



Technical University of Crete

**Electronic and Computer
Engineering Department**

**Small and Large Signal Modelling
of MOSFETs at High Frequencies**

**Diploma Thesis
of
Chalkiadaki Maria-Anna**

Chania, October 2008



Technical University of Crete
Electronic and Computer Engineering Department

Small and Large Signal Modelling of MOSFETs at High Frequencies

Diploma Thesis
of
Chalkiadaki Maria-Anna

Committee:

Matthias Bucher, Assistant Professor (Advisor)
Konstantinos Kalaitzakis, Professor
Konstantinos Balas, Associate Professor

Chania, October 2008

Acknowledgements

Pursuing the diploma in Electronic and Computer Engineering field and after lots of endeavors while working with my diploma thesis in MOSFET Modelling, I realized that this work would never be fulfilled without the support of some people.

First, I am grateful to my advisor Professor Matthias Bucher, who gave me the opportunity to enter the unknown world of MOSFET modelling.

I would also like to express my sincere gratitude and appreciation to Dr. Antonios Bazigos. Without his continuous guidance and encouragement, my research work towards this thesis would never be possible.

Finally I would like to thank Dr. Sadayuki Yoshitomi who provided all the necessary measurements that I needed in order to implement this thesis.

Abstract

In modern integrated circuits (IC) the MOS transistor is a basic component. Nowadays, CMOS technologies offer transistors with channel lengths that do not exceed a few decades of

nanometers. Furthermore, much more complicated structures of multigate devices tend to gain ground. Advanced CMOS technologies are very attractive for IC design due to their flexibility, availability, low cost and high integration. CMOS technology covers the needs of high-volume production of consumer electronics, and in recent years, especially also in the domain of wireless radio communication products.

However, advanced CMOS technology is highly complex, and presents intense short-channel and parasitic effects. An analytical MOSFET model should be able to predict with accuracy the physical behaviour of the transistor, and should also be adequately simple for the designers who wish to have an insight into the device behaviour.

This diploma thesis deals with CMOS modelling at radio frequencies (RF). A short introduction to the operation of the MOSFET device will be presented. Basic definitions of modelling and compact modelling are given. Furthermore, the increased demand of fully integrated RF transceivers calls for implementation of power amplifiers within the same chip. This poses the problem of the correct representation of non-linear effects in the devices. Differences of modelling between low and high frequencies as well as small and large signal are described.

The EKV3 MOSFET compact model, which is used throughout this thesis, is based on the charge sheet theory. The main equations of the core of the model will be presented. Moreover the EKV3 model includes all the phenomena that appear in deep submicron CMOS technologies, so a general description of them will be done.

The EKV3 model has been designed for computational efficiency and ease of parameter extraction. A complete model must offer a methodology that allows the fitting of the electrical characteristics of a certain technology, with the use of the measurements over a range of devices of this specific technology. In this thesis, the principles of model parameter extraction are presented. This work is based on data from advanced 110 nm and 90 nm CMOS technologies, with static and RF data (S-parameters) in the range of 50MHz to 20.5GHz, as well as load-pull RF measurement, at 5.8 GHz.

TABLE OF CONTENTS

Acknowledgements

Abstract

Chapter 1: Introduction

- 1.1 The MOSFET device*
 - 1.1.1 Physical Structure of MOSFET*
 - 1.1.2 MOSFET's Operation*
 - 1.1.3 CMOS Circuits and Threshold Voltage*
- 1.2 MOSFET Modelling*
 - 1.2.1 Compact Modelling*
- 1.3 Low and High Frequencies*
- 1.4 RF Modelling*
 - 1.4.1 Small Signal Modelling*
 - 1.4.2 Large Signal Modelling*

Chapter 2: EKV3 Model

- 2.1 A Charge-Based Model*
 - 2.1.1 General Equations*
- 2.2 High and Low Frequency Macromodels*
- 2.3 Phenomena and Related Parameters*
 - 2.3.1 Effects in Wide and Long Transistors*
 - 2.3.2 Short Channel Effects*
 - 2.3.3 Narrow Channel Effects*
 - 2.3.4 Effects related to Temperature*
 - 2.3.5 Other Channel Effects*
 - 2.3.6 Noise*

Chapter 3: Parameter Extraction of EKV3 MOSFET model

- 3.1 Static and IV Analysis for Model Parameters Extraction at LF*
 - 3.1.1 Long and Wide Channel Device*
 - 3.1.2 Short and Wide Channel Device*
 - 3.1.3 Length Scaling for Wide Channel Devices*
 - 3.1.4 Long and Narrow Channel Device*
 - 3.1.5 Width Scaling for Long Channel Devices*
 - 3.1.6 Short and Narrow Channel Devices*
- 3.2 Model Parameters Extraction related to Temperature*
- 3.3 Y-Parameter Analysis for Model Parameters Extraction at RF*
 - 3.3.1 General Equations*
 - 3.3.2 Extraction of RF Parameters*
- 3.4 Large Signal Modelling*
 - 3.4.1 Theory of Load-pull Measurement*
 - 3.4.2 Load-pull Measurement Setups*
 - 3.4.3 Load-pull Results*

Chapter 4: Evaluation of the EKV3 MOSFET model

4.1 Evaluation for NMOS devices

4.1.1 Static Analysis

4.1.2 Y-Parameter Analysis

4.1.3 Load-pull Analysis

4.2 Evaluation for PMOS devices

4.2.1 Static Analysis

4.2.2 Y-Parameter Analysis

4.2.3 Load-pull Analysis

4.3 Evaluation Results

Chapter 5: Conclusions

5.1 Future Work

Appendix A: Model Cards for NMOS and PMOS Transistors

References

Chapter 1

Introduction

1.1 The MOSFET device

Nowadays, transistors are the basic elements of integrated circuits (ICs). The major difference between transistor and passive elements (resistors, capacitors, inductors and diodes) is that the transistor's current and voltage characteristics depend on the voltage (or current) on a control terminal. There are two types of transistors with different physical principles: the bipolar transistors and field effect transistors (FETs). One type of bipolar transistor exists and specifically the bipolar junction transistor (BJT), and two types of FET transistors-the junction field effect transistor (JFET) and the metal oxide semiconductor field effect transistor (MOSFET). Today the BJT transistors are used in specific digital technologies and more generally in analog circuits, the JFETs have specific application and are not used in digital applications but MOSFETs appear in more than 90% percent of digital ICs application.

MOSFETs have three signal terminals: gate (G), source (S) and drain (D), plus a bulk terminal (B) to which the gate, drain and source voltages are referenced. Figure 1.1 (a) shows a three dimensional representation with its four terminals and its insulator (with thickness T_{ox}) between gate and bulk and Figure 1.1 (b) shows the commonly used symbols for MOSFETs, where the bulk terminal (B) is either labeled or implied. There are two types of MOSFET transistor, the NMOS and PMOS, depending on the polarity of the carriers responsible for conduction.

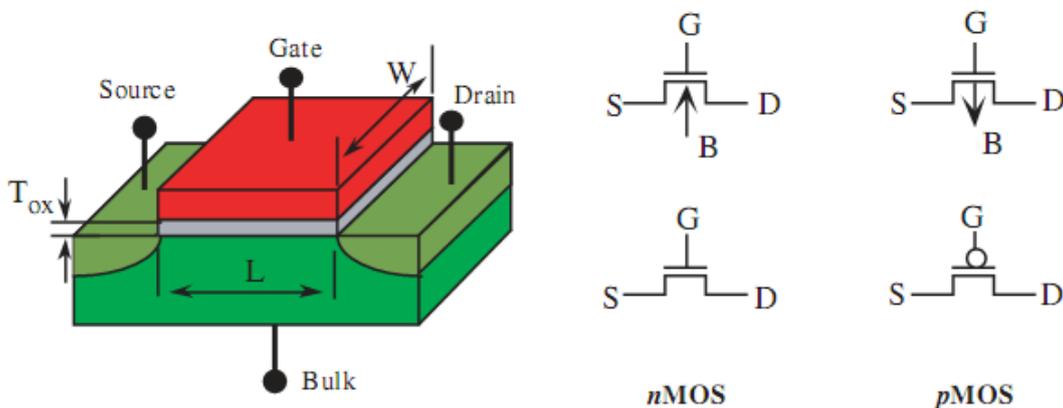


Figure 1.1 (a) MOS structure (b) Symbols used at the circuit level

1.1.1 Physical Structure of MOSFET

The MOSFET transistors are made from a crystalline semiconductor that forms the host structure called the substrate or bulk of the device. The thin oxide of the transistor electrically isolates the gate from the semiconductor crystalline structures underneath. Drain and source regions are made from crystalline silicon by implanting a dopant with polarity opposite to that of the substrate.

The region between the drain and source is called the channel and the distance from the drain to the source is a geometrical parameter called channel length (L) of the device. There is also the transistor channel width (W). These two parameters are defined by the circuit designer and are shown in Figure 1.1(a). However some other parameters, such as the transistor oxide thickness, threshold voltage, and doping levels, depend on the fabrication process, and cannot be changed in the design level-they are technology parameters.

The gate is the control terminal, and the source provides electron or hole carriers that are collected by the drain. Often, the bulk terminals of all transistors are connected to the ground or power rail that is often the source and, therefore, not explicitly drawn in most schematics. The gate is electrically isolated from the drain, source, and channel by the gate oxide insulator.

Figure 1.2 shows NMOS and PMOS transistor structures. The NMOS transistor has a p-type silicon substrate with opposite doping for the drain and source. PMOS transistors have a complementary structure with an n-type silicon bulk and p-type doped drain and source regions. Since drain and source dopants are opposite in polarity to the substrate (bulk), they form pn junction diodes (Figure 1.2) that in normal operation are reverse-biased.

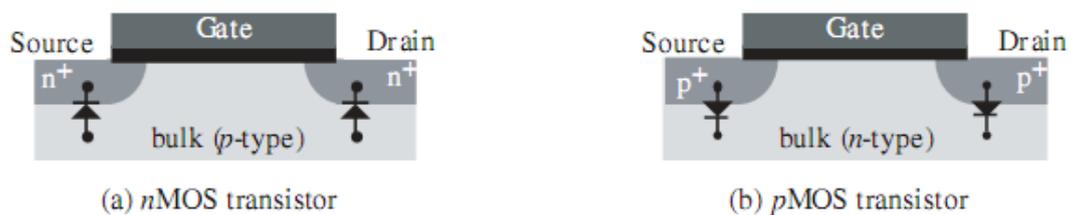


Figure 1.2 Relative doping and equivalent electrical connections between device terminals for (a) NMOS and (b) PMOS transistors

1.1.2 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. It is assumed that the bulk and source terminals are connected in order to simplify the description. (Figure 1.3)

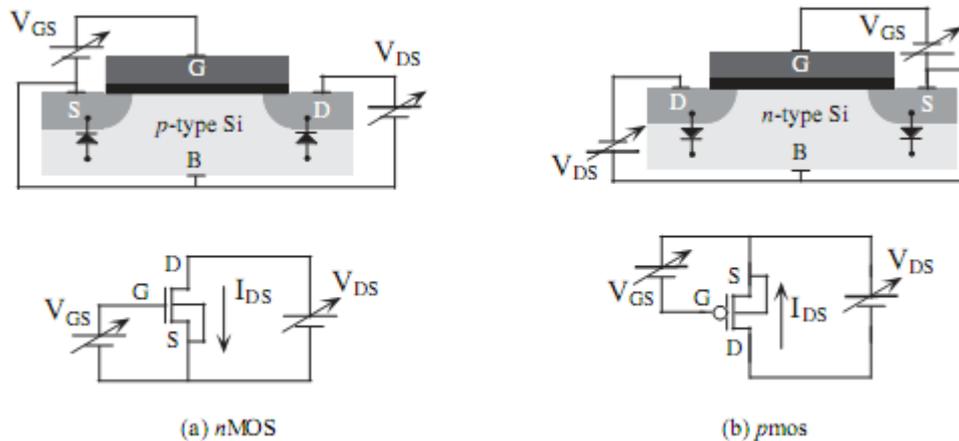


Figure 1.3 Normal transistor biasing (a) NMOS and (b) PMOS.

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 (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 (Figure 1.3), and there are no free charges between the drain and source. As a result, there is no current when $V_{GS} = 0$ for NMOS devices (the same hold for PMOS devices). This is the off, or nonconducting, state of the transistor.

At first the transistor operation when the source and substrate are at the same voltage is analyzed. When the voltage at the gate terminal an NMOS (PMOS) device is slightly increased (decreased), a vertical electric field starts to exist between the gate and the substrate across the oxide. In NMOS (PMOS) transistors, the holes (electrons) of the p-type (n-type) substrate close to the silicon–oxide interface initially “feel” this electrical field, and move away from the interface. As a result, a depletion region forms beneath the oxide interface or this small gate voltage (Figure 1.4). 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.

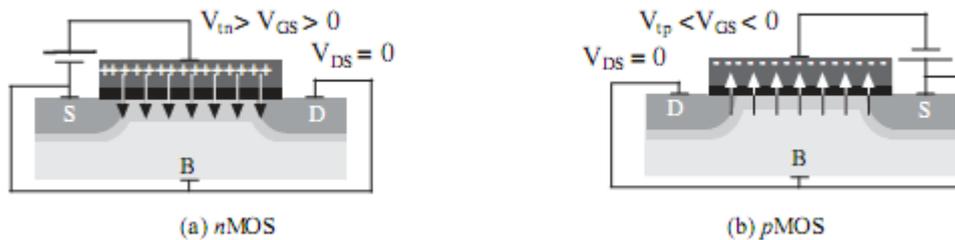


Figure 1.4 (a) Depleting the NMOS channel of holes with small positive values of gate–source voltage, and (b) depleting the PMOS channel of electrons with small negative values of gate–source voltage.

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 MOS 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 p-MOS device). These carriers form the inversion region or conducting channel, which can be viewed as a “short circuit” to the drain/source–bulk diodes. This connection is shown in Figure 1.5.

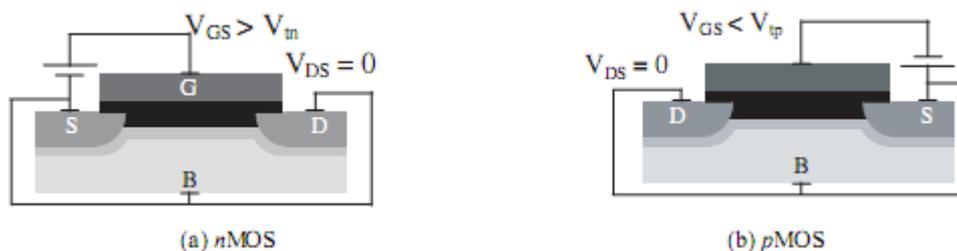


Figure 1.5 Creating conducting channel for (a) NMOS and (b) PMOS transistors.

Since the drain and source 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, referred to as V_{th} . As an approximation, V_{th} can be considered constant for a given technology. The threshold voltage of a NMOS transistor is positive, while for a PMOS transistor it is negative. Since NMOS and PMOS transistors have different threshold voltages, V_{thn} refers to the NMOS transistor threshold voltage 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 creates a horizontal electric field, moving the channel carriers toward 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:

$$\begin{aligned} V_{DS} < (V_{GS} - V_{thn}) & \quad \text{NMOS transistor} \\ V_{DS} > (V_{GS} - V_{thp}) & \quad \text{PMOS transistor} \end{aligned} \quad (1.1)$$

The above conditions state that the vertical electric field dominates the horizontal one. The transistor is in its linear region, also called the ohmic or nonsaturated region. If the drain voltage increases beyond the limit of Equation (1.1), the horizontal electric field becomes stronger than the vertical 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 1.6.

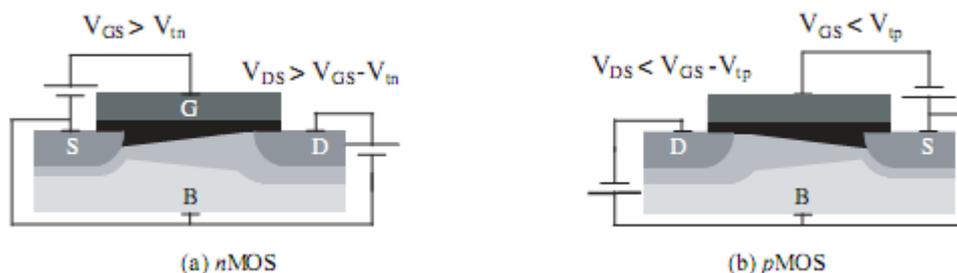


Figure 1.6 Channel pinch-off for (a) NMOS and (b) PMOS transistor devices.

Although there are no inversion charges at the drain end of the channel, the drain region is still 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 described here: off-state, saturated state, and the linear state.

1.1.3 CMOS circuits and threshold voltage

In the above paragraph it was assumed that the transistor source and substrate terminals were connected to the same voltage. This may be correct for isolated transistors, but in CMOS circuits where both PMOS and NMOS transistor are used, this condition may not hold for all devices. Figure 1.7 is a circuit cross

section of two NMOS and one PMOS transistors fabricated in a CMOS process. All devices are constructed on the same p-type silicon substrate but since PMOS transistors are formed on n-type substrates, there must be a region of that is oppositely doped to the initial bulk, forming what is called a well. The p-type substrate for NMOS transistors is connected to zero (or ground, GND), whereas the n-type well is connected to V_{DD} , since it forms the bulk of the PMOS transistors.

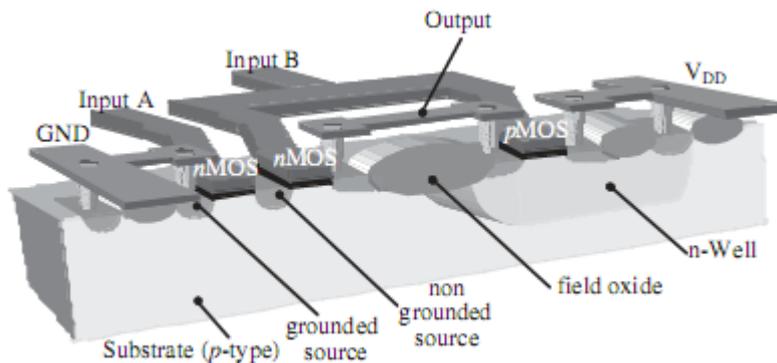


Figure 1.7 Structure for two series-connected NMOS transistors and one PMOS transistor in a CMOS technology.

When the source and substrate voltages differ, the gate-source voltage is not fully related to the vertical electric field responsible for creating the channel. The effect of the higher source voltage above (below) the substrate for an NMOS (PMOS) transistor is to lower the electric field induced from the gate to attract carriers to channel. The result is an effective raising of the transistor threshold voltage.

The threshold voltage can be estimated as:

$$V_t = V_{t0} \pm \gamma \sqrt{V_{SB}} \quad (1.2)$$

where V_{t0} is the threshold voltage when the source and the substrate are at the same voltage, and γ is a parameter dependent on the technology. The parameter γ is called the body effect coefficient. The positive sign is used for NMOS transistors, and the negative sign for PMOS transistors. When the source and substrate are tied together, $V_{SB} = 0$, and the threshold voltage is constant.

1.2 MOSFET Modelling

Nowadays, with the explosive growth of the IC industry and modern large-scale integrated circuits, the modelling of MOSFETs has raised in a big issue among the scientific and industry communities.

The models are used widely by the IC designers, who rely on simulations of their design before building a prototype and thus the models' role, which provide a mathematical description of the element behaviour in the circuit, becomes very important. Specifically, the accuracy of the results which come from the simulations depends on the quality of the element models, which means that the integrated circuit will behave as the designed ones.

For non critical digital circuits, this model may in principle be very simple. It would be sufficient to model each transistor as an on-off switch in order to design purely logic circuits. However, there are critical races among transitions, so the model must be extended to describe the dynamic behaviour of the device, in order to obtain the rise and fall time of these transitions. This dynamic behaviour is also needed when the frequency of operation approaches its maximum limit.

Analog circuits may contain a smaller number of transistors than digital ones, but they are more dependent on the exact behaviour of each transistor. The design of high-performance analog circuits requires a very detailed model of the transistor. This model must include a precise description of the voltage-current relationships, including the effect of the source that is often not grounded, and of the dynamic behaviour of the device. Its behaviour with respect to noise and to temperature variations must also be accounted for.

At this point, the two categories of device models should be differentiated, namely the numerical device simulation models and compact models. Numerical device simulators are used to study the device physics and to predict the electrical, optical, and thermal behaviour of the device. Numerical device simulators solve a set of (partial differential) equations associated with the physics involved in device operation but their requirements of intensive computation and huge amount of memory prevent them from being used for circuit simulations.

On the other hand, compact models describe the terminal properties of the device by means of simplified set of equations or by an equivalent circuit model.

From the above it can be concluded that the conditions for a model to be regarded as good is how correct and accurate can predict the behaviour if a MOSFET and how demanding can it become, in the simulation level. A good model must be accurate and fast.

1.2.1 Compact Modelling

The purpose of a compact model is to obtain simple, fast and accurate representations of the device behaviour. Compact transistor models are needed to evaluate the performance of

integrated circuits containing a large quantity of transistors. MOSFET compact models fall into three categories: (a) Physical Models, based on device physics, (b) Table Lookup Models, in the form of tables containing device data for different bias points and (c) Empirical Models, which represent the device characteristics through equations that fit data. Physical models take considerable time to develop but once they become mature, their advantages are significant: parameters have physical meaning, effects of device geometry, technological parameters, and temperature can be accounted for, statistical modelling can be applied to predict ranges of expected performance and, in many cases, the model can be applied to different generation technologies by simply changing parameters.

1.3 Low and High Frequencies

When transistors operate at low frequencies, we assume that the frequency does not influence their behaviour. This assumption, wants the charges and currents of the transistor to respond instantly to changes of the potential at the terminals. This is called quasi-static operation and has an upper frequency limit of validity. Beyond that limit, though, the transistor operation is called non-quasi static and the charges need time to adapt to voltage changes.

The frequencies used for radio communication cover a wide range. Usually RF modelling of MOSFETs refers to modelling at frequencies up to the transition (cutoff) frequency of the devices. For such application the non-quasi static modelling of the intrinsic part becomes indispensable. However the extrinsic part must be considered too and can be modeled using lumped approximations.

More detailed, let us consider the intrinsic part of a transistor, where in general each terminal consists of a dc bias and a small-signal sinusoidal component. Assume that only one small-signal voltage is nonzero at a time. If v_S is varying very slowly, the inversion layer charge has time to follow with practically no delay. However, if the variation of v_S is fast, the "inertia" of the inversion layer becomes no negligible, and the effect (get charge change) will lag behind the cause (source voltage change). A similar effect will be observed between the drain and the gate and similar conclusions can be drawn for the effect of the source and drain on the substrate. Consider now the effect of the gate voltage. If v_G is varying very fast the inversion layer charge does not have enough time to respond fully, and thus $|y_{dg}|$, which models this response will be small and the angle of y_{dg} is significant and negative, because of the delay between the cause (the variation in the gate voltage) and the effect (the variation in the gate current). Finally, similar observations hold for the effect of the substrate voltage on

the inversion layer charge. The above effects will be observed if the operating frequency exceeds the upper limit of the quasi-static modelling.

1.4 RF Modelling

1.4.1 Small Signal Modelling

In order to model phenomena that appear when transistors operate at high frequencies, a simple way, replaces the transistor with a structure of transconductances, with real and imaginary part. The same structure can be used when operating at low frequencies but the transconductances and transcapacitances will have zero imaginary part. The above solution may be used in a small signal analysis, since it replaces the transistor with a circuit that depends on the polarization of the transistor. Furthermore, it is an analytical solution which helps in better understanding of the phenomena that appear at high frequencies. Figure 1.8 shows the two small signal models that can replace the transistor at quasi-static and non-quasi-static region. These models replace the intrinsic part of the transistor and do not contain the extrinsic part.

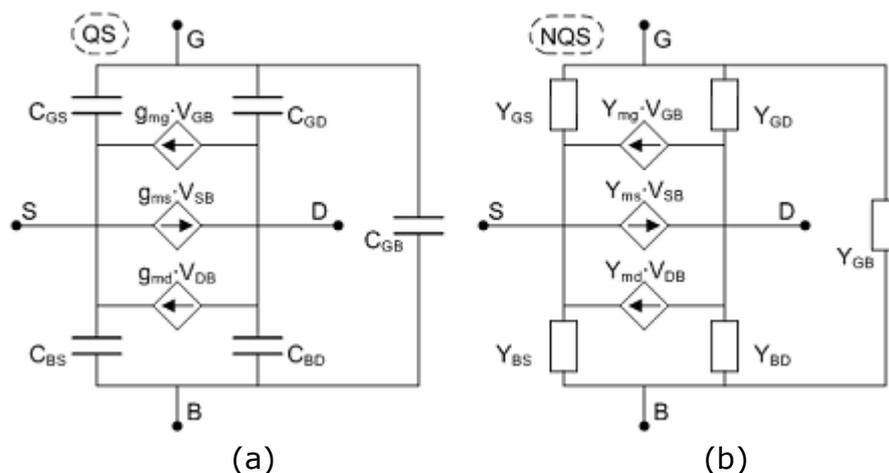


Figure 1.8 (a) quasi-static small signal model that can replace a MOSFET at low frequencies, (b) non-quasi static small signal model that can replace a MOSFET at high frequencies. Their difference is that the NQS model has complex transadmittances and transcapacitances in contrast to QS where these are real.

1.4.2 Large Signal Modelling

The previous approach, described a model that can replace the non-quasi-static behaviour of a transistor but only for small-signal analysis. In order to have a model that will operate in every analysis (small and large signal), a general solution should be

adopted. This solution could be a technique that is called channel segmentation, and replaces the MOSFET with more than one shorter MOSFETs in series with equal total gate length.

The main motivation to adopt channel segmentation scheme is to provide a large-signal NQS model consistent with the small-analysis. Of course, when segmenting the MOSFET channel into shorter parts, nodes are introduced between each channel segment increasing the computational effort, so the number of segments should be selected in a suitable trade-off among accuracy and speed. The goal is the channel segmented transistor to have the same static behaviour as a single-channel transistor to ensure consistency among QS and NQS models, so the length of each segment is chosen such that, for it, quasi-static model can be used. The combination of the models of all sections will then be a valid model for the whole transistor at the frequency of interest. The higher this frequency, the shorter the length needed for each segment.

Here it should be noted that each segment is a shorter part of the entire channel and not a shorter transistor. Extrinsic elements, such as series external resistors and overlap/fringing capacitances, are naturally placed only at the ends of the channel and not between the segments of the channel. On the other hand the length related parameters need to be scaled down, according to the shorter length of the segments.

It is obvious, that in static and low frequencies the quasi-static model and the non quasi-static model should give the same results. Figure 1.9 shows a channel segmented non-quasi-static transistor.

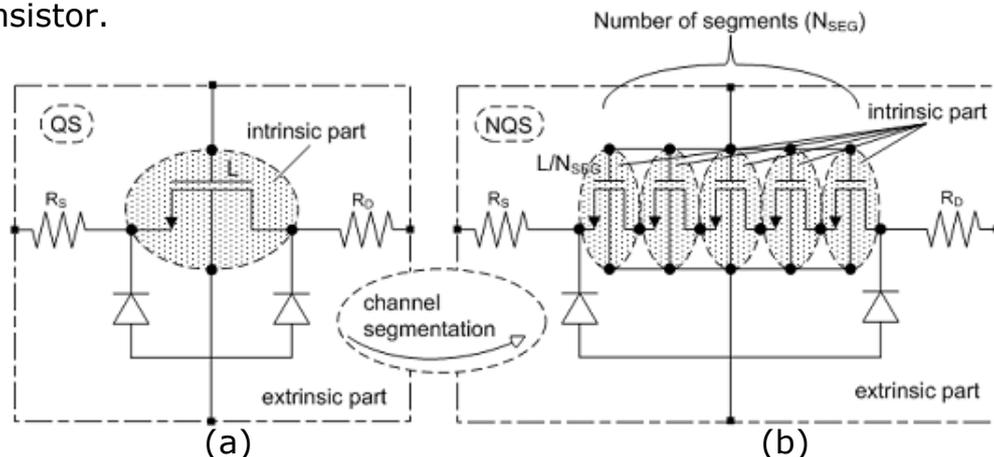


Figure 1.9 (a) a quasi-static model, (b) a channel segmented non-quasi-static model

Chapter 2:

EKV3 Model

2.1 A charge-based model

The requirements for good MOS analog simulation models such as accuracy and continuity of the large and small signal characteristics are well established. The EKV3 MOSFET model has been designed in order to have computational efficiency, ease of parameter extraction and the designer's need for insight into the device behaviour.

The 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_{iS} and Q_{iD} at the source and drain ends of the channel, to calculate the drain current and to model all aspects of the device behaviour. Below an analysis without going into much detail follows.

For zero electric field at the silicon surface, the source to drain structure of Figure 2.1 corresponds to two back-to-back diodes connected in series; thus, no current other than the junction leakage current can flow as long as V_S and V_D are positive. The situation remains qualitatively the same when more holes are attracted at the surface by applying a negative gate voltage V_G .

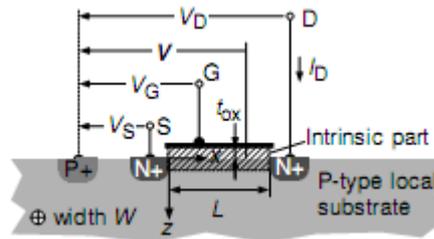


Figure 2.1 Cross Section of a MOS Transistor

On the contrary, if a positive voltage is applied to the gate, the holes are forced away from the surface, leaving the negatively charged P-doping atoms. As shown schematically in Figure 2.2, this corresponds to a negative charge of density Q_b per unit area. This charge is fixed and therefore cannot carry any current. By further increasing V_G , negative electrons are attracted to the surface thereby forming an N-type channel. It is this negative mobile inversion charge, of density Q_i per unit area, which will carry the drain-to-source current by a combination of drift and diffusion mechanisms of electrons. For the N-channel device, this current I_D will be defined positive if it enters the drain terminal.

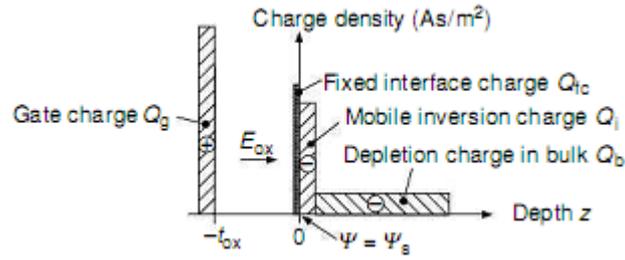


Figure 2.2 representation of various local charge densities

The total net charge induced underneath the surface of silicon per unit area of channel is given by

$$Q_{si}^{\Delta} = Q_b + Q_i \quad (2.1)$$

As depicted in Figure 2.2, an additional component of charge Q_{fc} is present at the silicon-oxide interface. This is fixed charge that includes the effect of charges trapped inside the oxide and weighted by their relative distance to the interface. This charge will be assumed to be independent of the gate voltage, although it might change very slowly in time at very high values of gate voltage.

