Jan Saijets

MOSFET RF Characterization Using Bulk and SOI CMOS Technologies

MOSFET RF Characterization Using Bulk and SOI CMOS Technologies

Jan Saijets

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Department of Electrical Communications Engineering for public examination and debate in Auditorium S4 at the Helsinki University of Technology (Espoo, Finland) on the 18th of June, 2007, at 12 o'clock noon.

Abstract

MOSFET radio-frequency characterization and modeling is studied, both with SOI CMOS and bulk CMOS technologies. The network analyzer measurement uncertainties are studied, as is their effect on the small signal parameter extraction of MOS devices. These results can be used as guidelines for designing MOS RF characterization layouts with as small an AC extraction error as possible. The results can also be used in RF model extraction as criteria for required optimization accuracy.

Modifications to the digital CMOS model equivalent circuit are studied to achieve better RF behavior for the MOS model. The benefit of absorbing the drain and source parasitic series resistances into the current description is evaluated. It seems that correct high-frequency behavior is not possible to describe using this technique. The series resistances need to be defined extrinsically. Different bulk network alternatives were evaluated using scalable device models up to 10 GHz. Accurate output impedance behavior of the model requires a bulk resistance network. It seems that good accuracy improvement is achieved with just a single bulk resistor. Additional improvement is achieved by increasing the number of resistors to three. At this used frequency range no further accuracy improvement was achieved by increasing the resistor amount over three. Two modeling approaches describing the distributed gate behavior are also studied with different MOS transistor layouts. Both approaches improve the RF characteristics to some extent but with limited device geometry. Both distributed gate models describe well the high frequency device behavior of devices not commonly used at radio frequencies.

Tiivistelmä

MOSFETin radiotaajuuskarakterisointia ja mallitusta tarkastellaan sekä SOI CMOS että bulk CMOS -teknologioilla. Piirianalysaattorien mittaus-epävarmuutta tarkastellaan ja niiden vaikutusta MOS-transistorin piensignaali-parametrien ekstraktointiin. Näitä tuloksia voidaan käyttää ohjenuorana RF MOS -karakterisointiin käytettävien piirikuvioiden suunnittelussa, kun halutaan AC-ekstraktoinnin virhe mahdollisimman pieneksi. Tuloksia voidaan käyttää myös RF-mallin ekstraktoinnissa halutun optimointitarkkuuden kriteerinä.

Digitaalisen CMOS-mallin vastinpiirimuunnelmia on tarkasteltu tarkoituksena saada MOS-mallille paremmat radiotaajuusominaisuudet. Kanavan kanssa sarjassa olevien parasiittisten vastusten vaikutusta on tarkasteltu, kun ne ovat joko erillisinä tai suoraan virtayhtälöön sisällytettyinä. Jälkimmäisen tavan hyötyä on arvioitu. Näyttää siltä, että oikeanlaatuisen suurtaajuuskäyttäytymisen kuvaaminen ei onnistu tällä tekniikalla. Kanavan kanssa sarjassa olevat vastuk-set on määriteltävä ulkoisiksi. Erilaisia substraattivastinpiirien vaihtoehtoja on arvioitu käyttäen skaalautuvia transistorimalleja 10 GHz:n taajuuteen asti. Tarkan ulostuloimpedanssin kuvaaminen edellyttää transistorille substraatti-vastinpiiriä. Näyttää siltä, että merkittävään mallin tarkkuuden parantumiseen riittää yksi ainoa substraattivastus. Tarkkuus paranee tästäkin lisättäessä vastusten määrää kolmeen. Käytetyllä taajuusvälillä ei saavutettu mallin tarkkuuden lisäparannusta, kun yritettiin nostaa substraattiverkon vastusten määrää yli kolmen. Erilaisilla MOS-transistorin piirikuvioilla tarkasteltiin myös kahta mallitustapaa, joilla voidaan kuvata jakautuneen hilan käyttäytymistä. Kummatkin lähestymistavat parantavat mallin radiotaajuusominaisuuksia johonkin rajaan asti, mutta vain tietyillä transistorigeometrioilla. Molemmat jakautuneen hilan mallit kuvaavat hyvin sellaisten transistoreiden suurtaajuus-ominaisuuksia, joita ei yleensä käytetä radiotaajuuksilla.

Preface

Contents

Appendix A: Complete Uncertainty Equations

List of Symbols and Abbreviations

3D three-dimensional
AC alternating current
A_v amplifier low-frequency gain
BAW bulk-acoustic-wave (resonator)
C_{bd} bulk-drain junction capacitance
C_{bs} bulk-source junction capacitance
C_{ds} drain-source capacitance
C_{dsi} drain-to-source intrinsic capacitance
C_{essw}, C_{edsw}, C_{esb}, C_{edb}, C_{bb}, C_{ge0} SOI-related substrate capacitances of BSIM3SOI model
C_{GA} area term of parasitic gate capacitance in parallel with R_G
C_{gb} gate-to-bulk capacitance
C_{GB0}, C_{gbi} gate-to-bulk zero-bias and intrinsic capacitance
C_{gd} gate-to-drain capacitance
C_{GD0}, C_{gdi} gate-to-drain zero-bias and intrinsic capacitance
C_{GP} perimeter part of parasitic gate capacitance in parallel with R_G
C_{gs} gate-to-source capacitance
C_{gs}' gate-to-source capacitance absorbed with source resistance effect
C_{GS0}, C_{gsi} gate-to-source zero-bias and intrinsic capacitance
C_{in} input capacitance
C_m gate-to-drain voltage dependent small signal transcapacitance
C_{mb} bulk-to-source voltage dependent small signal transcapacitance
CMOS complementary metal oxide surface (transistor)
COL MOS Model 9 overlap capacitance parameter
C_{out} output capacitance
DC direct current
ΔC_{in} input capacitance error
ΔC_{out} output capacitance error
DIBL drain-induced barrier lowering (short channel effect)
ΔR_{in} input resistance error
ΔR_{out} output resistance error
ΔS_{11m} measurement error of S_{11} magnitude
ΔS_{12m} measurement error of S_{12} magnitude
ΔS_{21m} measurement error of S_{21} magnitude
ΔS_{22m} measurement error of S_{22} magnitude
$\Delta\theta_{11}$ measurement error of S_{11} phase
$\Delta\theta_{12}$ measurement error of S_{12} phase
$\Delta\theta_{21}$ measurement error of S_{21} phase
$\Delta\theta_{22}$ measurement error of S_{22} phase
DUT device under test

FD	fully-depleted (SOI device)
FET	field-effect-transistor
f_{max}	maximum frequency of oscillation
f_t	cut-off frequency
GaAs	gallium arsenide
GHz	gigahertz
g_{ds}	drain conductance
g_{ds}'	drain conductance absorbed with source resistance effect
g_{in}	input conductance
g_m	gate voltage dependent transconductance
g_m'	transconductance absorbed with source resistance effect
g_m''	the g_m' transconductance referred for intrinsic gate voltage
g_{mb}	bulk-to-source voltage dependent transconductance
IC	integrated circuit
IF	intermediate frequency
I_{bd}	bulk-to-drain junction diode current
I_{bs}	bulk-to-source junction diode current
I_{ds}	drain-to-source current
L	channel length
L_d	parasitic drain inductance
L_g	parasitic gate inductance
L_s	parasitic source inductance
MOS	metal-oxide-surface
MOSFET	metal-oxide-surface-field-effect-transistor
n_f	number of parallel devices in a multifinger layout
NFD	non-fully-depleted (SOI device)
NMOS	N type MOS transistor
NQS	non-quasi-static (channel)
nm	nanometer
PD	partially-depleted (SOI device)
θ_{11}	S_{11} parameter phase
θ_{12}	S_{12} parameter phase
θ_{21}	S_{21} parameter phase
θ_{22}	S_{22} parameter phase
Q_B	bulk charge
Q_{BD}	bulk-to-drain junction charge
Q_{BS}	bulk-to-source junction charge
Q_D	drain charge
Q_{GB}	gate-to-bulk charge
Q_I	inversion charge
Q_S	source charge
RF	radio frequency
R_B	bulk resistance

R_{bds}, R_{bdd}, R_{bsb}, R_{bpb}, R_{bdb} Substrate network resistance components of the BSIM4 model

R_{BD} substrate resistance between substrate drain node and bulk node

R_{BS} substrate resistance between substrate source node and bulk node

R_D drain resistance

R_{ddiff} drain diffusion parasitic series resistance component of BSIM4 model

R_{ds} small signal drain-to-source resistance $= 1/g_{ds}$

R_{DS} parasitic drain to source resistance in parallel with the MOS channel

R_{DSB} substrate resistance between internal bulk node and substrate node

R_{ed} excess diffusion channel resistance

R_G gate resistance

$R_{g,elt}$ extrinsic input resistance of BSIM4 model

R_{Gtot} total resistance between the whole length of gate polysilicon

R_{GX} additional parasitic series resistance at the gate

R_i GaAs FET small-signal resistance between gate and source

R_{ii} intrinsic input resistance of BSIM4 model

R_{in} input resistance

$R_{in,i}$ intrinsic input resistance

R_j GaAs FET small-signal resistance between gate and drain

R_{juns}, R_{jund}, R_{bulk}, R_{well} Substrate network resistance components of the PSP model

R_{out} output resistance

R_S source resistance

R_{sdiff} source diffusion parasitic series resistance component of BSIM4 model

R_{subd} substrate resistance related to the drain

R_{subd2} substrate resistance related to the drain

R_{subs} substrate resistance related to the source

R_{subs2} substrate resistance related to the source

R_{st} quasi-static channel resistance

s Laplace term

S scattering parameters (two port case: S_{11}, S_{12}, S_{21}, S_{22})

S_{11m} S_{11} parameter magnitude ($=|S_{11}|$)

S_{12m} S_{12} parameter magnitude ($=|S_{12}|$)

S_{21m} S_{21} parameter magnitude ($=|S_{21}|$)

S_{22m} S_{22} parameter magnitude ($=|S_{22}|$)

SCBE   substrate current-induced body effect

SOI   silicon-on-insulator

SOLT   short-open-load-thru (network analyzer calibration method)

SOS   silicon-on-sapphire

σ_x^2 S parameter variance matrix

τ GaAs FET small-signal time constant

μm micrometer

v GaAs FET small-signal voltage over C_{gs} capacitance

V_{bd}	body-to-drain voltage
V_{dg}	drain-to-gate voltage
V_{gi}	intrinsic gate voltage
V_{gs}	gate-to-source voltage
V_{gsi}	intrinsic gate-to-source voltage
V_{in}	input voltage
V_{out}	output voltage
VTT	Technical Research Institute of Finland
W	channel width
ω	angular frequency
W_{eff}	effective channel width
W_f	width of one device in a multifinger RF layout
Y	admittance parameters (two port case: Y_{11}, Y_{12}, Y_{21}, Y_{22})
Y_{ds}	drain-to-source admittance
Y_{gs}	gate-to-source admittance
Y_{in}	input admittance
Y_L	load admittance
Y_{out}	output admittance
Y_P	parallel de-embedding parameters (two port case: Y_{p11}, Y_{p12}, Y_{p22})
Y_S	drain node source admittance in S_{12} calculation
Z_0	reference impedance level (50 Ω in this thesis)
Z_{in}	input impedance
Z_L	series impedance de-embedding parameters (two port case: Z_{L1}, Z_{L2}, Z_{L3})
Z_{out}	output impedance

1. Introduction

1.1 Importance of Models

The rapid growth of mobile telecommunication markets emphasizes the need for reliable analog integrated circuit (IC) design at high frequencies. Nowadays IC technology is the most practical solution for the mass production of electronic circuits, due to cheap silicon processing and the small size and weight of IC circuits.

