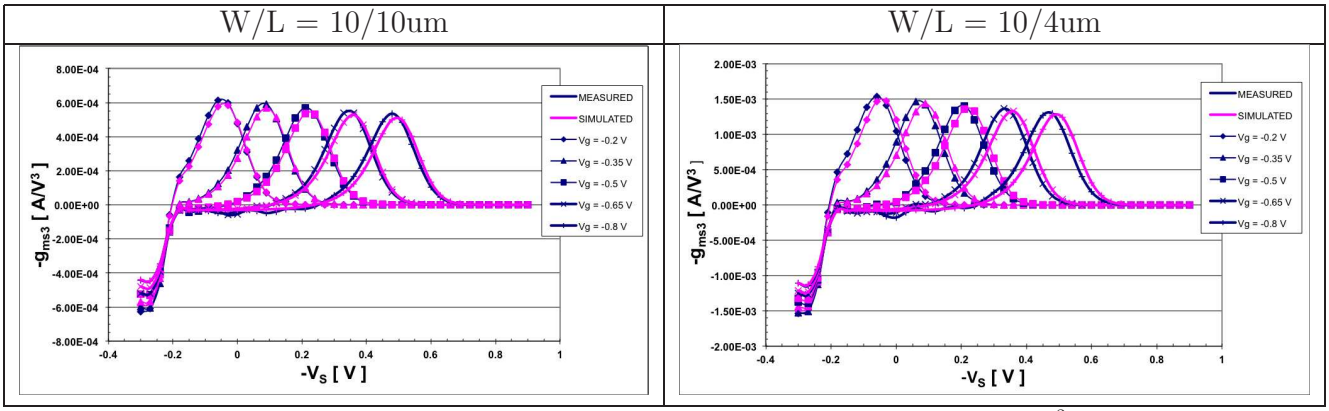


Figure 5.2.33: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_S^3}$ (g_{ms3}) versus V_S, $-V_S = -300mV, -270mV, ..., 0.9V, -V_G = 200mV, 350mV, ..., 0.8V, -V_D = 1V, V_B = 0V$. PMOS transistor, $N_f = 1$. (Cont.)

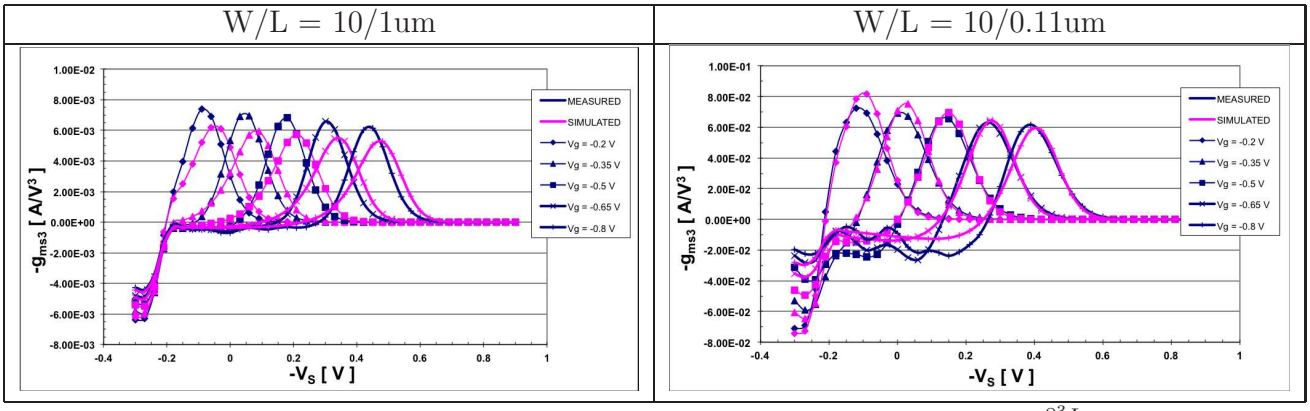


Figure 5.2.34: Measurement and simulation with EKV3.0 for $\frac{\partial^3 I_D}{\partial V_S^3}$ (g_{ms3}) versus V_S, $-V_S = -300mV, -270mV, ..., 0.9V, -V_G = 200mV, 350mV, ..., 0.8V, -V_D = 1V, V_B = 0V$. PMOS transistor, $N_f = 1$.

In Section 4.5 we mentioned that the source transconductance is affected by the same second-order effects which affect the gate and bulk transconductance (i.e velocity saturation, VFMR and STI).These effects are mainly noticeable in strong inversion and usually affect short channel transistors. For the purposes of this thesis the simulation results that were presented above are completely satisfactory.

More specifically, for DUT's with channel length 10um and 4um, the EKV 3.0 model predicts quite accurately all high-order derivatives of the drain current. Nevertheless, improvements should be made for better fitting of g_{ms}, g_{ms2} and g_{ms3} for both NMOS and PMOS devices of gate length 4um. As the gate length becomes shorter, the second-order effects become more obvious. The EKV3.0 model predicts the peaks of the transistor with channel length 1um, but the higher-order derivatives are not yet modeled very well. Finally, from the short channel transistor results, i.e 0.11um, we can observe that our model has a very good performance, although improvements should be made to fit accurately g_{ms2} and the negative values of g_{ms3}.