The 0-reference of electrostatic potential ψ is that of the bulk of silicon, at a distance from the surface where it is not affected by the gate voltage. At the silicon surface ($z = 0$), ψ takes the particular value ψ_s called the surface potential. The electric field E_{ox} in the oxide depends on $V_G - \psi_s$, but is modified by Φ_{ms} (contact potential of body material to gate material), the difference between the extraction potentials of gate and channel materials. It corresponds to the barrier of potential that would be created at their interface if the oxide thickness t_{ox} would be zero. The electric field in the oxide is therefore given by:

$$E_{ox} = \frac{V_G - \Phi_{ms} - \psi_s}{t_{ox}} \quad (2.2)$$

2.1.1 General Equations

As shown in Figure 2.1, the active region of the transistor located between source and drain is limited to the gate-to-surface capacitor plus a thin layer of silicon in which the potential and the charge distribution are modified by the effect of the gate. It is called the intrinsic part of the transistor. All the rest is the extrinsic part. It includes the source and drain diodes, series access resistors or inductors to the four terminals, and all external parasitic capacitors, in particular those of the D and S junctions and the direct overlap capacitors from gate electrode to source and drain islands.

When voltage is applied either to the source or to the drain terminal, electrons and holes are forced out of equilibrium, splitting

their respective quasi-Fermi potential by V_S at the source end of the channel and V_D at the drain end. This splitting propagates along the channel, and can be characterized by a channel voltage V that varies monotonically from V_S at $x = 0$ (source end) to V_D at $x = L$ (drain end). For an N-channel device in normal operation (potential increased at the surface by the voltage applied to the gate), the quasi-Fermi potential of holes can be assumed to be constant throughout the structure. Thus V is (within a constant) the quasi-Fermi potential of electrons in the channel.

Another important voltage is the thermodynamic voltage given as:

$$U_T = \frac{kT}{q} \quad (2.3)$$

where k is the Boltzmann constant and q is the elementary charge. Proportional to the absolute temperature T , it is a measure of the thermal energy of electrons. Since it appears ubiquitously in MOS modelling equations, it is a more natural unit of voltage for devices and circuits than the standard unit of 1 V. Its value is 25.8 mV at 300 K or 27° C.

Below the main equations of the compact model's core, will be described.

Part 1:

Since the model is charge-based the equation which describes the general charge-voltage relationship for a long-channel transistor is:

$$2q_i + \ln q_i = v_p - v \quad (2.4)$$

This equation is dimensionless, it is valid in every point of the channel where u is the voltage and q_i the inverse charge of the point, and uses normalized variables. Only three model parameters and one physical parameter are needed to obtain from it the relation between applied voltages V_G and V , and the resulting inverted charge density Q_i .

The physical parameter, used to normalize all voltages in the dimensionless equation is U_T .

The following are the three device parameters:

1. The slope factor n which is defined as:

$$n = \frac{dV_{TB}}{d\Psi_s} = 1 + \frac{\Gamma_b}{2\sqrt{\Psi_s}} \quad (2.5)$$

where Γ_b is the substrate modulation factor. The slope factor n should be evaluated to best fit the range of operation of the transistor and n would be best evaluated as the slope of the secant of the threshold function V_{TB} between a particular value of surface potential Ψ_s and its pinch-off value Ψ_p given by:

$$n = n_{opt} = \frac{V_{TB}(\Psi_P) - V_{TB}(\Psi_S)}{\Psi_P - \Psi_S} = 1 + \frac{\Gamma_b}{\sqrt{\Psi_P} + \sqrt{\Psi_S}} \quad (2.6)$$

2. The threshold voltage V_{T0} which is the equilibrium threshold voltage is defined as:

$$V_{T0} \stackrel{\Delta}{=} V_{TB}(V=0) = V_{FB} + \Psi_0 + \Gamma_b \sqrt{\Psi_0} \quad (2.7)$$

It is very slightly dependent on V_G through Ψ_p but can be considered bias independent in practice.

These first two parameters relate the gate voltage V_G to the pinch-off voltage V_P according to:

$$V_P \cong \frac{V_G - V_{T0}}{n} \quad (2.8)$$

3. The oxide capacitance per unit area C_{ox} . It is combined with n and U_T to obtain the specific charge Q_{spec} defined by the following equation and is used to normalize the charge density:

$$Q_{spec} \stackrel{\Delta}{=} -2nU_T C_{ox} WL \quad (2.9)$$

Introducing these parameters in (2.4) provides the non normalized general charge-voltage relation:

$$\frac{-Q_i}{C_{ox}} + nU_T \ln \frac{-Q_i}{2nC_{ox}U_T} = V_G - V_{T0} - nV \quad (2.10)$$

It is obvious that U_T plays an important role in weak inversion where the logarithmic term dominates. This role disappears in strong inversion when the logarithmic term becomes negligible.

Part 2:

The equation (2.11) describes the general current-voltage relationship for a long-channel transistor.

$$v_p - v_{s,d} = \sqrt{1 + 4i_{f,r}} + \ln(\sqrt{1 + 4i_{f,r}} - 1) - (1 + \ln 2) \quad (2.11)$$

It is continuously valid from weak to strong inversion. Similarly to the first part, only three model parameters and one physical parameter are needed to obtain from this dimensionless equation the relation between bias voltages V_G and $V_{S,D}$, and the resulting current component $I_{F,R}$. The physical parameter is U_T . It is used to normalize all voltages in the dimensionless equation. The three device parameters are:

1. The slope factor n .
2. The threshold voltage V_{T0} .

These first two parameters were already introduced to obtain the charge-voltage relationship in Part 1.

3. The transfer parameter β given by:

$$\beta = \mu C_{ox} \frac{W}{L} \quad (2.12)$$

where μ is equivalent mobility of electrons in the channel. It is combined with n and U_T to obtain the specific current I_{spec} defined as:

$$I_{spec} = \mu U_T \frac{W}{L} (-Q_{spec}) = 2n\mu C_{ox} \frac{W}{L} U_T^2 = 2n\beta U_T^2 \quad (2.13)$$

is used to normalize components I_F and I_R of the drain current. The drain current in all the modes of operation of the transistor is represented in Figure 2.3 and can be obtained by subtracting $I_R(V_D, V_G)$ from $I_F(V_S, V_G)$ according to

$$I_D = \beta \int_{V_S}^{\infty} \frac{-Q_i}{C_{ox}} dV - \beta \int_{V_D}^{\infty} \frac{-Q_i}{C_{ox}} dV = I_F - I_R \quad (2.14)$$

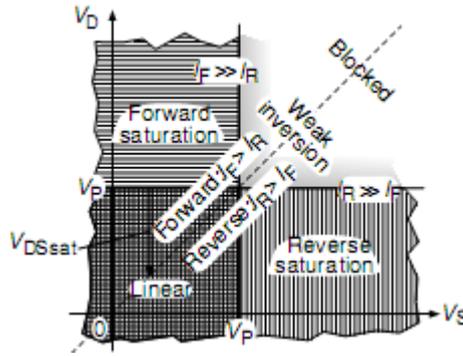


Figure 2.3 Modes of operation of a MOS transistor

The normalized inversion charges at source and drain (q_s and q_d , respectively) may be calculated according to the difference between V_P and V_{SB} and V_{DB} , respectively.

$$q_s = F^{-1}\left(\frac{V_P - V_{SB}}{U_T}\right) \quad , \quad q_d = F^{-1}\left(\frac{V_P - V_{DB}}{U_T}\right) \quad (2.15)$$

$$F(q) = 2q + \ln q$$

Here, $F(q)$ is the normalization factor of the potentials.

Finally, the drain current is given by the following equation:

$$I_{DS} = I_{SPEC} * (q_s^2 + q_s - q_d^2 - q_d) \quad (2.16)$$

The equations for inversion charge (2.15) and drain current (2.16) may be considered as the core of a MOSFET model that has only two parameters, V_P and I_{SPEC} , and two inputs, source and drain potentials. This core model produces the drain current according to the source and drain potential.

The rest of the model is responsible for describing all the higher-order effects. The EKV3 includes extensions that cover all the phenomena that appear in modern CMOS technologies. Macroscopically, this part of the model takes as input the geometry of the device and the potentials at the nodes, and feeds the core model with the effective values that needs as input. Note that the core charge model equations actually are a minor portion of the overall computational effort.

2.2 High and Low Frequency Macromodels

Although the fundamental behaviour is indeed dictated by the intrinsic part, the extrinsic part also plays an increasingly important role when either dimensions are reduced and/or the operating frequency is increased. Except from the intrinsic part of the MOSFET around it there are other natural structures which must be also modeled in order to describe accurately the transistor's behaviour.

To access the source and drain intrinsic terminals requires the source and drain extensions (SDE) as well as the source and drain diffusions which are covered with a silicide and contacted by vias. All these parts add some parasitic access resistances which are modeled by the source and drain resistances R_S and R_D .

The gate is made of polysilicon which is usually covered by silicide in order to lower the gate resistance. Although this resistance is small should be accounted particularly in RF IC design, where even small series resistances can play an important role. The access to the gate can also be modeled by a simple gate resistance R_G .

The modelling of the substrate access is a bit more difficult since it strongly depends on the device layout. Modelling it by a simple substrate series resistance R_B is usually sufficient in most cases however more accurate models are used for RF IC design.

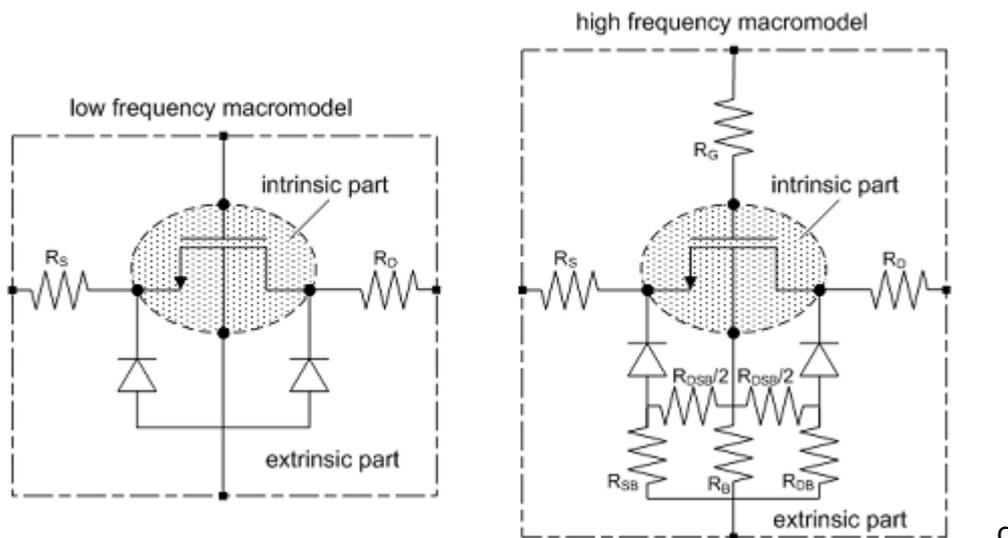
In addition to the four series access resistances R_S , R_D , R_G , and R_B , there are also additional parasitic capacitances. The overlap capacitances between gate and source C_{GS0} and between gate and drain C_{GD0} are due to the overlap of the gate and gate oxide over the SDE. These overlap arise after forming the SDE, by lateral diffusion of the SDE dopants under the gate.

There is also a fringing capacitance C_{GB0} which appears between the gate electrode above the field oxide and on top of the substrate.

In addition, the source and drain junctions and their extensions are modeled by the diodes SB between the bulk and the source and DB between the bulk and the drain which in dynamic operation they are modeled by two junction capacitances C_{BSj} and

C_{BDj} . The latter are obviously bias dependent and are also made of several parts.

Although it is always possible to model the device in great detail taking into account every little series resistances and capacitances, this results in an accurate but usually very complex equivalent circuit. Furthermore, all the components of this equivalent circuit can most of the time not be extracted from experiments in an accurate way, or some not at all. It is therefore important to find the right trade-off between the accuracy required by the circuit designers, which always depends on the circuit application, and the complexity of the equivalent circuit used for the design and the simulations. Also note that most of the parasitic components are distributed resistances and capacitances, which are then modeled by lumped elements. The equivalent circuit shown in Figure 2.4(a) models the extrinsic part of the MOS transistor is usually accurate enough for most of the circuit design applications and Figure 2.4(b) models the extrinsic of the MOS transistor for RF circuits, where an even more elaborate equivalent circuit is required, particularly for the substrate network.



extrinsic part (a) when operating at low frequencies, (b) when operating at high frequencies.

In the high frequency macromodel we can notice the symmetry at the substrate network resistance, in relation to the edges of the channel, which allows the model to represent with more accuracy the transistor's behaviour. Moreover, this structure, despite its complexity, is closer to the natural structure and thus the model is better in a wider range of frequencies.

2.3 Phenomena and Related Parameters

To characterize a MOSFET model as accurate it must be able to be used by all the possible technologies and dimensions. For this reason it must include all the effects that appear in modern CMOS sumicron technologies. In this part, all these effects and the related model parameters will be briefly described.

2.3.1 Effects in Wide and Long Transistors

In this part effects that refer to large devices will be described.

Vertical Field Mobility Effect

Firstly, there is dependence of the mobility of the vertical field effect, since the inverse charge is non homogenous in the channel. The mobility is influenced from the intensity of the vertical field.

a) Surface scattering effect

However, the maximum intensity of the field does not define the mobility, but a smaller one due to the existence of the surface scattering effect. The parameters related to this effect (surface scattering) are **E0** (First Order Coefficient for Mobility Reduction due to Vertical Field), **E1** (Second Order Coefficient for Mobility Reduction due to Vertical Field) and **ETA** (Mobility Reduction due to Vertical Field Factor).

b) Coulomb Scattering

Another mechanism that influences the mobility of electrons in the bulk is Coulomb Scattering and is related to the electrically charged ionized impurity atoms. Coulomb scattering increases at lower temperatures and higher doping densities, and is therefore important at low-temperature operation and/or with highly doped substrates as in advanced CMOS. The related model parameters are **ZC** (Coulomb Scattering Factor) and **THC** (Coulomb Scattering coefficient for normalized inversion charge).

Gate Depletion

In modern CMOS technologies the standard material for the gate formation is polycrystalline silicon heavily doped, so as to behave almost like a metal. However it is not possible to dope the polysilicon gate to arbitrarily high concentrations. As a result, depending on the doping type of the gate relative to the channel, the gate may deplete somewhat when an inversion layer is formed in the channel beneath it. Usually in NMOS transistors the gate is

doped n-type and p-type in PMOS. So the parameters which model this effect are **GAMMAG** (Body Effect Coefficient for Gate) and **TG** (Type of Gate). The latter parameter equals to -1 if the gate is doped opposite from the substrate, +1 if the gate is doped the same as the substrate and 0 for no polysilicon depletion effect.

Quantum Effects

As the thickness of the gate oxide becomes smaller and smaller, the quantum effects which appear in the gate plays an important role in the transistor. The smaller thickness of the gate leads to a more intense electric field in the channel. This results in an increase of the band gap that a carrier needs in order to exist at zero energy in the conduction. Quantum effects cause a lower effective value for the oxide capacitance, in both accumulation and inversion, while pinch-off voltage and surface potential are affected, too. The parameters related to this effect are: **AQMA** (Quantum Effect Coefficient for Accumulation Region), **AQMI** (Quantum Effect Coefficient for Inversion Region) and **ETAQM** (Quantum Effect Factor).

Drain Induced Threshold Shift

In short channel transistor, it has been observed, a form of breakdown, known as punchthrough. This phenomenon refers to the connection of the source and drain not due to the channel but due to their expansion which is related to the increase of V_{SB} and particularly of V_{DB} . To prevent punchthrough, ion implantation is used to selectively increase the substrate doping at the edges of the channel, at a depth slightly larger than the effective source and drain regions. These depletion regions are called pocket implants and result to a non homogenous channel and to drain induced threshold voltage shift (DITS) (since V_{TH} will be bigger at the pocket implants) which must be modeled. The parameters used are: **FPROUT** (Output resistance for DITS effect), **PDITS** (DITS parameter), **PDITSL** (DITS dependence on length), **PDITSD** (DITS dependence on drain bias) and **DDITS** (Smooth factor of DITS effect).

2.3.2 Short Channel Effects

Till now only transistors with large dimensions have been studied. However, transistors short channel are the most important in the design of analog and digital circuits.

Velocity Saturation

The gate in short channel devices does not have full control of the channel due to the longitudinal field that exists and which is not negligible compared to the transverse field caused by the gate over a large portion of the channel. The velocity of the carrier should be proportional to the longitudinal field; however there is a maximum velocity that the carriers can obtain. So the linear relationship between the longitudinal field and the velocity of the carriers is preserved only for low intense field and for greater values the velocity will tend to the maximum value. This is the velocity saturation effect and the parameters related are: **UCRIT** (Critical Velocity of Electrons) and **DELTA** (Order of velocity saturation model).

Channel Length Modulation

As mentioned above the channel may be divided into two parts, one with velocity saturation and one linear. So except from the voltage at the end of the linear part, its length must be also calculated. This is called channel modulation and the parameters related to this are: **LAMBDA** (Early effect factor) and **ACLM** (Channel Length Modulation Factor).

Reverse Short Channel Effect

The non-homogenous longitudinal doping of the channel, due to the higher doped pocket implants at the ends of the channel, which was mentioned in the large devices, becomes more important for short channel devices since the pocket implants cover a bigger percent of the channel. This has a direct influence on the threshold voltage, the Fermi voltage and the body effect coefficient. This effect is called reverse short channel effect (RSCE) and the parameters which model it are: **LR** (Length Factor for RSCE), **QLR** (Threshold Voltage Factor of RSCE), **NLR** (Body Effect Coefficient Factor of RSCE) and **FLR** (Bulk Fermi Potential of RSCE).

Charge Sharing Effect

Another effect, which appears both in short and narrow devices is the charge sharing effect. The charge in the channel region of a long-channel device is mainly influenced by the gate and the substrate. However at the ends of the channel the charge is also influenced by the source and the drain. This becomes important in short-channel devices since the regions of the channels which are influenced by the source and the drain cover a bigger percent of the channel. Equivalently, this happens also along the width of the

device. This results in a decrease of the body effect γ with the increase of the V_{SB} and V_{DB} . The parameters related to the charge sharing effect are: **LETA** (Short Channel Charge Sharing Coefficient), **LETAO** (Long Channel Charge Sharing Coefficient), **LETA2** (Short Channel Scaling Coefficient), **WETA** (Narrow Channel Charge Sharing Coefficient), **NCS** (Slope Factor Dependence from Charge Sharing).

Drain Induced Barrier Lowering

Except from the body coefficient, the voltages at the terminals at the ends of the channel also influence the surface potential throughout the channel. Since usually the voltage at the drain is higher than voltage at the source, the effect is called drain induced barrier lowering and results in a decrease in the pinch-off voltage V_P when the voltage at the drain increases. The parameters related to this effect are: **ETAD** (Primary DIBL Coefficient) and **SIGMAD** (Secondary DIBL Coefficient).

Overlap and Fringing Capacitances

The oxide under the gate creates the capacitance between the substrate and the gate, which creates the channel. However this oxide is usually bigger than the distance of the source and drain and thus the edges of it are between the gate and source and drain terminals. In these regions, overlap capacitances are created whose value, in short channel devices cannot be negligible.

The model parameters for the overlap capacitances are: **GAMMAOV** (Body effect coefficient of the overlap area), **GAMMAGOV** (Body effect coefficient of the gate of the overlap area), **VFBOV** (Flat-band voltage of the overlap area), **LOV** (Length of the overlap area), **VOV**, **CGSO** (Bias-independent gate to source overlap capacitance), **CGDO** (Bias-independent gate to drain overlap capacitance) and **CGBO** (Bias-independent gate to bulk overlap capacitance).

Except from the overlap capacitances just mentioned, there is another parasitic capacitance between the gate and the source and drain which is called fringing capacitance. There is the inner fringing capacitance which exists due to the electric field between the gate and the source and drain, whose lines do not pass only through the channel but connect the gate with the source and gate sides to the channel. The value of the inner fringing capacitance depends on the charge concentration in the channel. Furthermore there is the outer fringing capacitance which exists between the gate and the source, drain and substrate upper sides. This type of fringing capacitance depends on the materials and not on the voltages at the terminals.

The parameters for inner fringing capacitances are: **KJF** (Fringing capacitance factor), **CJF** (Fringing capacitance bias factor), **VFR** (Built-in correction for fringing capacitance) and **DFR** (Smooth factor of fringing capacitance model)

Series Resistance

Despite the fact that for long-channel devices, the source and drain terminals may be considered of negligible-resistance, in short channel devices, the resistance which appears in the inner of the effective source and drain regions due to the metal contact in the terminals, is important. The series resistance, results in a reduction in the effective drain-source voltage. The model parameters follow: **RSH** (Square resistance of active area), **RS** (LDD Source series resistance) and **RD** (LDD Drain series resistance).

2.3.3 Narrow Channel Effects

Inverse Narrow Channel Effect

Real devices for most fabrication processes are along channel's width further insulated so as to achieve channel electric isolation from the adjacent ones. In these processes a thick oxide, called field oxide, becomes thinner and defines the channel width as equal to the extent of the thin oxide region. There are two different processes: LOCOS (local oxidation of silicon) isolation, which results in the gradual transition from thick (field) oxide to thin (gate) oxide and an increase in the threshold voltage and the shallow-trench isolation (STI), which is a fully recessed isolation structure and leads to a reduce in the threshold voltage especially in narrow devices. This effect is equivalent to the reverse short channel effect and is called inverse narrow width effect (INWE). These two techniques are shown in Figure. 2.5.

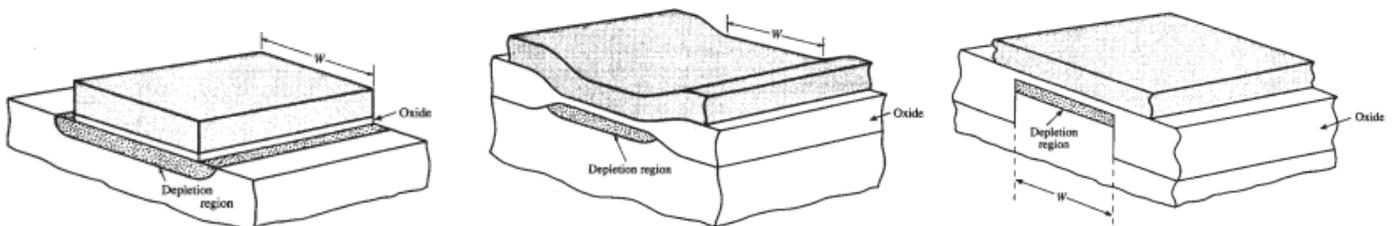


Figure. 2.5 Cross Sections along the width of MOSFET (a) simplified structure, (b) LOCOS-isolated device, (c) STI-device

Edge Conductance

In narrow channel devices, the charge sharing effect also exists as mentioned before and also edge conductance appears. The channel cannot be granted as homogeneous, which means that the

edges of the channel have a smaller threshold voltage than the main part of the model. In weak inversion this edges conduct first and therefore its current dominates. The parameters related to the edge conductance effect are: **WEDGE** (Width of edge conduction area), **DGAMMAEDGE** (Difference of body effect coefficient of edge conduction area with respect to the main part of the channel) and **DPHIEDGE** (Difference of fermi potential of edge conduction area with respect to the main part of the channel).

2.3.4 Effects related to Temperature

MOS transistor characteristics are strongly temperature dependent. The Fermi voltage has a linear dependence from the temperature and this influences the flat band voltage- V_{FB} , and thus the threshold voltage. Specifically, with the increase of the temperature the threshold voltage reduces and as a result there is an increase of the current for small gate voltages. Moreover the temperature increase, leads to a decrease to the effective mobility of the carriers in the channel.

The parameters related to temperature are: **TNOM** (Nominal Temperature), **TETA** (Temperature dependence of ETA), **TLAMBDA** (Temperature dependence of LAMBDA), **TCV** (Temperature dependence of VTO), **BEX** (Temperature dependence of mobility), **UCEX** (Temperature dependence of UCRIT), **TEOEX** (Temperature dependence of E0), **TE1EX** (Temperature dependence of E1), **TR** (First order temperature coefficient of resistors) and **TR2** (Second order temperature coefficient of resistors).

2.3.5 Other Channel Effects

Shallow Trench Isolation Stress Effect

Till now we have mentioned that the channel has a different behaviour at the edges. When the shallow trench isolation technique exists in a narrow transistor this behaves differently from a wide one. Furthermore, the STI structures stress the substrate and thus influencing its electric properties. This stress is proportional to the distance from the STI structure. This effect is more complicated in transistors with more than one finger. The stress of the STI structures has impact on the carriers' mobility and as a result on the DIBL, velocity saturation and some short channel effects. Moreover the change of the doping at the edges of the channel influences the threshold voltage and body effect

The parameters related to the STI stress effect are: **SA**, **SB** (Distance from STI), **KKP** (Mobility dependence on STI) and **KVTO** (Threshold voltage dependence on STI).

Impact Ionization Current

As the carriers accelerate due to the potential difference between the source and the drain, maximize their velocity. Near the drain terminal these fast moving carriers can impact on silicon atoms and ionize them, producing free current carriers. This is referred as impact ionization and these free carriers, due to V_{DB} , may create a current from the drain to the substrate (I_{DB}). This current is called impact ionization current and the parameters that model this effect are: **IBA** (I_{DB} coefficient), **IBB** (I_{DB} exponential factor) and **IBN** (I_{DB} factor of V_{SAT}).

Gate Current

The gate oxide in MOSFET forms a potential energy barrier that prevents carries from been injected into the gate from the silicon surface. However in modern MOSFET technology, where the oxide thickness becomes very small, the oxide becomes leaky enough and gate current becomes nonnegligible. This gate current is related to the tunneling effect and beside it there is another current which goes through the oxide in the region between gate and effective source and drain regions at the edges of the channel and is called overlap gate current. The parameters which model these currents are: **XB** (Silicon to Silicon oxide tunneling barrier height), **EB** (Characteristic electrical field), **KG** (Mobility for Gate Current) and **LOVIG** (Overlap Length for Gate current).

2.3.6 Noise

Noise plays an important role to the operation and the behaviour of circuits. The correct prediction of its value and its relationship of the various effects is a complicated procedure.

The noise that may appear in a MOSFET can be divided in four categories, shortly described below. (a) thermal noise, whose value is the same for every frequency and influenced by various effects, (b) flicker noise, which is conversely proportional to a positive power of the frequency, (c) induced gate noise, which is important in high frequencies analysis and is the noise which transfers to the gate due to non quasi static effects, and (d) shot noise, which is related to the leakage current of the gate.

Chapter 3

Parameter Extraction of EKV3 MOSFET model

The implementation of a MOSFET compact model lays on the description of the behaviour of the MOSFET, through a set of equations which combine the terminal properties, namely the terminal voltages and the currents through them, the charge and the noise at the terminals. The natural effects that happen inside the MOSFET are determined with the use of parameters, related to the technology used for the construction and the materials used.

There are two categories of parameters. The first refers to the instance parameters of the transistor, which describe each device separately through its physical dimensions. The most common are the gate length and width, the number of fingers in a multi-finger transistor etc. These parameters are defined by the designer in a circuit and depict the physical structure of the element in the integrated circuit.

The second category, the model parameters, characterize a specific CMOS technology and consequently all the transistor of this technology and independent of specific dimensions. These parameters are not defined by the user but are extracted through a wide range of measurements. These measurements must include static, in low or high frequency, for a variety of device dimensions, with or without noise and for different temperatures. Model parameters are related to the materials used and to the procedure of the construction.

The procedure of parameter extraction of a specific CMOS technology is not an effortless or a totally straight-forward procedure. Although EKV3 model, has the fewer parameters to extract, comparatively to other models, lots of phenomena that influence parameters' optimal value exist making this procedure quite a hard task if the steps to be followed are not determined. Due to the need of such a procedure determination, the guidelines of parameter extraction methodology will be displayed.

The model parameters extraction must be done separate for NMOS and PMOS transistors since they are made of different materials and through a different construction process. Here, the parameter extraction for NMOS transistor will be presented which is equivalent to the one followed for PMOS transistors. The model which was used is EKV301.01.

3.1 Static and IV Analysis for Model Parameters Extraction at LF

Throughout this analysis the different colors in the plots describe:

red: measurements
pink: increased value
blue: extracted value
green: decreased value

All measurements and simulations refer to T:27° C.

3.1.1 Long and Wide Channel Device

The procedure starts by describing long and wide devices, here a device with maximum **Length=10u** and maximum **Width=10u** has been chosen, which do not comprise many of the phenomena that appear in small devices.

C_{GG} vs. V_G analysis

Firstly in the procedure of parameter extraction, some parameters that describe some core physical properties will be extracted through the CV analysis. The capacitance per unit area of the oxide under the gate of the MOSFET (**COX**) is of primal importance in both the static and dynamic behaviour of the device. The nominal value of this parameter is equal to the ratio of the permittivity of the oxide (ϵ_{ox}) over the thickness of the oxide (t_{ox}), equation (3.1).

$$COX = \frac{\epsilon_{ox}}{t_{ox}} \quad (3.1)$$

The **VTO** is the reference parameter for threshold voltage in saturation and under $V_{SB} = 0$ bias for devices with no short or narrow channel effects. The Fermi potential of the body is given by **PHIF**, which nominally depends on the thermal potential (U_T), the doping of the body (N_A) and the intrinsic carrier concentration (n_i) and is given by equation (3.2).

$$PHIF = U_T * \ln \frac{N_A}{n_i} \quad (3.2)$$

The body effect coefficient (**GAMMA**) is also connected with N_A , and here it is used as a free parameter in order to allow more degrees of freedom during the parameter extraction procedure and obtain a better fit of the model on the measurements, equation (3.3).

$$GAMMA = \frac{\sqrt{2 * q * \epsilon_{si} * N_A}}{COX} \quad (3.3)$$

In case the material of the gate is highly doped polysilicon, and not metal, the **GAMMAG** reflects its doping (N_G), and an extreme high value for this parameter would remove the influence of the effect, making the gate material resemble an ideal metal, equation (3.4).

$$GAMMAG = \frac{\sqrt{2 * q * \epsilon_{si} * N_G}}{COX} \quad (3.4)$$

The value of these first parameters can be extracted from the C_{GG} vs. V_G , where C_{GG} is the capacitance between the gate terminal and the rest three terminals sort-circuited. The shape of this plot depends mostly on these parameters (see: fig.1, fig.2, fig.3, fig.4, fig.5). It should be noted here, that the effective value of the COX is influenced by quantum effect, whose parameters (AQMA, AQMI, ETAMQ) are also extracted by this analysis. As shown by the plots the value of COX is extracted either in strong inversion region or deep accumulation, the values of GAMMA is related with the minimum value of C_{GG} , the value of VT0 defines the transition point between the weak and strong inversion and the value of PHIF defines the value of balance voltage in relation to V_{th} .