Previously, the only noteworthy radio frequency processes were GaAs-based field-effect transistors (FETs), as silicon-based transistors were slow. Rapid technological improvements lead to fast silicon bipolar transistors usable at the frequencies used by mobile communication. Somewhat later the CMOS technology caught up and went beyond bipolar technology due to significant reductions in the channel length of the devices.

In the early days of electronics, simple voltage and current equations could be calculated by hand as the circuits were kept simple and accuracy was not as crucial. Higher demands on circuit complexity and the need to maximize the number of working products at a smaller cost required the use of a circuit simulator with models that map the real world accurately. Modern IC design is heavily dependent on accurate device models used in circuit simulation. Complex device behavior is practically impossible to calculate by hand, even in the case of a single active device in the circuit. This is true even at DC but at high frequencies device behavior is even more challenging to simplify or model.

Modern simulators are based mainly on the Newton-Raphson-algorithm, where the operating point as well as transient currents, charges and voltages at every node are calculated in an iterative manner. The model itself consists of a tabular set of model parameters which describe the FET technology at hand, and the device characteristics are described with a collection of mathematical equations implemented in the circuit simulator. Today, a state-of-the-art general compact MOS model consists of 300 to 400 equations. These equations are calculated using the provided model parameter values. Compact modeling is the only noteworthy approach to modeling semiconductor devices and is the chosen

approach for mainstream circuit simulators. The conflicting requirements of computational efficiency and model accuracy have ruled out the possibility of using physical 3D numerical models of the semiconductors. This approach would lead to accurate and physical simulation results but it would be computationally very inefficient leading to very long simulation times.

One challenge of the compact circuit model is to be able to describe all the operation modes of a semiconductor device. These operation regions or modes have to be described correctly and combined smoothly in order to avoid breaks in model continuity. This continuity is required by the Newton-Raphson algorithm used in the circuit simulators.

Another problem of using a compact model is its inaccurate behavior at high frequencies. For instance, modeling a geometry-dependent capacitance with a single lumped element is the conventional approach, as it is computationally the lightest method. A single lumped element is still just an approximation of continuous physical situations that usually behave less abruptly. A compact model does not take into account other geometry-dependent couplings that have been described in the model. For instance, a 3D model could take into account substrate couplings between different devices. Taking this into account in a circuit simulator would require a lot of modeling knowledge from the circuit designer and still an approximate lumped coupling model.

One typical problem of semiconductor devices is that the DC and AC models differ at high frequencies. This dispersion effect is luckily not present with MOSFETs, except for a special silicon-on-insulator MOSFET case where the output impedance has a frequency-dependent kink.

Modern deep submicron CMOS technologies have reached speeds not possible to imagine ten years ago. 40 and 60 GHz amplifier blocks have been demonstrated with 130 nm channel length CMOS devices [1]. Simulating CMOS devices at such frequencies is not possible using traditional digital MOSFET models. Even the parasitic component values for a scalable device model are very hard to describe accurately [1]. Without accurate models the circuit design is demanding, if not impossible. Any circuit simulator is only as good as its models.

CMOS is nowadays a real choice for RF circuitry due to its fast devices and low cost, which is due to the simple process. Although lithography is a very tough challenge, CMOS manufacturing does not require many process steps. Large cellular systems can now be processed on one chip containing most of the transceiver blocks at RF and baseband. Even digital circuitry can be put on the same chip, reducing product size and weight. An important driving force for RF CMOS modeling development has been the RF transceiver technology shift from super heterodyne to direct conversion architectures. Bulky and large IF-filters are not needed anymore, but the demands on computational power at baseband have increased a lot. Trends to co-integrate BAW filters on CMOS emerged a few years ago [2],[3], but they have not yet been utilized due to yield problems.

State-of-the-art CMOS technologies are well below 100 nm gate lengths. Technologies with minimum features of 90 nm are commercially available from the foundries of, for instance, Fujitsu, Texas Instruments, TMSC, IBM and UMC. Smaller gate lengths of 65 nm are also commercially available, whereas the state-of-the-art technologies are below 50 nm. The 90 nm technologies have transistor cut-off frequencies and f_{max} over 160 GHz. 65 nm processes should be as much as 30 to 40% faster. These technologies enable less challenging CMOS designs up to 20 GHz. In order to use these devices at 20 GHz or higher, the models are required to describe the behavior up to the cut-off frequency to account for the wide band characteristics as well as non linearities. Large signal modeling requires wide band operation of the MOS models. As these technological advances have enabled faster devices and RF circuitry is used basically by everyone, accurate MOSFET modeling is definitely required.

1.2 General Modeling Problems

With shrinking linewidths, scalable device sizes and the requirement of higher-frequency operation has put more and more weight on accurate device characterization. The models have to be very accurate and the extraction of model parameters must be done carefully. In case of CMOS technology the downscaling has reached gate lengths shorter than 100 nm, requiring the models to account for different kinds of small geometry effects accurately, as short channel effects occur even in the 1 μm region. In addition to this the model

should cover the whole set of interesting device geometries from short to long and from narrow to wide gates.

There are many approaches to device modeling, such as table-look-up models, 3D numerical models and different types of compact models based on various mathematics and process-dependent model parameters. The most accurate approach to most of the modeling problems would be the table-look-up model that is based on a huge table of experimental data that directly tells the circuit simulator the currents, the voltages and the charges with the specified requirements.

In CMOS modeling the older MOSFET SPICE models like Berkeley SPICE Model Level 1, 2, 3 and BSIM1 had difficulties even characterizing the DC properties of MOS devices properly. In addition to inaccuracies, a lot of current and capacitance discontinuities made the circuit simulation suffer from convergence errors. Newer mainstream models like BSIM3, MOS Model 9, EKV and BSIM4 describe the current and charge behavior much more accurately, but inaccuracies still exist. Except for EKV and PSP they still have some discontinuities in the zero drain voltage region, which is mainly a concern for digital circuit designers.

The accuracy of DC currents and DC voltage dependent charges are crucial for AC or high frequency simulation accuracy. The AC model is a linearization of the DC model at a specific operating point. In the case of the common source coupled MOSFET amplifier shown in Fig. 1 a) the somewhat simplified large signal equivalent circuit in b) is linearized to that of c) to get the AC equivalent circuit in d). All of the amplifier properties are dependent on the DC model, for instance input and output impedance match and the low frequency gain. The input impedance match depends on the input capacitance accuracy which can be simplified as,

$$C_{in} \approx \frac{\partial(Q_S + Q_D)}{\partial V_G} \qquad (1)$$

The input impedance match also depends on the impedance real part in,

$$Z_{in}(\omega) = R_{in} + j\omega C_{in} \qquad (2)$$

The reason for the input resistance R_{in} is not shown in Fig. 1, but it is partly a series-connected parasitic resistance, as well as partly affected by the non-quasi-static channel effects. Low frequency gain is approximately

$$A_v \approx \frac{g_m}{g_{ds}} = \frac{\dfrac{\partial I_{ds}}{\partial V_{gs}}}{\dfrac{\partial I_{ds}}{\partial V_{ds}}}$$ (3)

determined by transconductance g_m and drain conductance g_{ds}. The output match depends on the output capacitance

$$C_{out} \approx \frac{\partial Q_D}{\partial V_{dg}} + \frac{\partial Q_{BD}}{\partial V_{bd}}$$ (4)

and the output impedance real part

$$R_{out} \approx \frac{1}{g_{ds}} = \frac{1}{\dfrac{\partial I_{ds}}{\partial V_{ds}}}$$ (5)

In Eq. (4), the last term is due to the bulk-drain junction diode practically seen as a capacitance between drain and source at low frequencies before the bulk resistance becomes a noteworthy impedance compared to the bulk-junction capacitance.

AC properties are thus very sensitive to current-voltage and charge-voltage curve slopes. Small differences in the I_{ds}–V_{ds} curves can result in a very large difference in the derivate or g_{ds}–V_{ds} *curve at some operation points.* In a MOSFET amplifier circuit this g_{ds} error would lead to a large simulation inaccuracy of gain and output impedance match. An example of a 20% g_{ds} error could lead to a large S_{22} magnitude error close to the Smith chart center in Fig. 2 a). At low frequencies the match error could differ by many dB, as is shown in Fig. 2 b). In this example an output impedance error of 16 Ω can be seen as a 4 dB error in the output match. Low frequency gain error would only be of a similar amount, being approximately 20%. Typically the g_m accuracy is much better than the g_{ds} accuracy, which thus usually causes the gain error. Luckily the largest g_{ds} inaccuracies occur in the deep saturation region which does not affect the output match error much. However, the low frequency gain error can

easily be 100% erroneous, which is typical of the drain conductance accuracy. In the linear region of operation the model-versus-measurement error can be lower than 5% or better.
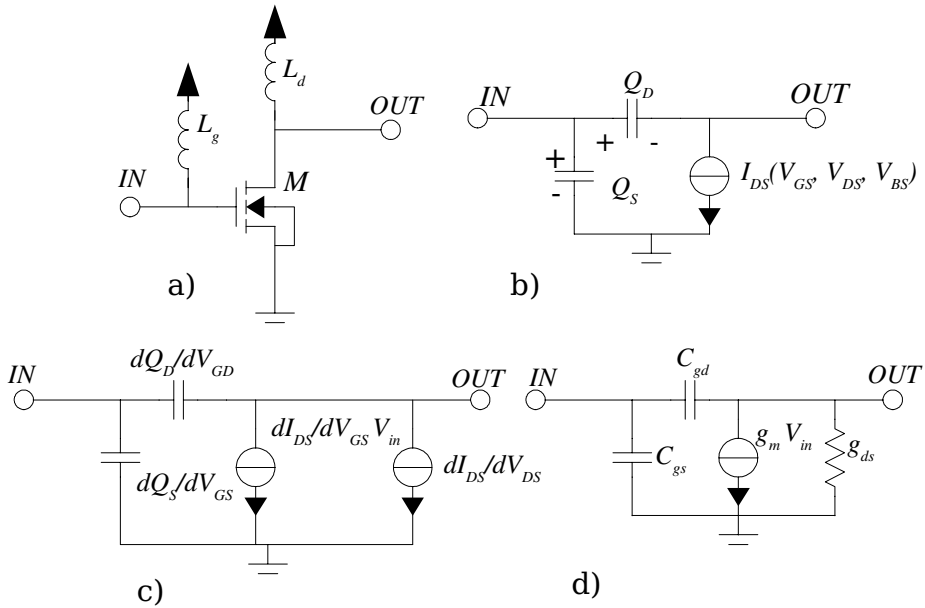
Figure 1. a) Common-source connected MOSFET b) DC equivalent circuit c) AC equivalent circuit d) simplified small signal equivalent circuit.
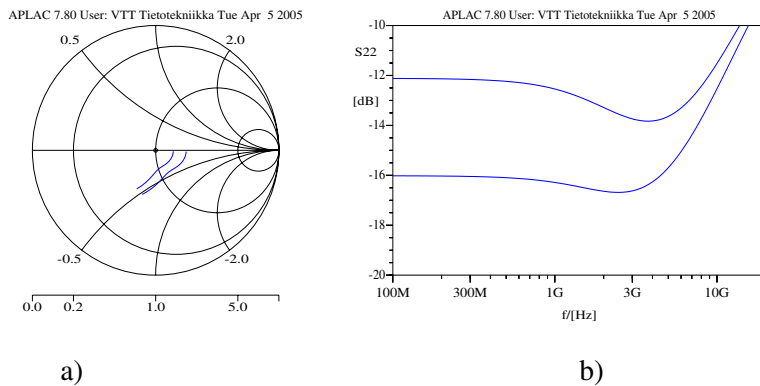
Figure 2. Output match difference due to a 20% error in drain conductance.