In Chapter 3, we paid special attention to the drain induced non-linearities. Thus, it is of great importance a good MOSFET model to accurately fit the high-order derivatives of the drain current with respect to the drain terminal. However, measurements of especially third-order derivative of the drain current (i.e g_{ds3}) are extremely sensitive and prone to measurement noise-particularly for g_{ds3} in saturation-which make them impractical when derived from DC measurements; they are therefore not shown here. As a result the correct modeling of these high-order terms is a challenging task. Nonetheless, in the following figures we present the measured and simulated results of up to g_{ds2}.

- I_D vs. V_D
  - **NMOS**

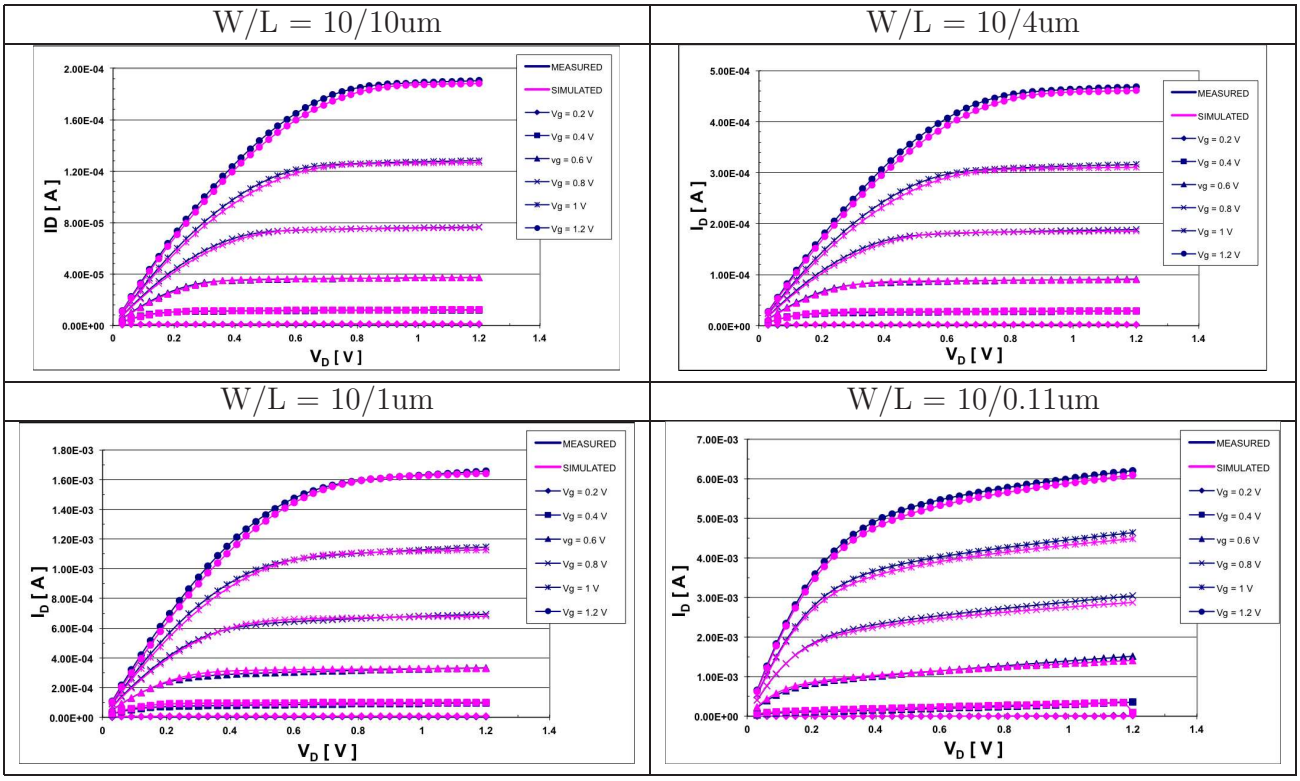


Figure 5.2.35: Measurement and simulation with EKV3.0 for I_D versus V_D, $V_D = 30mV, ..., 1.2mV, V_G = 200mV, ..., 1.2V, V_S = 0V, V_B = 0V$. NMOS transistor, $N_f = 1$.