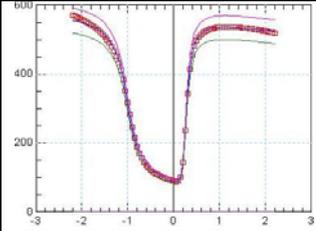
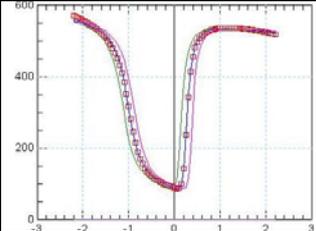
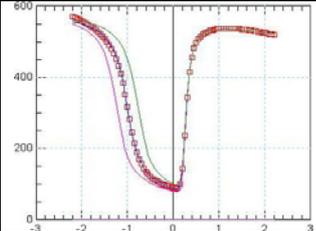
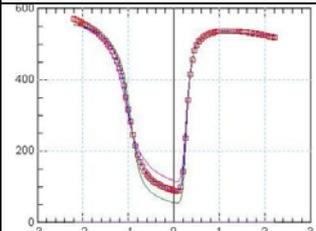
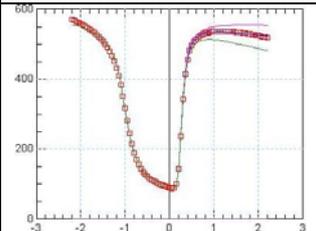
 <p>extracted value: 12.02m</p> <p>fig.1</p>	 <p>extracted value: 166.6m</p> <p>fig.2</p>	 <p>extracted value: 461.9m</p> <p>fig.3</p>
 <p>extracted value: 228.2m</p> <p>fig.4</p>	 <p>extracted value: 6.524</p> <p>fig.5</p>	

Table 3.1. V_G vs. C_{GG} , referred technology 110nm,
 $V_G = -2.2 \dots 2.2V$, $V_S = 0.0V$, $V_D = 0.0V$, $V_B = 0.0V$

I_D, g_m vs. V_G analysis

The next step is to use IV analysis for further parameter extraction. Through the logarithmic scale of I_D vs. V_G plot in the linear operation a small adjustment or verification of the extracted parameters **GAMMA**, **VT0** can be done when focusing in the weak inversion (see: fig.6, fig.7).

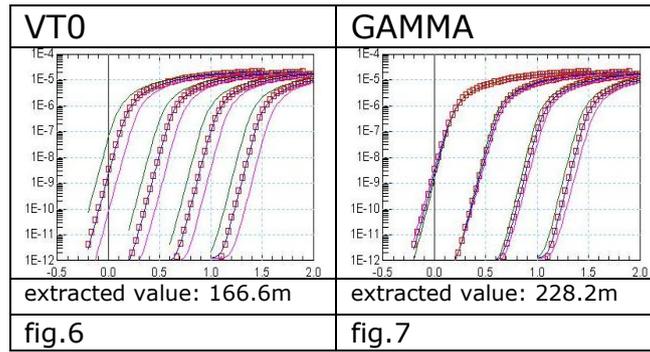
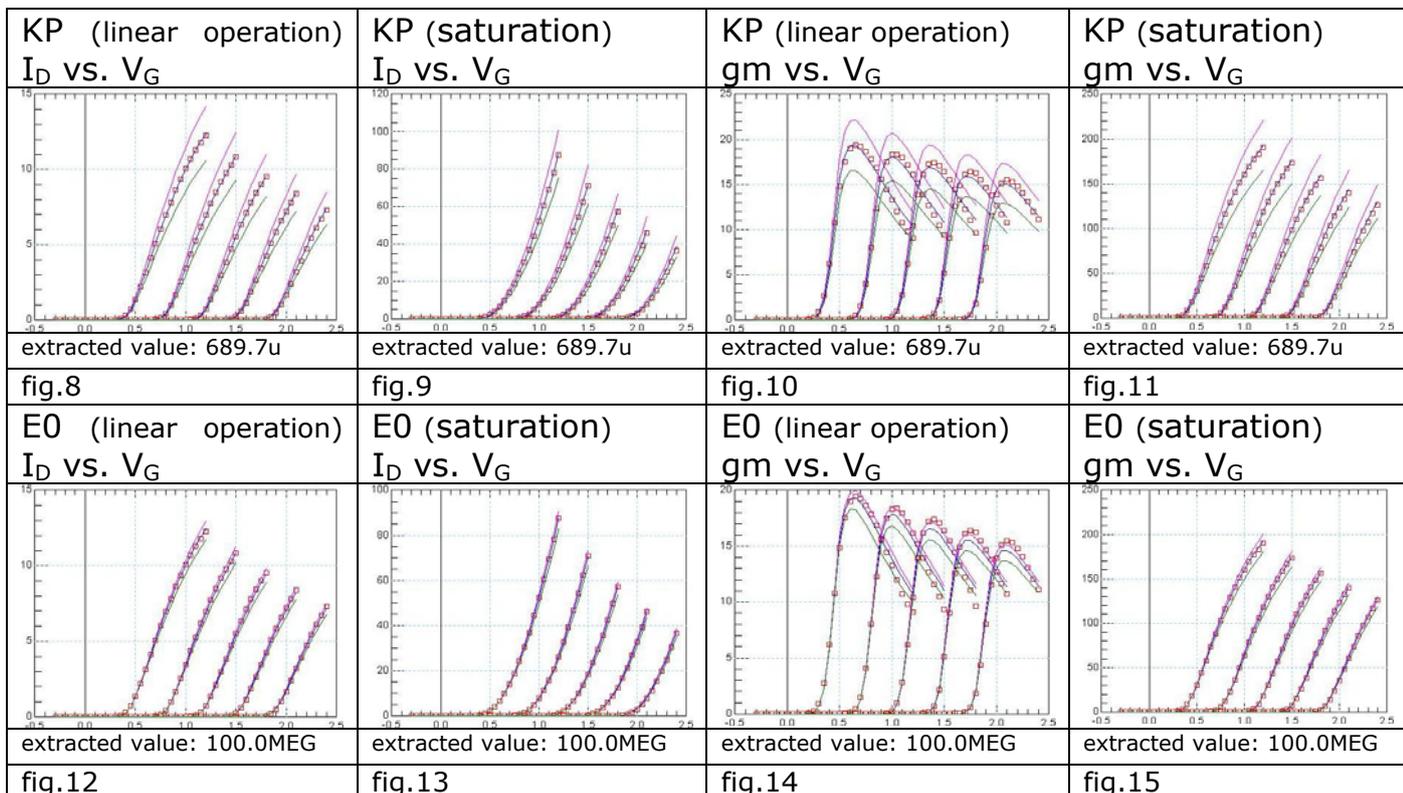


Table 3.2. $\log(I_D)$ vs. V_{GB} , referred technology 110nm, $V_{GB} = -0.2 \dots 1.5V$, $V_B = 0 \dots -1.2V$, $V_s = 0.0V$, $V_D = 50mV$

Mobility (μ) is passed into the model parameters via its product with the oxide capacitance per unit area (COX), forming the parameter **KP**, equation (3.5).

$$KP = \mu * COX \quad (3.5)$$

The effective mobility of the device illustrates an important dependence on the vertical field of the channel. **E0** and **E1** are, first and second order coefficients, respectively, for the mobility reduction due to vertical field. The **ETA** (Reduction due to Vertical Field Factor) nominally is 1=2 for the NMOS transistors and 1=3 for the PMOS, but is left as a parameter for best fitting. The above parameters may be extracted from the I_D vs. V_G and g_m vs. V_G analysis through the strong inversion in linear operation and saturation (see: fig.8, fig.9, fig.10, fig.11, fig.12, fig.13, fig.14, fig.15, fig.16, fig.17, fig.18, fig.19, fig.20, fig.21, fig.22, fig.23).



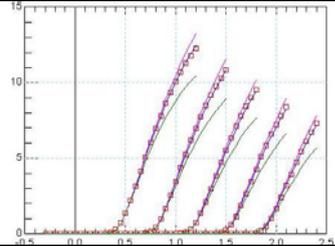
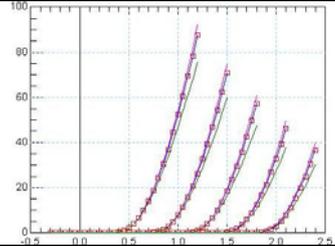
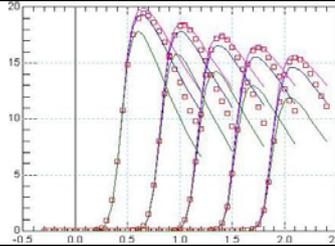
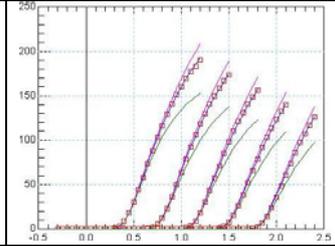
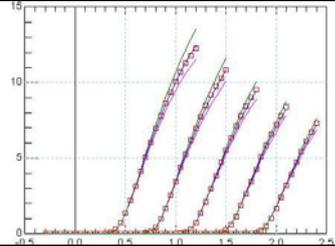
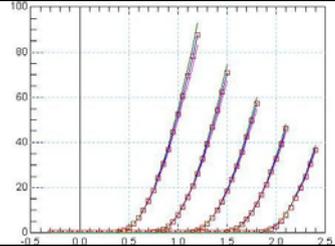
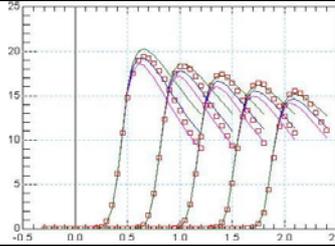
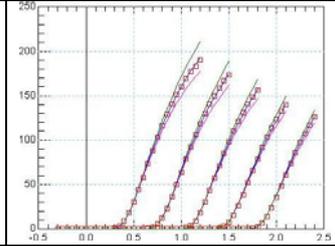
E1 (linear operation) I_D vs. V_G	E1 (saturation) I_D vs. V_G	E1 (linear operation) g_m vs. V_G	E1 (saturation) g_m vs. V_G
			
extracted value: 207.8MEG	extracted value: 207.8MEG	extracted value: 207.8MEG	extracted value: 207.8MEG
fig.16	fig.17	fig.18	fig.19
ETA (linear operation) I_D vs. V_G	ETA (saturation) I_D vs. V_G	ETA (linear operation) g_m vs. V_G	ETA (saturation) g_m vs. V_G
			
extracted value: 455.9m	extracted value: 455.9m	extracted value: 455.9m	extracted value: 455.9m
fig.20	fig.21	fig.22	fig.23

Table 3.3. referred technology 90nm, $V_G = -0.3...1.2V$,
 $V_B = 0...-1.2V$, $V_s = 0.0 V$, $V_D = 50mV$ (linear), $V_D = 1,2V$ (saturation)

Moreover through I_D vs. V_G analysis in linear operation we can extract the Coulomb Scattering Coefficients **ZC** and **THC**. The Coulomb Scattering phenomenon, affects the mobility of the carriers when the vertical electrical field is low. Note that the effect is more intense at low temperatures, while at room temperatures might be ignored for certain technologies. In order to show the influence of the ZC and THC parameters and due to the nature of the Coulomb Scattering Effect we used two parameter sets a) ZC=10m, increased THC=500m, decreased THC=-500m and b) THC=500m, increased ZC=50m, decreased ZC=5m (see: fig.24, fig.25).

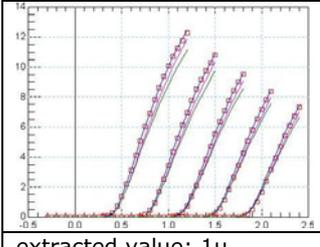
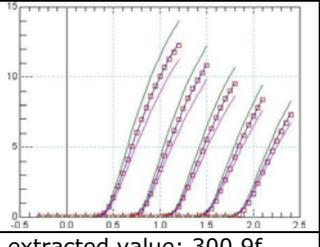
ZC	THC
	
extracted value: 1u	extracted value: 300.9f
fig.24	fig.25

Table 3.4. I_D vs. V_G referred technology 90nm,
 $V_G = -0.3...1.2V$, $V_B = 0...-1.2V$, $V_s = 0.0 V$, $V_D = 50mV$

Finally, through I_D vs. V_G analysis in linear operation we can extract all the parameters related to the Edge Conductance Effect. This effect appears in some modern CMOS technologies and due to it, the parts of the channel that are on the sides of the channel, are characterized by different properties with respect to the middle and main part of the channel, forming this way some parallel devices to the main one. This devices having, a lower pinch-of-voltage (**DPHIEDGE**), thus lowers threshold voltage, and body effect coefficient (**DGAMMAEDGE**) make their appearance sooner in weak inversion, hardly noticed in strong inversion. This parallel combination forms a step-shaped curve in weak inversion, with the current in logarithmic scale. The summed width of the two edge devices is given to the **WEDGE** (see: fig.31, fig.32, fig.33, .34, fig.35, fig.36).

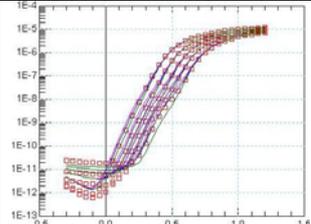
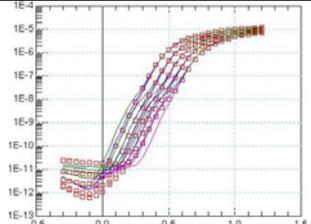
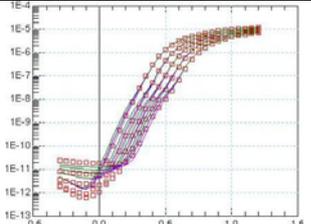
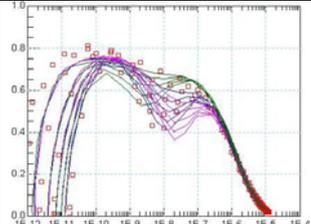
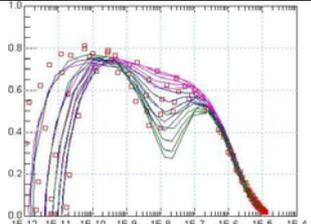
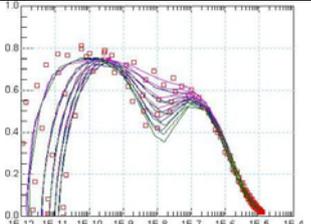
WEDGE ($\log(I_D)$ vs. V_G) 	DGAMMAEDGE ($\log(I_D)$ vs. V_G) 	DPHIEDGE ($\log(I_D)$ vs. V_G) 
extracted value: 36.53n	extracted value: -113.3m	extracted value: -86.38m
fig.30	fig.31	fig.32
WEDGE (g_{m_id} vs. V_G) 	DGAMMAEDGE (g_{m_id} vs. V_G) 	DPHIEDGE (g_{m_id} vs. V_G) 
extracted value: 70.0n	extracted value: -150.0m	extracted value: -40.0m
fig.33	fig.34	fig.35

Table 3.6. referred technology 90nm,
 $V_G = -0.3...1.2V$, $V_B = 0...-1.2V$, $V_s = 0.0 V$, $V_D = 50mV$

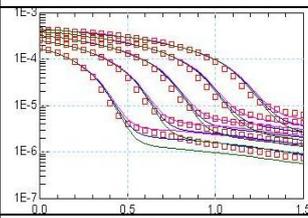
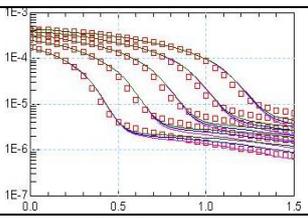
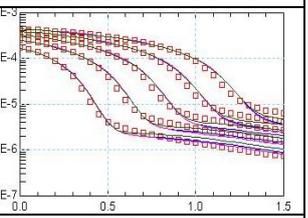
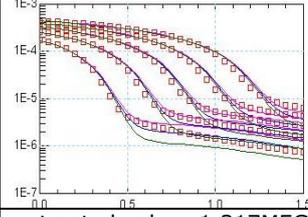
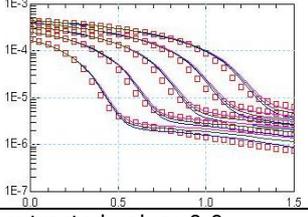
I_D , g_{ds} vs. V_D analysis

The pocket implants, that are created at the source and drain ends of the channel, protect mainly the short channel MOSFETs from certain short channel effects. However, their existence affects also the behaviour of long channel MOSFETs, and most importantly the output conductance ($g_{ds} = \frac{\partial I_D}{\partial V_D}$) in saturation. (3.6)

So the next step is to use g_{ds} vs. V_D analysis to extract parameters that are related with the influence of the horizontally non-homogeneous doping (pocket implants) **PDITS** (DITS

Parameter), **PDITS** (DITS dependence on drain bias), **DDITS** (Smooth Factor of DITS Effect), **FPROUT** (Output Resistance for DITS Effect), (see: fig.26, fig.27, fig.28, fig.29).

Also through this analysis is a charge sharing effect parameter, according to which the part of the channel that is close to its source and drain ends, is not modulated strictly by the gate potential, but also the source and drain potentials play a role. The effect is naturally more important in short channel devices but due to longitudinally non uniform doping, it also appears in long channel devices (**LETA0**), degrading the output conductance of the transistor (see: fig.30).

 <p>extracted value: 14.79u</p> <p>fig.26</p>	 <p>extracted value: 676.1m</p> <p>fig.27</p>	 <p>extracted value: 10.00</p> <p>fig.28</p>
 <p>extracted value: 1.317MEG</p> <p>fig.29</p>	 <p>extracted value: 0.0</p> <p>fig.30</p>	

note: in LETA0 graph both pink and green represent increased values

Table 3.5. g_{ds} vs. V_D , referred technology 110nm,
 $V_D = 0 \dots 1.5V$, $V_G = 0.6 \dots 1.5V$, $V_B = 0.0V$, $V_S = 0.0V$

I_G vs. V_G analysis

In modern CMOS technologies, with the downscaling of the minimum gate length, the oxide gets thinner and thinner. As a result the gate current that flow due to tunneling effect through the oxide gets to be more important. The model of the gate current is based on the calculation of the probability of a carrier to get across the oxide, which is calculated after the silicon to oxide silicon barrier height (**XB**) and a characteristic electrical field (**EB**). The estimation of the gate current uses also the **KG** which is directly proportional to the gate current, (see: fig.37, fig.38, fig.39). Note that setting the latter at zero value, disables the effect.

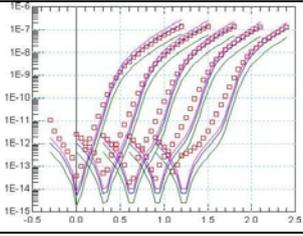
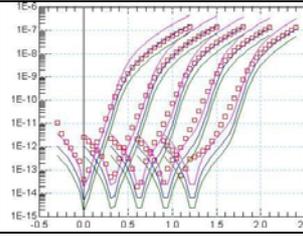
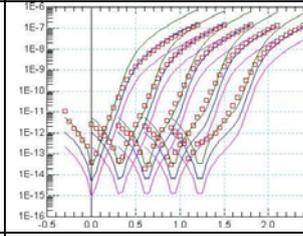
KG	XB	EB
		
extracted value: 1.569u	extracted value: 2.989	extracted value: 31.25G
fig.37	fig.38	fig.39

Table 3.7. I_G vs. V_G , referred technology 90nm,
 $V_G = -0.3 \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_S = 0.0V$, $V_D = 0,5V$

3.1.2 Short and Wide Channel Device

After extracting the parameters related to wide and long devices, the wide and short devices parameters follow. For this analysis, a device with minimum channel **Length=70n** and maximum channel **Width=10u** is chosen.

C_{GC} vs. V_{GB} analysis

The non-homogenous doping of the channel, due to the higher doped pocket implants at the ends of the channel, results into a dependence of the effective doping of the channel on its length, since the spread of the pockets is fixed for all geometries. This increase of the doping is reflected on an increase of the threshold voltage, the body effect coefficient and the Fermi potential. The relating parameters of this effect which is called, reverse short channel effect (**RSCE**) are firstly extracted.

More specifically, an arbitrary value for the length coefficient for the length scaling of this effect is given to the **LR** and then the values for the threshold voltage and the body effect coefficients, **QLR** and **NLR** respectively, are extracted so to have a good fitting. The values of **LR**, **QLR** and **NLR** which are related with the reverse short channel effect (**RSCE**) will be further tuned on length scaling extraction.

The model uses a set of equation for calculating the effective dimensions of the gate, from the drawn dimensions. The offset is rather small, hence it is important only for devices with small length, width or both. The **DL** is the difference between the effective length of the gate (L_{eff}) and the drawn length (L_{drawn}), while the **DLC** allows some further fine tuning for the effective length for the dynamic behaviour ($L_{eef,c}$), see equations (3.7) and (3.8). So the short device is used for the extraction of a first value of Effective Length Parameter-**DL** (later fined tuned).

$$L_{eff} = L_{drawn} + DL \quad (3.7)$$

$$L_{eff.c} = L_{drawn} + DLC \quad (3.8)$$

The parameters for the overlap capacitances, which refer to the overlap areas that exist between the gate and the nodes at the ends of the channel, are also extracted in this step. These parasitic capacitances are proportional to the overlap length (**LOV**). A body effect coefficient (**GAMMAOV**) is defined for the source and drain region which depends on their doping.

A parasitic bias dependent capacitance that appears in the MOSFET is the inner fringing capacitance. The element refers to the capacitance that appears between the gate node on the one side and the source and drain nodes on the other side. The value of this inner fringing capacitance is proportional to the total width of the device (via the **KJF**) and depends on the level of inversion and the amount of the charges that will appear between the end nodes. Its maximum value appears in depletion and weak inversion while it diminishes in strong inversion and accumulation. Moreover there is a fringing Capacitance Bias Factor (**CJF**) which must also be extracted.

Below, the above parameters in CV aspect through C_{GC} vs. V_{GB} plot are extracted. Here CJF is not extracted because CJF modulates the fringing capacitances for VSB and VDB values other than zero. (see: fig.40, fig.41, fig.42, fig.43, fig.44, fig.45).

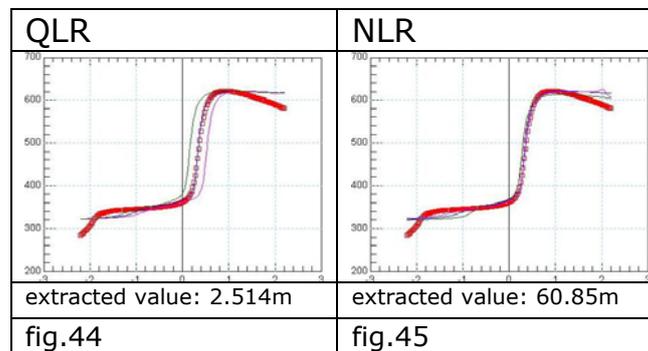


Table 3.8. C_{GC} vs. V_{GB} , referred technology 110nm, $V_G = -2.2 \dots 2.2V$, $V_s = 0.0V$, $V_D = 0.0V$, $V_B = 0.0V$

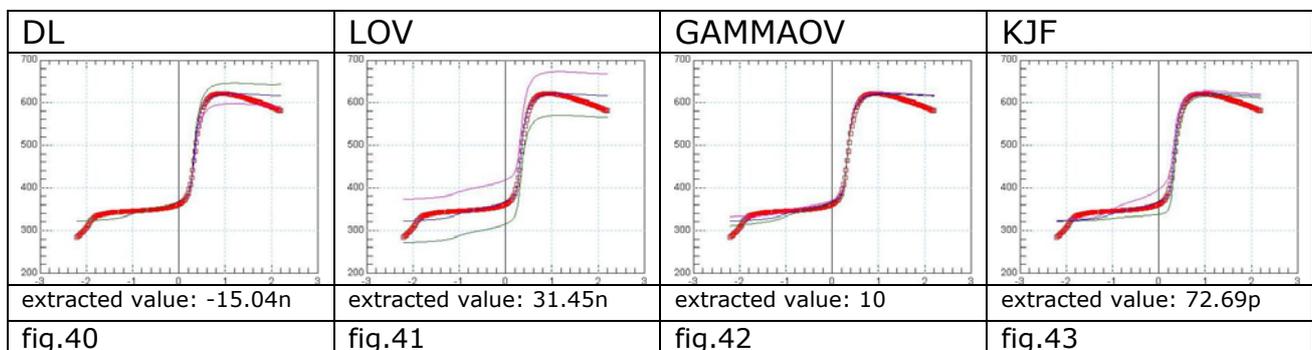


Table 3.9. C_{GC} vs. V_{GB} , referred technology 110nm, $V_G = -2.2 \dots 2.2V$, $V_s = 0.0V$, $V_D = 0.0V$, $V_B = 0.0V$

I_D vs. V_G analysis

Then from I_D vs. V_G and the g_m vs. V_G analysis in linear operation the parameter related to the Series Resistance (**RLX**) RLX, which defines the resistance per width units, can be extracted and the Effective Length Parameter (**DL**) can be verified. Some small differences may be tuned by the use of Effective Length Parameter for Capacitance **DLC** parameter when using C_{GC} vs. V_{GB} analysis. (see: fig.46, fig.47, fig.48, fig.49).

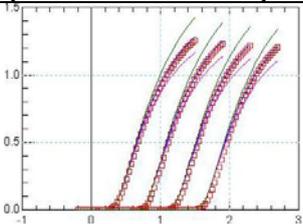
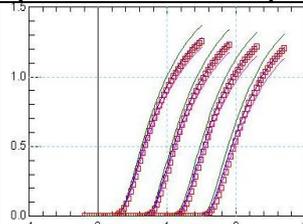
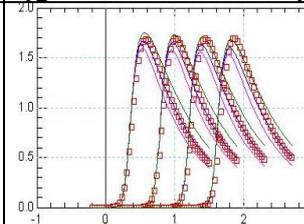
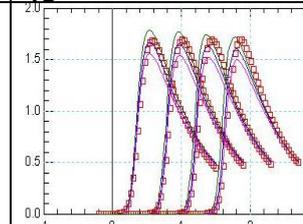
RLX (I_D vs. V_G linear)	DL (I_D vs. V_G linear)	RLX (g_m vs. V_G linear)	DL (g_m vs. V_G linear)
			
extracted value: 48.99u	extracted value: -20.00n	extracted value: 48.99u	extracted value: -20.00n
fig.46	fig.47	fig.48	fig.49

Table 3.10. referred technology 90nm,
 $V_G = -0.3 \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_S = 0.0V$, $V_D = 50mV$

The drain potential, also affects importantly the pinch-off voltage (V_P). This results directly into a dependence of the threshold voltage (V_{TH}) on V_{DS} , and into a lower value of V_{TH} in saturation with respect to linear inversion. This difference is controlled using the **ETAD**, while the **SIGMAD**, allows the modulation of the effect for non-zero V_{SB} values.

It can be noted here that the Drain induced barrier lowering effect (**DIBL**) and the DITS effect are closely connected. Within the formulation of this model, the part of the model that refers to DITS effect deals with the output conductance degradation, mainly in long channel devices, while the DIBL effects covers the threshold voltage dependence on the drain potential for the short channel devices.

So for the DIBL the I_D vs. V_G analysis in saturation is used. Moreover in this analysis appears partially charge sharing effect so we can extract the related parameter Short Channel Charge Sharing Coefficient in short channel devices (**LETA**) (see: fig.50, fig.51, fig.52, fig.53, fig.54, fig.55).

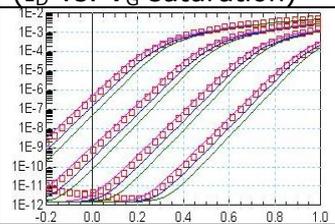
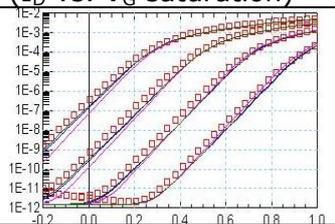
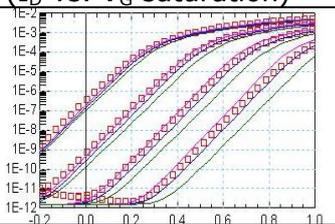
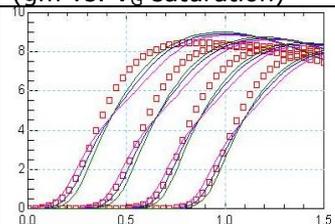
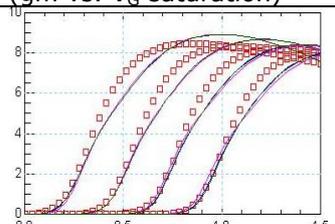
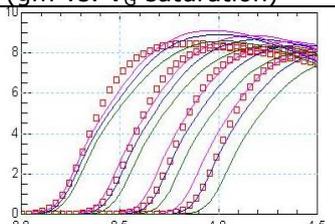
ETAD (I_D vs. V_G saturation) 	SIGMAD (I_D vs. V_G saturation) 	LETA (I_D vs. V_G saturation) 
extracted value: 925.0m	extracted value: 200.0m	extracted value: 903.1m
fig.50	fig.51	fig.52
ETAD (g_m vs. V_G saturation) 	SIGMAD (g_m vs. V_G saturation) 	LETA (g_m vs. V_G saturation) 
extracted value: 925.0m	extracted value: 200.0m	extracted value: 903.1m
fig.53	fig.54	fig.55

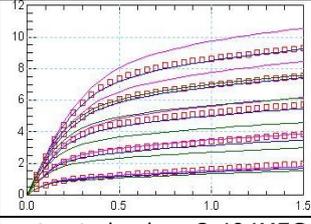
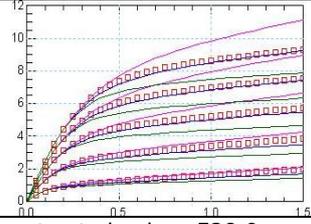
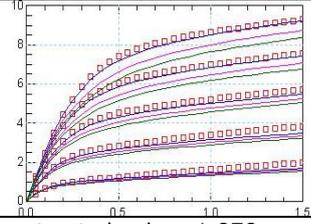
Table 3.11. referred technology 110nm,
 $V_G = -0.2...1.5V$, $V_B = 0V$, $V_S = 0...0.6V$, $V_D = 1.5V$

I_D vs. V_D analysis

The saturated operation of short channel transistors is dominated by the velocity saturation effect. The critical longitudinal electrical field, for which the phenomenon appears, is assigned to the **UCRIT**.

The phenomenon is accompanied by a modulation of the effective length of the channel of the MOSFET, which is controlled by the **LAMBDA**. The model implements a formation of variable order, defined by **DELTA**, which may take values, strictly, between 1 and 2. Ideally the order of the model should be 2 for the NMOS devices, and 1 for the PMOS case, but in any case it is left as a parameter for optimum t. The order of the effect changes the shape of the current curves during the transition from linear operation to saturation, but it also affects the length scaling of the saturation current.

Then the velocity saturation and channel length modulation parameters may be extracted by I_D vs. V_D analysis and specifically in saturation region (see: fig.56, fig.57, fig.58).

UCRIT	LAMBDA	DELTA
		
extracted value: 2.494MEG	extracted value: 528.9m	extracted value: 1.878
fig.56	fig.57	fig.58

note: in DELTA graph both pink and green represent decreased values

Table 3.12. I_D vs. V_D , referred technology 110nm,
 $V_D = 0 \dots 1.5V$, $V_G = 0.6 \dots 1.5V$, $V_B = 0V$, $V_s = 0.0 V$

I_G vs. V_G analysis

For the short channel devices the part of the oxide that overlaps with the source and drain nodes is relatively more important. The length of these overlap regions, fitted for overlap gate current aspects (LOVIG), plays an important role for such geometries. So the Overlap Length for Gate Current (**LOVIG**) may be extracted through the I_G vs. V_G analysis (see: fig.59).