Inaccuracy in the voltage dependences of current and charge are not the only source of simulation errors; over-simplified equivalent circuits at high frequencies are also a problem. Real circuits or semiconductor devices contain more poles and zeros than their models do, resulting in inaccurate frequency behavior. For instance, the simple MOSFET equivalent circuit of Fig. 1 d) fails to describe the measured output resistance magnitude. In general, RF modeling is more complex than DC or even low-frequency AC simulation, as devices behave in a more distributed way at higher frequencies. At low frequencies a simple *RC*-equivalent circuit may well describe the MOSFET input, whereas at RF the gate resistance resembles more of an *RC*-ladder, and parasitic coupling from extrinsic metalizations and through the substrate become substantial. Substrate coupling is also the main reason why the simulated output impedance real part in Fig. 3 differs from the measured curve. At DC the MOSFET bulk or substrate can be considered to be the node to which it is connected, but at RF the substrate behaves in a distributed manner.
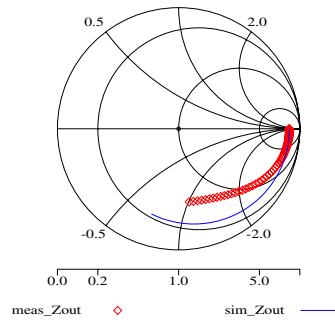
Figure 3. Output impedance difference due to inaccurate bulk resistance network.

Thus far the challenges have been easy in the linear world, but as the models are required to be used in nonlinear simulation they face even more problems. Accurate derivation of current and charges with the respect to gate or drain voltages more than once is required in nonlinear circuit analysis. For instance, an accurate spectrum magnitude for the second harmonic of the wanted signal requires the second derivative of currents and charges with respect to voltages to be accurate. Luckily this problem is out of the scope of this thesis.

Another aspect of a model is its speed in simulator use. The model should be as simple as possible and as accurate as possible. This is often a challenging trade-off and many mainstream MOSFET models are quite complex. Usually this is not a severe problem for a simple RF/analog circuit with only a few devices, but it can make circuit simulation very slow with large digital circuits.

1.3 Model Extraction Accuracy

Accurate circuit simulation does not require only that the model equations are accurate; it also requires that the model parameters used in the model equations are accurate. The process of obtaining these model parameters is referred to as parameter extraction. Model parameters are determined both from vendor specifications and measurement data of the device at hand.

MOSFET DC extraction routines are quite well established as the DC measurement equipment is quite accurate, reaching current uncertainties as small as 1 fA. Yet the currents must be measured carefully to also obtain the slope of current voltage curves accurately, in order to be able to extract the parameters describing conductance behavior. A suitable set of different-sized devices must be used for the extraction of the different geometry effects. AC extraction also has some well-established routines, but parasitic resistance extraction is still very troublesome. The extraction of small signal equivalent circuit component values is very inaccurate. For instance, our measurement of a transistor input resistance calculated from scattering parameters showed a peculiar frequency dependence in Fig. 4. Such behavior is not typical of RF MOSFETs with many parallel fingers; rather, it is characteristics of wide single-finger devices. In general, a good knowledge of AC extraction accuracy limits is needed not to trial for unphysical combinations of parameters. However, very little attention has been paid to this subject.
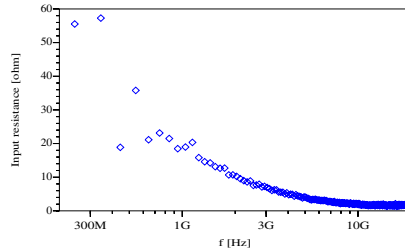
Figure 4. Out-of-the-ordinary input resistance behavior of a 60 x 6.8 μm x 0.5 μm SOI NMOS transistor.

1.4 Scope of This Thesis

The scope of this thesis is to cover the different aspects of MOSFET modeling, including both bulk and SOI CMOS technologies. In many cases bulk and SOI CMOS behavior differ very little and most of the modeling issues apply to both technologies, but a few differences exist. Characterization of the AC model is considered, as are the effects of the measurement uncertainties. Different equivalent circuit alternatives are considered and compared to the conventional modeling approach. This includes improvements to the equivalent circuits like the bulk resistance network, as well as the simplification of or modifications to the external resistances at the gate or drain and source.

Chapter 2 presents the state of the art in RF CMOS modeling. Chapter 2.1 describes general modeling issues and briefly introduces the different mainstream simulator models. The typical equivalent circuits implemented in circuit simulators are presented in Chapter 2.2, after which the distinct features of bulk and SOI technologies are considered in the two subsequent subchapters. Small signal equivalent circuit high-frequency behavior is studied in Chapter 2.4. The equations are formulated by me, and they follow an approach published only a few years back. Chapter 2.5 presents one of the most important high-frequency effects studied in recent years – the bulk effect and its equivalent circuits. The different mainstream models for bulk and SOI CMOS modeling are presented in more detail in Chapter 2.6.1.

The general extraction and measurement methods used in this thesis are presented in Chapter 3. None of the methods have real contributions from me. In

Chapter 3.1.1 both the DC and AC characterization measurements and calculations are discussed, along with the choice of measurement devices and de-embedding techniques, which is a sort of a post-calibration method. Chapter 3.2.1 discusses AC direct extraction approaches with different ways of simplifying the equivalent circuit and how the small signal values are calculated from it.

In Chapter 4 the measurement uncertainty effect on small signal parameter characterization is presented. Typical network analyzer uncertainties are chosen as the basis of the analysis. The total differential error of the MOSFET small signal parameters are calculated to get a realistic view of the extraction accuracy. These results can be used as guidelines to design suitable test layouts and for transistor biasing for AC extraction. All of the results are my own.

Chapter 5 discusses the typical series resistance approach in modern MOS models, where the drain and source resistances have been absorbed into the drain current equation. The effect on AC behavior is studied, and it is compared to the conventional way of modeling series resistances. All of the calculations and results are mine.

Chapter 6 discusses the similarities and differences between bulk and SOI CMOS devices from the modeling point of view and concentrating on input impedance. The views presented are general findings and results from the study. All of the results are my own.

In Chapter 7 the bulk effect on AC accuracy is studied. The AC accuracy of published and modified substrate networks are compared using experimental data. The improvements achieved with different substrate networks are presented on the S parameter fits, as well as small signal parameter fits. All of the results are my own. Most of the equivalent circuits studied are not developed by me.

Chapter 8 discusses other modifications to the MOSFET models and their improvements to AC accuracy. Basically, the distributed gate effect is studied with two approaches to modeling it. The first one is a small signal approximation by putting a parallel C_G capacitance in parallel with the gate resistance, and the second approach is to model the wide MOSFET with many

subdevices in parallel, but with their gates in series having resistances in between. The S parameter fits of both approaches are evaluated. All of the results are my own. The suggested modifications to the equivalent circuit are not mine.

Chapter 9 discusses the usability of the different results presented in this thesis. These are also compared with previous publications.

Chapter 10 concludes the results of this thesis.

2. Review of State of the Art High Frequency Characterization of MOSFETs

2.1 General MOSFET Modeling Issues

For a long time CMOS was not considered as a serious alternative for RF use and modeling efforts were mainly put into DC characterization and capacitance modeling to be used in low-frequency analog simulations. As linewidths shrank, RF CMOS modeling began to be studied more at the beginning of the 1990s.

One of the first comprehensive studies of MOSFET high-frequency behavior was done by Y. Tsividis [4] in 1987. The book characterizes MOSFET AC behavior by different equivalent circuits, depending on the required model accuracy. Both quasi-static and non-quasi-static operations have been considered. The approach by Tsividis is very general and theoretical to avoid differences caused by process parameters such as channel length. CMOS vendors kept distributing their design kits with MOS models, not paying very much attention to the extrinsic parasitic components like resistances or unwanted capacitances. The MOSFET design kits were designed merely for digital IC purposes. Later, at the beginning of the 1990s, the parasitic gate resistance was found to be a crucial RF component [5]. Not until the last half of the 1990s was serious RF CMOS design research begun [6]–[22]. A lot of impressive results with RF CMOS circuits were achieved with surprisingly long device lengths [19]. However, there were not many studies on RF CMOS modeling available. Our comparison of four different mainstream MOS model accuracies was reported in [23] and [24], which showed a lot of modeling challenges for MOSFETs at RF, mainly in achieving a scalable model accurate in all regions of operation. RF MOSFET modeling work mostly started to appear in the late 1990s. The main focus was put on output impedance modeling with the substrate resistance network [25]–[48], [90]. Modified and simple networks were also developed for the extrinsic gate [26], [28], [32], [37], [87], [95]. Many vendors did not include even a single resistor at the gate of their own MOS models in the 1990s.

As a result of our studies [23] and [24], it seemed that extracting scalable models of mainstream CMOS models and comparing them at RF did not clearly

present the sole RF accuracy of the models. The RF comparison possibly showed more differences caused by the DC model than differences caused by the AC properties of the models. The studies also showed that a lot of skills are required in the DC extraction of the models. As the first derivatives of the DC model components describe the linearized AC model, the DC model accuracy is crucial for low and high-frequency simulations [50]. In that sense the heavy emphasis on DC modeling before RF modeling was crucial. Along with the DC model and active charge description, most of the RF properties are already set in the MOS model, as was presented in Section 1.2. The current description derivatives with respect to different node voltages define the different conductances at the drain, and that in turn defines the first-order frequency response of the model. A small difference in simulated currents may result in large differences in its derivatives, as is shown in Fig. 5. This can result in large inaccuracies in circuit simulation.
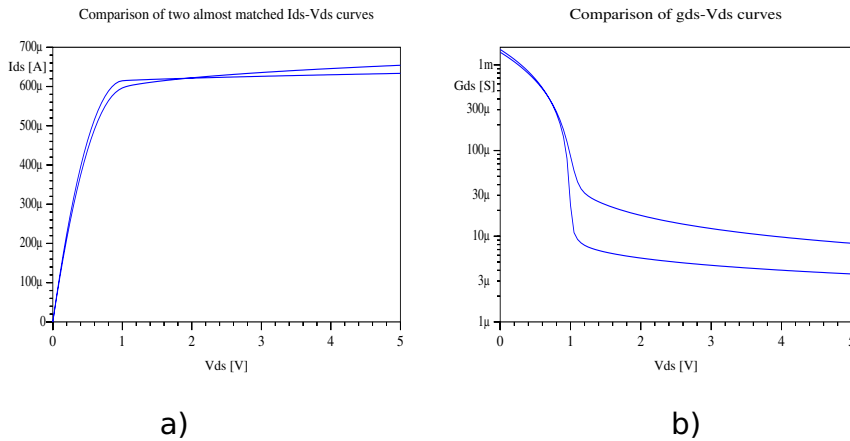
Figure 5. The small difference in currents can result in large derivative differences.