  - **PMOS**

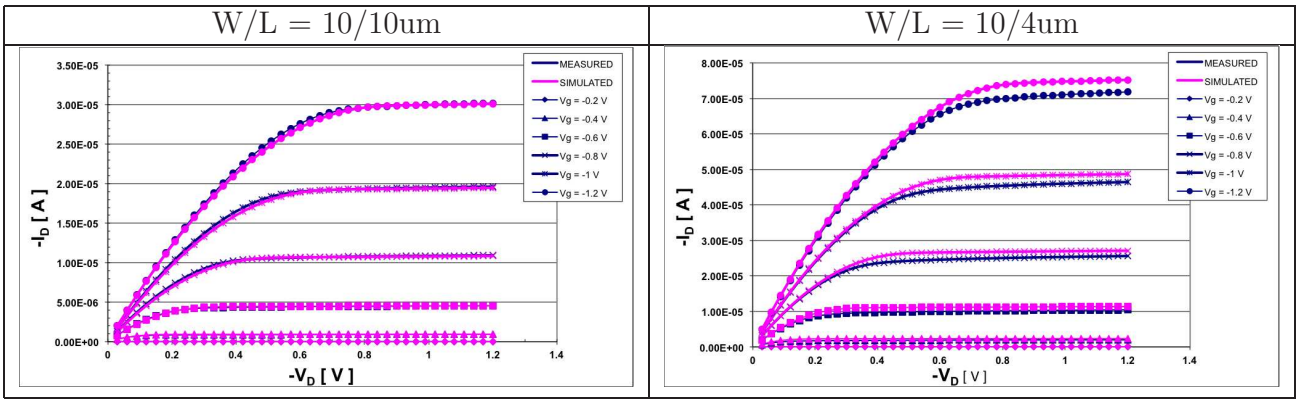


Figure 5.2.36: Measurement and simulation with EKV3.0 for I_D versus V_D, $-V_D = 30mV, ..., 1.2mV, -V_G = 200mV, ..., 1.2V, -V_S = 0V, -V_B = 0V$. PMOS transistor, $N_f = 1$. (Cont.)

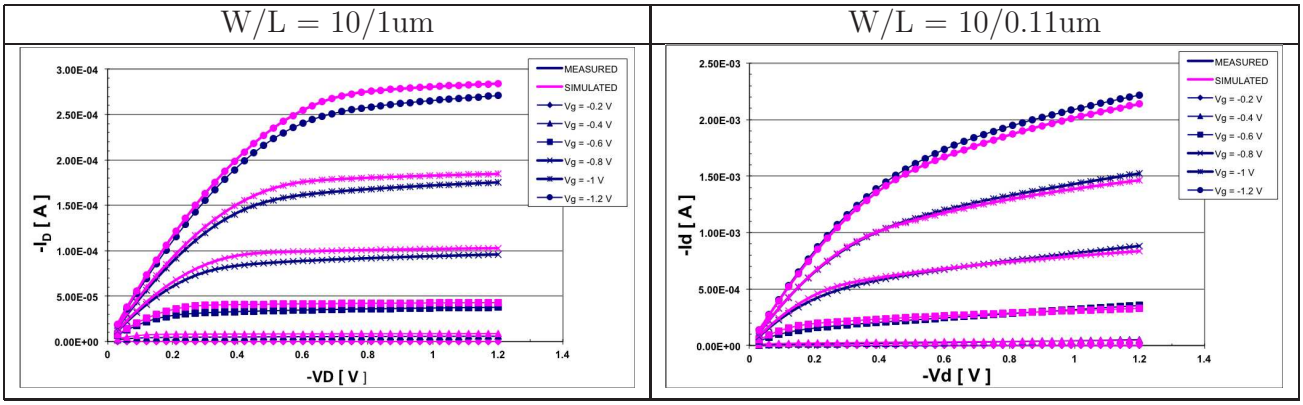


Figure 5.2.37: Measurement and simulation with EKV3.0 for I_D versus V_D, $-V_D = 30mV, ..., 1.2mV, -V_G = 200mV, ..., 1.2V, -V_S = 0V, -V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_{ds} vs. V_D
  - **NMOS**