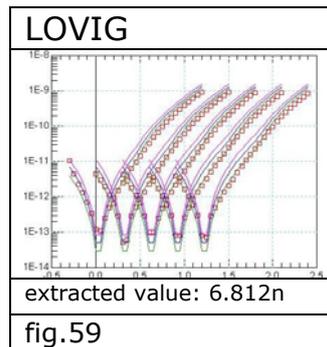


Table 3.13. I_G vs. V_G , referred technology 90nm,
 $V_G = -0.3V \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_s = 0.0 V$, $V_D = 50mV$

I_B vs. V_G analysis

The saturated carriers that reach the drain side of the channel in maximum speed are able to ionize the area. This effect results into a current from the drain node to the bulk (I_B). Note that this current is added to the drain current that flows through the drain node. So, the last step for short and wide devices is to extract the parameters related to impact ionization current (**IBA**, **IBB** and **IBN**), through the I_B vs. V_G analysis (see: fig.60, fig.61, fig.62).

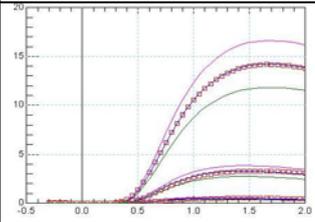
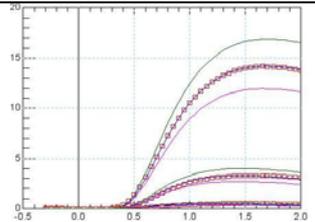
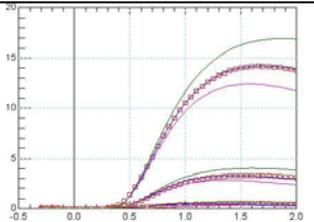
IBA	IBB	IBN
		
extracted value: 5.516G	extracted value: 1.676G	extracted value: 96.72m
fig.60	fig.61	fig.62

Table 3.14. I_B vs. V_G , referred technology 90nm, $V_G = -0.3V \dots 2V$, $V_D = 1.5 \dots 2V$, $V_S = 0.0V$, $V_B = 0.0V$

3.1.3 Length Scaling for Wide Channel Devices

Most of the above parameters are extracted after one specific device with certain geometry. Many of these parameters reflect properties of the whole technology so it is unnecessary to use other geometries for their extraction. Unfortunately, it is observed that even these properties of the technology level show some dependence on the scaling of the transistor.

This formulation follows a scheme inversely proportional to the gate dimensions, leaving this way the behaviour of transistors with large dimensions unchanged, while it allows the fine tuning of the behaviour of short, narrow and combined short and narrow transistors.

So, when all parameters that are related to short-wide and long-wide channel devices are extracted, devices with channel length between these two corner values must be examined. So the next step is to cover length scaling for wide channel devices.

I_D vs. V_G analysis

As already mentioned, the higher doped pocket implants at the ends of the channel affect the length scaling of threshold voltage and body effect coefficient. So, the final values of the parameters related to short channel effect (**QLR**, **NLR**, **LR**) can be extracted through the I_D vs. V_G analysis in linear operation for various values of V_{SB} and channel length devices (see: fig.63, fig.64, fig.65).

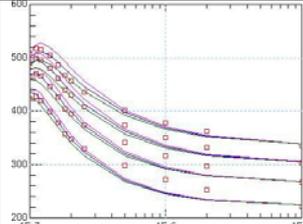
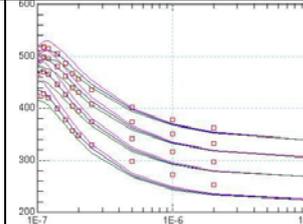
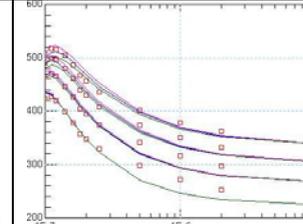
LR	QLR	NLR
		
extracted value: 67.08n	extracted value: 1.500m	extracted value: 33.73m
fig.63	fig.64	fig.65

Table 3.15. V_{th} vs. L , referred technology 110nm, $V_{SB}=0...1,2V$, $V_D=0.5V$, $W=10\mu$, $L=100n,110n,120n,140n,160n,180n,200n,250n,500n,1\mu,2\mu,10\mu$

In order to extract the parameters related to the mobility (First Factor for Mobility Length Scaling-**KA**, First Critical Length for Mobility Length Scaling-**LA**, Second Factor for Mobility Length Scaling-**KB**, Second Critical Length for Mobility Length Scaling-**LB**) the length scaling analysis in strong inversion in both linear operation and saturation is used (see: fig.66, fig.67, fig.68, fig.69, fig.70, fig.71, fig.72, fig.73).

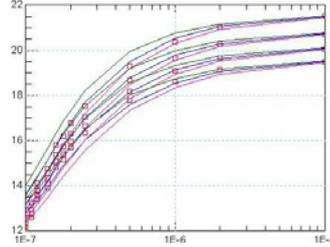
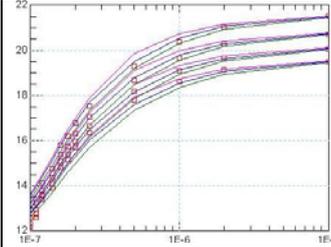
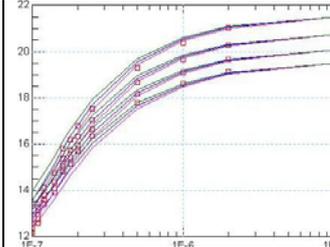
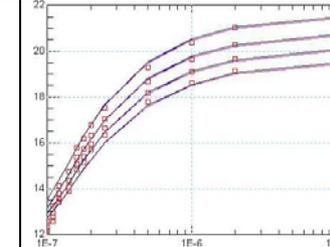
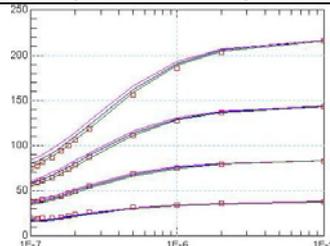
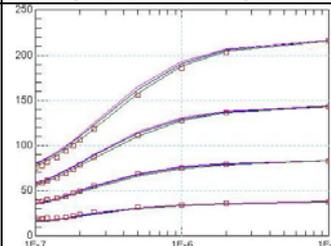
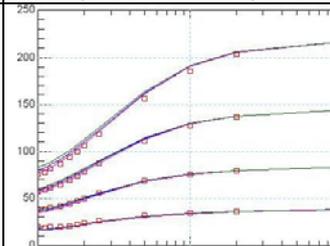
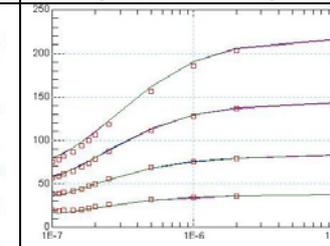
KA (linear operation)	LA (linear operation)	KB (linear operation)	LB (linear operation)
			
extracted value: -141.0m	extracted value: 180.4n	extracted value: 3.121m	extracted value: 444.1p
fig.66	fig.67	fig.68	fig.69
KA (saturation)	LA (saturation)	KB (saturation)	LB (saturation)
			
extracted value: -141.0m	extracted value: 180.4n	extracted value: 3.121m	extracted value: 444.1p
fig.70	fig.71	fig.72	fig.73

Table 3.16. I_D vs. L , referred technology 110nm, $V_{GB}= 1.5, 1.9, 2.4, 2.7V$, $V_s= 0.0 V, V_D=0.5V$ (linear), $V_D=1.5V$ (saturation), $W=10\mu$, $L=100n,110n,120n,140n,160n,180n,200n,250n,500n,1\mu,2\mu,10\mu$

3.1.4 Long and Narrow Channel Device

After the length scaling is concluded, phenomena that appear in narrow and long devices are studied. Here a device with maximum **Length=10 μ** and small **Width=200n** is chosen.

I_D vs. V_G analysis

Similarly to the reverse short channel effect there is the inverse narrow width effect (**INWE**). The phenomenon is connected with the isolation that is used for the MOSFETs, and changes its behaviour according to the type of the isolation, like LOCOS or STI. The formulation used by the model is the similar to the RSCE case. The **WR** defines a critical width for the scaling of the effect across the gate width, the **QWR** is connected with the dependence of the threshold voltage on the effect, while **NWR** covers the dependence of the body effect coefficient. At first, an evaluation of their values is carried out, before a later step where those values will be fine tuned.

From the narrowest channel device the parameter related to the Effective Width (**DW**), which has similar role to DL for the gate width, and scaling parameters (Width Scaling of E0-**WE0**, Width Scaling of E1-**WE1**) which affect the mobility in such devices, are extracted through the linear operation in strong inversion. In the following plots we can observe the relationship which exists between the gate current and the width. Such parameters have not yet been defined but this work will be done in the future (see: fig.74, fig.75, fig.76, fig.77, fig.78).

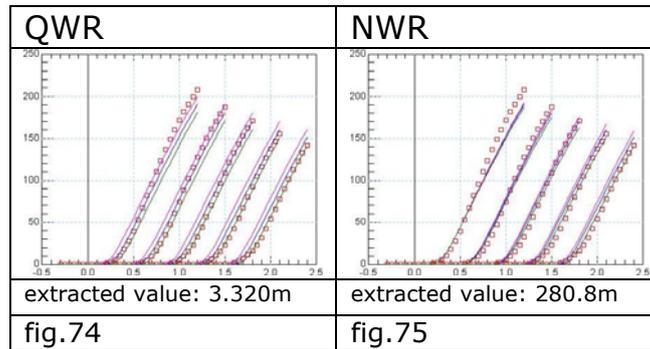


Table 3.17. I_D vs. V_G , referred technology 90nm, $V_G = -0.3 \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_S = 0.0V$, $V_D = 0.5V$

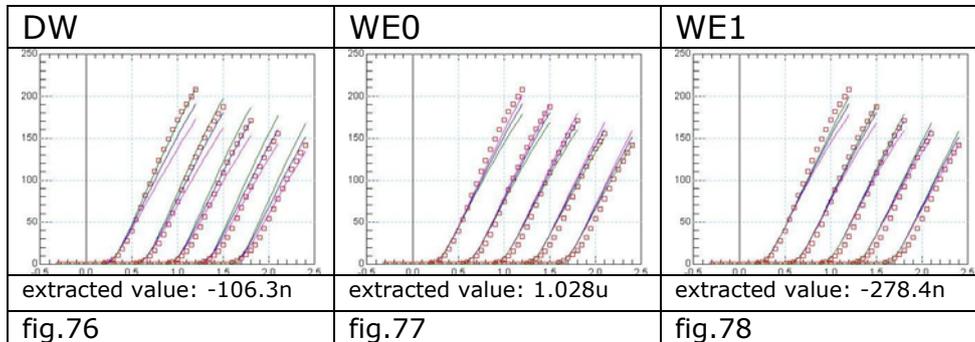


Table 3.18. I_D vs. V_G , referred technology 90nm, $V_G = -0.3 \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_S = 0.0V$, $V_D = 0.5V$

Moreover, in long and narrow devices, charge sharing effect related to the width appears, so through I_D vs. V_G analysis in saturation mode, the Narrow Channel Charge Sharing Coefficient (**WETA**) can be extracted (see: fig.79).

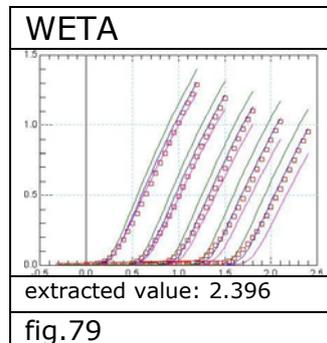


Table 3.19. I_D vs. V_G , referred technology 90nm, $V_G = -0.3 \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_S = 0.0V$, $V_D = 1.5V$

3.1.5 Width Scaling for Long Channel Devices

I_D vs. V_G analysis

The next step is to complete the extraction of the parameters related to the inverse narrow channel effect **QWR**, **NWR**, **WR**. The procedure is equivalent to the one we followed to extract QLR, NLR, WR but this time we use V_{th} for various widths (see: fig.80, fig.81, fig.82).

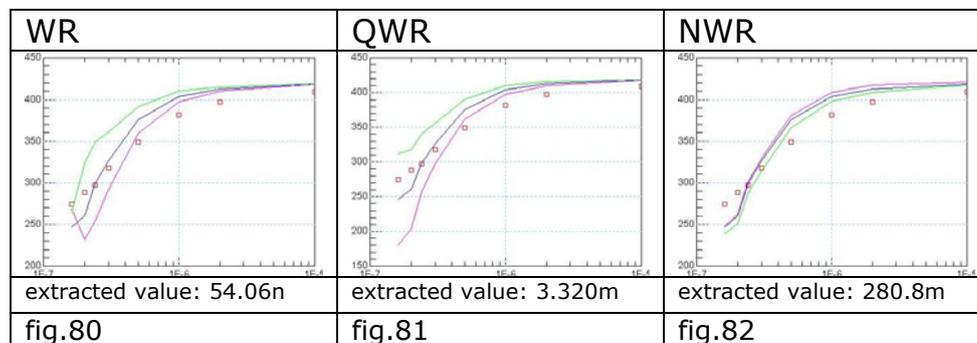


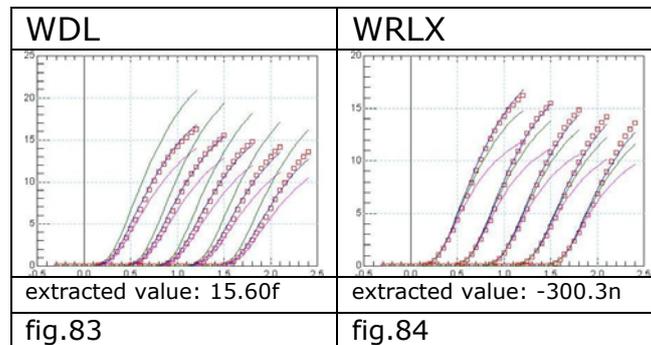
Table 3.20. V_{to} vs. W , referred technology 90nm $V_{SB} = 0V$, $L = 5\mu$, $W = 160n, 200n, 240n, 300n, 500n, 1\mu, 2\mu, 10\mu$

3.1.6 Short and Narrow Channel Devices

Till now, we have studied only devices which were either long or wide. The last step is to check combined narrow and short channel devices. Here a device with minimum **Length=70n** and minimum **Width=160n** is chosen. From such devices fine tuning of some scaling parameters might be required.

I_D vs. V_G analysis

The **WDL** is a scaling parameter of effective length for narrow and short channel MOSFETs and **WRLX** is the Width scaling of RLX. So, the above parameters can be extracted from I_D vs. V_G and g_m vs. V_G analysis in linear operation for devices with the minimum channel size (see: fig.83, fig.84).



note: in WRLX graph both pink and green represent increased values

Table 3.21. I_D vs. V_G , referred technology 90nm
 $V_G = -0.3 \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_S = 0.0V$, $V_D = 0.5V$

I_D vs. V_D analysis

Then, Width scaling of UCRIT (**WUCRIT**) can be defined if there are differences in velocity saturation between the short and narrow device and the long and narrow device, through the I_D vs. V_D analysis (see: fig.85).

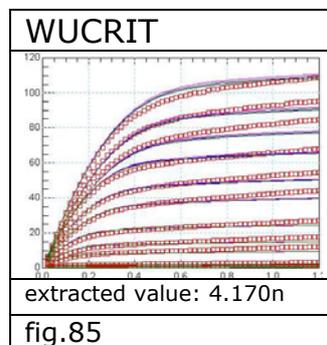


Table 3.22. I_D vs. V_D , referred technology 90nm
 $V_D = 0.02 \dots 1.2V$, $V_G = 0.3 \dots 1.2V$, $V_B = 0 \dots -1.2V$, $V_S = 0.0V$

3.2 Model Parameters Extraction related to Temperature

Below follows a theoretically description of the procedure for extracting parameters which show the dependence of the technology behaviour on the temperature. The same analysis in terms of geometry is kept.

First, a long and wide device is studied where a dependence of the threshold voltage on temperature exists. So through the V_{TH}

vs. T analysis we can extract Temperature Dependence of VTO parameter (threshold voltage) (**TCV**). Next the I_D vs. V_G analysis at strong inversion of linear operation and saturation for various temperatures follows, which allows us to extract mobility and vertical field mobility effect parameters related to temperature, Temperature Dependence of KP (mobility) (**BEX**), Temperature Dependence of E_0 (**TEOEX**) and Temperature Dependence of E_1 (**TE1EX**).

Lastly, a short and wide device should be used in order to extract the Length dependence of TCV (**TCVL**) parameter which can fine-tune the dependence of the threshold voltage on temperature for short devices. Also, First Order Temperature Coefficient of Resistors (**TR**) which shows the dependence of external resistors on temperature can be extracted. Then from the I_D vs. V_D analysis for various temperature we can extract the velocity saturation dependent on temperature parameters, which are Temperature Dependence of LAMBDA (**TLAMBDA**) and Temperature dependence of UCRIT (**UCEX**).

3.3 Y-Parameter Analysis for Model Parameters Extraction at RF

3.3.1 General Equations

Most of the commercially available MOS transistor models have been originally developed for digital and low-frequency analog circuit design, which focus on the dc drain current, conductances, and intrinsic charge/capacitance behaviour in the megahertz range. However, as the operating frequency increases to the gigahertz range, the importance of the extrinsic components rivals that of the intrinsic counter-parts.

For efficient circuit design and simulation, the simplicity of the MOS transistor model should be maintained. Here, the intrinsic quasi-static model, which is suitable for submicrometer CMOS high-frequency applications, is used. The gate series resistance, significantly affects the input admittance at RF and needs to be carefully considered. In addition, substrate coupling effects through the drain and source junctions and the substrate resistances, which play an important role for the output admittance, are also very critical. This effective admittance of the substrate network can reach 50% of the total output admittance. Transcapacitances in the channel of the complete quasi-static model have to be included for accurate prediction.

All the extrinsic elements influence the operation equations of the transistor. Assuming the transistor as a two-port circuit, with common reference terminal for input and output the short circuit

between the substrate and the source, with input the port between the gate and the source and output the port between the drain and the source, we can define Y-parameters. The Y-parameters of a two port are defined as:

$$\begin{aligned} y_{11} &= \left. \frac{I_1}{V_1} \right|_{V_2=0} & y_{12} &= \left. \frac{I_1}{V_2} \right|_{V_1=0} \\ y_{21} &= \left. \frac{I_2}{V_1} \right|_{V_2=0} & y_{22} &= \left. \frac{I_2}{V_2} \right|_{V_1=0} \end{aligned} \quad (3.9)$$

These equations may be resolved in such a way that the values of the extrinsic elements may be extracted. These measurements are very hard and demand careful procedures of deembedding so that only the Y-parameters of the actual device, without the influence of the extrinsic or parasitic elements, are calculated. More generally the Y-parameters of the transistor are:

$$\begin{aligned} Y_{11} &\approx \omega^2 R_G C_{GG}^2 + j\omega C_{GG} \\ Y_{12} &\approx -\omega^2 R_G C_{GG} C_{GD} - j\omega C_{GD} \\ Y_{21} &\approx g_m - \omega^2 R_G C_{GG} (\tau_{qs} g_m + C_{GD}) - j\omega (\tau_{qs} g_m + C_{GD} + g_m R_G C_{GG}) \\ Y_{22} &\approx g_{ds} - \omega^2 (R_G C_{GD} (C_{GD} + g_m R_G C_{GG}) + R_B C_{BD} (C_{BD} + g_{mb} R_B (C_{BD} + C_{BS} + C_{GB}))) \\ &+ j\omega (C_{GD} (1 + g_m R_G) + C_{BD} (1 + g_{mb} R_B)) \\ C_{GG} &= C_{GS} + C_{GD} + C_{GB} \end{aligned} \quad (3.10)$$

We must note here that, C_{GS} , C_{GD} , C_{GB} , C_{BS} , C_{BD} are the equivalent capacitances of the Figure 3.1 (a) with the values of the equivalent overlap and fringing capacitances and parasitic diodes counted in them. And the resistance R_B refers to that of the Figure 3.1 (b), in a simpler look of the circuit where R_{DSB} is zero and R_{SB} and R_{DB} infinite.

Especially the gate resistance can be given by the following equations:

$$R_G = \frac{\text{Re}(Y_{11})}{\text{Im}(Y_{11})^2}, \quad R_G = \frac{\text{Re}(Y_{11})}{\text{Im}(Y_{11}) * \text{Im}(Y_{12})}, \quad R_G = \text{Re}(Y_{11}^{-1}) \quad (3.11)$$

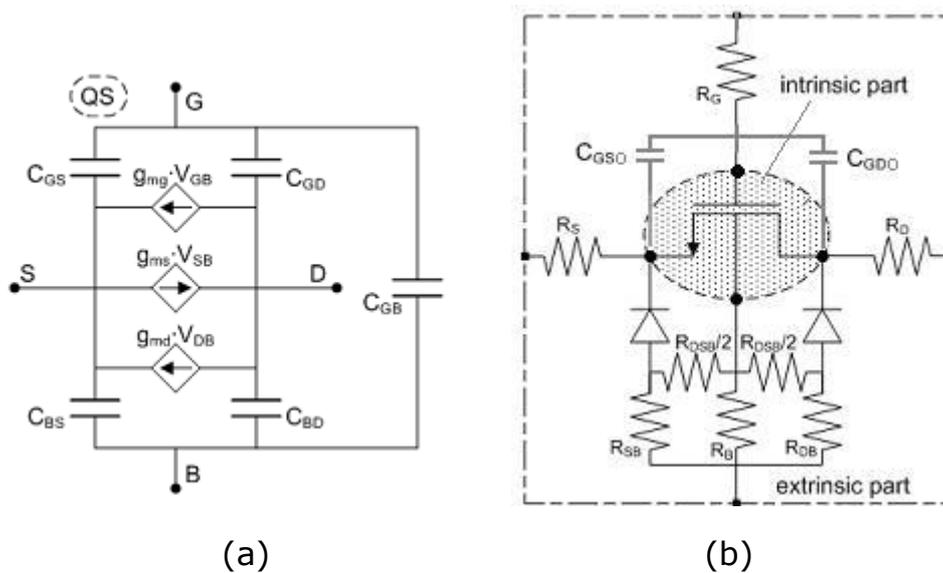


Figure 3.1 (a) quasi-static small signal model of the intrinsic part of a transistor, (b) high frequency macromodel of a transistor

3.3.2 Multifinger and One/Two-Sided Gate Contact Transistor

RF MOS transistors are usually designed as large devices in order to achieve the desired transconductance required to meet the RF requirements. As shown in Figure 3.2 they are usually laid out as multifinger devices, because in deep submicron CMOS processes, the maximum finger length (corresponding to the finger width W_F) is limited. This is due to the so-called narrow-line effect increasing the silicided polysilicon sheet resistance as the finger width (corresponding to the transistor gate length L_F) decreases due to grain boundary problems. Typical devices have up to 10 or more fingers. The total transistor effective width is then simply:

$$W_{eff} = N_F W_F \quad (3.12)$$

where N_F is the number of the fingers.

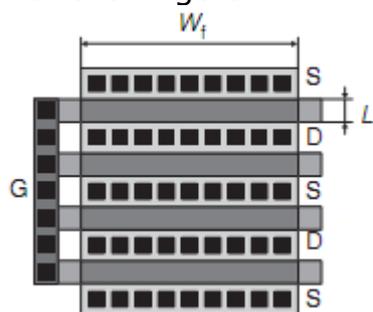


Figure 3.2 Layout of a typical RF MOS transistor

A multifinger transistor allows the minimization of the gate resistance for a specific width. In Figure 3.2, we can see an one-sided gate contact device which means that the gate connects only with the one side of the fingers. It is possible, that the gate connects with both sides of the fingers and thus having a two-sided

gate contact device. In the last case the gate resistance is decreased to the 1/4 of the value of the gate resistance of the one-sided gate contact device. The above can be described by the following equation, which describes the gate resistance depending on the gate contact:

$$R_G = \begin{cases} \frac{1}{3} \frac{W_F}{N_F L_F} R_{G,sq} & \text{one - sided} \\ \frac{1}{4} \frac{1}{3} \frac{W_F}{N_F L_F} R_{G,sq} & \text{two - sided} \end{cases} \quad (3.13)$$

where $R_{G,sq}$ is the sheet resistance of the gate.

3.3.3 Extraction of RF Parameters

The EKV3 model uses the macromodel as shown in Figure 3.1 (b). The gate resistance depends on the parameter **RGSH**, which stands for the resistance of the gate per square, and also on the gate's width, length and number of fingers. The gate resistance is given by the equation below:

$$R_G = k \frac{1}{3} \frac{W_f}{N_f L_g} RGSH \quad (3.14)$$

where k is 1 for one-sided gate contacts and 1/4 for both-sided gate as described before.

It is important to notice the symmetry between drain and source. The values of R_{SB} , R_{DB} and R_B scale with the width of the device and, depending on the shape of the substrate contact around the device and on the number of fingers N_f . R_{DSB} scales proportionally with L_g , $1/W_f$ and $1/N_f$.

The procedure of parameter extraction continues from the point that CV, Static and IV-analysis stopped. At first, we will show how the change in the value of the extrinsic resistances (gate, substrate) and overlap capacitances, influences all the Y-parameters. Next, we will result in those Y-parameters that can be used in order to extract the value of the model parameters (extrinsic resistances and capacitances).

For the analysis we will use a 90nm technology device. The device characteristics are: **type=NMOS**, **Length** (minimum) **=70nm**, **Width=2um**, number of fingers **NF=10**. The measurements and simulations are held in temperature **T=27° C**, for frequencies **Freq=50MHz...20.5GHz** and terminal voltages **V_D=0.2...1.2V**, **V_G=0.8V**, **V_S=0V**.

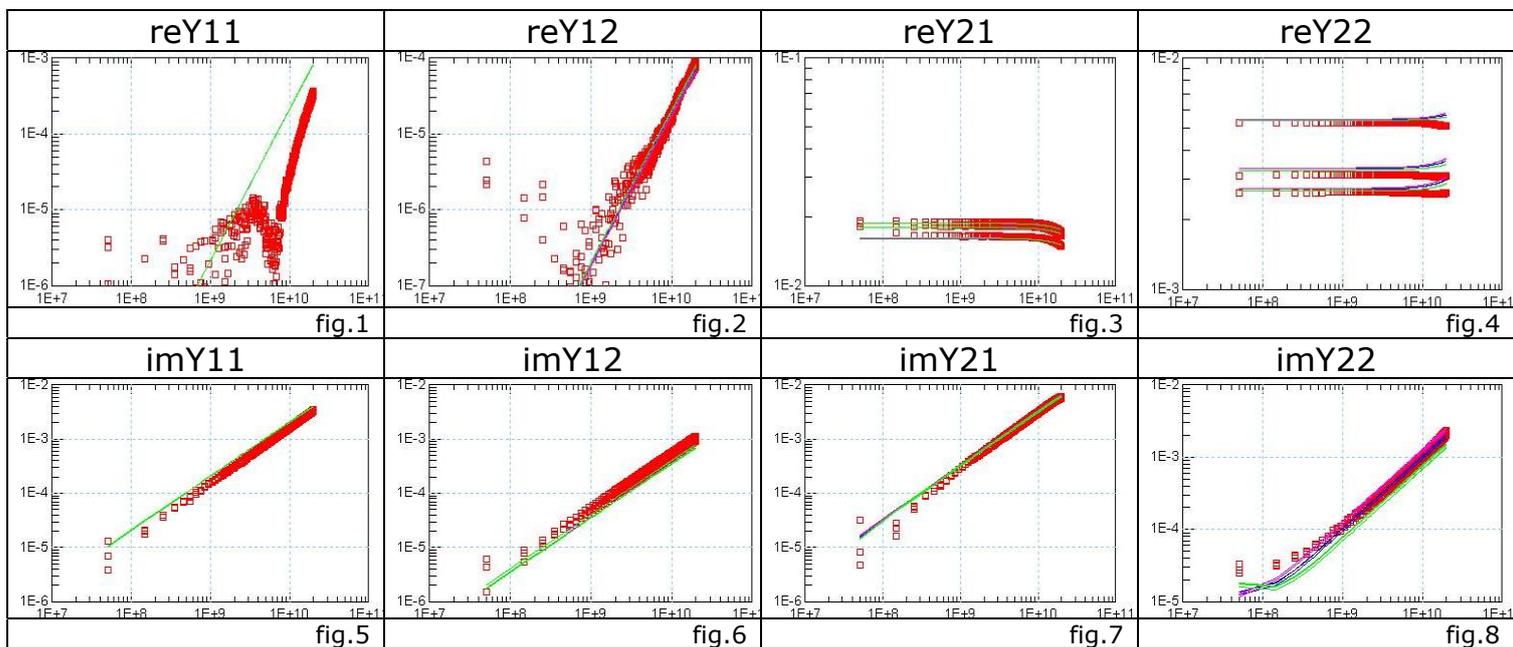
Throughout this analysis the plots represent Y-parameter vs. F and the different colors in the plots describe:

red: measurements
pink: increased value
blue: extracted value
green: decreased value

The Overlap capacitances (LOV, GAMMAOV) and junction capacitances of external diodes (CJ and CJSW), are extracted from CV measurements, but their values may be different in the RF transistor structure. Possibly, they may need to be tuned when big deviation is observed in the Y-parameter at the first place. However, care should be taken to avoid destroying the original capacitance-voltage behaviour.

Firstly we will show the impact of junction capacitances on Y-parameters (fig.1, fig.2, fig.3, fig.4, fig.5, fig.6, fig.7, fig.8).