At the beginning, the great efforts put into DC modeling (a lot of references are available in [53]) were meant to correct unwanted features of the model, like discontinuities in the derivatives of the current descriptions and the voltage-dependent active charge model. Practically all current FET models are based on the simple long and wide MOSFET analytical equations [51], [52] developed in the 1960s. The newer models try to describe the old current model in a simulator-friendly way by introducing small geometry effects using a lot of mathematical conditioning and empirical fitting parameters [53]. The first-

generation models, like the Berkeley MOSFET Model Levels 1, 2 and 3, failed to describe the current derivative continuity producing drain conductances presented in Fig. 6 a) and b) using the Level 3 model [53]. The current itself in Fig. 6 a) is continuous and is seemingly quite smooth, but the derivative in Fig. 6 b) is apparently much worse. Another discontinuity was found in the transition region between subthreshold and moderate inversion of the I_{ds}–V_{gs} curve. Other discontinuities also exist, for instance in the transition regions of charges. Newer mainstream models, like BSIM1, BSIM2, BSIM3, MOS Model 9 and EKV, corrected many of the problems caused by the discontinuities, namely the convergence errors in the circuit simulations. Convergence errors are due to the fact that most circuit simulators use iteration methods to decide the next iteration step by using the derivative information. One of the most important such methods is the Newton-Raphson algorithm commonly used in many circuit simulators. Most of the models still suffer from the discontinuity of zero drain voltage where the model tries to decide which of the symmetric terminals is the drain and which is the source.

a)   b)
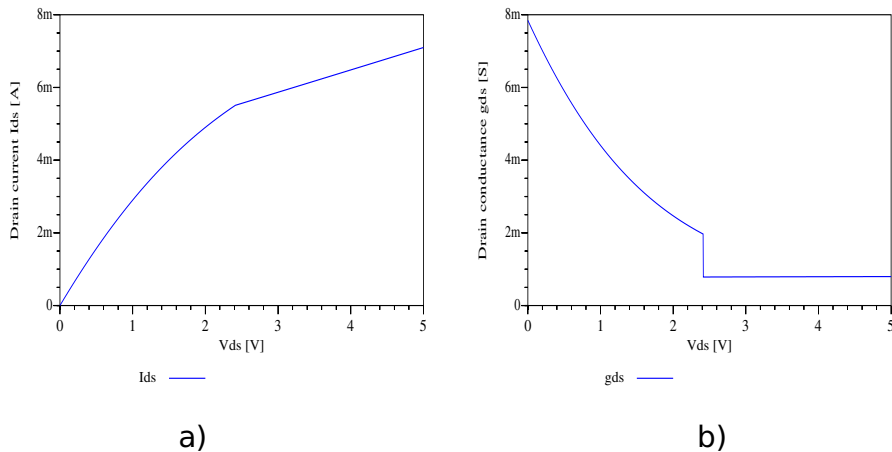
Figure 6 Current and conductance discontinuity of Berkeley SPICE Level 3 model.

Later, at the beginning of the new millennium, linearity also became a more thoroughly studied subject [47], [54]–[63]. The importance of DC model accuracy grew even more. The higher-order derivatives of the drain current and voltage-dependent active charges define the model nonlinearity behavior [55]. For example, the drain current second and third derivatives are important factors

26

in defining the first and second harmonic responses respectively of the model. According to the studies, the major sources of nonlinearity in MOS transistors are due to current nonlinearity [54]. The drain conductance nonlinearity was found to be an important factor at low frequencies, as was the transconductance nonlinearity. At higher frequencies the drain conductance effect is reduced by the feedback of the capacitances, leading to a dominant nonlinearity source by the transconductance. Capacitances, as well as the substrate network of MOSFETs, have been found to be less important nonlinearity sources [54]. Even with the newer models like BSIM3 and BSIM4, nonlinearity is a problem at zero-drain voltage [60]. The third-order intermodulation product is especially inaccurate in simulation due to the model's failure to describe the second order derivative at $V_{ds} = 0$ V.

2.2 Small Signal Equivalent Circuit of a MOSFET

The typical MOSFET large signal equivalent circuit implemented in mainstream circuit simulators is presented in Fig. 7. The model consists of the current model, bulk junction diodes, active charges, parasitic capacitances and parasitic series resistances. Intrinsic nodes are marked with encircled characters, whereas the extrinsic nodes are written. Parasitic series resistances R_G, R_D, R_S and R_B are between these intrinsic and extrinsic nodes. Parasitic capacitances or overlap capacitances are from the gate to the other nodes: gate to bulk, C_{GB0}, gate to drain, C_{GD0} and gate to source, C_{GS0}. The drain and source doping areas in the substrate surfaces form diodes to the bulk, which are modeled as current sources I_{bs} and I_{bd} and also diode capacitances C_{bs} and C_{bd}, which depend on the voltages over the bulk diodes. The active charge on the gate, channel and bulk is described with the active charges Q_{GB}, Q_S and Q_D, which all depend on all node voltages. R_{DS} is the unwanted resistance in parallel with the channel. Usually it is very large, being in the GΩ and TΩ region.
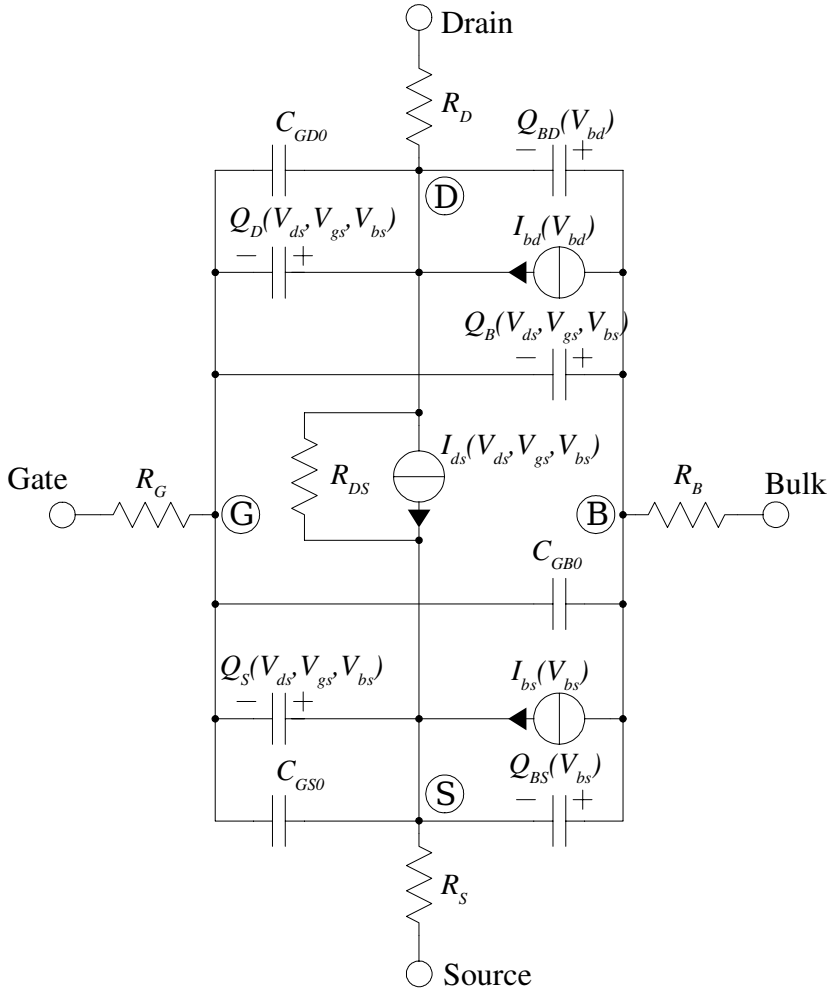
Figure 7. Equivalent circuit of a MOSFET as implemented in a circuit simulator.

When the equivalent circuit of Fig. 7 is linearized with a quasi-static approximation we get the small signal equivalent circuit of Fig. 8. In quasi-static approximation it is assumed that the channel transit time of electrons is negligibly small or zero. In the equivalent circuit of Fig. 8 only the very large R_{DS} resistance is neglected, which can be included with the drain conductance, g_{ds}. The "i" subscript stands for the intrinsic capacitance of the active charge model. All transconductances are complex or they include the transcapacitances of the active charge model. For instance, C_m is the difference between C_{dgi} and C_{gdi}. Without the transcapacitances the capacitances of the device would be

reciprocal, which is not true for a MOSFET. For example, in saturation the channel is in pinch-off and a small signal voltage at the gate modulates the inversion charge, whereas a voltage at the drain has a much smaller effect on the charge (mostly due to the channel length modulation effect). Thus, C_{dgi} is much larger than C_{gdi} which is virtually zero. C_{dsi} is a very small capacitance and it is negligible compared to the load capacitances at the drain. Usually C_{dsi} is a very small negative capacitance.
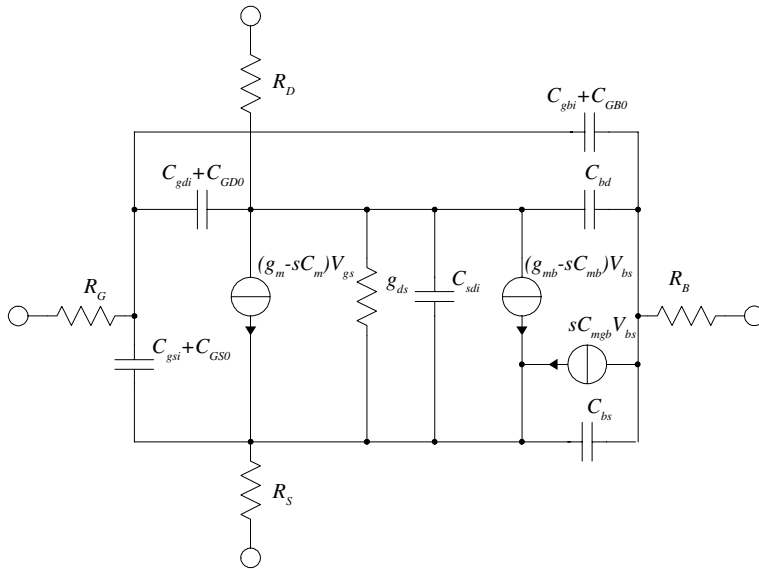
Figure 8. The small signal equivalent circuit of a MOSFET is achieved by linearizing the circuit of Fig. 7.

2.3 SOI CMOS

The SOI CMOS equivalent circuit is different to the bulk CMOS case as the substrate or bulk node is different. The equivalent circuit depends on the type of SOI that is used and some realization will be explained in the SOI mainstream model chapter later. The simplified case is very similar to the bulk CMOS equivalent circuit. The most critical differences lie in the drain current description. The kink-effect of the drain conductance (described later) depends on the floating body behavior modeled in the current equations. The active charge model also has to be determined based on the floating body effect.