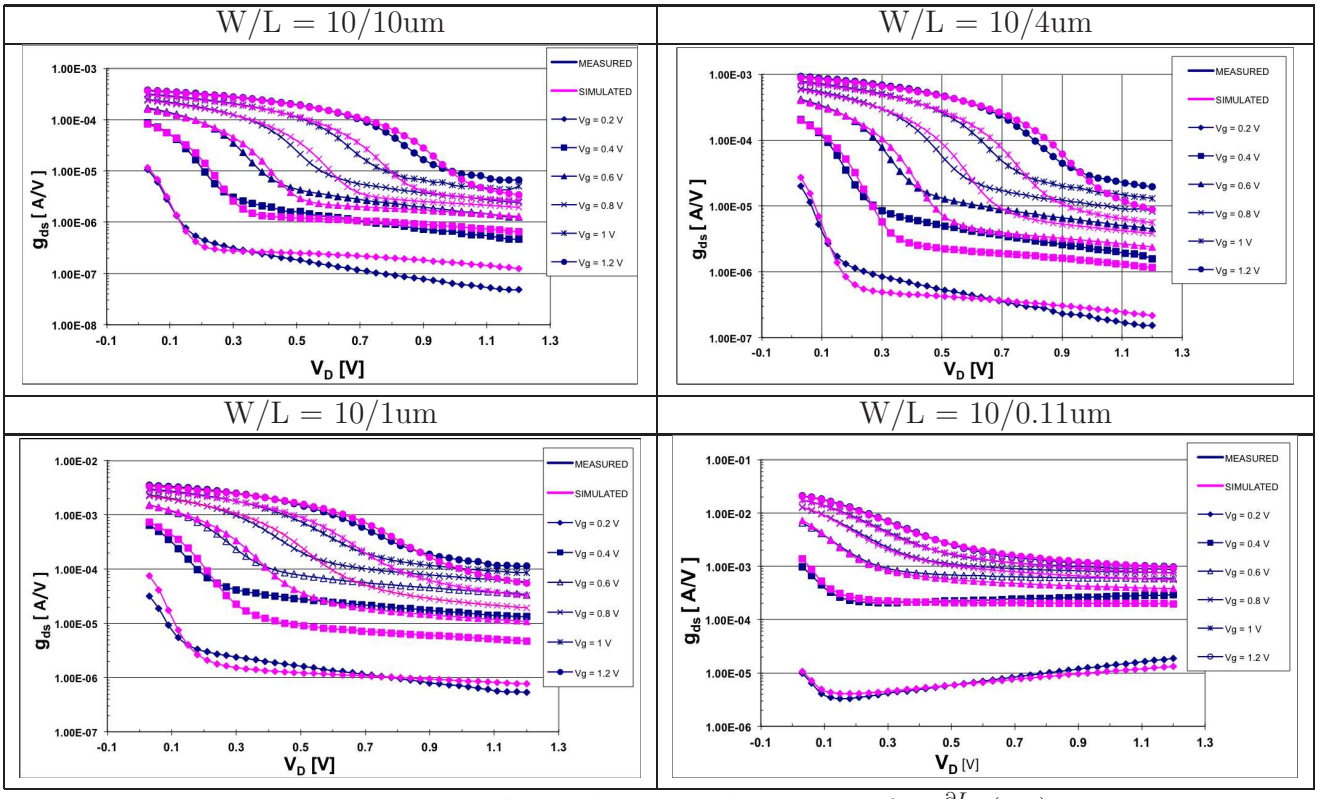


Figure 5.2.38: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_D}(g_{ds})$ versus V_D, $V_D = 30mV, ..., 1.2mV, V_G = 200mV, ..., 1.2V, V_S = 0V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– **PMOS**

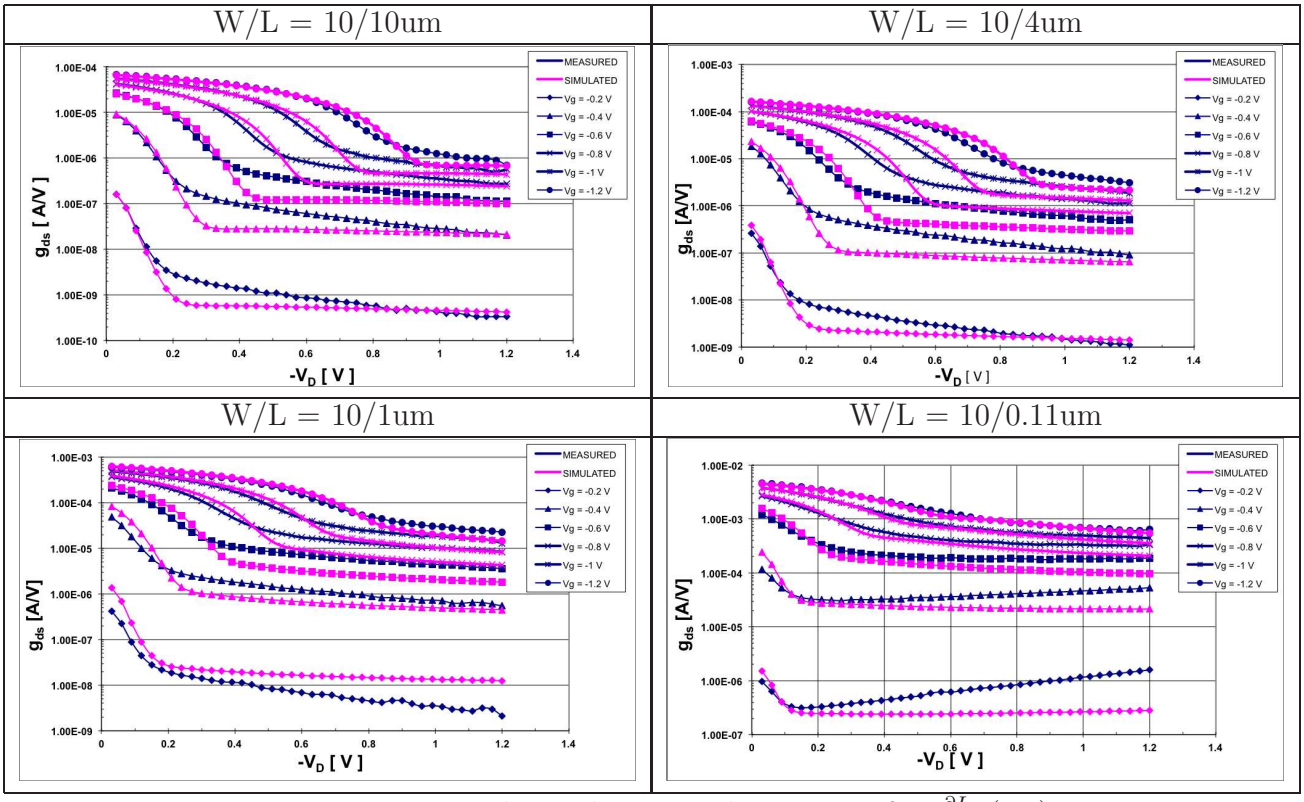


Figure 5.2.39: Measurement and simulation with EKV3.0 for $\frac{\partial I_D}{\partial V_D}(g_{ds})$ versus V_D, $-V_D = 30mV, ..., 1.2mV, -V_G = 200mV, ..., 1.2V, -V_S = 0V, -V_B = 0V$. PMOS transistor, $N_f = 1$.