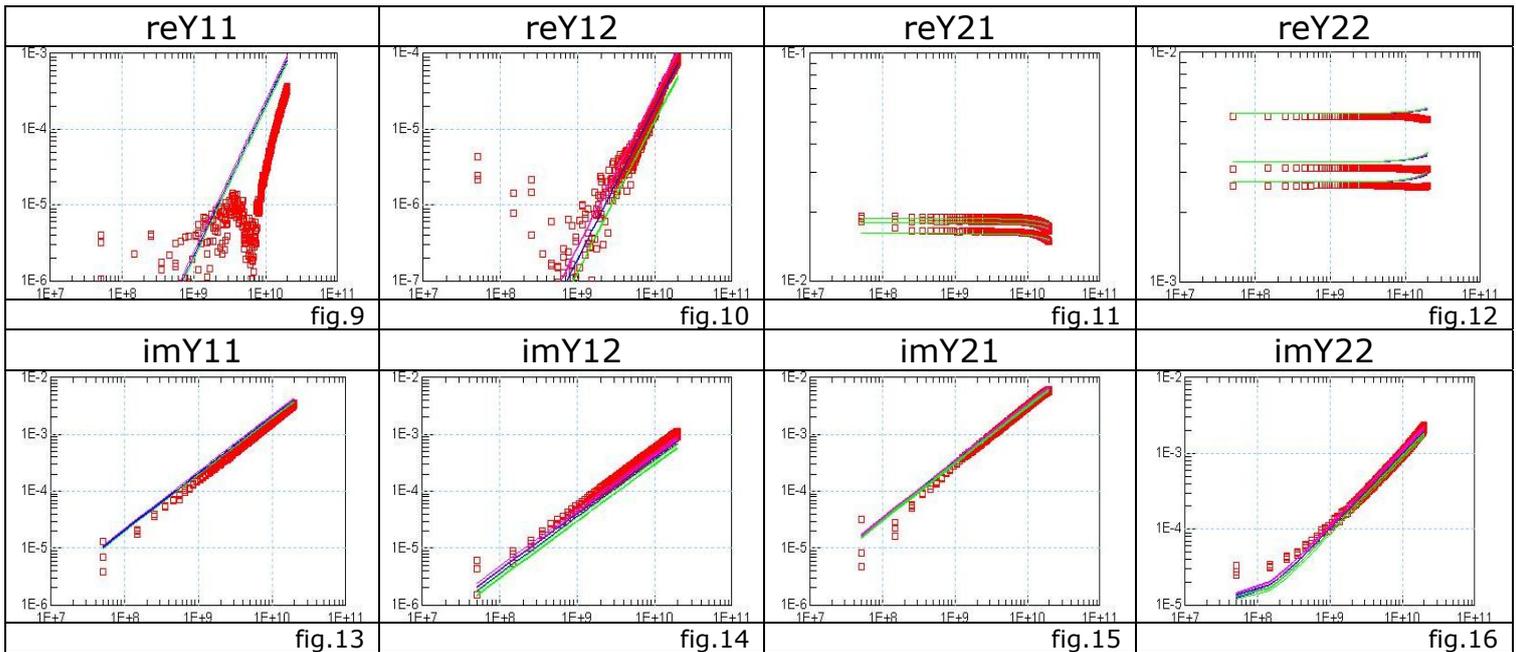
Table 3.23, **CJ(S/D)** (Area Source and drain diode capacitances) – extracted value: **4.000m**



The plots obtained when we changed the value of the source and drain diode capacitances showed that there was an impact mainly imY22 and secondarily on reY22. Thus imY22 can be used for the above model parameters extraction.

Below the influence of the overlap capacitances on the Y-parameters can be seen (fig.9, fig.10, fig.11, fig.12, fig.13, fig.14, fig.15, fig.16, fig.17, fig.18, fig.19, fig.20, fig.21, fig.22, fig.23, fig.24,).

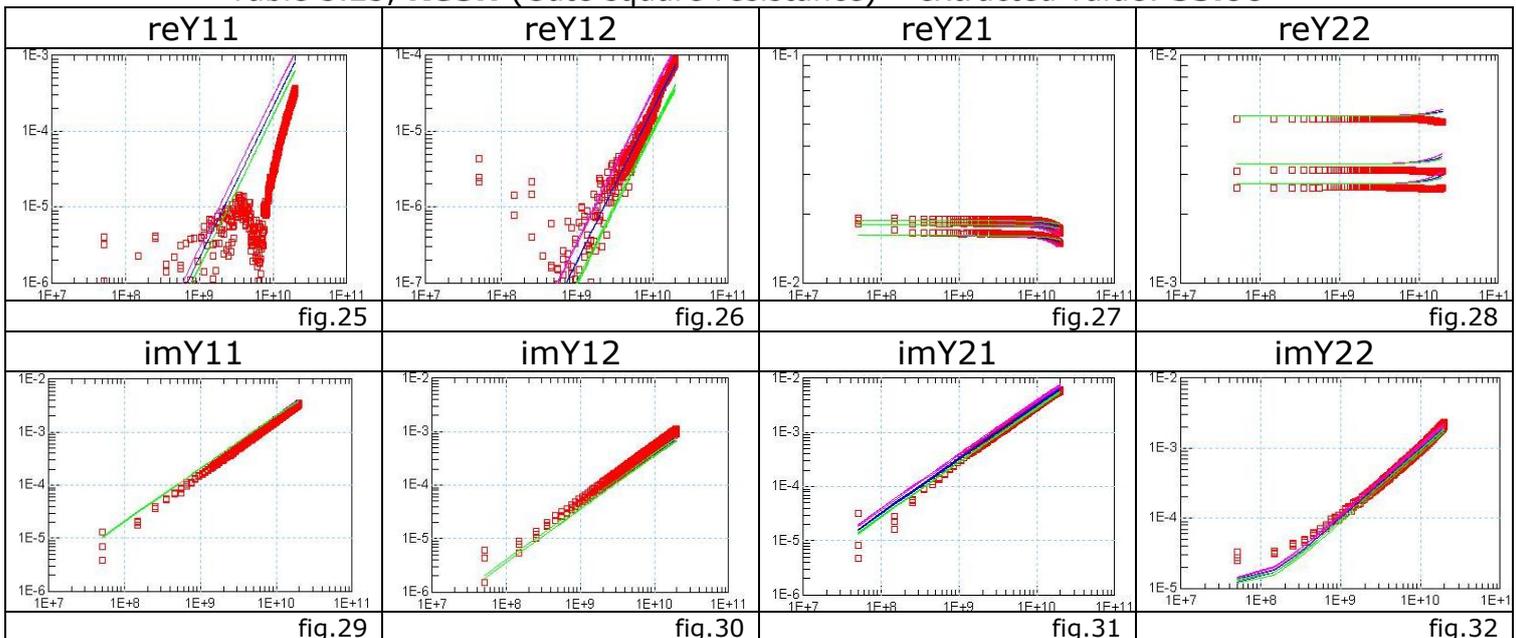
Table 3.24, **LOV** (Length of the overlap area) – extracted value: **21.00n**



From the plots obtained through changes in the values of LOV we assume that LOV needs to be tuned to manipulate Y12 (both real and imaginary part) and imY22. A further fine tuning may be done through the reY11, imY11, imY21 and imY22 as we can see by the plots. We notice that LOV influences the imaginary parts of the Y-parameters for the whole range of frequencies.

The gate sheet resistance parameter (**RGSH**) will follow (fig.25, fig.26, fig.27, fig.28, fig.29, fig.30, fig.31, fig.32).

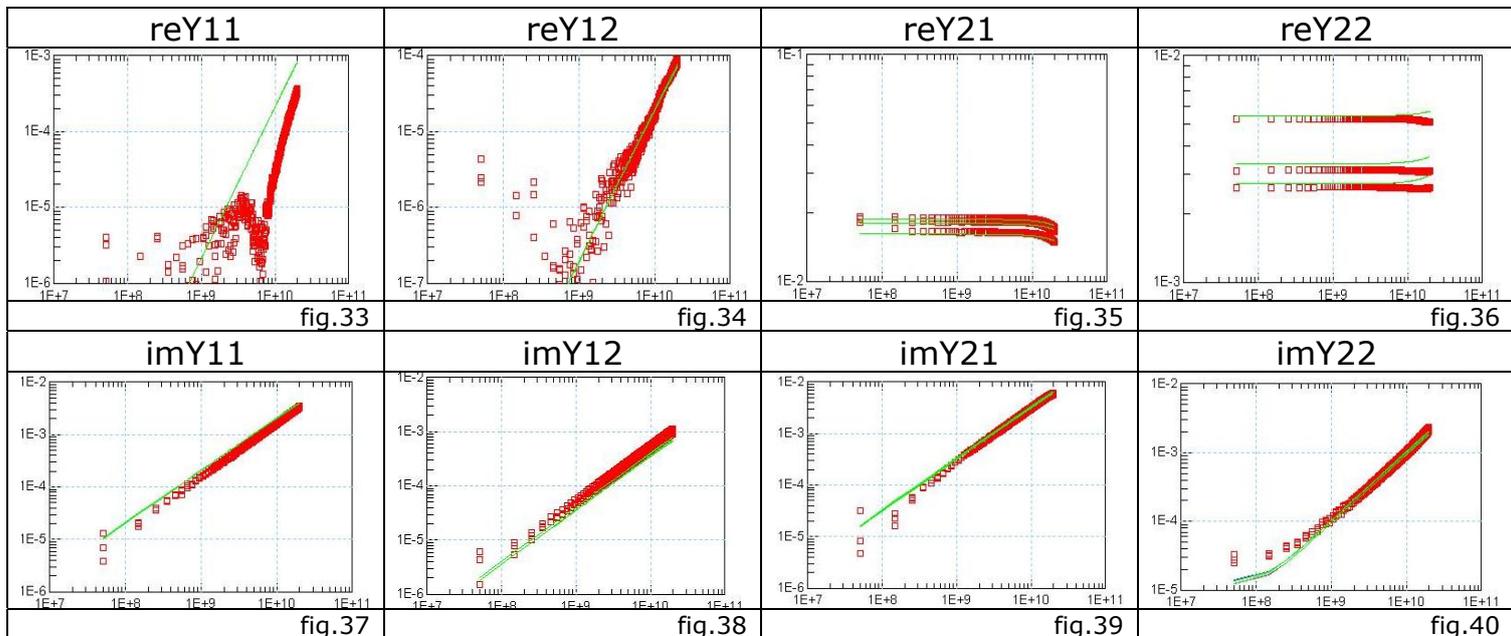
Table 3.25, **RGSH** (Gate square resistance) – extracted value: **53.00**



The plots that we obtained when we changed the **RGSH** show that this parameter has strong influence on reY_{11} , reY_{12} , imY_{21} and imY_{22} so it can mainly be extracted from those Y-parameters. A further tuning could be done through the reY_{21} , reY_{22} plots. Iteration between this step and the previous one may be required to find best value.

As far as substrate resistances are concerned, it would be wise, as a first step, to study the behaviour of the transistor when the model has only one substrate resistance R_B . For this reason, bellow we show how **RBWSH** parameter influences Y-parameters. When the **RBWSH** is extracted, an optimization for finding explicit optimal values for all the elements of the substrate resistance network can be used (fig.33, fig.34, fig.35, fig.36, fig.36, fig.38, fig.39, fig.40).

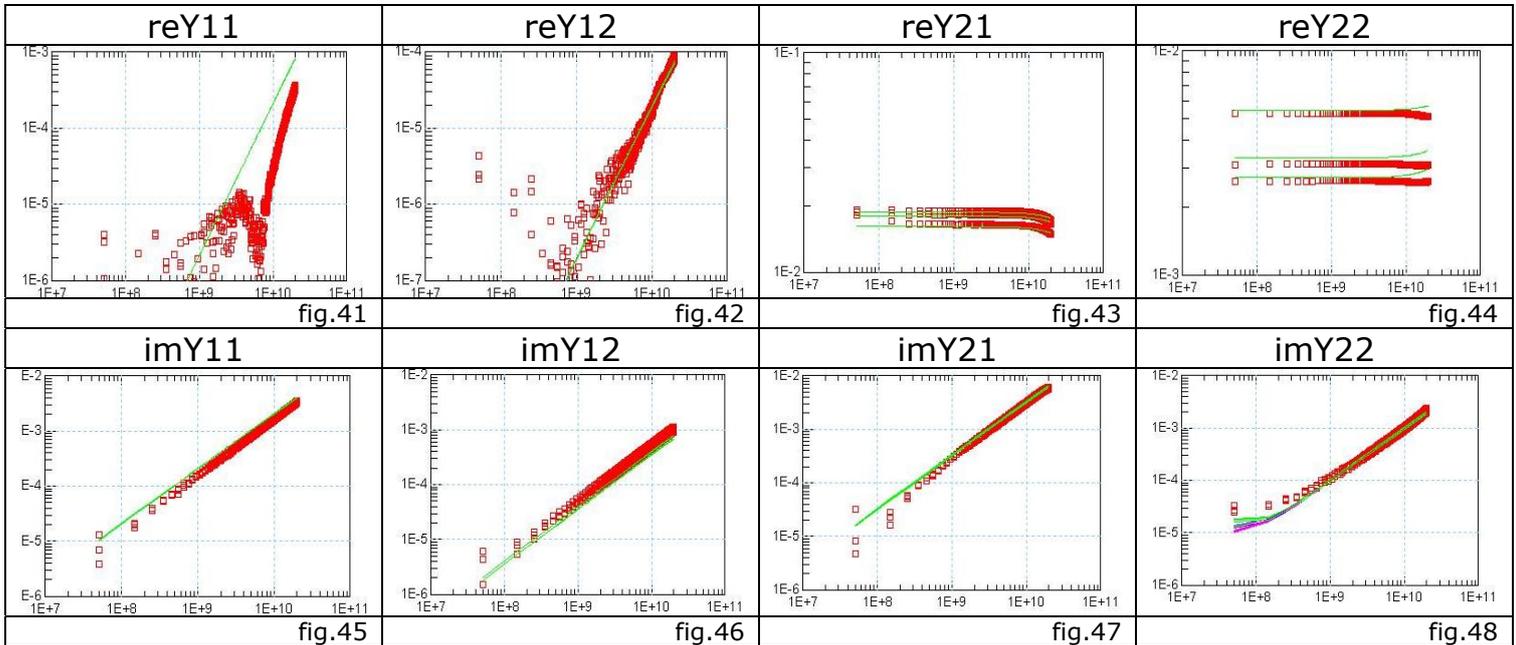
Table 3.26, **RBWSH** (Inner-bulk to external-bulk resistance) – extracted value: **1.00K**



The above plots show that a change in the **RBWSH** has almost no impact in the RF behaviour of the device, since it hardly influences Y-parameters. A fine tuning could be hardly done through reY_{12} and imY_{22} but this is not of special interest.

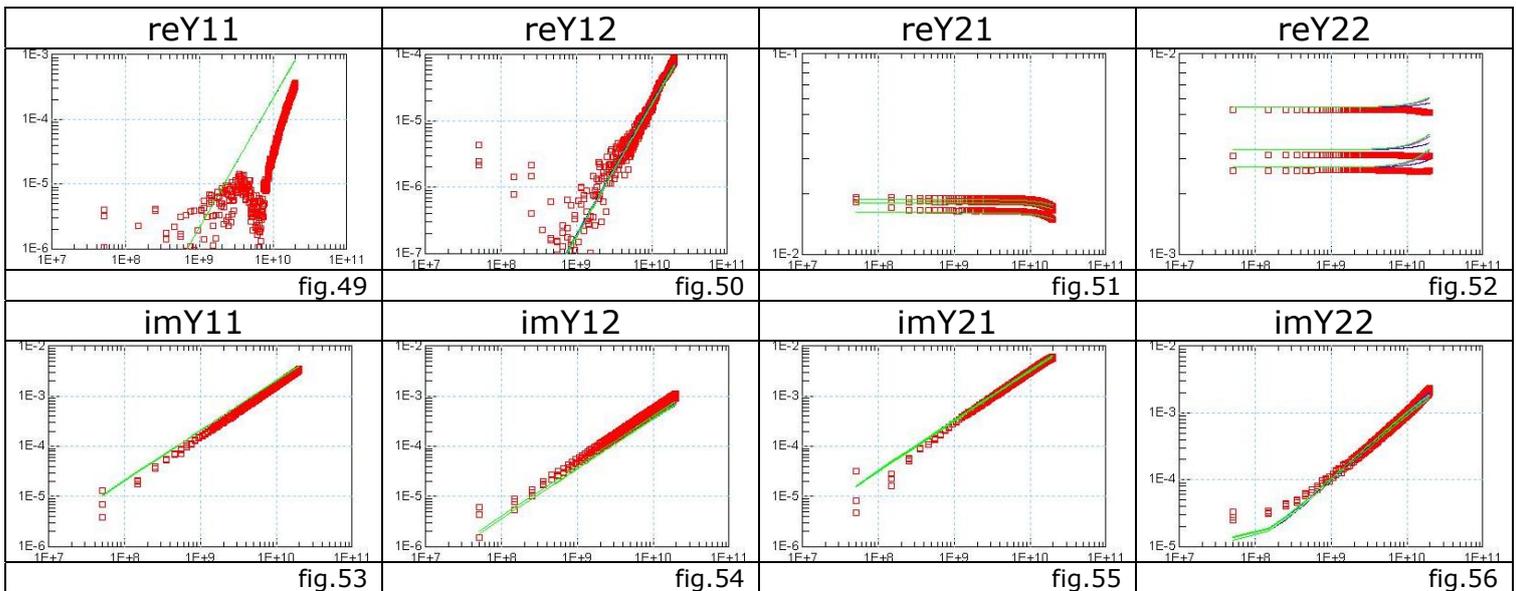
Then we will show the rest three parameters for the substrate resistance network (**RDSBSH**, **RSBWSH**, **RDBWSH**, and **RBWSH**) (fig.41, fig.42, fig.43, fig.44, fig.45, fig.46, fig.47, fig.48, fig.49, fig.50, fig.51, fig.52, fig.53, fig.54, fig.55, fig.56).

Table 3.27,
RSBWSH (Inner-bulk source side to external bulk resistance),
RDBWSH (Inner-bulk drain side to external bulk resistance)-
 extracted value: **50.00**



From the plots above we can see that a change in the **RSBWSH** and **RDBWSH**, influences **imY22** and as a result can be extracted through this Y-parameter's plots

Table 3.28, **RDSBSH** (RDSBSH) – extracted value: **1.500K**

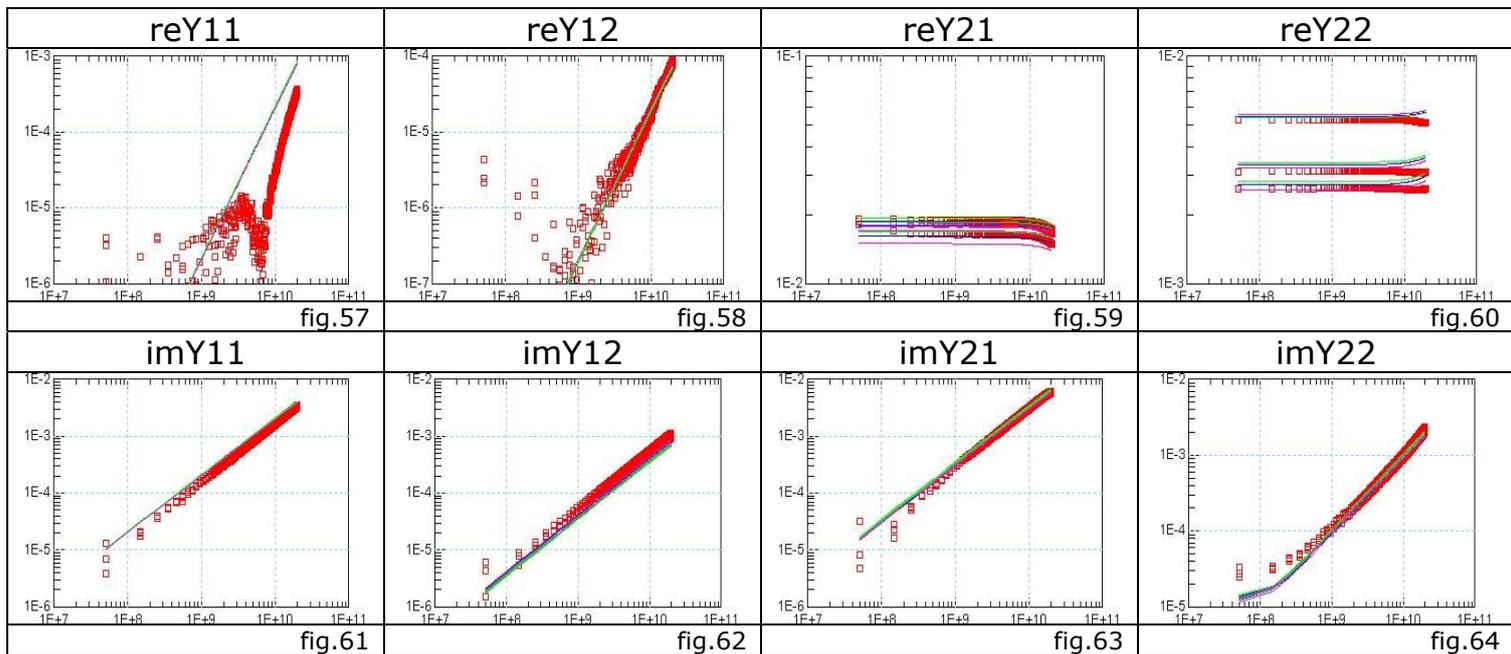


note: in RDSBSH graph both pink and green represent increased values

RDSBSH has strong effect on reY22 at higher frequencies, where we notice flexibility, so its value can be extracted through the real part of Y22.

The series resistances refer to the resistance of the active areas of the source and drain nodes that appear between the external metallic contact and the end of the actual channel. So, the series resistance **RLX**, which defines the resistance per width units, will be studied (fig.57, fig.58, fig.59, fig.60, fig.61, fig.62, fig.63, fig.64).

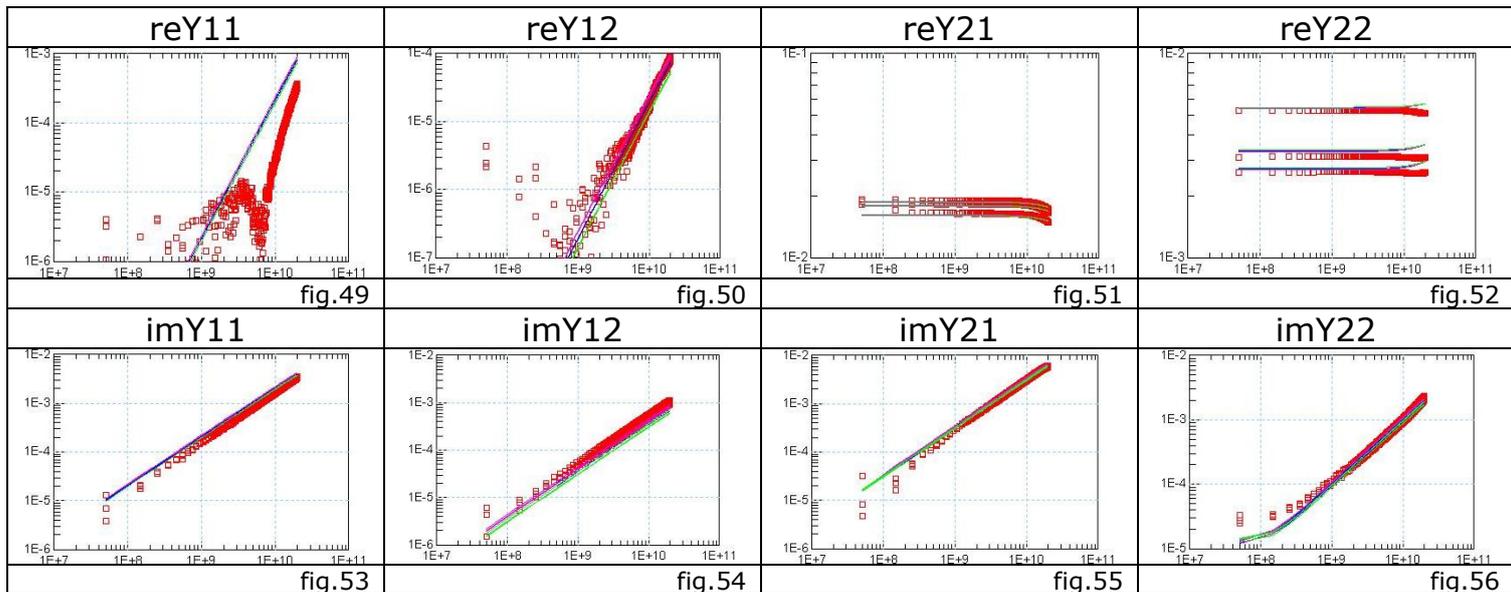
Table 3.29, **RLX** (Series resistance (symmetric model)) - extracted value: **75.00u**



From the above plots, we can see that **RLX** influences mostly reY21, reY22, imY22 so it can mainly be extracted from those Y-parameters. Of course since it has a small impact on imY12, imY21 the extracted value should be further fine tuned so as to have a good fit in those plots too.

CGBO (Gate to bulk outer fringing capacitance) is a parameter that may be used for devices with a large number of gate fingers. Introduction of CGBO enables us to model the increase of input capacitance and its effect on the reduction of cut-off frequency (f_T).

Table 3.30, **CGBO** (Gate to bulk outer fringing capacitance) –
 extracted value: **1.00f**



From the plots above we see that CGBO affects mostly Y12 in both real and imaginary parts and less the real and imaginary parts of the rest Y-parameters.

3.4 Large Signal Modelling

RF applications and wireless systems continue to gain in popularity and CMOS technology is becoming increasingly popular for the realization of the radio frequency components. As well as the front-end, applicability of the low-cost, high-density digital CMOS devices for Power Amplifier may be expected in the near future. From this point of view, a compact model should predict the distortion performance with acceptable accuracy. One way to check the models capability in large signal conditions is to measure of Load-pull characteristics. Then we can tune model parameters to fit load pull behaviour.

3.4.1 Theory of Load-pull Measurement

Load pull consists of varying or “pulling” the load impedance seen by a device-under-test (DUT) while measuring the performance of the DUT (source pull is the same as load pull except that the source impedance is changed instead of the load impedance).

Load (and source) pull is used to measure a DUT in actual operating conditions. This method is important for large-signal, nonlinear devices where the operating point may change with power

level or tuning. Load (or source) pull is not usually needed for linear devices, where performance with any load can be predicted from small signal S-parameters.

The performance of an active device is a function of many things: Frequency, Bias point, Temperature, Source/load impedance at fundamental frequency, Load impedance at harmonic frequencies, Power level. At small-signal analysis, when we measure active devices we are looking at response in a fifty-ohm system, as a function of frequency and bias point, and maybe temperature. In such case we accurately predict the small-signal response if the device sees impedances other than fifty ohms.

However it's more difficult to predict performance under large-signal conditions. So we use load pull so to gather all of data we need to model large-signal behaviour. In load pull analysis we can examine data at one frequency at a time.

3.4.2 Load-pull Measurement Setups

In order to test the behaviour of an NMOS device a load pull simulation setup was used for two different measurements. Firstly, **Pout** at operating frequency **F=5.8GHz** when the load was varying in a specific range shown in Figure 3.3 was studied and then the **gain vs. Pin** when load (Z_L) was about **50Ω**. The difference of the two cases is that at the first one Pin and terminal voltages are set but the load Z_L is varying and at the second, Z_L is fixed at about 50Ω and terminal voltages are set and Pin is varying from -20dBm to +5dBm. The two circuits used in the two cases are shown in Figures 3.4 and 3.5

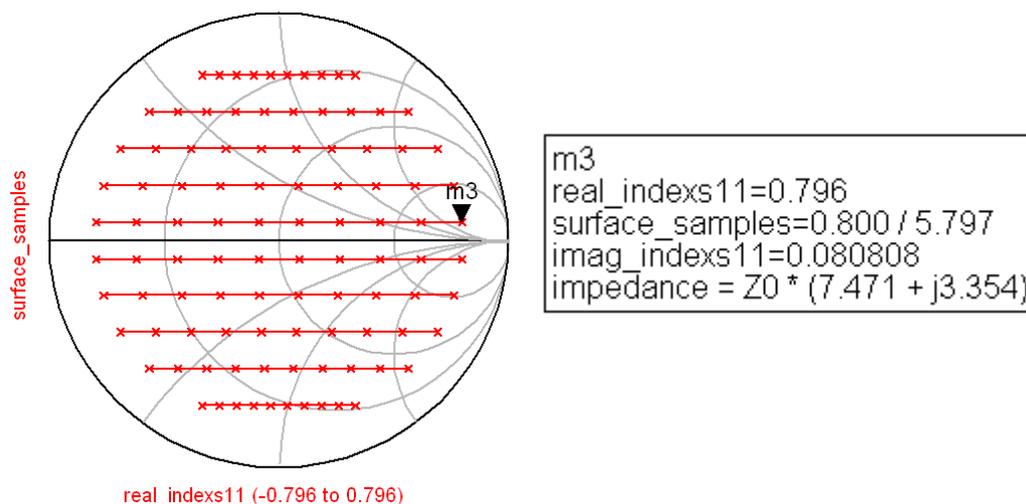


Figure 3.3 Smith plot of the simulated load coefficients, m3 is a complex sample of the circle's surface and is multiplied with Z_0 and in such way we get the values for the simulated loads used in the load pull analysis. Z_0 is about 50Ω.