Basically, SOI CMOS devices can be divided into three types based on depletion characteristics, as shown in Fig. 9. Non-fully depleted (NFD) device behavior is depicted on the left, partially depleted (PD) in the middle and fully depleted (FD) device behavior is shown on the right. The type of device characteristics are mainly determined by the SOI film thickness, and by the bias voltages in some cases. The device characteristics differ a lot between the three types of transistors. NFD devices resemble most conventional bulk CMOS devices as the depletion region never reaches the buried oxide. Backgate coupling is negligible and the differences in bulk CMOS behavior are subtle. At the other extreme the transistor body will be fully depleted of charge, resulting in the FD device on the right. FD devices have a strong backgate effect, basically with a constant body charge. FD device I–V characteristics resemble bulk CMOS devices, although the physics behind them are quite different. When the depletion region reaches the buried oxide surface, as in the middle of Fig. 9, the device has PD characteristics. There is a floating body charge that has a strong effect on the current and charge behavior of the device that is very different to the bulk CMOS case. The most notable effect resulting from this floating body is the drain conductance kink seen in Fig. 10. Although slightly similar to the MESFET kink-effect of the output impedance, the cause is in SOI technology different. In MESFETs the kink-effect is affected by the impact ionization under the gate as in the SOI case it is caused by the floating body effect.

Due to the depletion region and source drain buried oxide geometry, the bulk junction diodes are different to bulk CMOS cases. In the NFD case the characteristics are qualitatively similar to bulk CMOS devices but the diode

NFD   PD   FD

gate   gate   gate

S   D   S   D   S   D

backgate   backgate   backgate
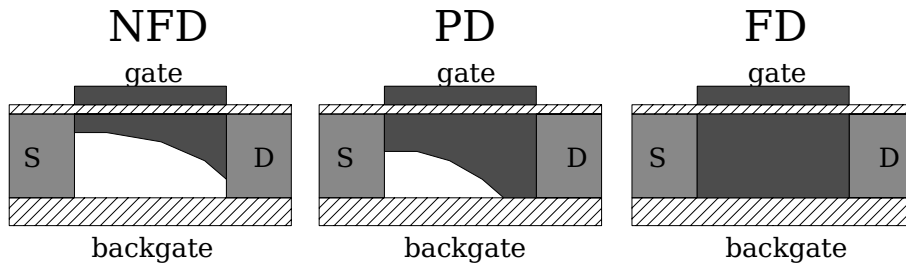
Figure 9. Cross-cut of different types of SOI CMOS devices. The dark color in the middle depicts the depletion region. The hatched areas depict gate and buried oxides. The types of devices are non-fully depleted, partially depleted and fully depleted devices.

capacitances are much smaller, consisting of sidewall components only. In FD devices the diodes are fully depleted, resulting in negligible bulk junction capacitances. In the PD case only the drain is fully depleted and the source end resembles bulk CMOS diode behavior.
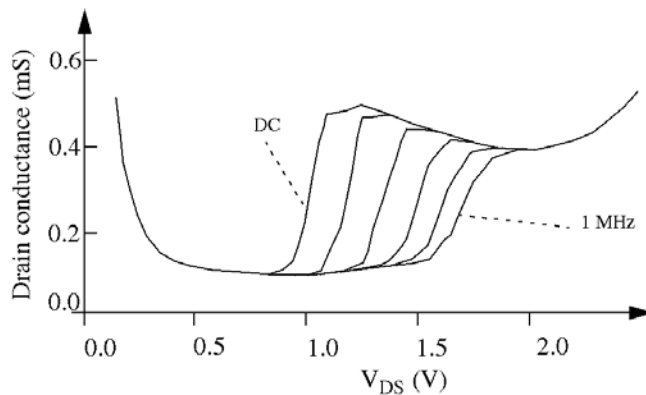
Figure 10. PD SOI kink-effect in the drain conductance and its frequency dependence. The highest kink is achieved at DC, whereas the kink flattens considerably at higher frequencies.

As in some cases there is the possibility of a change of device characteristics between NFD, PD and FD behavior depending on bias, it is necessary to model

this transitional behavior. Especially for FD devices a transition to PD behavior may occur with some backgate bias, resulting in accumulation on the buried oxide surface.

A lot of the physical effects are common for both bulk and SOI CMOS. These effects are, for instance, the short channel effect, polysilicon depletion, velocity saturation, drain-induced barrier lowering in the subthreshold, the narrow width effect and mobility degradation, as well as source and drain resistance.

2.4 High-Frequency Small Signal Behavior of MOSFETs

To get an understanding of the high-frequency properties of a MOSFET and to analyze them the equivalent circuit should be simplified a lot. Quasi-static behavior of the channel charge is assumed where the channel transit time of electrons is negligibly small. The transcapacitances are also neglected.

In the case of input-related parameter analysis, the circuit of Fig. 7 can be simplified to that of Fig. 11. The bulk or substrate node has been neglected as well as the g_{mbs}, C_{gb} and R_B components. Although the circuit is quite simple, it still contains three nodes, resulting in a very complex input impedance equation. This circuit can be simplified even further by using local series-series feedback [75] by absorbing R_S into C_{gs}, g_m and g_{ds} [76], as is done in Fig. 12. The source and load impedances have also been included, as has the input signal source. When analyzing the input impedance, the voltage source and source impedance must be replaced by an input current source, which is not presented here. The modified circuit elements of Fig. 12 are:

$$C_{gs}' = \frac{C_{gs}}{1 + g_m R_S} \tag{6}$$

$$g_{ds}' = \frac{g_{ds}}{1 + g_m R_S} \tag{7}$$

and

$$g_m' = \frac{g_m}{1 + g_m R_S} \tag{8}$$
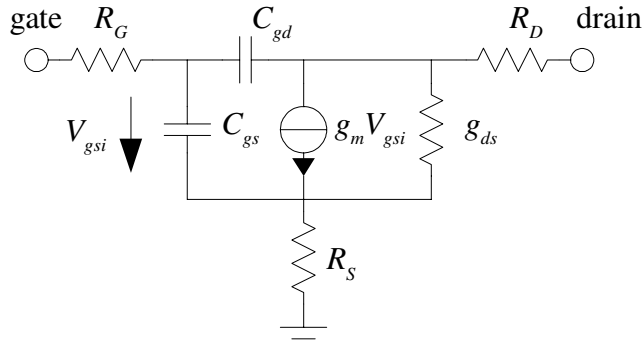
Figure 11. Simplified small signal equivalent circuit of a MOSFET.
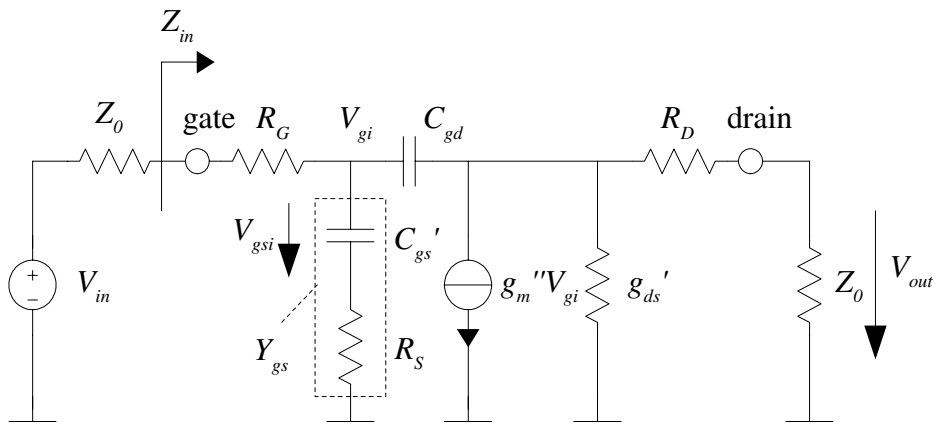
Figure 12. Simplifying circuit of Fig. 11 further using local series-series feedback.

The transistor can be considered as a two-node circuit if the C_{gs}–R_S series connection is handled as a Y_{gs} admittance and the g_m' is replaced with a modified transconductance g_m'' taking the V_{gs} voltage division into account over the effective C_{gs} capacitance. Thus we get the control voltage of the voltage-

dependent current source to be the intrinsic gate node voltage V_{gi} to keep the device intrinsic node count to only two. The modified transconductance is

$$g_m{''}=\frac{g_m{'}}{1+sR_SC_{gs}{'}} \tag{9}$$

The current $g_m{'}V_{gsi}$ in Fig. 12 is simply replaced by $g_m{''}V_{gi}$.

The input impedance Z_{in} can be calculated when the voltage source V_{in} is changed to a current source ($V_{in}/(Z_0 + R_G)$) having in parallel an impedance of $Z_0 + R_G$. Writing current node matrices and solving for V_{gi} results in Z_{in} of

$$Z_{in}(s)=R_G+\cfrac{1}{Y_{gs}+sC_{gd}+\cfrac{sC_{gd}(g_m{''}-sC_{gd})}{Y_L+sC_{gd}}} \tag{10}$$

here the load admittance and the Y_{gs} are defined as

$$Y_L=g_{ds}{'}+\frac{1}{R_D+Z_0} \tag{11}$$

and

$$Y_{gs}(s)=\frac{sC_{gs}{'}}{1+sC_{gs}{'}R_S} \tag{12}$$

It is interesting to note that according to Eq. (10), Z_{in} has quite a clear transconductance dependence although a quasi-static channel approximation has been used. This phenomenon is related to the Miller-capacitance effect present in basic amplifier configurations.

From Z_{in}, the S parameter S_{11} can be defined to be

$$S_{11}(s)=\frac{Z_{in}(s)-Z_0}{Z_{in}(s)+Z_0} \tag{13}$$

The S_{11} geometry dependence of a MOSFET is demonstrated in the Smith Chart of Fig. 13 with four different device sizes doubling in width each step. The frequency is swept from 300 MHz up to 30 GHz. The real part can be seen to decrease by increasing the device width. The smallest mos1 device in Fig. 13 is

a 2 x 20 µm x 0.18 µm MOSFET, not very practical for many RF purposes. The largest transistor has 16 parallel devices with the same finger geometry. By increasing the number of fingers the input resistance decreases, whereas the input capacitance increases. The bias dependence of S_{11} is less dramatic and is mainly caused by the input capacitance bias dependence.
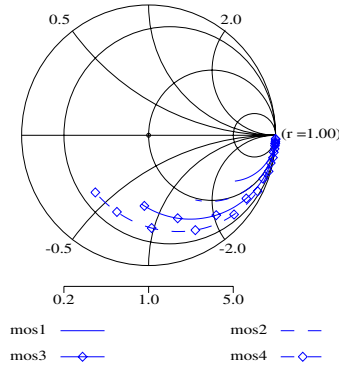
Figure 13. Input reflection geometry dependence of MOSFETs in saturation region with the same finger layout but with a different number of parallel fingers. The frequency sweep is from 300 MHz up to 30 GHz;. mos1, mos2, mos3 and mos4 have 2, 4, 8 and 16 parallel devices respectively.