- g_{ds2} vs. I_D

– **NMOS**

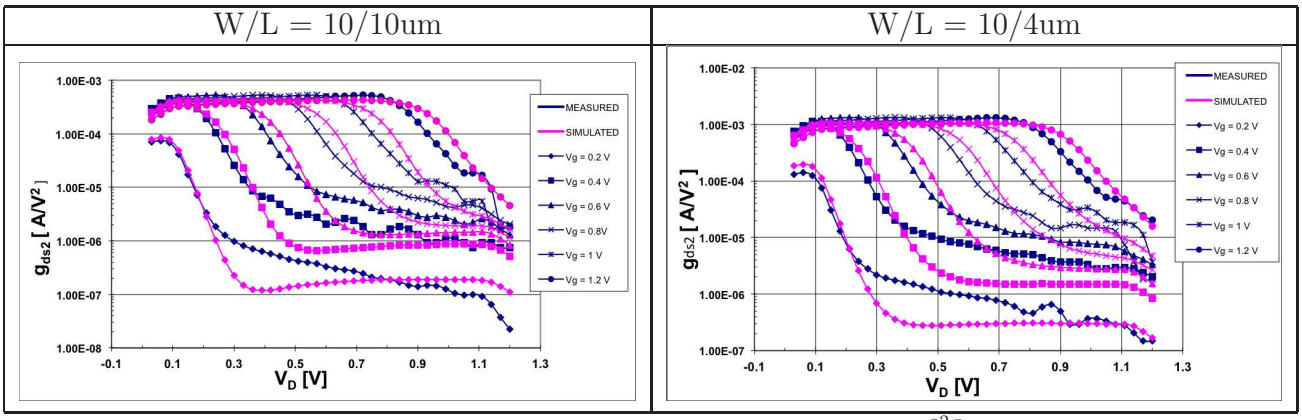


Figure 5.2.40: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_D^2}(g_{ds2})$ versus V_D, $V_D = 30mV, ..., 1.2mV, V_G = 200mV, ..., 1.2V, V_S = 0V, V_B = 0V$. NMOS transistor, $N_f = 1$. (Cont)

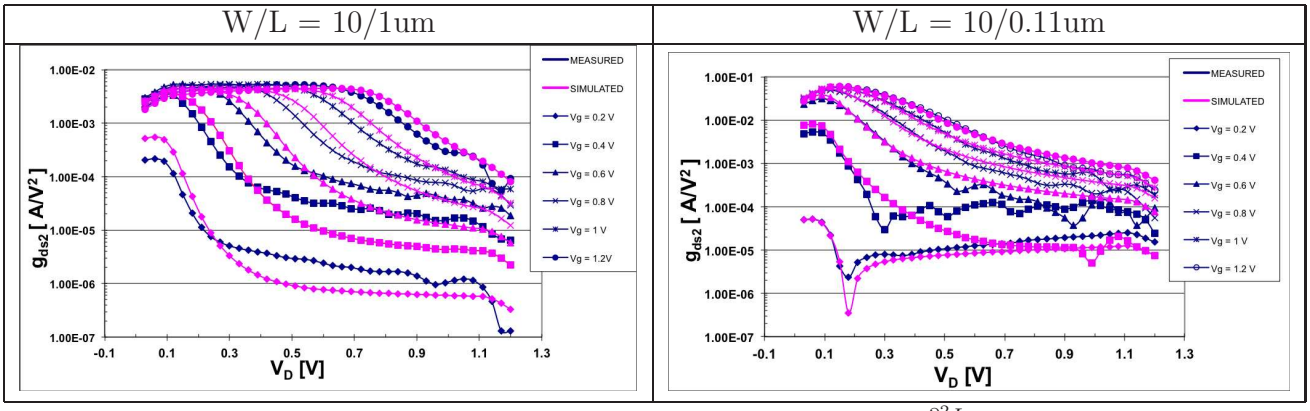


Figure 5.2.41: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_D^2}(g_{ds2})$ versus V_D, $V_D = 30mV, ..., 1.2mV, V_G = 200mV, ..., 1.2V, V_S = 0V, V_B = 0V$. NMOS transistor, $N_f = 1$.