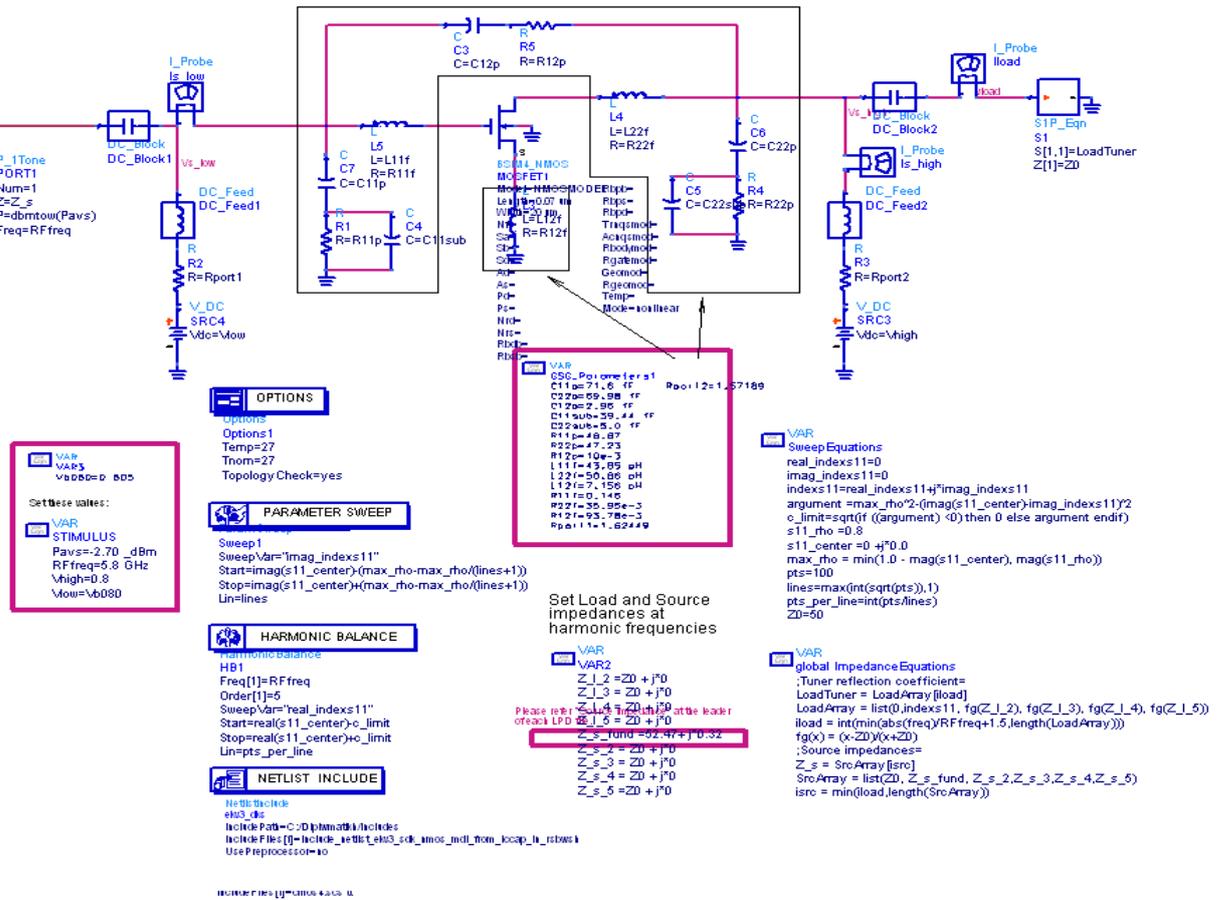


Figure 3.4 Load-pull simulation setup, pulling Z_L

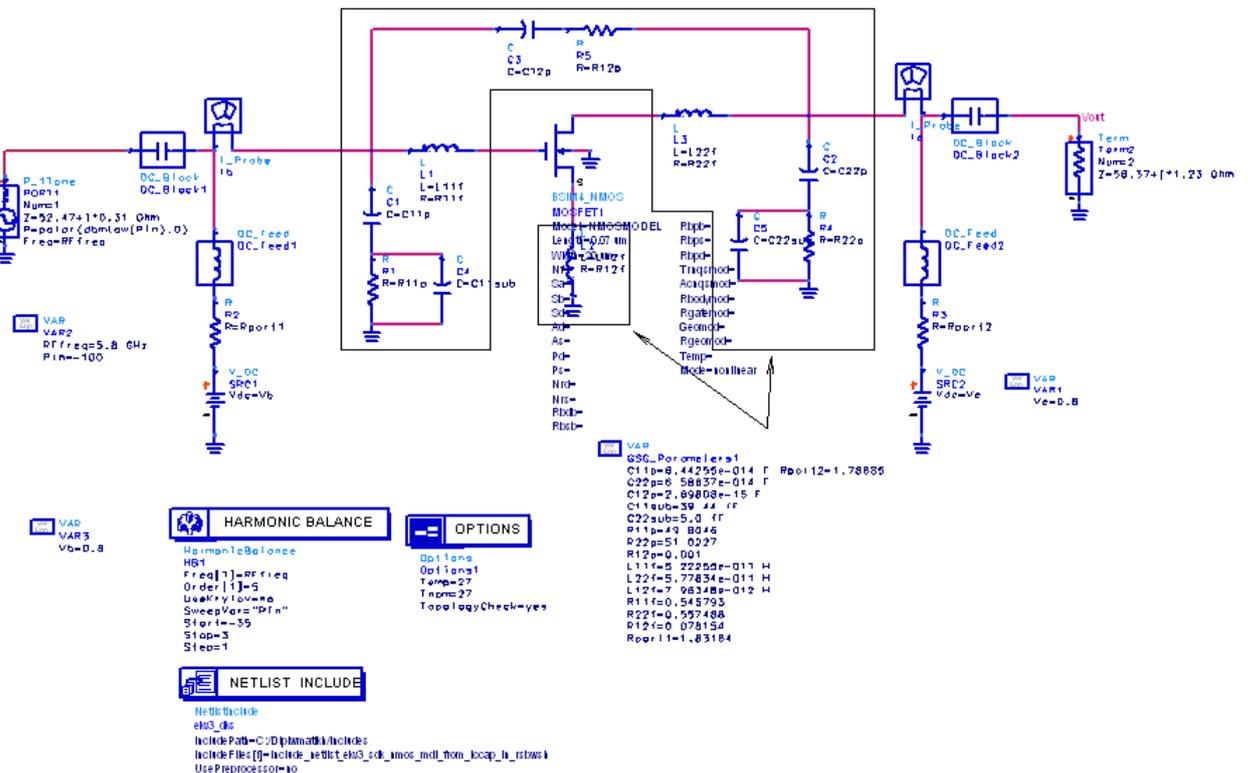


Figure 3.5 Load-pull simulation setup, $Z_L \approx 50 \Omega$, varying P_{in}

3.4.3 Load-pull Results

Below, the influence of changes in some model parameters on the RF large-signal behaviour of a specific device under specific conditions will be shown. Of course, those model parameters have already been extracted by CV, IV, Y-parameters analysis but a first step for understanding large-signal behaviour is to show the impact of some of the model parameters.

Here, a transistor of a technology with nominal length, 90nm is chosen. The transistor has gate Length $L_g=0.07\mu$, width of a finger $W_F=2\mu$ and number of fingers $N_F=10$. The voltages at the terminals are $V_{GS} = 0.8V$, $V_{DS} = 0.8V$ and $V_{SB} = 0V$ and frequency is $f=5.8GHz$, as already mentioned. For the load pull analysis $P_{in}=-2.73dBm$.

The model parameters that have been chosen are some basic model parameters such as COX, VTO, PHIF, GAMMA, KP, LAMBDA etc and RF parameters (since $f=5.8GHz$) that model substrate resistances and overlap and fringing capacitance. The changes on their values have been chosen, so as their values remain between physical boundaries.

Throughout this analysis the different colors in the contours describe:

red: measurements
green: increased value
blue: extracted value
yellow: decreased value

Firstly, P_{out} contours, which represent the P_{out} data divided into contours levels, will be shown on a Smith Chart at $F=5.8GHz$ and secondly gain vs. P_{in} at $Z_L=58.37+j1.23\Omega$ will be displayed.

The Smith Chart is a graphical aid designed for electrical and electronics engineers specializing in radio frequency (RF) engineering to assist in solving problems with transmission lines and matching circuits. Use of the Smith Chart utility has grown steadily over the years and it is still widely used today, not only as a problem solving aid, but as a graphical demonstrator of how many RF parameters behave at one or more frequencies. It allows the relatively complicated mathematical calculations, which use complex algebra and numbers, to be replaced with geometrical constructs, and it allows us to see at a glance what the effects of altering the transmission line (feed) geometry will be. It gives the practitioner a really good feel for the behaviour of transmission lines and the wide range of impedance that a transmitter may see.

	load pull analysis – Pout contours	gain vs. Pin
COX extracted value: 14.50m		
KP extracted value: 1.100m		
LOV extracted value: 21.00n		
RLX extracted value: 75.00u		

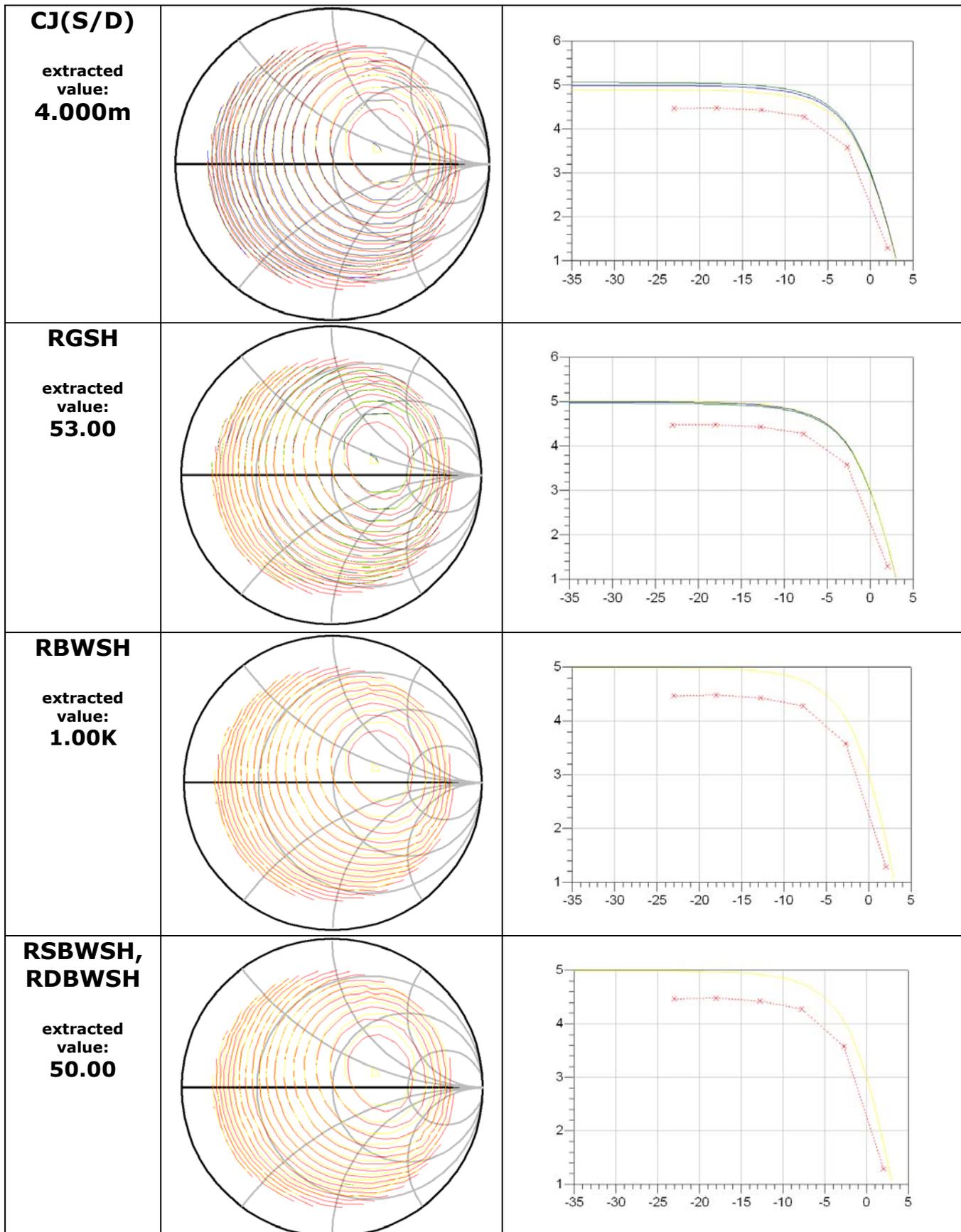


Table 3.31, Load-pull analysis for various model parameters

As was expected the oxide, substrate and gate doping related parameters that depend on the core's physical properties of the technology such as COX (VTO, PHIF, GAMMA) greatly affect large signal analysis. Mobility related parameter KP also affects the model's capability of prediction in the case of large signal.

Lastly, from the above plots we can see that the model parameters for the extrinsic elements such as overlap capacitances and substrate and gate resistance have almost no impact on the large signal behaviour. We can notice a small impact by the source and drain fringing capacitances, however, a tuning could be done through the series resistance RLX, maintaining though the accuracy of the model in small signal analysis.

Chapter 4:

Evaluation of the EKV3 MOSFET model

This chapter is dedicated to the evaluation of the EKV3 MOSFET model. The purpose of a good model is to be able to predict accurately a transistor's behaviour of a specific technology under all conditions and independently of the structural characteristics of the transistor such as length, width and number of fingers.

For EKV3's evaluation, analyses such as static through IV characteristics, small signal through Y-parameters and large signal through load pull measurements, have been used. Also four different devices of a CMOS technology with nominal Length of 90nm have been selected and shown in the Table 4.1.

DUT	Length	Width of Finger	Number of Finger	Total Width
NMOS	70nm	2um	10	20um
NMOS	70nm	2um	40	80um
PMOS	70nm	2um	10	20um
PMOS	70nm	2um	40	80um

Table 4.1. Devices under test for the evaluation of EKV3 MOSFET model

The same evaluation procedure will be followed for all the different devices and is divided into three parts: i) Static Analysis, ii) Y-Parameter Analysis and iii) Load-pull Analysis

4.1 Evaluation for NMOS devices

Firstly the accuracy of the EKV3 MOSFET model when having the NMOS devices shown in Table 4.1. will be evaluated. Throughout the evaluation the different colors in the plots describe:

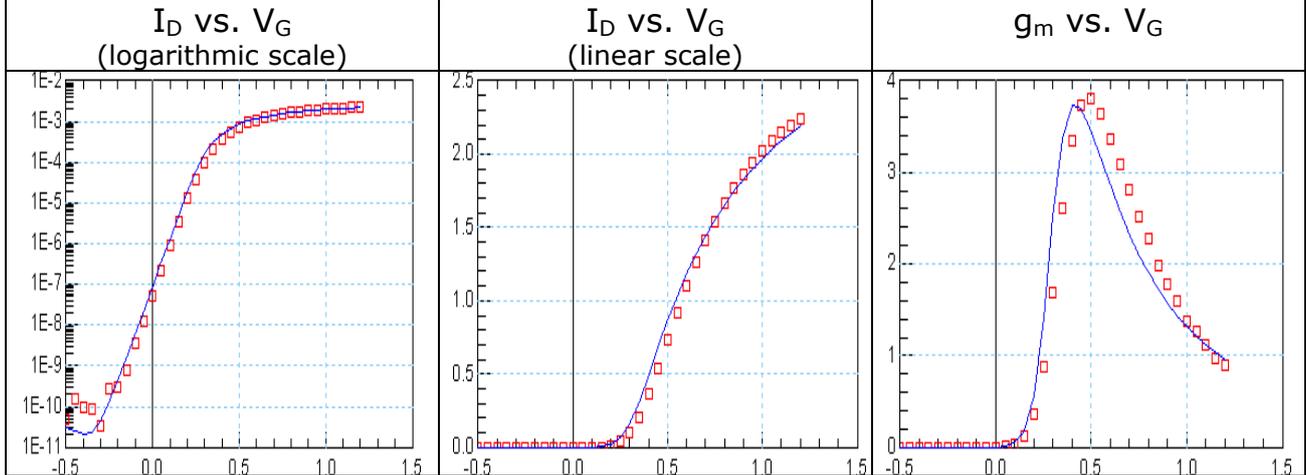
red: measurements blue: simulations
--

Then the same procedure for the evaluation of the EKV3 MOSFET model for the two PMOS devices shown in Table 4.1. will be followed in Chapter 4.2.

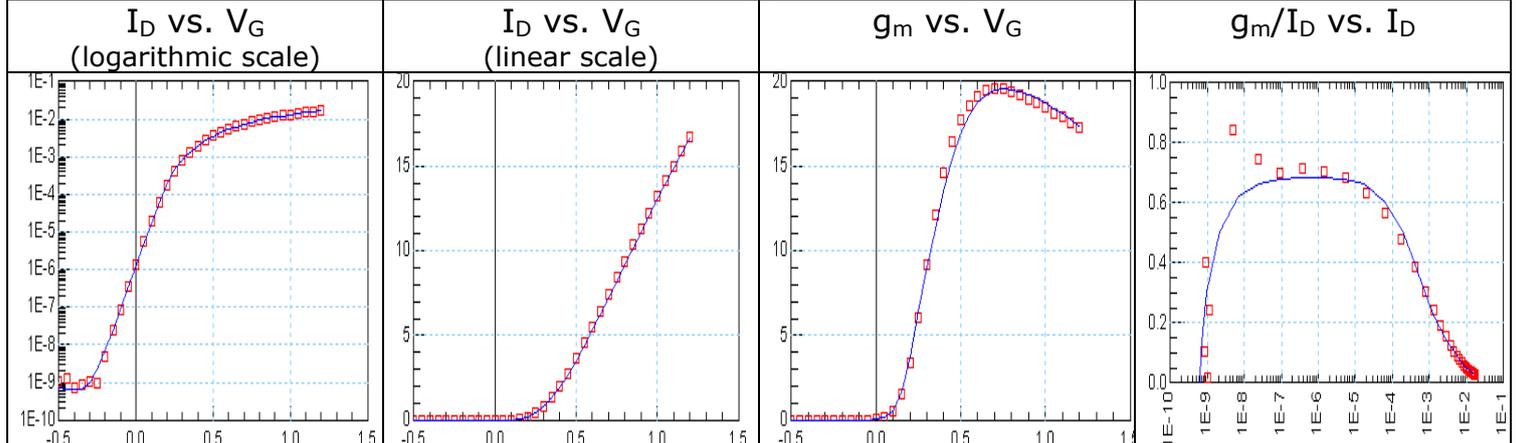
4.1.1 Static Analysis

DUT: NMOS, WF = 2um, LF = 70nm, NF = 10

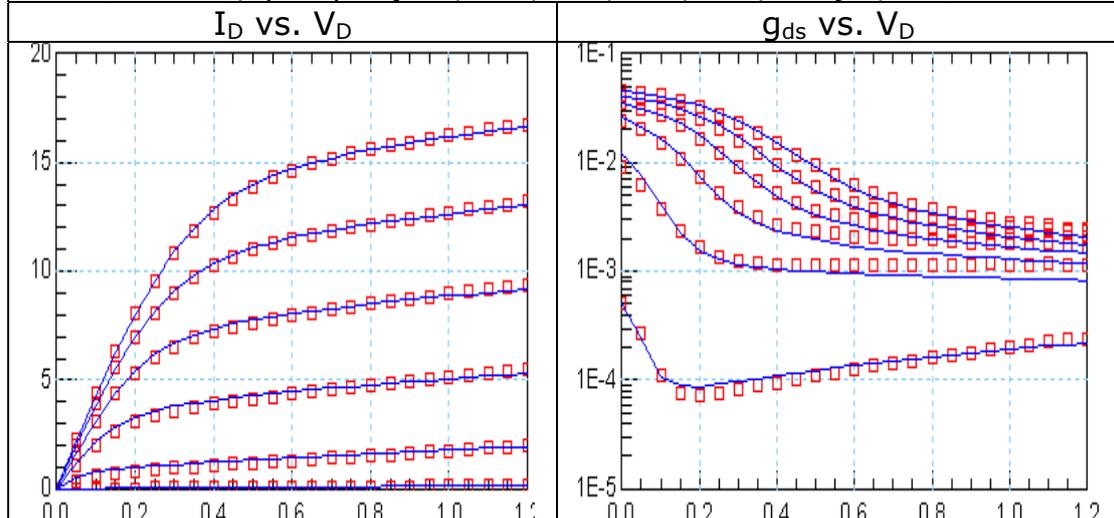
I_D vs. V_G (linear); $|V_{DS}| = 50mV$; $V_{SB} = 0V$



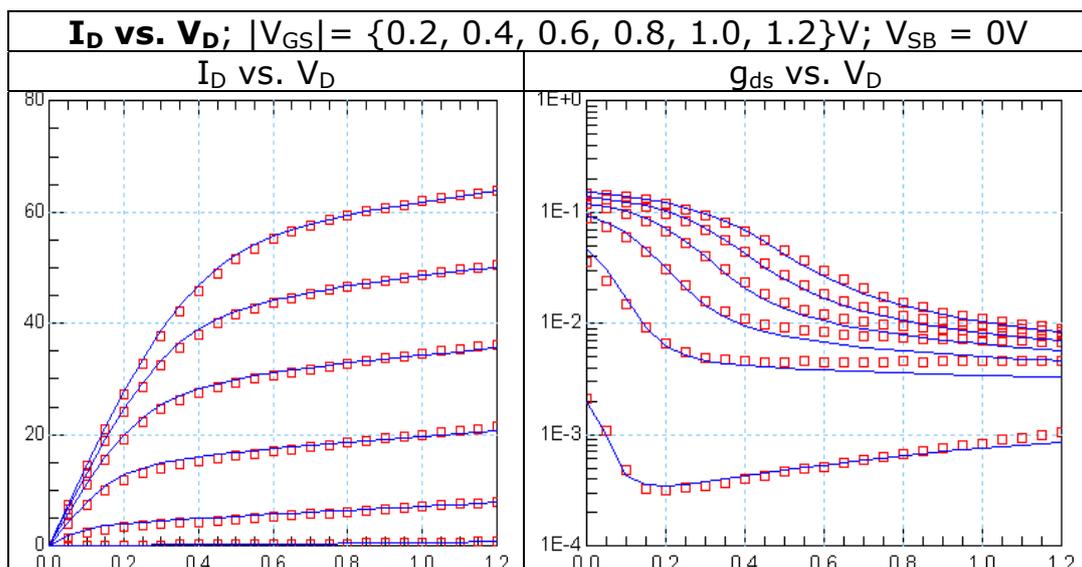
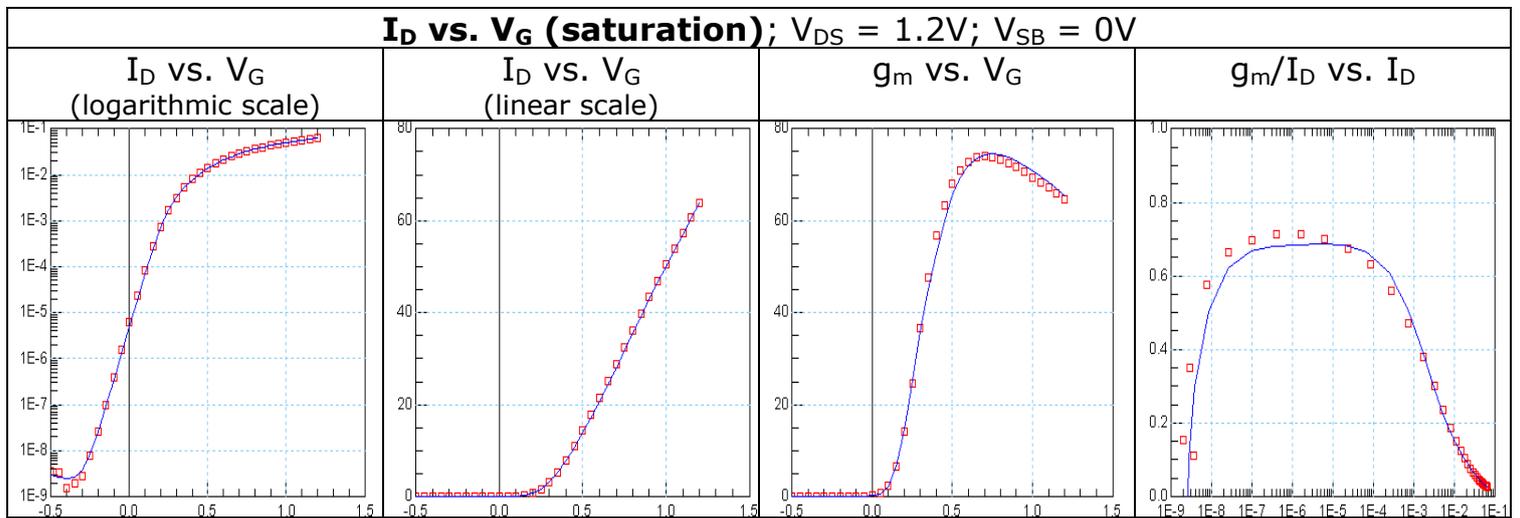
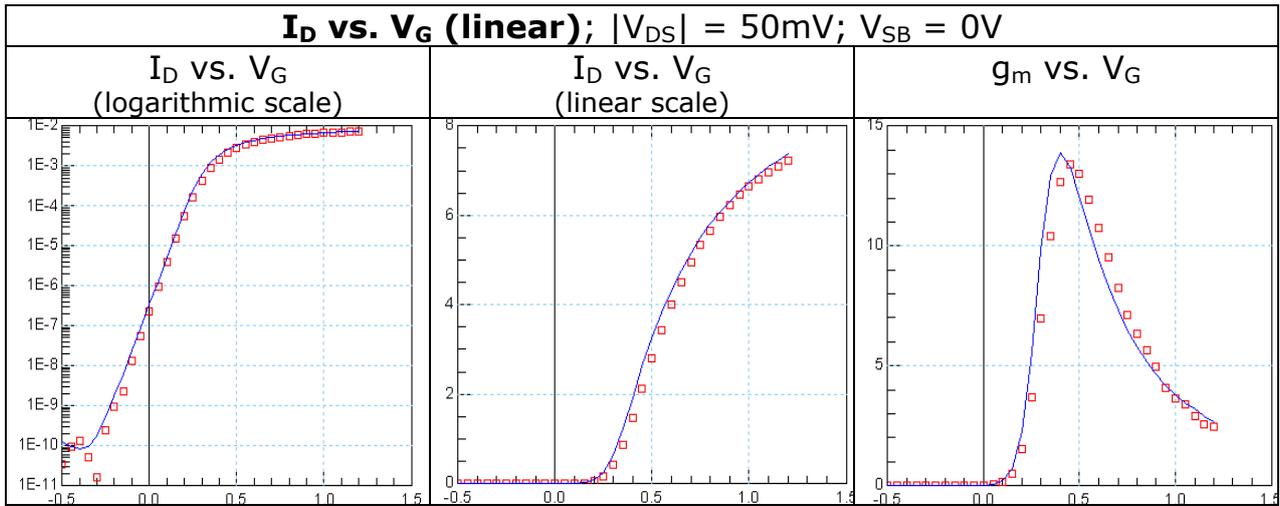
I_D vs. V_G (saturation); $V_{DS} = 1.2V$; $V_{SB} = 0V$



I_D vs. V_D ; $|V_{GS}| = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2\}V$; $V_{SB} = 0V$