Studying Equations (12) and (13) it can be seen that in the cut-off region the MOSFET input impedance is simply the series connection of R_G and input capacitance C_{in}, which is

$$C_{in} = C_{gs}' + C_{gd} \tag{14}$$

This is the result of zero transconductance and a very high drain conductance value. A simple S_{11} approximation in the cut-off region of operation can be drawn as

$$S_{11}(j\omega) \approx \frac{R_G + \dfrac{1}{j\omega C_{in}} - Z_0}{R_G + \dfrac{1}{j\omega C_{in}} + Z_0} \tag{15}$$

At low frequencies S_{11} is thus practically unity decreasing with increasing frequency. The geometry dependence of equation (15) very much resembles that of Fig. 13.

Transistor gain or S_{21} can be calculated as twice the voltage gain of the circuit in Fig. 12

$$S_{21}(s) = 2 \frac{V_{out}(s)}{V_{in}} \tag{16}$$

The voltage gain can be calculated similarly to Z_{in} but solving for V_{out} instead of V_{gi} of the current node matrices. Thus we get

$$S_{21}(s) = \frac{-2Z_0(g_m'' - sC_{gd})}{[(\frac{1}{Z_0+R_G} + Y_{gs} + sC_{gd})(Y_L + sC_{gd}) + sC_{gd}(g_m'' - sC_{gd})]} \tag{17}$$
$$\cdot \frac{1}{(Z_0+R_D)(Z_0+R_G)}$$

S_{21} is mostly affected by the transconductance. It can also be seen that the transfer function inverts the input as there is a negative sign. At low frequencies the transconductance dependence is especially apparent, which can be seen by the S_{21} DC value when s is put to zero in Eq. (17), resulting in

$$S_{21,DC} = \frac{-2Z_0 g_m'}{(Z_0+R_D)(g_{ds}' + \frac{1}{Z_0+R_D})} \tag{18}$$

For a large RF device the DC value is easier to understand qualitatively

$$S_{21,DC,RFdev} \approx \frac{-2g_m'}{g_{ds}' + \frac{1}{Z_0}} \tag{19}$$

Here it has been approximated for a large RF device $Z_0 \gg R_D$ or R_S. With a large RF device the parasitic series resistances are typically only a few ohms. At DC, S_{21} depends only on g_{ds}, g_m and the impedance level, Z_0, of the system.

If the device is in the cut-off region of operation, S_{21} depends mostly on the feed forward capacitance, C_{gd}. This effect is seen if zero transconductance and drain conductance values are put into Eq. 17.

$$S_{21,cut-off}(s) = \frac{2Z_0}{(Z_0+R_D)(Z_0+R_G)}$$

$$\cdot \frac{sC_{gd}}{(\frac{1}{Z_0+R_G}+Y_{gs}+sC_{gd})(\frac{1}{Z_0+R_D}+sC_{gd})-s^2 C_{gd}} \qquad (20)$$

As expected, there is no negative sign as the coupling is passive through the feed forward capacitance in the device. This can also be seen in Fig. 14, where S_{21} bias dependence is shown for a 20 x 10 μm x 0.18 μm MOSFET. The low-frequency value of the S_{21} magnitude is quite well predicted, with (19) in the saturation region of operation, whereas in cut-off it is dominated by the feedforward capacitance C_{gd}. In the saturation region of operation the phase is shifted 180 degrees at lower frequencies, whereas there is no phase shift in cut-off region.
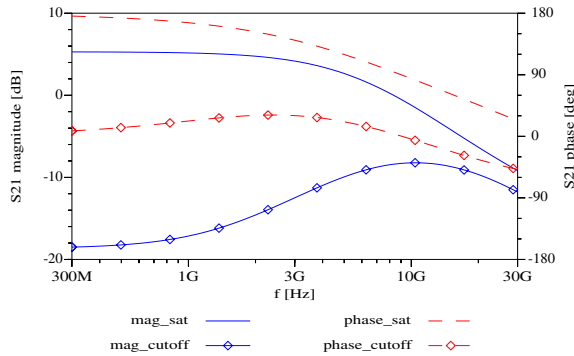
Figure 14. S_{21} bias dependence of a 20 x 10 μm x 0.18 μm MOSFET. The saturation and cut-off region magnitude and phase curves show the clear bias effect.

To calculate the backward gain, output impedance, S_{12} and S_{22} it is better to include the effect of drain source capacitance in the circuit, as shown in Fig. 15. This capacitance is actually mostly due to drain bulk capacitance caused by the diode bulk-junction capacitance. As with the case of the input, the circuit can be

simplified by using local feedback theory, enabling the removal of one node. As in the case of the input, the analysis becomes easier if the voltage source and its series impedance Z_0+R_D at the output is replaced by its respective current source with a parallel impedance of the same magnitude when calculating S_{12}.
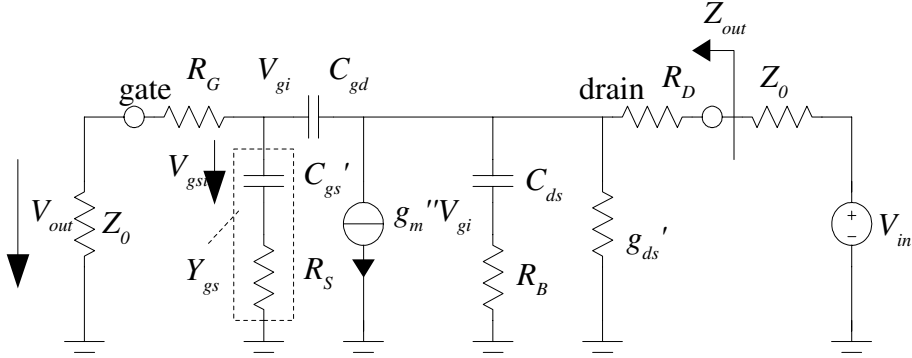
Figure 15. Simplified MOSFET equivalent circuit for calculation of Z_{out}, S_{22} and S_{12}.

When calculating Z_{out} or S_{22} the voltage source and the Z_0 impedance must be replaced with a current source. The output impedance can be calculated with current node equations as

$$Z_{out}(s)=R_D+\cfrac{1}{Y_{ds}+sC_{gd}\left(1+\cfrac{g_m{}''-sC_{gd}}{g_{in}+Y_{gs}+sC_{gd}}\right)} \tag{21}$$

Here the input admittance, g_{in}, is simply the series connection of Z_0 and R_G

$$g_{in}=\frac{1}{Z_0+R_G} \tag{22}$$

Y_{ds} is the admittance between drain and source. In the simple case of Fig. 15 it is determined as

$$Y_{ds}(s)=g_{ds}{}'+\cfrac{1}{\cfrac{1}{sC_{ds}}+R_B} \tag{23}$$

Y_{gs} is the series connection of C_{gs}' and source resistance, R_S, as calculated by Eq. (12). Again, the reflection coefficient S_{22} can be determined from the port impedance as

$$S_{22}(s) = \frac{Z_{out}(s) - Z_0}{Z_{out}(s) + Z_0} \tag{24}$$

The geometry and bias dependence of S_{22} magnitude is shown in Fig. 16 a) and b) respectively. The geometry dependence is shown with four different-sized devices in the saturation region of operation. In Fig. 16 a) the smallest device is a 2 x 10 µm x 0.18 µm transistor with the largest real part at the lower frequencies. As the number of parallel devices is increased, the low-frequency real part decreases. The bias dependence in Fig. 16 b) is qualitatively similar in that the largest effect is on the low-frequency real part. As the device is in the linear region of operation the real part is very low, but as the drain bias is increased above the saturation voltage the real part increases, moving the whole S_{22} curve to the right. The device is a 16 x 10 µm x 0.18 µm device in Fig. 16 b).
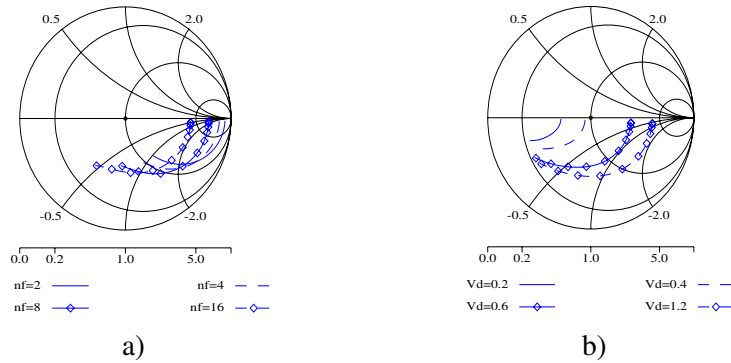
Figure 16. a) Geometry and b) bias dependence of MOSFET output reflection between 300 MHz and 30 GHz.

A quick look at Eq. (21) does not reveal a lot of output impedance behavior and it is better to study it in slightly special situations, as at low and high-frequency cases and in different regions of operation. As with the S_{21} case, the output impedance, Z_{out}, or S_{22} also has a DC value or a non-zero value at $f = 0$ *Hz*. If we define *s* to be zero in Eq. (21), we get the Z_{out} DC value:

$$Z_{out,DC} = R_D + \frac{1}{g_{ds}'} = R_D + \frac{1 + g_m' R_S}{g_{ds}} \qquad (25)$$

Now we can see that at DC or low frequencies the output impedance, Z_{out}, is real and depends on parasitic series resistances, transconductance and drain conductance.

In the cut-off state, when transconductance and drain conductance are zero, the output impedance is

$$Z_{out,cut-off} = R_D + \frac{1}{Y_{ds} + sC_{gd}\left(1 - \frac{sC_{gd}}{g_{in} + Y_{gs} + sC_{gd}}\right)} \qquad (26)$$

The behavior is in an off-state similar to the input impedance behavior, at least at low frequencies where the output impedance resembles an RC series connection.

In general, the output impedance equation (21) can be approximated at low frequencies as

$$Z_{out}(f = f_{low}) \approx \frac{1}{g_{ds} + \dfrac{1}{\dfrac{1}{sC_{ds}} + R_B}} \qquad (27)$$

This approximation is especially valid for large and typical RF device layouts. At higher frequencies the output impedance can be approximated as

$$Z_{out}(f = f_{high}) \approx \frac{1}{g_{ds} + \dfrac{1}{R_B} + sC_{gd}} \qquad (28)$$

Equations (27) and (28) represent the equivalent circuits of Fig. 17 a) and b), respectively.
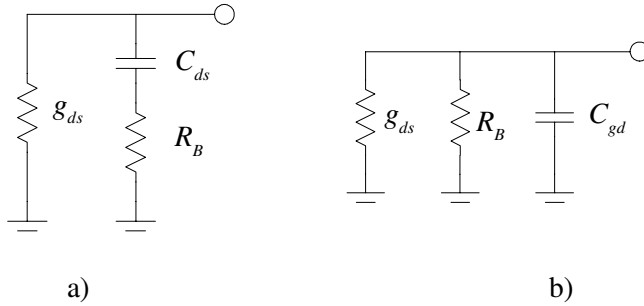
Figure 17. Approximation of the output impedance at a) low frequencies (27) and b) high frequencies (28).

Equations (27) and (28) form two circles that approximate the more accurate equation (21). In a special case the circles are very clear, as in Fig. 18.
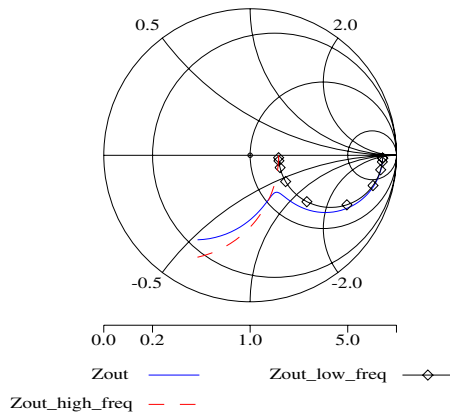
Figure 18 Output impedance approximations at low and high frequencies.