– PMOS

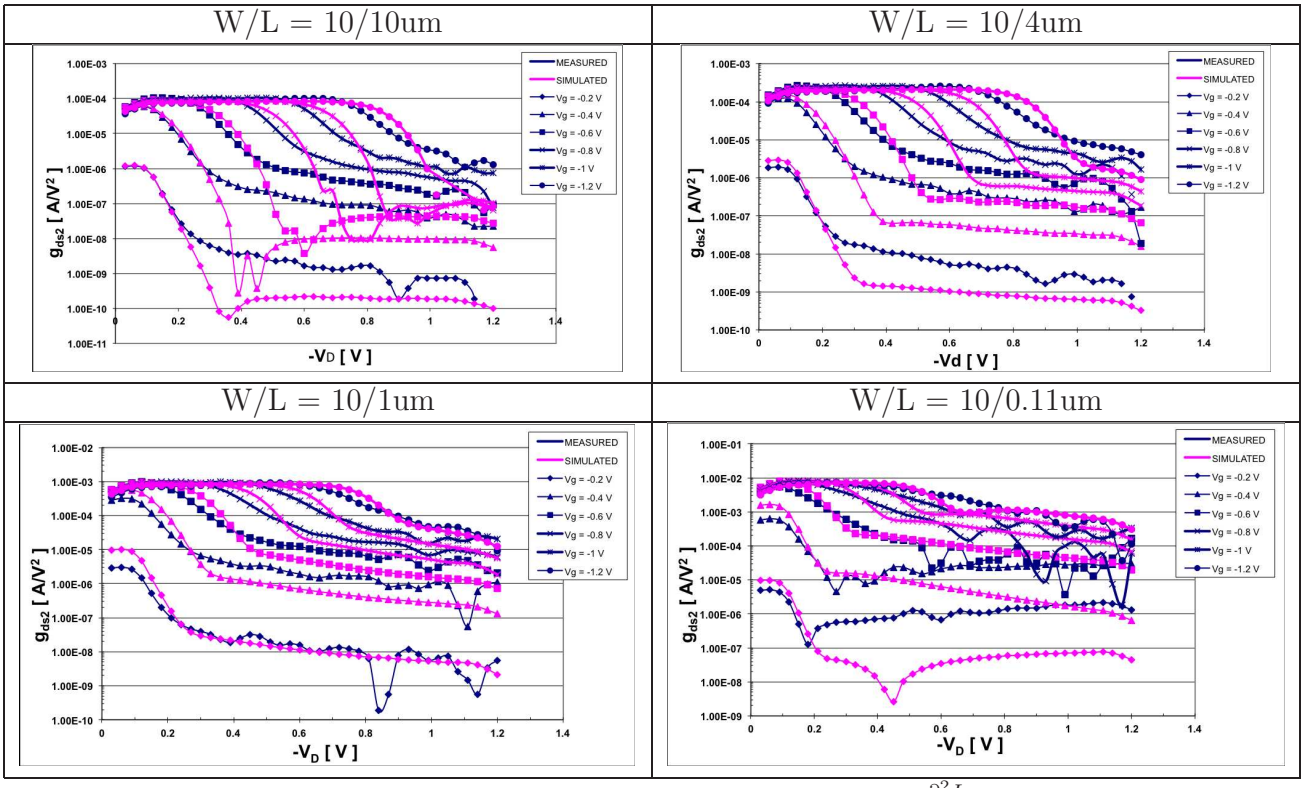


Figure 5.2.42: Measurement and simulation with EKV3.0 for $\frac{\partial^2 I_D}{\partial V_D^2}(g_{ds2})$ versus V_D, $-V_D = 30mV, ..., 1.2mV, -V_G = 200mV, ..., 1.2V, -V_S = 0V, -V_B = 0V$. PMOS transistor, $N_f = 1$.

In general, drain induced non-linearities of a MOSFET is mainly determined by velocity saturation in the ohmic region and channel length modulation in the saturation region. In order to predict the non-linear behavior an accurate description of the above-mentioned physical effects must be implemented in the drain current expression. Apart from this, for correct output conductance modeling in saturation, static feedback has to be incorporated in a MOS model.

From the above figures, we can observe that the EKV3.0 model fits very accurately the drain current for a wide range of channel lengths, drain and gate voltage values. As far as the output conductance (i.e g_{ds}) is concerned, our model illustrates a satisfactory behavior. Finally, further improvements should be made in order the EKV model to predict accurately the second derivative of the drain current(g_{ds2}) .

Chapter 6

Conclusion and Future Work

This thesis dealt with the investigation of the non-linearities in MOS transistors. With the growing importance of analog applications in CMOS technology, the demands for circuit design regarding subjects such as signal-to-noise ratio and linearity become more and more stringent. For a correct modeling of the non-linear behavior of a MOS transistor, not only an accurate description of I-V, Q-V and the corresponding first-order derivatives (transconductance, conductance and capacitances) is essential but also a precise description of the higher-order derivatives of the drain current with respect to any terminal, especially gate and drain terminals. Many contemporary MOSFET models struggle to predict these higher-orders of the drain current because of an inaccurate modeling and incorporation of the second-order effects to the model's equations. Since an accurate description of the DC behavior is a prerequisite for a good high-frequency behavior, this thesis focused on the DC behavior of MOS transistors. The investigation and modeling of the non-linear sources of the MOSFET's AC behavior is reserved for future work.