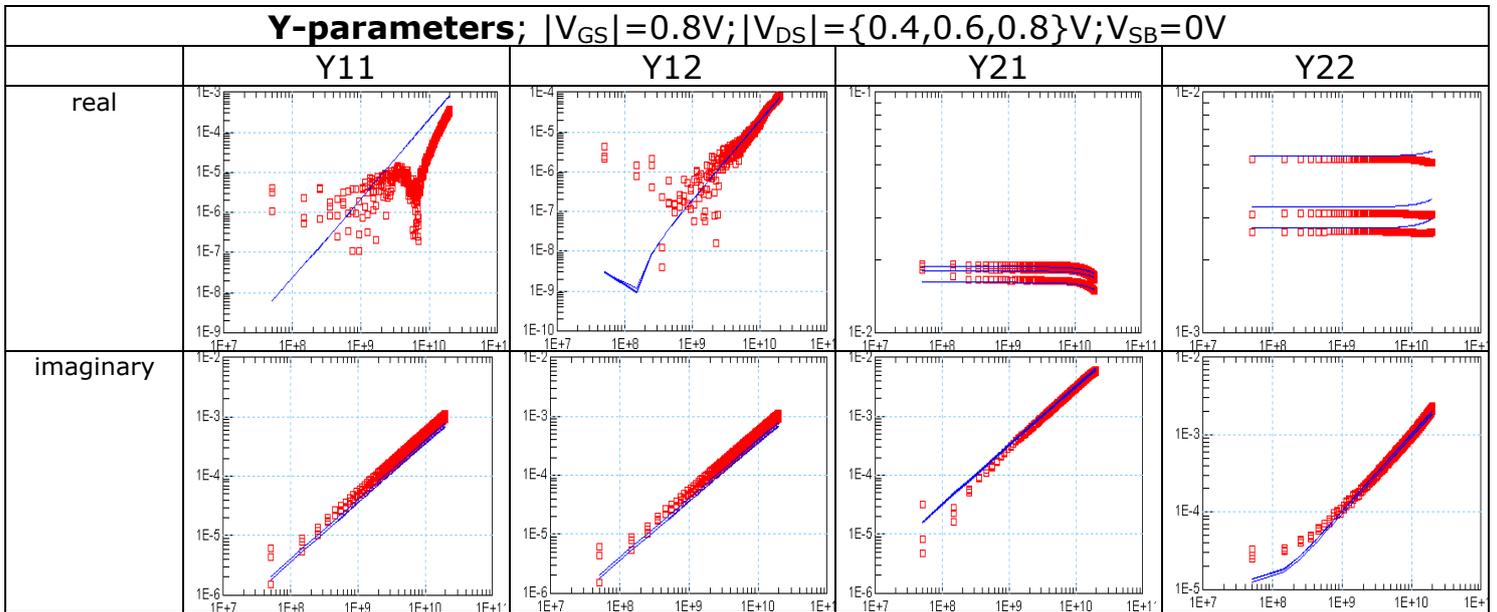


DUT: NMOS, WF = 2um, LF = 70nm, NF = 40

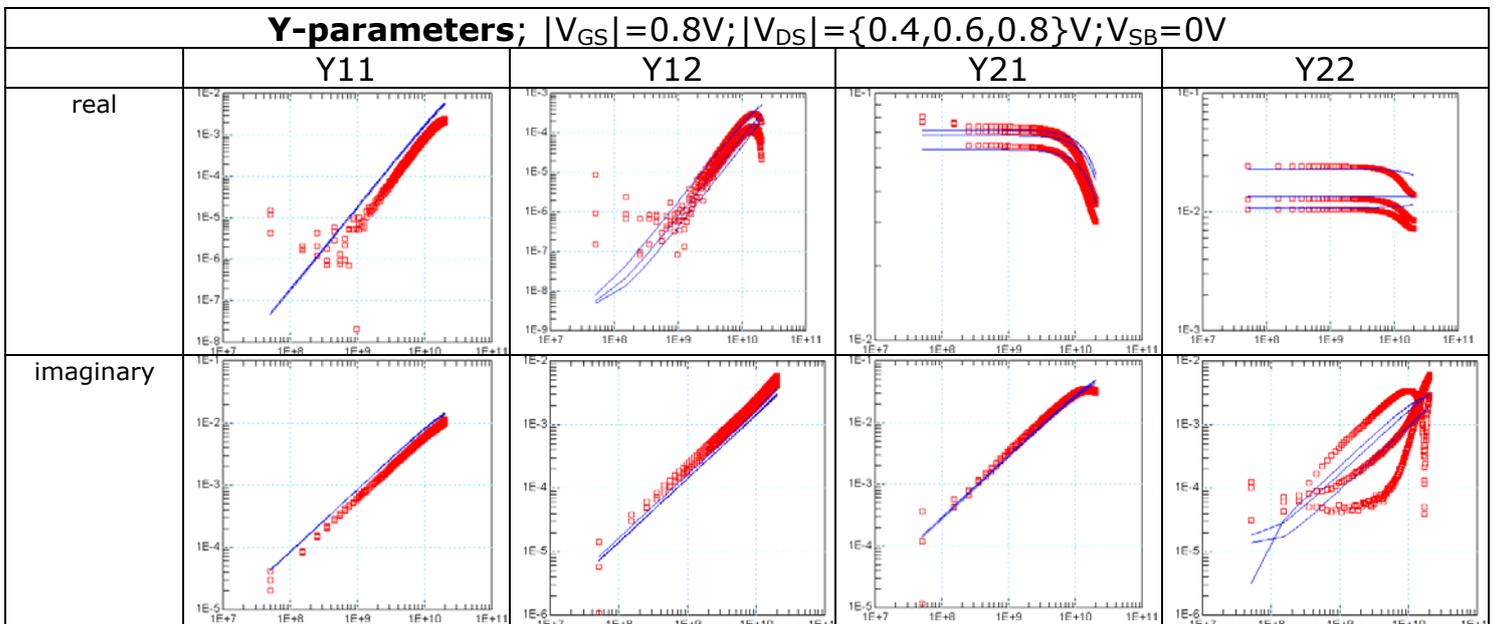


4.1.2 Y-Parameter Analysis

DUT: NMOS, WF = 2um, LF = 70nm, NF = 10

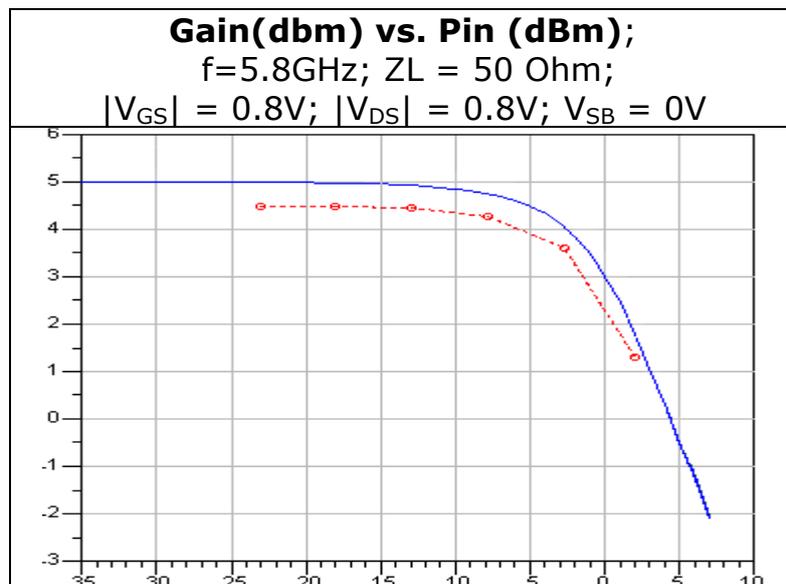
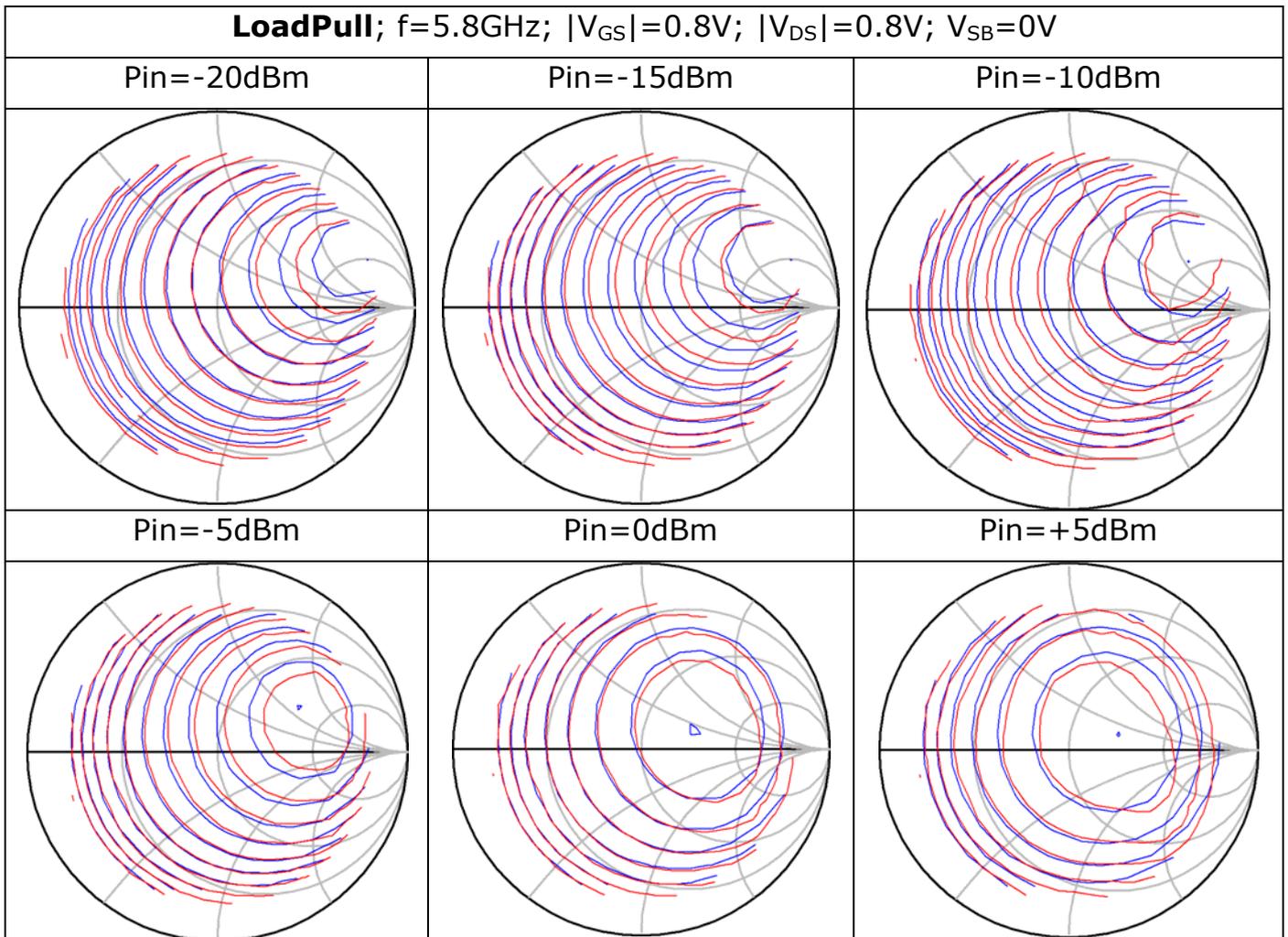


DUT: NMOS, WF = 2um, LF = 70nm, NF = 40

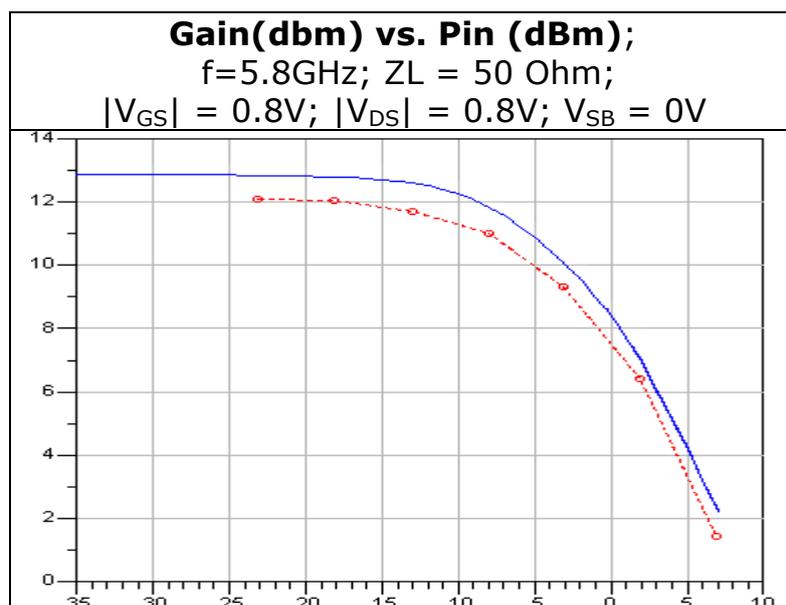
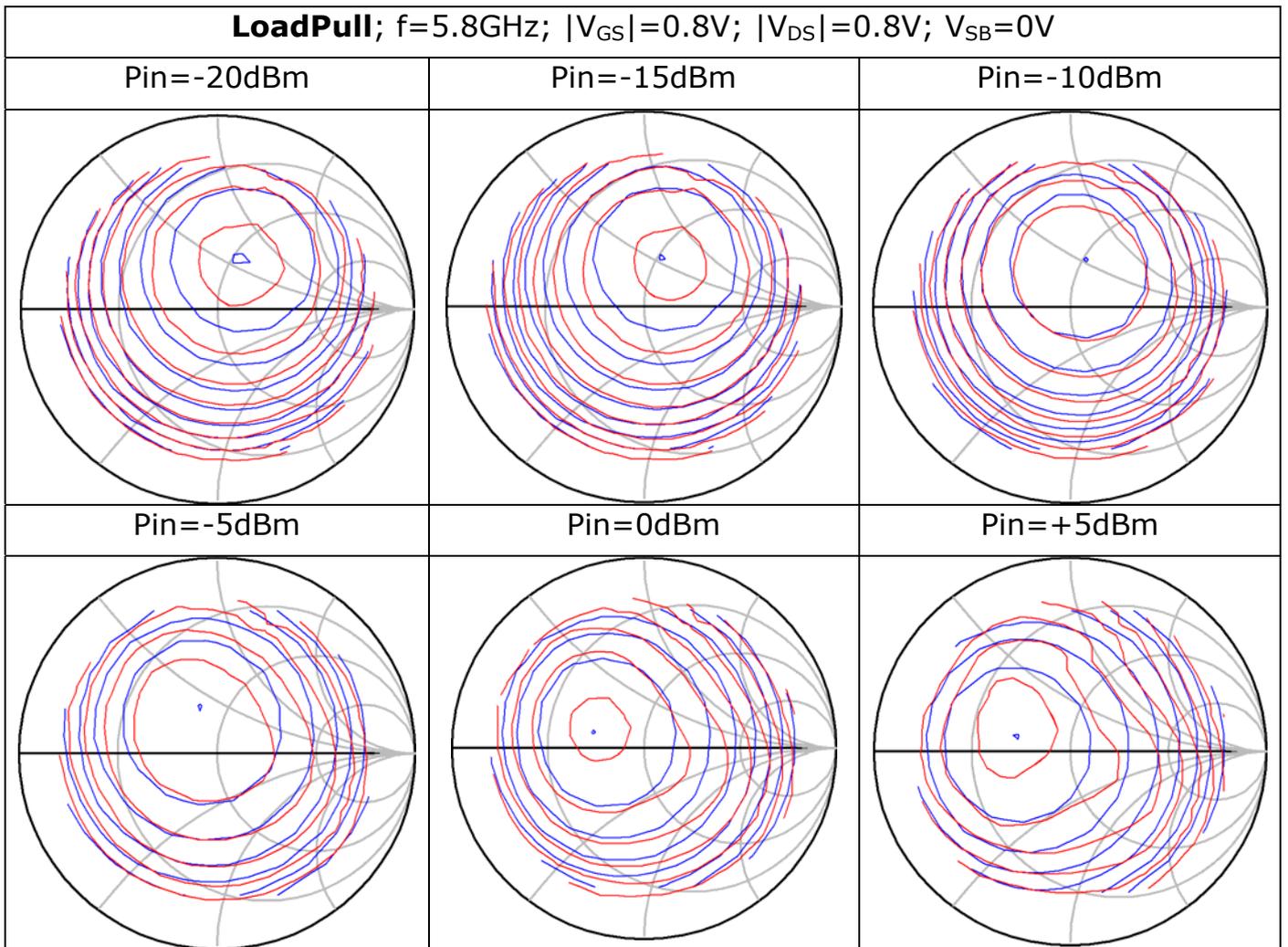


4.1.3 Load-pull Analysis

DUT: NMOS, WF = 2um, LF = 70nm, NF = 10



DUT: NMOS, WF = 2um, LF = 70nm, NF = 40

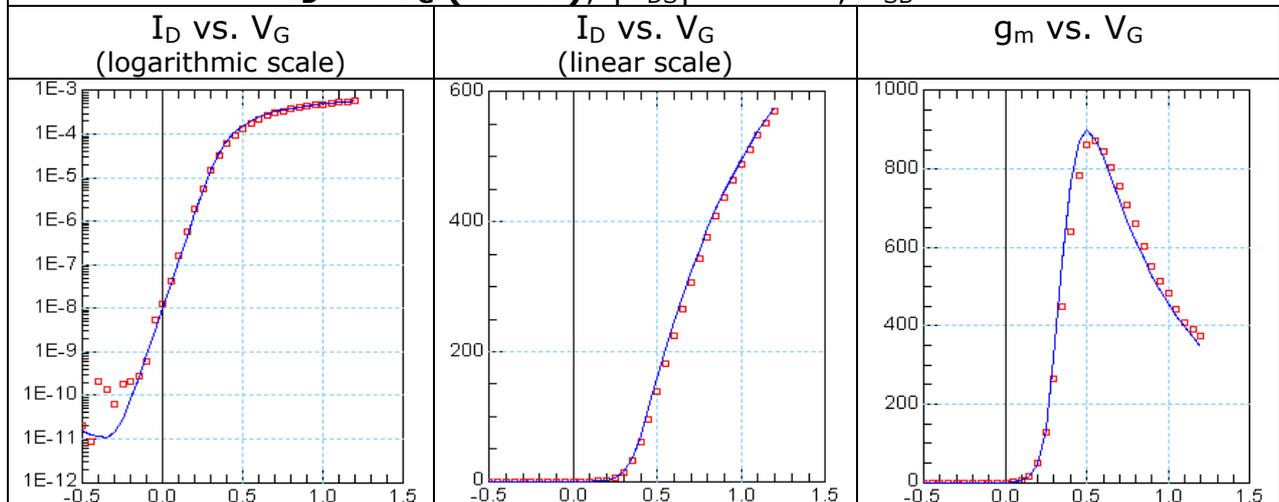


4.2 Evaluation for PMOS devices

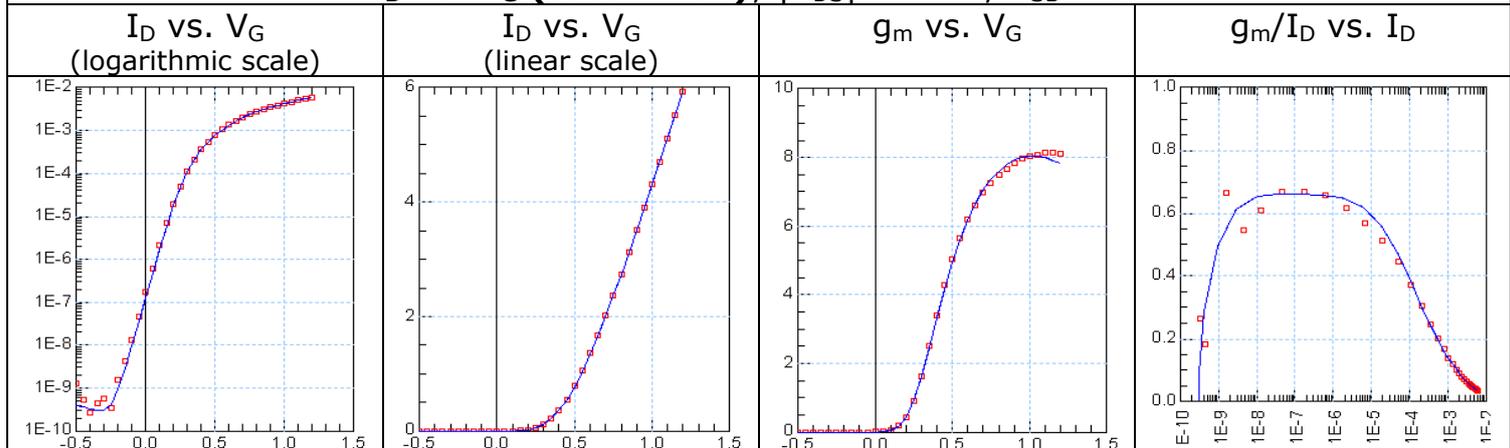
4.2.1 Static Analysis

DUT: PMOS, WF = 2 μ m, LF = 70nm, NF = 10

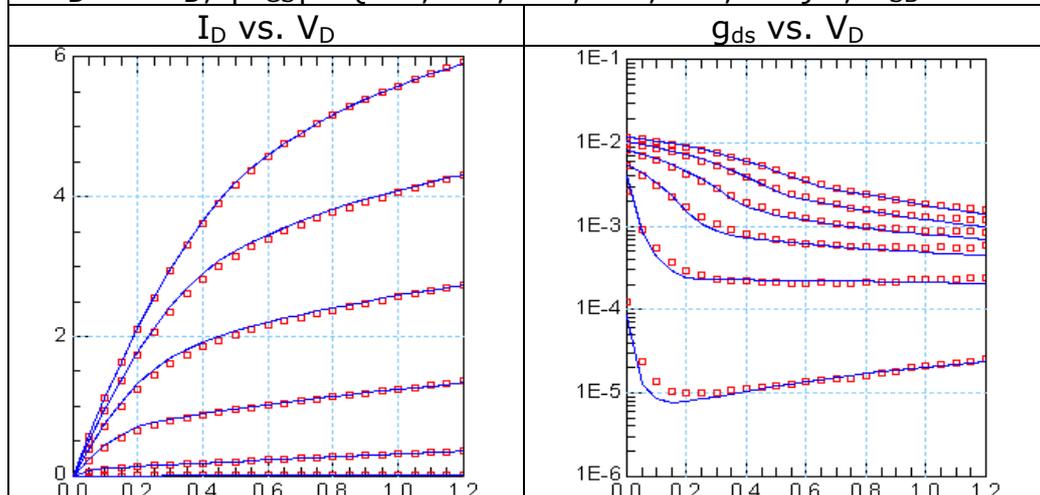
I_D vs. V_G (linear); $|V_{DS}| = 50\text{mV}$; $V_{SB} = 0\text{V}$



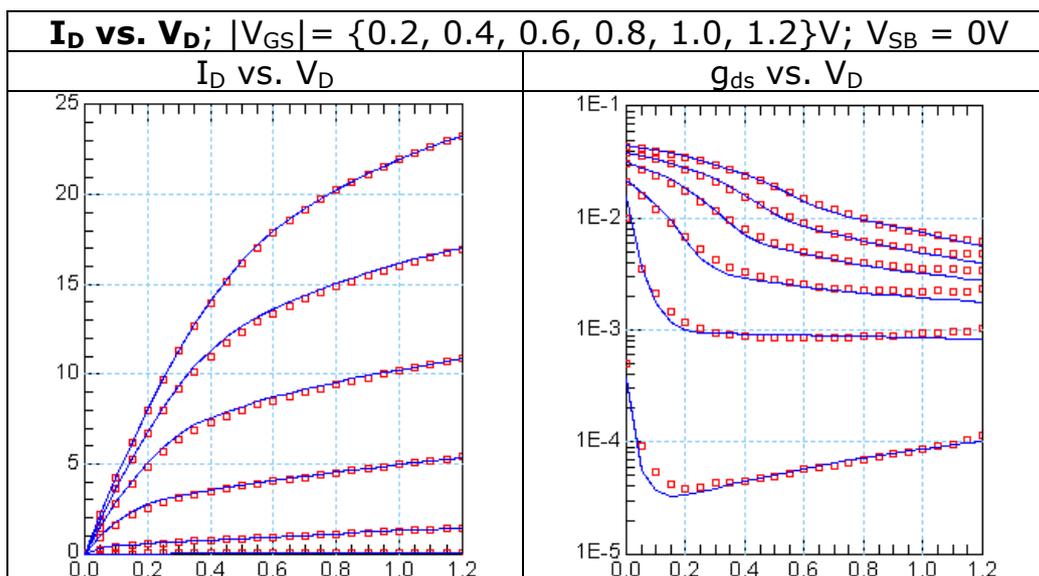
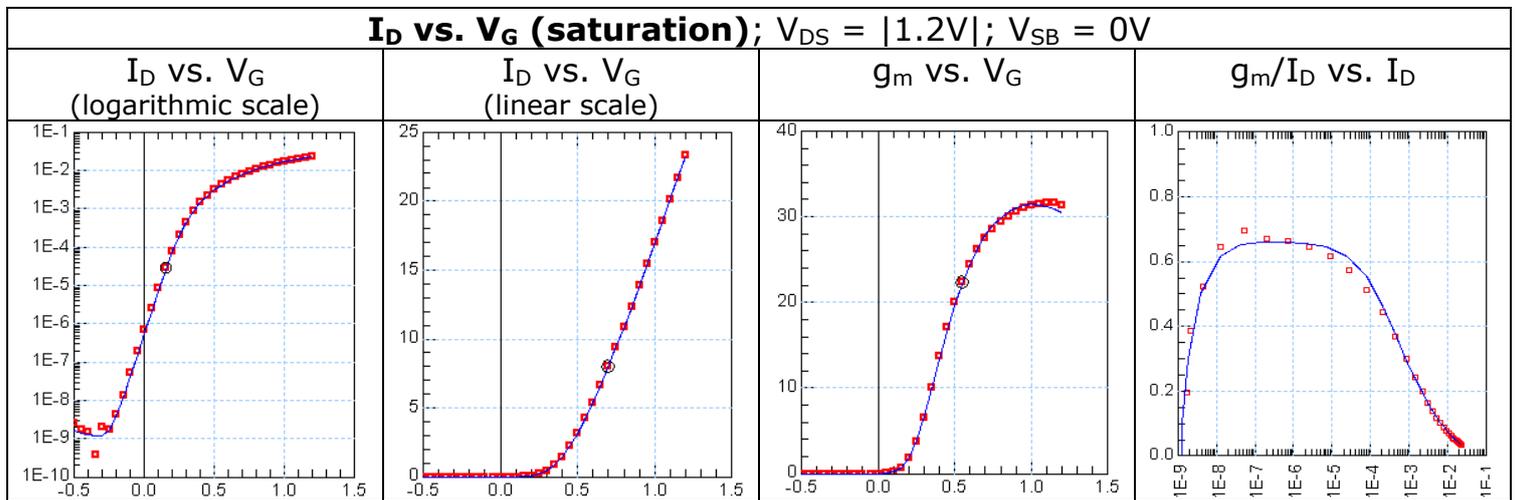
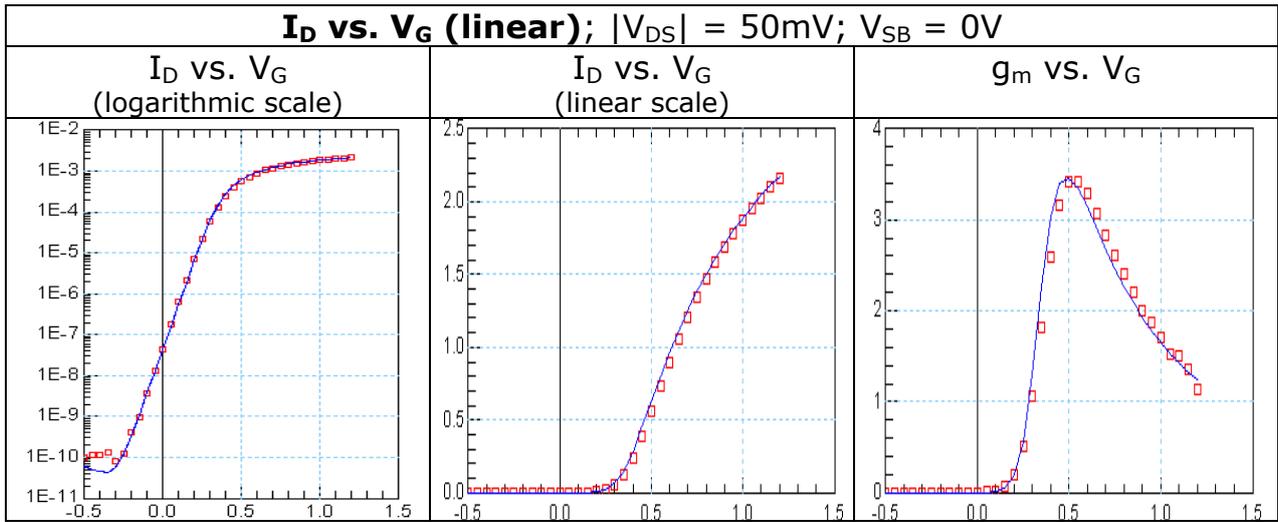
I_D vs. V_G (saturation); $|V_{DS}| = 1.2\text{V}$; $V_{SB} = 0\text{V}$



I_D vs. V_D ; $|V_{GS}| = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2\}\text{V}$; $V_{SB} = 0\text{V}$

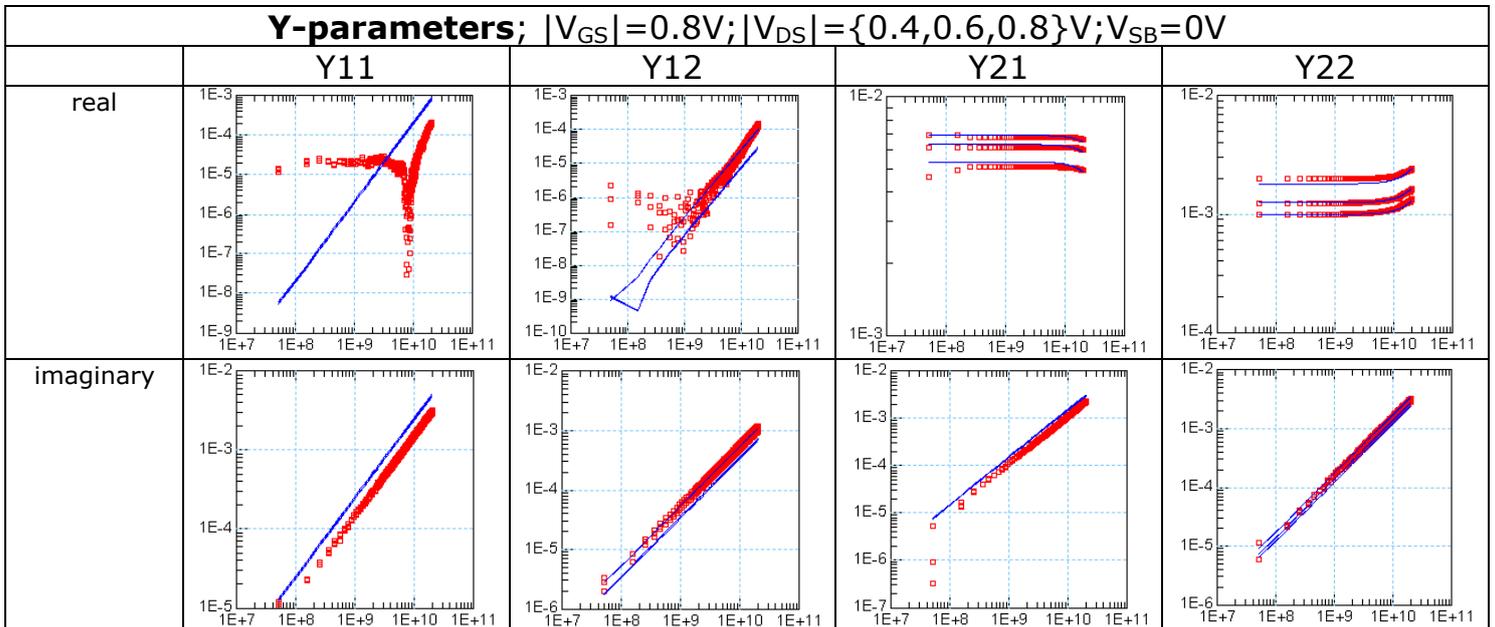


DUT: PMOS, WF = 2um, LF = 70nm, NF = 40

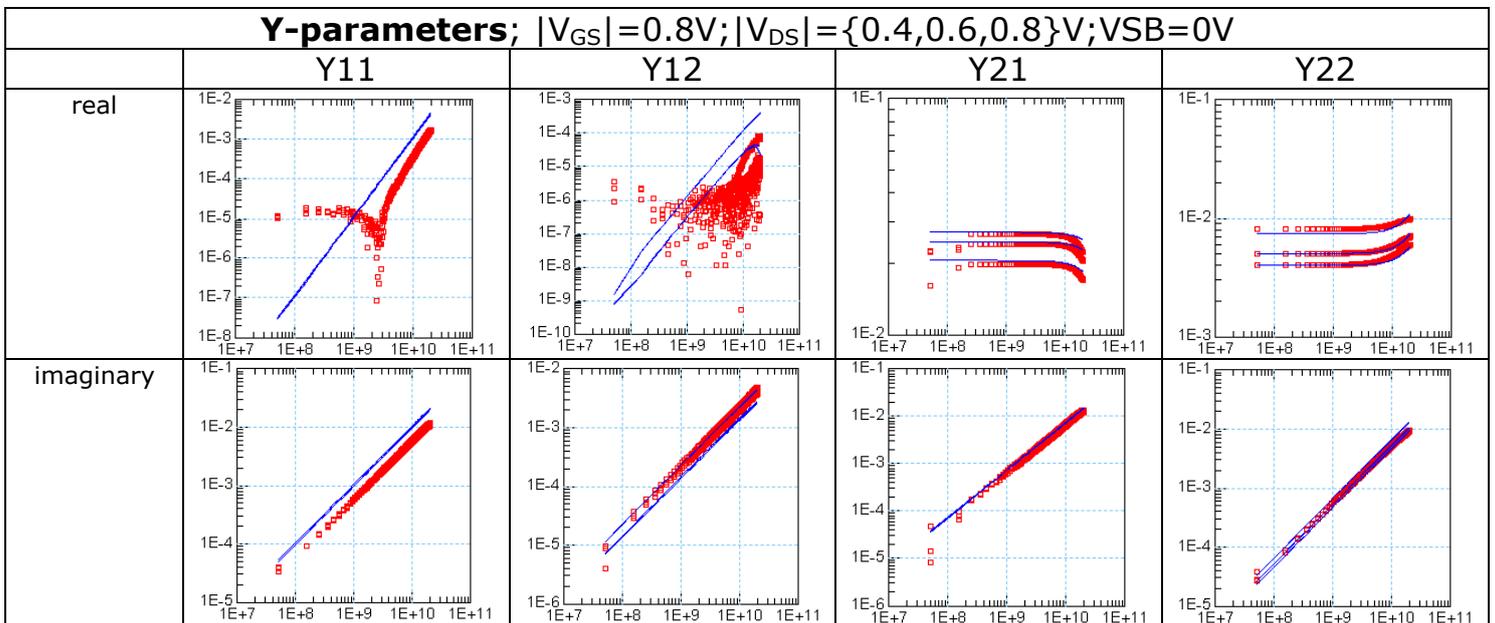


4.2.2 Y-Parameter Analysis

DUT: PMOS, WF = 2 μ m, LF = 70nm, NF = 10

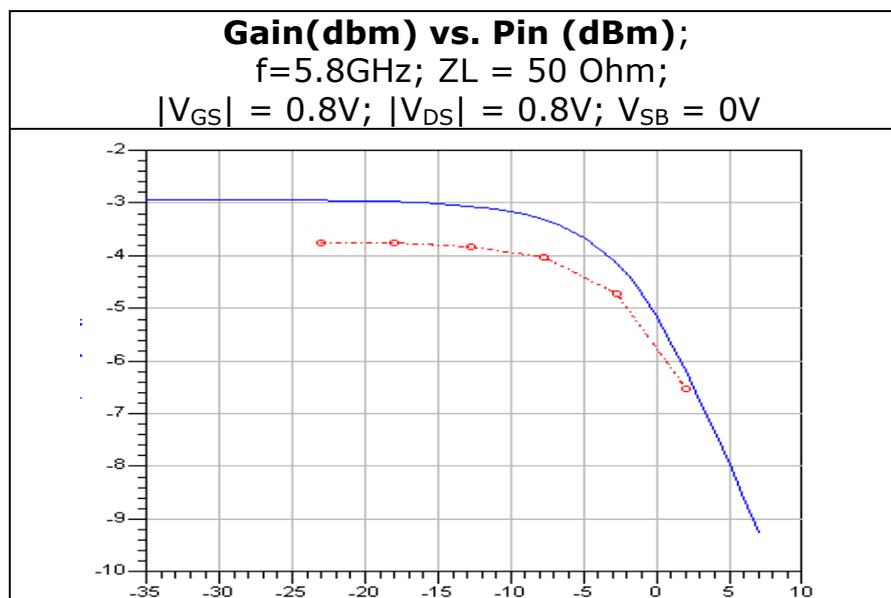
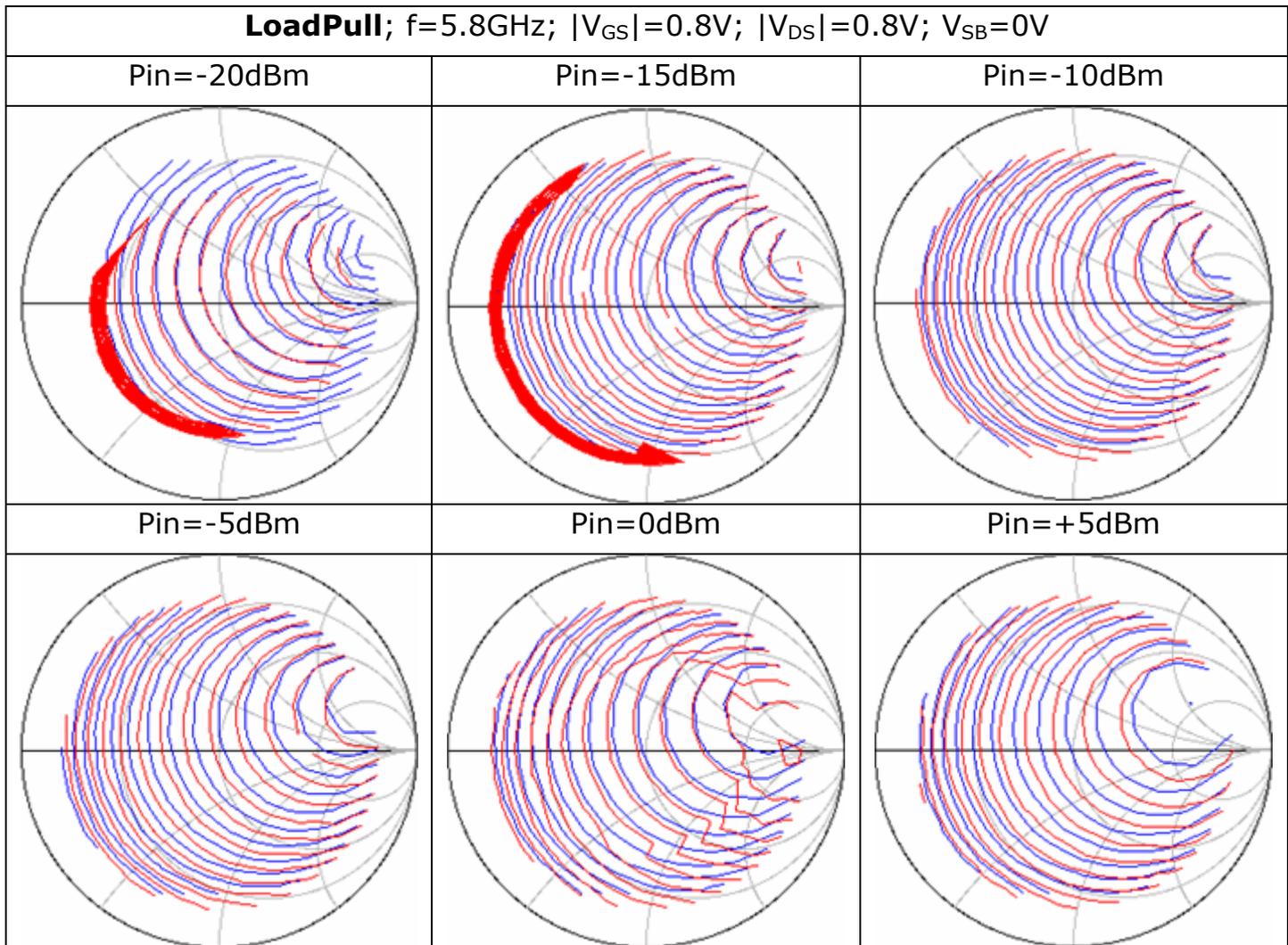