Analogous with the S_{21} case, the backward gain S_{12} can be determined by calculating the backward voltage gain from the output to the extrinsic gate node by Eq. (16). We get the backward gain or inverted isolation from output to input as

41

$$S_{12}(s) = \frac{sC_{gd}}{(g_{in}+Y_{gs}+sC_{gd})(g_{ds}'+Y_{ds}+Y_S+sC_{gd})+sC_{gd}(g_m''-sC_{gd})} \qquad (29)$$
$$\cdot \frac{2Z_0}{(Z_0+R_D)(Z_0+R_G)}$$

Again g_{in} is defined with Eq. (22), Y_{gs} is the gate-source admittance defined earlier with Eq. (23) and Y_s is the drain node source admittance for voltage gain calculation and is simply the series connection of R_D and the source impedance.

$$Y_S = \frac{1}{Z_0+R_D} \qquad (30)$$

S_{12} in Eq. (29) is mostly affected by the C_{gd} feedback capacitance value, and the signal is coupled without a phase shift of negative sign in front of Eq. (29). At lower frequencies this C_{gd} effect is more apparent and Eq. (29) can be simplified more:

$$S_{12}(s) \approx \frac{1}{1+\dfrac{1}{2Z_0\,sC_{gd}}} \qquad (31)$$

This approximation is more valid for a device in the saturation region of operation below frequencies of 1 to 5 GHz, depending on the device geometry. The larger the device the lower the frequency where Eq. (31) holds. The bias dependence could be refined by including the effect of g_{ds} and g_m. What is apparent from Eq. (31) is that S_{12} can be determined solely from C_{gd} and the impedance level, Z_0. S_{21} behavior is similar in the cut-off region.

Again, for a large RF device with multiple fingers Eq. (29) can be simplified even more to gain a better insight into S_{12} dependence. It can be approximated that parasitic series resistances at the gate, drain and source are zero. We get Eq. (29) as

$$S_{12,\,RFdev}(s) = \frac{sC_{gd}}{(\dfrac{1}{Z_0}+sC_{in})(g_{ds}+Y_{ds}+\dfrac{1}{Z_0}+sC_{gd})+sC_{gd}(g_m''-sC_{gd})} \qquad (32)$$

In cut-off the S_{12} of the large RF device is even simpler

$$S_{12,RFdev,cut-off}(s) = \frac{sC_{gd}}{(\frac{1}{Z_0}+sC_{in})(\frac{1}{Z_0}+Y_{ds}+sC_{gd})-s^2C_{gd}^2} \qquad (33)$$

The geometry dependence of S_{12} magnitude and phase is shown in Fig. 19 a) and b). The device is in the saturation region of operation. In Fig. 19 a) it can be seen that the smallest device (2 x 10 μm x 0.18 μm) has the largest backward gain, increasing linearly with frequency. This is a result of signal feedthrough through the gate to drain capacitance. Every time the device size or the number of parallel fingers is doubled, the linear part of the S_{12} magnitude jumps upward about 6 dB. The S_{21} phase in Fig. 19 b) decreases more steeply at higher frequencies , but qualitatively the behavior is not altered.
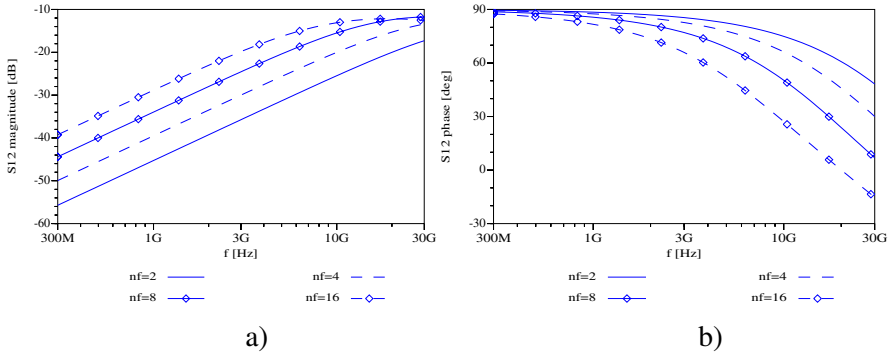
Figure 19. Geometry dependence of S_{12} a) magnitude and b) phase when the device is at the same operation point (saturation).

These calculations assumed reciprocal C_{gd} and C_{dg} capacitances which are not accurately true [4], [53]. This assumption can be fixed in the equations by changing C_{gd} to C_{dg} in S_{21}, S_{22} and Z_{out} equations.

2.5 Substrate Resistance Network Modeling

A lot of emphasis was put on substrate resistance network modeling when it was discovered that a single resistor is not sufficient [25]. According to measurements it seemed that the output impedance real part for a MOSFET in the off-state is quite constant with increasing frequency, but the models were

not able to describe such behavior. The intuitive approach of Fig. 20 a) is not correct as the output impedance real part would be 50 Ω. This approach neglects the large drain-bulk junction diode capacitance which short circuits the drain to bulk at high frequencies, as shown in Fig. 20 b). With the intrinsic bulk tied to ground this approach would lead to a 0 Ω real part at high frequencies, which cannot be correct either. A better model would include the substrate resistor, as in Fig. 20 c), from the intrinsic bulk to the grounded substrate connection. At very high frequencies this network is also inaccurate due to the source-bulk junction diode capacitance, which short circuits the R_B, as is the case in Fig. 20 d). More complicated substrate resistance networks have been studied than a simple resistor [25]–[27], [30], [31], [33], [41], [45], [48], [87], [91], [93]–[95].
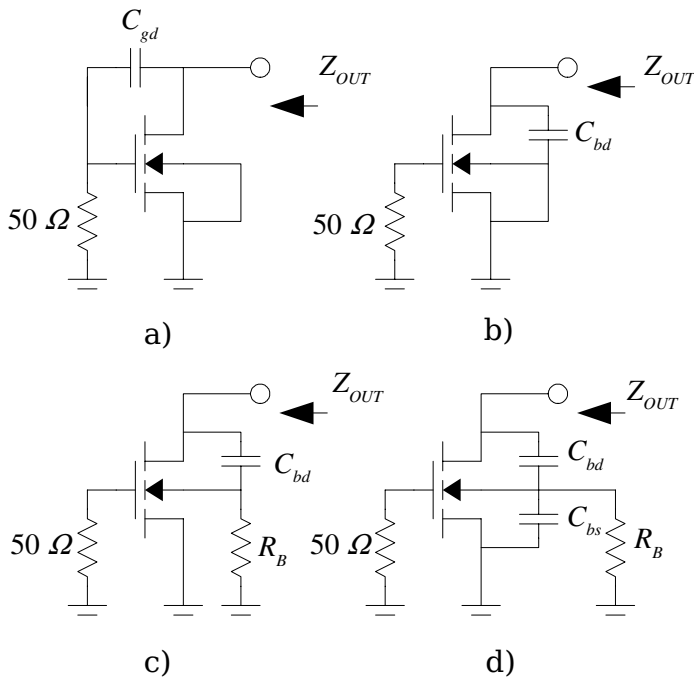
Figure 20. Substrate resistance model has an substantial effect on the output impedance accuracy.

Later it was discussed that the output impedance real part is not constant with increasing frequency [89], [38] as in hetero-junction bipolar transistors [49]. It

was found that S_{22} has a notch at very high frequencies. This was also considered in the S_{22} calculation in the previous subchapter, 2.4, where the notch was analyzed as being caused by the gate effect where the reference impedance at the input is in parallel with C_{gs} capacitance. Later studies showed that this notch may not only be the cause of the gate effect. It can also possibly be caused by the bulk effect [38]. The need for a more complicated substrate resistance model may not be required any more if the model usage is below 15 GHz.

2.6 Current State-of-the-Art Models

Practically, most models use the same equivalent circuit to describe the MOSFET and the main difference is in how accurately the drain current and the active charge model describe the device characteristics. Most models use a similar equivalent circuit to Fig. 7 as their basis, but many of the mainstream models, like BSIM3, MOS Model 9 and BSIM4, have the alternative of describing the parasitic series resistances internally by absorbing them into the drain current equation. This has the advantage of reducing the amount of circuit nodes, but reduces the high-frequency accuracy. Absorbing the R_S and R_D effects results in simpler high-frequency behavior, as can be seen by comparing the equivalent circuits with and without the parasitic series resistances. In many cases the R_G and R_B components are not present in circuit simulators' model implementations, as is usually the case with the newer models (BSIM3, MOS Model 9 and EKV). Those parasitics are usually included in the models supplied by the vendor.

2.6.1 BSIM3

BSIM3 has been the most-used mainstream MOSFET model since the end of the 90s. There are three main versions of it, but basically only the third version has been used, in other words BSIM3v3. This last version has a number of subversions as well. BSIM3 is attempting to have a model with a physical basis and still maintain the mathematical fitness of the previous model generations. BSIM3 is said to be a third-generation model [64]. The number of model parameters is very large; for instance in the APLAC circuit simulator implementation there are 168 parameters without the binning option being taken into account. The model description includes many small geometry effects, like drain-induced barrier lowering (DIBL), charge sharing, substrate current-

induced body effect and narrow channel effects. They affect mainly the threshold voltage and drain conductance descriptions, and the model has been used successfully on very deep submicron devices. The drain conductance model is also affected by channel modulation at lower drain voltages. Interestingly, the drain conductance model of BSIM3 is closely related to the lambda model used in the older Berkeley MOS models, like Level 1 and 2, for the saturation region. A small geometry effect like polysilicon gate depletion is also included by applying an effective gate voltage. As mentioned in the previous chapter, the parasitic series resistances are included in the drain current model to reduce the number of nodes required. The current is described as

$$I_{DS,lin} = \frac{V_{DS}}{R_{chan} + R_{DS}} \qquad (34)$$

Where R_{chan} is the active channel resistance and R_{DS} is due to the parasitic series resistance.

The active charge model is related to the commonly used approach [65] developed by Yang, Epler and Chatterjee. In contrast to the older models, the active charge model has parameters of its own to calculate the effective channel width and length. This slightly increases the possibility of extracting or fitting the model more easily compared to the situation where the bias-dependent charges are set along with the DC fit. Zero-bias capacitances from the gate to other nodes are included along with a bias-dependent fringe capacitance, increasing the fitting possibilities even further. For very-high-frequency simulations a non-quasi-static (NQS) model is even available. This is performed by an Elmore resistance put in series with gate-to-source and gate-to-drain capacitances.