A small-signal analysis was attempted for the calculation of the normalized gate and source transconductances. Since the gate transconductance and its higher-order derivatives of the drain current are considered to be one of the main causes of the MOSFETs non-linear behavior, special attention was paid to the experimental results and the second-order effects. Furthermore, the Inversion Coefficient (IC) was used as a criterion for the distinction of the levels of inversion, giving an insight look on how these high-order are affected due to the level of inversion, the channel length and the different bias conditions. Finally, we attempted a discussion of the experimental results by a design perspective since we came to the conclusion that between high moderate and low strong inversion regions should be the most preferable for low distortion operation.

In Chapter 5, we provided an overview of the basic principles of the EKV MOSFET model. EKV is an analytical model based on charge sheet theory using relatively small number of parameters with the use of simple equations. EKV3 incorporates all the modern phenomena that appear in modern submicron technologies. Additionally, we presented the measured and simulated results of the EKV3.0 MOSFET model which illustrate its excellent capabilities to fit accurately the I-V characteristics for a set of representative transistor gate lengths. Furthermore, the evaluation results prove the accurate prediction of the gate induced non-linearities especially for long and semi-short transistors; though improvements should be made for short transistors. As far as the drain induced non-linearities are concerned, simulation results show that the EKV model has the potential to accurately predict the non-linear behavior.

6.1 Future Work

In this thesis we presented the measured and simulated results for a representative number of same width and number of fingers (NFs) transistors which were shown in Fig. 4.4.1. However, for the same bias conditions which were mentioned in Chapter 4, measurements were also performed on the following DUTs:

Table 6.1: DUT's shorted by Number of Fingers

DUT	Width	Length	Number of Fingers
NMOS\PMOS	2u	1u	1
	2u	5u	1
	5.2u	180n	2
	5.2u	140n	2
	5.2u	110n	2
	2.5u	180n	2
	2.5u	140n	2
	2.5u	110n	2
	1.2u	180n	2
	1.2u	140n	2
	1.2u	110n	2
	5.2u	180n	4
	5.2u	140n	4
	5.2u	110n	4
	5.2u	180n	6
	5.2u	140n	6
	5.2u	110n	6
	5.2u	180n	12
	5.2u	140n	12
	5.2u	110n	12
	5.2u	180n	24
	5.2u	140n	24
	5.2u	110n	24

For the above mentioned DUTs, we calculated up to the 3^{rd} derivative of the drain current with respect to the gate, source and drain terminal. In the future it would be interesting to study how these high-order derivatives vary as a function of the transistor's dimensions and NFs. Additionally, with the downscaling of modern CMOS transistors, an evaluation of the EKV3 MOSFET model would be mandatory.

Furthermore, it would be interesting to investigate the AC non-linear behavior by making a low frequency measurement setup (up to 50 MHz) and extract the relevant parameters of the EKV3 model. Finally, RF measurements, where not only the gate and drain induced non-linearities are dominant, but also parasitic elements (capacitances etc.) play an important role, would be desirable.

Appendix

The equipment we used to perform the measurements was the following:

- PC with IC-CAP 2008 installed
- AGILENT USB/GPIB Interface
- AGILENT 4142B DC Analyzer
- SUMMIT 10600 Thermal Probe Station
- Wafer
- TRIAX cables

The measurements were performed on a 110nm CMOS technology covering both RF NMOS and PMOS transistors. The procedure we followed is presented in the following diagram.

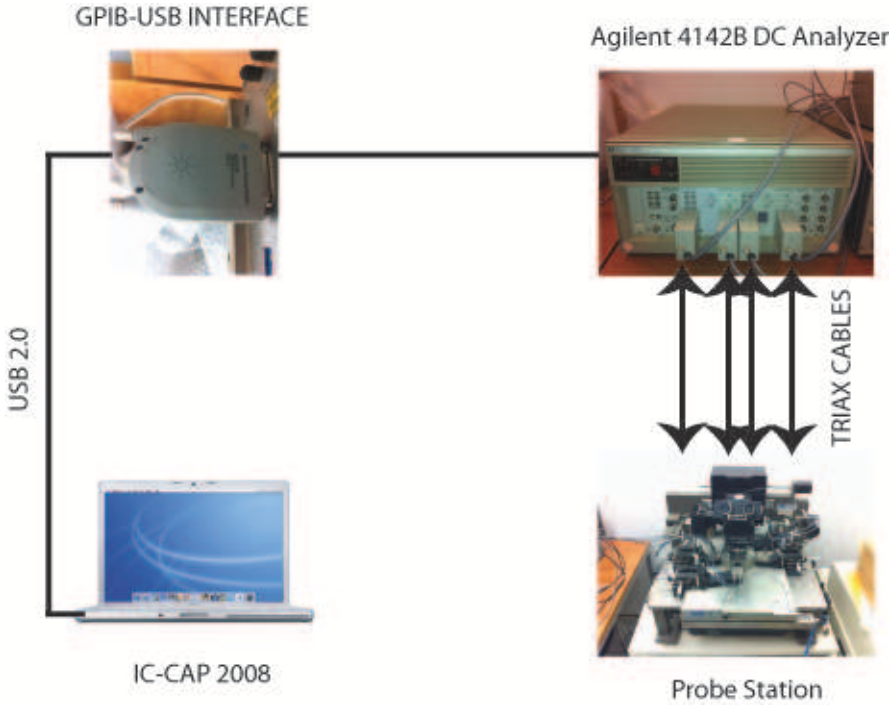


Figure 6.1.1: Measurement Setup

The computer was connected through a USB 2.0 cable to an AGILENT USB/GPIB Interface which was attached to the DC Analyzer HP4142B Modular DC Source Monitor. The USB/GPIB Interface was responsible to apply the bias conditions which were defined on the IC-CAP

2008 software, to the DC Analyzer. The latter was responsible to apply these bias conditions through the SMUs to the SUMMIT 10600 Thermal Probe Station where the wafer was located. By the reverse procedure, the results(drain current) were presented in our computer where we were able to perform our calculations (transconductance, derivatives of the drain current etc.) and the simulation using the EKV3.0 model.