DUT: PMOS, WF = 2 μ m, LF = 70nm, NF = 40

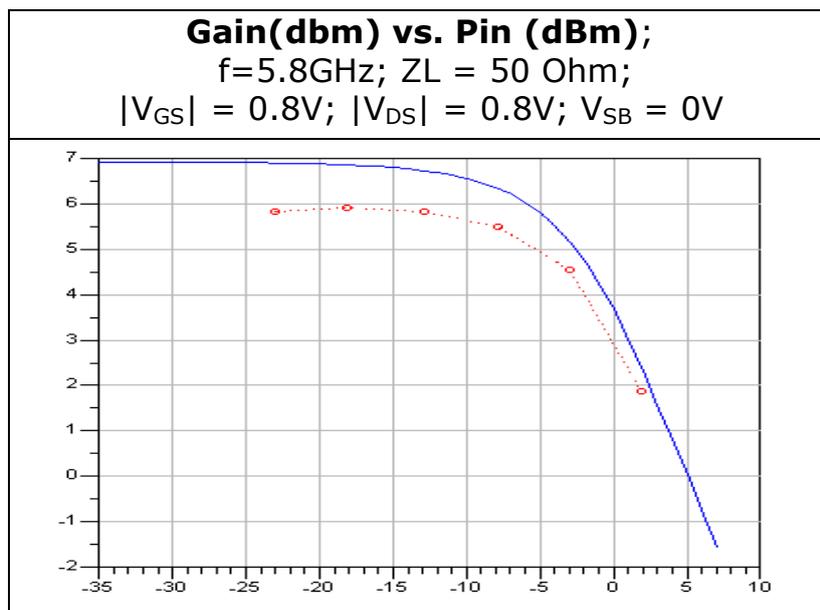
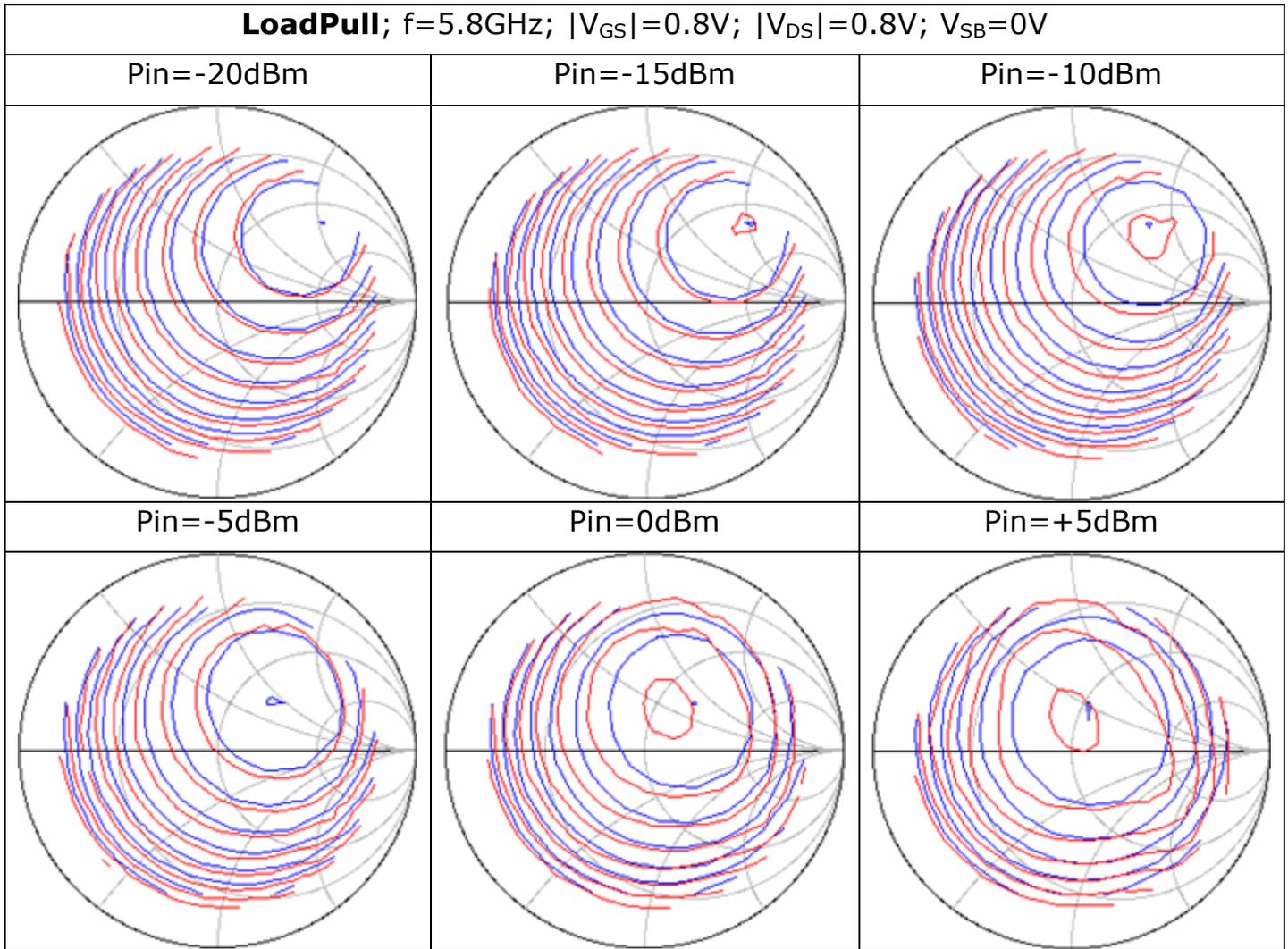


4.2.3 Load-pull Analysis

DUT: PMOS, WF = 2um, LF = 70nm, NF = 10



DUT: PMOS, WF = 2 μ m, LF = 70nm, NF = 40



4.3 Evaluation Results

In this chapter and as well as in Chapter 3, throughout description of the parameter extraction procedure, several plots have been displayed which depict the good capability of the EKV3 model to predict with a very good accuracy the behaviour of a transistor in a wide range of measurements.

EKV3 incorporates physical aspects which are necessary to model deep submicron CMOS technology and thus responds to the needs of transistor small-signal behaviour at low and high frequencies. Continuity of the large and small signal characteristics from weak through moderate to strong inversion and from conduction to saturation is one of the characteristics of the EKV3 MOSFET model. Yet further investigation should be worked out to achieve better accuracy in large-signal conditions where self-heating and other phenomena appear.

The parameter sets for the cases of NMOS and PMOS transistors which were used for the evaluation of the EKV3 model are presented in Appendix A.

Chapter 5

Conclusions

The target of this thesis is to display the MOSFET modelling procedure in a comprehensive way. The model which is used is a newly developed one and specifically the EKV3 MOSFET model.

CMOS technology gains more and more ground in integrated circuits due to its small size and low fabrication cost. Nowadays, MOSFETs that have gate length a few decades of nanometers have appeared. Due to the downscaling of supply voltages, transistors are used in low consumption circuits and thus operate in weak or moderate inversion, contrary to the past when they were used mostly in strong inversion.

MOSFET models should take into account the new conditions that have risen and be able to predict the behaviour of transistors accurately. In particular, the scaling of multi-finger layout that is typically used in RF design, needs to be addressed. However, a model, except from accuracy should give the designers further insight into the device behaviour. The EKV3 model is very well suited for this. EKV3 is an analytical model based on charge sheet theory using a relatively small number of parameters with the use of simple equations. EKV3 incorporates all the modern phenomena that appear in modern submicron technologies. Thanks to the minimization of the model parameters of EKV3 the guidelines for parameter extraction are displayed in this thesis.

This thesis has specifically contributed to:

- ❖ Evaluating the EKV3 MOS transistor compact model versus DC, CV, and RF data from network analyzer and load-pull equipment, for two CMOS technologies, 110nm and 90nm.
- ❖ Establish and further refine parameter extraction procedures for DC and RF.
- ❖ Comparison among small- and large-signal performance for advanced CMOS technology.

The EKV3 model shows accuracy and consistency to modern CMOS technologies, which is depicted in the evaluation of the model with the use of a 90nm CMOS technology. Continuity of the large and small signal characteristics from weak through moderate to strong inversion and from conduction to saturation and under any bias conditions is one of the characteristics of the EKV3 MOSFET model.

5.1 Future Work

Nowadays, IC designers rely on MOSFET models for their designs. With the development of wireless applications a MOSFET model should be able to physically describe the behaviour of a transistor not only at low but also at high frequencies too.

The extraction of RF parameters remains a highly interesting but also demanding field of research. One specific goal of further research is the establishment of a more concise RF parameter extraction method. Furthermore, MOSFETs are already used in fully integrated RF transceivers and as a result operate under large signal conditions. Achieving consistency among small and large signal modelling remains a challenge. The present work can also serve to evaluate other advanced features, such as low- and high-frequency noise, which have not been addressed.

Finally, the compact model itself has shown its strengths but also certain weaknesses which should be overcome. The model should in particular evolve with future CMOS technology generations, such as 45nm technology and beyond. Each generation of technology is expected to add new challenges to the above.

Appendix A

Model Cards for NMOS and PMOS Transistors

In this appendix the whole set of the values of the parameters (model card) that used for the evaluation of the EKV3 model are presented. These parameters refer to 90nm CMOS technology. The first set covers the case of an NMOS transistor and the second the case of a PMOS transistor.

NMOS model card

```
-----  
-- FLAGS AND SETUP PARAMETERS --  
-----  
  
SIGN          = 1.000  
TNOM          = 27.00  
TG            = -1.000  
QOFF          = 0.000  
XL            = 0.000  
XW            = 0.000  
SCALE         = 1.000  
  
-----  
-- OXIDE, SUBSTRATE AND GATE DOPING RELATED PARAMETERS --  
-----  
  
VTO           = 55.00m  
COX           = 14.50m  
XJ            = 20.00n  
PHIF          = 428.0m  
GAMMA         = 155.0m  
GAMMAG        = 15.00  
NO            = 1.050  
  
-----  
-- QUANTUM EFFECTS --  
-----  
  
AQMA          = 2.000  
AQMI          = 500.0m  
ETAQM         = 750.0m  
  
-----  
-- VERTICAL FIELD MOBILITY EFFECTS PARAMETERS --  
-----  
  
KP            = 1.100m  
E0            = 140.0MEG  
E1            = 10.00G  
ETA           = 1.000
```

```

-----
-- COULOMB SCATTERING --
-----

ZC          = 1.000u
THC         = 0.000

-----
-- DRAIN INDUCED THRESHOLD VOLTAGE SWIFT --
-----

FPROUT      = 3.000MEG
PDITS       = 60.00u
PDITSL      = 0.000
PDITSD      = 200.0m
DDITS       = 10.00

-----
-- MOBILITY RELATED GEOMETRICAL PARAMETERS --
-----

KA          = 400.0m
LA          = 300.0n
KB          = 500.0m
LB          = 20.00n
WKP1       = 240.0n
WKP2       = 600.0m
WKP3       = 500.0m

-----
-- LONG AND WIDE CHANNEL VTO AND GAMMA CORRECTION --
-----

LVT         = 1.000
WVT         = 100.0u
AVT         = 35.00m
LGAM        = 1.000
WGAM        = 100.0u
AGAM        = 25.00m

-----
-- GEOMETRICAL PARAMETERS --
-----

DL          = 25.00n
DW          = -10.00n
DLC         = 0.000
DWC         = 15.00n
LL          = 0.000
LLN         = 1.000
WDL         = 7.000f
LDW         = 2.500f

-----
-- EXTERNAL SERIES RESISTANCE --
-----

RLX         = 75.00u

```

```

-----
-- REVERSE SHORT CHANNEL EFFECT --
-----

LR          = 67.08n
QLR         = 3.300m
NLR         = 75.00m
FLR         = 1.000

-----
-- INVERSE NARROW WIDTH EFFECT --
-----

WR          = 100.0n
QWR         = 5.000u
NWR         = 5.000m

-----
-- CHARGE SHARING EFFECT --
-----

NCS         = 1.500
LETA0       = 1.00MEG
LETA        = 1.050
LETA2       = -25.00n
WETA        = 0.000

-----
-- VELOCITY SATURATION & CLM RELATED PARAMETERS --
-----

UCRIT       = 3.800MEG
LAMBDA      = 480.0m
DELTA       = 2.000
ACLM        = 830.0m

-----
-- DRAIN INDUCED BARRIER LOWERING --
-----

ETAD        = 1.100
SIGMAD      = 300.0m

-----
-- OVERLAP CAPACITANCES --
-----

LOV         = 21.00n
GAMMAOV     = 4.000
VFBOV       = 0.000
VOV         = 1.000

-----
-- FRINGING CAPACITANCES --
-----

KJF         = 0.000
CJF         = 0.000

```

```

-----
-- IMPACT IONIZATION CURRENT --
-----

IBA          = 0.000
IBB          = 900.0MEG
IBN          = 100.0m

-----
-- GATE CURRENT --
-----

KG           = 15.00u
XB           = 6.000
EB           = 40.00G
LOVIG        = 25.00n

-----
-- FLICKER NOISE PARAMETERS --
-----

AF           = 1.000
KF           = 0.000

-----
-- TEMPERATURE PARAMETERS --
-----

TCV          = 300.0u
BEX          = -1.200
TETA         = 0.000
UCEX         = 1.500
TLAMBDA     = 200.0m
TE0EX       = 0.000
TE1EX       = 0.000
IBBT         = 900.0u

-----
-- TEMPERATURE AND GEOMETRY SCALING --
-----

TCVL         = 0.000
TCVW         = 80.00p
TCVWL        = 0.000

-----
-- RESISTANCES TEMPERATURE SCALING --
-----

TR           = 0.000
TR2          = 0.000

-----
-- JUNCTION DRAIN - BULK AND SOURCE-BULK AREA, CURRENT, CAPACITANCE -
-----
+ SOURCE SIDE

NJS          = 1.250
JSS          = 10.00u
JSSWS        = 0.000

```

JSSWGS	= 0.000
MJS	= 150.0m
MJSWS	= 468.3m
MJSWGS	= 468.3m
PBS	= 598.8m
PBSWS	= 1.457
PBSWGS=	= 1.457
CJS	= 4.000m
CJSWS	= 70.00p
CJSWGS	= 70.00p
NJTSS	= 2.000
NJTSSWS	= 1.000
NJTSSWGS	= 1.000
VTSS	= 300.0m
VTSSWS	= 300.0m
VTSSWGS	= 300.0m
XJBVS	= 1.000
BVS	= 3.000
XTIS	= 3.000
TNJTSS	= 2.500
TNJTSSWS	= 2.000
TNJTSSWGS	= 2.000

+ DRAIN SIDE

NJD	= 1.250
JSD	= 10.00u
JSSWD	= 0.000
JSSWGD	= 0.000
MJD	= 150.0m
MJSWD	= 468.3m
MJSWGD	= 468.3m
PBD	= 598.8m
PBSWD	= 1.457
PBSWD	= 1.457
CJD	= 4.000m
CJSWD	= 70.00p
CJSWGD	= 70.00p
NJTSD	= 2.000
NJTSSWD	= 1.000
NJTSSWGD	= 1.000
VTSD	= 300.0m
VTSSWD	= 300.0m
VTSSWGD	= 300.0m
XJBVD	= 1.000
BVD	= 3.000
XTID	= 3.000

TNJTSD = 2.500
TNJTSSWD = 2.000
TNJTSSWGD = 2.000

+ BOTH SIDES

GMIN = 10.00p
TCJ = 867.4u
TCJSW = 687.3u
TCJSWG = 687.3u

TPB = 0.000
TPBSW = 0.000
TPBSWG = 0.000

-- SPICE GEOMETRICAL MODEL --

RSH = 0.000
LDIF = 0.000
HDIF = 240.0n

-- WIDTH SCALING PARAMETERS --

WE0 = 50.00n
WE1 = 0.000
WRLX = 50.00n
WUCRIT = -40.00n
WLAMBDA = 0.000
WETAD = 0.000
WUCEX = 0.000
WLR = 0.000
WQLR = 0.000
WNLR = 130.0n
WDPHIEDGE = 0.000

-- LENGTH SCALING PARAMETERS --

LWR = 0.000
LQWR = 1.500u
LNWR = 0.000
LDPHIEDGE = 0.000

-- EDGE DEVICE --

DGAMMAEDGE = 0.000
DPHIEDGE = 0.000
WEDGE = 0.000

-- COMBINED SCALING PARAMETERS --

WLDPHIEDGE = 0.000
WLDGAMMAEDGE = 0.000

-- STI STRESS EFFECT --

SAREF = 1.460u
SBREF = 1.460u
KVTO = 0.000
KKP = -35.00n
KGAMMA = 0.000
KETAD = 30.00n
KUCRIT = -9.000

-- GATE RESISTANCE --

GC = 1.000
RGSH = 53.00

-- SUBSTRATE RESISTANCE NETWORK --

RDSBSH = 1.500K
RBWSH = 1.000K
RSBWSH = 50.00
RDBWSH = 50.00

PMOS model card

-- FLAGS AND SETUP PARAMETERS --

SIGN = -1.000
TNOM = 27.00
TG = -1.000
QOFF = 0.000
XL = 0.000
XW = 0.000
SCALE = 1.000

-- OXIDE, SUBSTRATE AND GATE DOPING RELATED PARAMETERS --

VTO = -88.00m
COX = 13.50m
XJ = 10.410n
PHIF = 420.0m
GAMMA = 25.96m
GAMMAG = 5.500
N0 = 1.100

-- QUANTUM EFFECTS --

AQMA = 590.8m
AQMI = 392.6m
ETAQM = 750.0m

-- VERTICAL FIELD MOBILITY EFFECTS PARAMETERS --

KP = 310.0u
E0 = 95.00MEG
E1 = 8.259G
ETA = 1.311

-- COULOMB SCATTERING --

ZC = 1.000u
THC = 0.000

-- DRAIN INDUCED THRESHOLD VOLTAGE SWIFT --

FPROUT = 10.00MEG
PDITS = 15.00u
PDITSL = 0.000
PDITSD = 800.0m
DDITS = 10.00

-- MOBILITY RELATED GEOMETRICAL PARAMETERS --

KA = 0.000
LA = 3.000u
KB = 0.000
LB = 170.0n
WKP1 = 300.0n
WKP2 = -270.0m
WKP3 = 1.500

-- LONG AND WIDE CHANNEL VTO AND GAMMA CORRECTION --

LVT = 1.000
WVT = 100.0u
AVT = 150.0m
LGAM = 1.500u
WGAM = 1.000
AGAM = -200.0m

```

-----
-- GEOMETRICAL PARAMETERS --
-----

DL          = 85.00n
DW          = 0.00n
DLC         = 0.000
DWC         = 100.0n
LL          = 0.000
LLN         = 1.000
WDL         = 0.000
LDW         = 0.000

-----
-- EXTERNAL SERIES RESISTANCE --
-----

RLX         = 50.00u

-----
-- REVERSE SHORT CHANNEL EFFECT --
-----

LR          = 38.26n
QLR         = 8.300m
NLR         = 7.494
FLR         = 0.000

-----
-- INVERSE NARROW WIDTH EFFECT --
-----

WR          = 40.00n
QWR         = 300.0u
NWR         = -30.00m

-----
-- CHARGE SHARING EFFECT --
-----

NCS         = 2.000
LETA0       = 2.700MEG
LETA        = 1.500
LETA2       = 0.000
WETA        = 0.000

-----
-- VELOCITY SATURATION & CLM RELATED PARAMETERS --
-----

UCRIT       = 16.50MEG
LAMBDA      = 1.450
DELTA       = 1.000
ACLM        = 830.0m

-----
-- DRAIN INDUCED BARRIER LOWERING --
-----

ETAD        = 1.750
SIGMAD      = 1.000

```

-- OVERLAP CAPACITANCES --

LOV = 24.00n
GAMMAOV = 3.000
VFBOV = 0.000
VOV = 1.000

-- FRINGING CAPACITANCES --

KJF = 0.000
CJF = 0.000

-- IMPACT IONIZATION CURRENT --

IBA = 0.000
IBB = 400.0MEG
IBN = 100.0m

-- GATE CURRENT --

KG = 24.61u
XB = 4.280
EB = 28.42G
LOVIG = 28.06nn

-- FLICKER NOISE PARAMETERS --

AF = 1.000
KF = 0.000

-- TEMPERATURE PARAMETERS --

TCV = -500.0u
BEX = -1.100
TETA = 0.000
UCEX = 3.500
TLAMBDA = 1.500
TE0EX = 500.0m
TE1EX = 1.000
IBBT = 700.0u

-- TEMPERATURE AND GEOMETRY SCALING --

TCVL = -30.00p
TCVW = 30.00p
TCVWL = 0.000

-- RESISTANCES TEMPERATURE SCALING --

TR = -5.000m
TR2 = 0.000

-- JUNCTION DRAIN - BULK AND SOURCE-BULK AREA, CURRENT, CAPACITANCE -

+ SOURCE SIDE

NJS = 1.000
JSS = 1.000
JSSWS = 0.000
JSSWGS = 0.000

MJS = 900.0m
MJSWS = 700.0m
MJSWGS = 700.0m

PBS = 800.0m
PBSWS = 600.0m
PBSWGS = 600.m

CJS = 12.00m
CJSWS = 0.000
CJSWGS = 0.000

NJTSS = 1.000
NJTSSWS = 1.000
NJTSSWGS = 1.000

VTSS = 0.000
VTSSWS = 0.000
VTSSWGS = 0.000

XJBVS = 0.000
BVS = 10.00
XTIS = 3.000

TNJTSS = 0.000
TNJTSSWS = 0.000
TNJTSSWGS = 0.000

+ DRAIN SIDE

NJD = 1.000
JSD = 1.000
JSSWD = 0.000
JSSWGD = 0.000

MJD = 900.0m
MJSWD = 700.0m
MJSWGD = 700.0m

PBD = 800.0m
PBSWD = 600.0m
PBSWGD = 600.m

CJD = 12.00m

```

CJSWD      = 0.000
CJSWGD     = 0.000

NJTSD      = 1.000
NJTSSWD    = 1.000
NJTSSWGD   = 1.000

VTSD       = 0.000
VTSSWD     = 0.000
VTSSWGD    = 0.000

XJBVD      = 0.000
BVD        = 10.00
XTID       = 3.000

TNJTSD     = 0.000
TNJTSSWD   = 0.000
TNJTSSWGD  = 0.000

```

+ BOTH SIDES

```

GMIN       = 0.000
TCJ        = 694.8u
TCJSW      = 1.124m
TCJSWG     = 1.124m

TPB        = 2.450m
TPBSW      = -1.360m
TPBSWG     = -1.360m

```

-- SPICE GEOMETRICAL MODEL --

```

RSH        = 0.000
LDIF       = 0.000
HDIF       = 240.0n

```

-- WIDTH SCALING PARAMETERS --

```

WE0        = 150.00n
WE1        = 0.000
WRLX       = 0.000
WUCRIT     = 130.0n
WLAMBDA    = 25.00n
WETAD      = -40.00n
WUCEX      = 0.000
WLR        = 0.000
WQLR       = -20.00n
WNLR       = -170.0n
WDPHIEDGE  = 0.000

```

-- LENGTH SCALING PARAMETERS --

```

LWR        = 0.000
LQWR       = 0.000
LNWR       = -50.00n

```

LDPHIEDGE = 0.000

-- EDGE DEVICE --

DGAMMAEDGE = 0.000
DPHIEDGE = 0.000
WEDGE = 0.000

-- COMBINED SCALING PARAMETERS --

WLDPHIEDGE = 0.000
WLDGAMMAEDGE= 0.000

-- STI STRESS EFFECT --

SAREF = 1.460u
SBREF = 1.460u
KVTO = 7.000n
KKP = 30.00n
KGAMMA = 0.000
KETAD = 0.000
KUCRIT = 0.000

-- GATE RESISTANCE --

GC = 1.000
RGSH = 50.00

-- SUBSTRATE RESISTANCE NETWORK --

RDSBSH = 10.00K
RBWSH = 1.000K
RSBWSH = 75.00
RDBWSH = 75.00

References:

1. J. SEGURA, C. F. HAWKINS, "Cmos Electronics – How it works, How it Fails", IEEE Press, 2004.
2. C. GALUP-MONTORO, M. C. SCHNEIDER, "MOSFET Modelling For Circuit Analysis and Design", World Scientific, 2007.
3. A. BAZIGOS, "Modelling of MOS Transistors at High Frequencies", PhD thesis, National Technical University of Athens, May 2008.
4. S. YOSHITOMI, A. BAZIGOS, M. BUCHER, "The EKV3 Model Parameter Extraction and Characterization of 90nm RF-CMOS Technology", Proc. 14th Int. Conf. on Mixed Design of Integrated Circuits and Systems (MIXDES 2007), pp. 74-79, Ciechocinek, Poland, June 21-23, 2007.
5. S. YOSHITOMI, "Challenges of Compact Modelling for Deep-Submicron RF-CMOS Devices", Proc. 12th Int. Conf. on Mixed Design of Integrated Circuits and Systems (MIXDES 2005), Krakow, Poland, June 22-25, 2005.
6. F. KRUMMENACHER, M. BUCHER, W. GRABINSKI, "HF MOSFET Model Parameter Extraction", European Project, Project No 25710, 2000.
7. S. HUNG-MIN JEN, C. ENZ, D. PEHLKE, M. SCHROTER, B. SHEU "Accurate Modelling and Parameter Extraction for MOS Transistors Valid up to 10GHz", IEEE Transactions on Electron Devices, Vol.46, No.11, November 1999.
8. C. ENZ, E. VITTOZ, "Charge Based MOS Transistor Modelling: The EKV model for low-power and RF IC design", John Wiley & Sons Ltd, 2006.
9. M. BUCHER, C. LALLEMENT, C. ENZ, F. KRUMMENACHER "Accurate MOS Modelling For Analog Circuit Simulation using the EKV Model", IEEE Symp. On Circuits and Systems (ISCAS 1996), pp. 703-706, Atlanta, Georgia, May, 1996.
10. M. BUCHER, A. BAZIGOS, "An Efficient Channel Segmentation Approach for Large-Signal NQS MOSFET Mode, Solid-State Electronics", Volume 52, Issue 2, , pp. 275-281, Feb. 2008.
11. Y. TSIVIDIS, "Operation and Modelling of the MOS Transistor", Second Ed. Oxford University Press, 1999.
12. A. BAZIGOS, M. BUCHER, F. KRUMMENACHER, A. ROY, C. ENZ, "EKV3 MOSFET Compact Model, Model's Documentation", Model Version: 301.02, Technical Report, Technical University of Crete, July 2008.
13. A. BAZIGOS, M. BUCHER, "The EKV3.0 Model Code and Parameter Extraction", EKV Model Users' Group Meeting and Workshop, Online: <http://legwww.epfl.ch/ekv/workshop>, EPFL, Lausanne, Switzerland, November 4-5, 2004

14. A. BAZIGOS, M. BUCHER, P. SAKALAS, M. SCHROTER, W. KRAUS, "High-frequency Compact Modelling of Si-RF CMOS", *Phys. Status Solidi (c), Current Topics in Solid State Physics*.
15. M. BUCHER, A. BAZIGOS, **M.-A. CHALKIADAKI**, "Aspects of High-Frequency Modelling of MOSFETs with EKV3", *European Solid-State Circuits/Device Research Conference (ESSCIRC/ESSDERC), MOS-AK Workshop, Edinburgh, Scotland, September 19, 2008*.