Bibliography

[1] Ronald van Langevelde, *A Compact MOSFET Model for Distortion Analysis in Analog Circuit Design*. -Eindhoven : Technische Universiteit Eindhoven, 1998

[2] Jaume Segura and Charles F. Hawkins, *CMOS electronics - How it works. How it fails* , 2004

[3] Behzad Razavi, *RF Microelectronics*, Prentice Hall, 1997

[4] Chien-Hsiung Feng, Fredrik Jonsson, Mohammed Ismail and Hakan Olsson, *Analysis of Nonlinearities in RF CMOS Amplifiers*. Proceedings of ICECS '99. The 6th IEEE International Conference on ,137-140 vol.1, 1999

[5] Bruun E. , *Harmonic distortion in CMOS current Mirrors* , Proc IEEE Int Symp Circ Syst, 1998;1:567-70

[6] Wambacq P., Sansen W., *Distortion analysis of analog integrated circuits* , Dordrecht: Kluwer;

[7] Shanghoon Kang, Byoungi Choi, and Bumman Kim, *Linearity Analysis of CMOS for RF Application*. IEEE Transactions on Microwave theory and techniques,VOL. 51, NO. 3, MARCH 20031998

[8] Georgios Rozakis, and Apostolos Samelis, *Symbolic/Numerical Nonlinear Circuit Analysis Using Volterra Series*, Proceedings of the 36*th* European Microwave Conference, Manchester UK, 2006

[9] Gerasimos Theodoratos, Athanasios Vasilopoulos, Georgios Vitzilaios, and Yannis Papananos, *Calculating Distortion in Active CMOS Mixers Using Volterra Series*, IEEE; 2006

[10] H.Hakan Kuntman and Ali Zeki, *Novel approach to the calculation of non-linear harmonic distortion coefficients in CMOS amplifiers*, Microelectronics Journal 29; VOL.43-48, 1998

[11] A. Cerdeira, M. Estrada, R. Quintero, D. Flandre, A. Ortiz-Conde, F.J. García Sánchez, *New method for determination of harmonic distortion in SOI FD transistors*, Solid-State Electronics 46; VOL.103-108, 2002

[12] Antonio Cerdeira, Miguel A. Alemán, Magali Estrada, Denis Flandre, *Integral function method for determination of nonlinear harmonic distortion*, Solid-State Electronics 48; VOL.2225-2234, 2004

[13] Behzad Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill Science/Engineering/Math; 1 edition (August 15, 2000)

[14] C.Enz, F.Krummenacher, E.Vittoz, *An analytical MOS Transistor Model Valid in all Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications*, Analog Int. Circ. Signal Proc. J., Vol. 8, pp. 83-114, 1995

[15] E. Vittoz, *Micropower Techniques*, Prentice Hall, 1994

[16] M. Bucher, D. Kazazis, F. Krummenacher, D. Brinkley, D. Foty, Y. Papananos, *Analysis of Transcoductance at all Levels of Inversion in Deep Submicron CMOS*, Proc. IEEE Int. Conf. on Electronics, Circuits and Systems (ICECS 2002), pp. 1183-1186, Dubrovnik, Croatia, Sept. 2002

[17] Matthias Bucher, George Diles, Nikos Makris, *Analog performance of advanced CMOS in weak, moderate, and strong inversion*, Mixed Design of Integrated Circuits and Systems (MIXDES), 2010 Proceedings of the 17th International Conference, p. 54-57, Warsaw, June 2010

[18] D.M. Binkley, *Tradeoffs and Optimization in Analog Circuit Design*, John Wiley & Sons, 2008

[19] R. van Langevelde, L.F. Tiemejier, R.J Havens, M.J Knitel, R.F/M Roes, P.H Woerlee, D.B.M Klaasen, *RF-Distortion in deeip-submicron CMOS technologies*, Electron Devices Meeting, 2000, IEDM Technical Digest. International, p.807-810, 2000

[20] A. BAZIGOS, M. BUCHER, *THE EKV3.0 Model Code and Parameter Extraction*, EKV Model Users' Group Meeting and Workshop, EPFL, Lausanne, Switzerland, November 4-5, 2004

[21] M. Bucher, A. Bazigos, N. Nastos, Y. Papananos, F. Krummenacher, S. Yoshitomi, *Analysis of harmonic distortion in deep submicron CMOS*, Proceedings of the 2004 11th IEEE International Conference on Electronics, Circuits and Systems, p.395-398, 13-15 Dec. 2004