2.6.2 MOS Model 9

MOS Model 9 is an industry-based model developed at Philips Research laboratories [66]. The model is somewhat simpler than BSIM3 and has some similarities in its approach. The current and charge descriptions are made continuous over transition regions by applying smoothing functions. In contrast to BSIM3, there are geometry dependences only in some parameters. The drain current model takes into account small geometry effects like channel length

modulation, DIBL and weak avalanche current, which has a similar effect on conductance to the SCBE effect in BSIM3. The drain and source parasitic series resistances are included in the current description by introducing drain voltage dependence into the mobility. The number of parameters is 128 in the Aplac implementation, including the geometry-dependent parameters. The charge model is quite simple, with a conventional quasi-static Yang, Epler and Chatterjee-like approach for the active part. The parasitic zero-bias capacitance model has only one parameter, COL, that describes the overlap capacitance for both the drain and source. The gate-to-bulk capacitance has been neglected.

2.6.3 MOS Model 11

MOS Model 11 [67] is the successor to MOS Model 9 from Philips Semiconductors, with many improvements in RF and nonlinearity behavior. The parasitic series resistances of the gate, drain, source and bulk are included in MOS Model 11. The substrate network approach is rather complicated, resembling that of BSIM4 in the next subchapter. The substrate network is similar to the PSP model approach in the following chapters. A non-quasi-static description has also been added using channel segmentation in order to work both in small and large signal simulations. The approach is to divide the channel length into a series connection of many channels, with the gates and bulks connected in parallel. More emphasis has been placed on geometry scaling to describe small geometry effects below the 100 nm regime, including with the possibility to bin the model. There are a lot of parameters; in the Aplac implementation there are slightly over 200 when the Level 11010 version is used!

2.6.4 BSIM4

The quite new BSIM4 bases on BSIM3 but has a lot of improvements and modifications. The number of model parameters according to BSIM4 documentation [68] is enormous: well over 300 (323 in Aplac)! 375 additional binning parameters are included. Thus almost 700 parameters need to be defined for a scalable BSIM4 model. Much efforts have to be put on parameter characterization which is a slow process. The intention of the model is to reach accurately to even smaller dimension than BSIM3: the sub 100 nm channel length regime. Most notable improvements from the RF perspective is the gate

resistance model options as well as the substrate resistance network as shown in Fig. 21. R_G resistance previously taken account by the extrinsic vendor models, for instance in BSIM3, has been absorbed into the core model description. An intrinsic input resistance, R_{ii}, intended to better account for the NQS effects is included. Also the older Elmore resistor approach of BSIM3 is used along with a new transient charge-deficit NQS-model. The substrate resistance network is quite complex and the resistance values are geometry independent. The parasitic source and drain series resistances can be described as extrinsic with a constant part and a bias dependent part as in Fig. 21. The constant part is due to diffusion resistance. In general many features previously modeled by the extrinsic vendor models have been absorbed into the core model. In addition to the RF equivalent circuit improvements also a lot of efforts have been put to develop multifinger device geometry dependences.
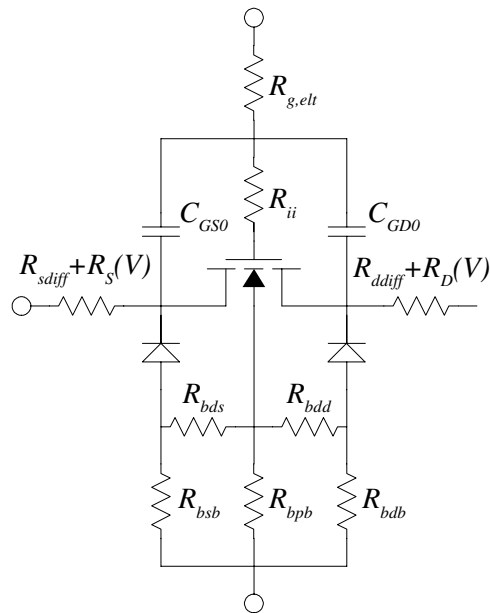

Figure 21. Parasitics of the equivalent circuit of BSIM4.

The drain conductance model is quite similar to BSIM3 but has some modifications to account for even smaller geometry effects. Also tunneling from gate to other nodes is considered for oxide thicknesses below 3 nm where the effect is notable.

The active charge model is similar to BSIM3v3. Also the zero-bias and fringe capacitance model is similar to the BSIM3v3 model except for the possible use of inner intrinsic gate resistance R_{ii} in Fig. 21. The zero-bias capacitances from gate to source and drain are connected from the outer side of R_{ii} instead of being in parallel with the intrinsic capacitances.

2.6.5 BSIM3SOI

Another derivative of BSIM3 is the SOI version of it: BSIM3SOI [69]. Again, the complicated BSIM3 model is made even more complex, but this time in order to account for the different SOI effects. These include the floating body effect, which has been made dynamic depending on the device geometry and operating voltages. This model can dynamically and continuously describe the transitions from FD to PD behavior affecting both the current and active charge models. The basic IV model is modified from BSIM3v3.1 but the charge model represents a lot of new formulation. Basically, the active charge model is related to the Yang, Epler and Chatterjee model, but more charge components in the substrate have been taken into account. The description of depletion charge is different to BSIM3. As in BSIM3, the charge model has a separate effective channel dimension. The zero-bias and fringe capacitances from the gate are similar to the basic BSIM3 model. However, the back-gate underneath the insulator or buried oxide requires additional fringe capacitances, as in Fig. 22 and Fig. 23. The physical sources of SOI MOSFET parasitic capacitances are shown in Fig. 22 and the resulting equivalent circuit in Fig. 23. There are notable capacitances formed between the source and the back-gate, as well as between the drain and the back-gate. These are the fringing side-wall components C_{essw} and C_{edsw}, as well as the extrinsic bottom area capacitances C_{esb} and C_{edb}, which are voltage dependent. The capacitance between the floating body and back-gate has also been taken into account.
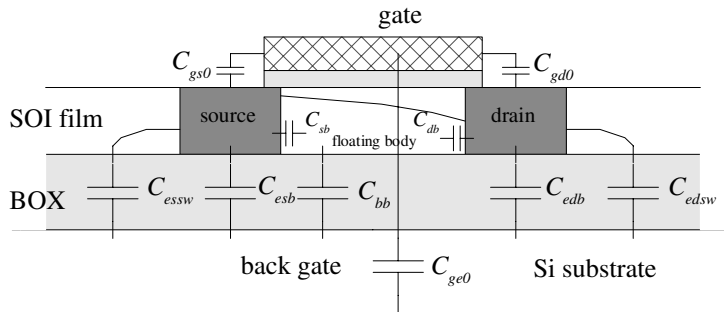
Figure 22. Sources of SOI MOS parasitic capacitances.
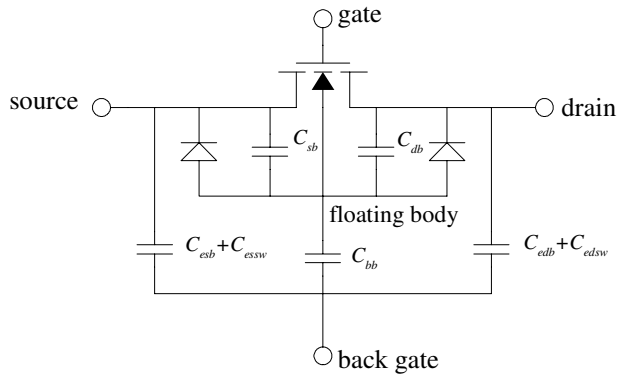
Figure 23. BSIM3SOI components between the floating body, source, drain and back gate.

BSIM3SOI is a rather heavy simulation model with its large number of parameters; there are over 200! This complexity makes BSIM3SOI a slow model in circuit simulation, being more than 10 times slower than the regular BSIM3 model. The large number of parameters makes parameter extraction a slow process which may take longer than a process run. This may lead to vendors using simpler models like BSIM3 for SOI purposes, trading off accuracy for speed.

2.6.6 EKV

The EKV model [70] developed at the Swiss Federal Institute of Technology (EPFL) represents a different approach compared to the other mainstream MOS models. Instead of having the source as the reference node, EKV takes a truly symmetrical approach by having the substrate node as the reference. The current is the sum of the forward and reverse direction currents and thus does not require the decision of which node is the drain and source. This has to be done in all older mainstream models. Model scaling has been validated even on 65 nm technology, and the model should include all pertinent effects down to 45 nm. Similar to the other models, EKV also has smoothing functions to ensure the continuity of derivatives of currents and charges. The model is rather light compared to the newer Berkeley-based BSIM models; there are just 69 parameters to describe scalable device behavior. It is mainly intended for low power and low current use. As with BSIM3, the core EKV model equivalent circuit does not include parasitic gate or bulk resistances. Parasitic series resistances at drain and source are included and they can optionally be described as internal or external. Series resistances can also be bias dependent. The charge model is not very different to the conventional intrinsic charge descriptions, except for having a symmetrical approach and that the zero-bias capacitances from gate to drain, source and bulk are included. These are bias dependent fringe capacitances as in BSIM3 and its derivatives. There is an NQS-model description in EKV with an approach via channel segmentation similar to MOS Model 11.

2.6.7 PSP-model

PSP [71] is the newest model, which has been approved by the Compact Modeling Council as the next standard model. It is has been developed by merging the best features of SP [72] and MOS Model 11. Both models are based on the charge calculated from the surface potential, in contrast to the previously so-typical approach of calculating the inversion charge. The current and charge equations are continuous and higher than just to first order, enabling more realistic distortion analysis. The approach of dividing the model into global and local levels is similar to MOS Model 11, where the local level describes the behavior of one device of a specific geometry. The global model with hundreds of parameters can be converted into the local model with only about 35

51

parameters. A lot of deep submicron device characteristics have been described in the PSP equations and reports suggest that is able to model transistor behavior down to 65 nm lengths and up to frequencies well above 50 GHz. One reference presents accurate behavior down to 40 nm channel lengths [73].

The bulk resistance network implementation of the PSP is presented in Fig. 24. The approach is somewhat similar to BSIM4, but with less elements. The bulk resistance values are bias independent and need to be known for each device geometry.
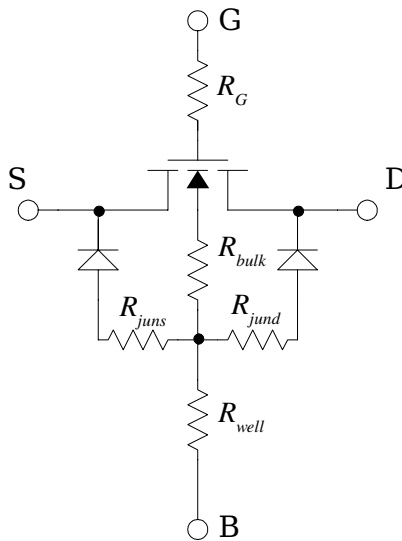
Figure 24. Bulk and gate resistance network of the PSP NQS model.

An NQS model is also included in PSP, where the channel has been partitioned into many sections of equal length and an approximation of the one-dimensional current continuity equation is applied [84].

2.6.8 Typical Vendor Modeling Approach

As the parasitic gate and bulk resistances are not available in most of the core models, semiconductor process vendors have taken the approach of building

their own model around an existing mainstream model. Typically, BSIM3 is taken as the basis and parasitic inductances, resistances and capacitances are added to better model high-frequency behavior. A typical example is shown in Fig. 25. It is common to define very carefully the allowed MOSFET geometries where the model is valid. Typically, only the extrinsic gate parasitic capacitances and the gate resistor are geometry dependent, whereas the drain-source capacitance and the parasitic inductances are constant for a certain range of device geometries.
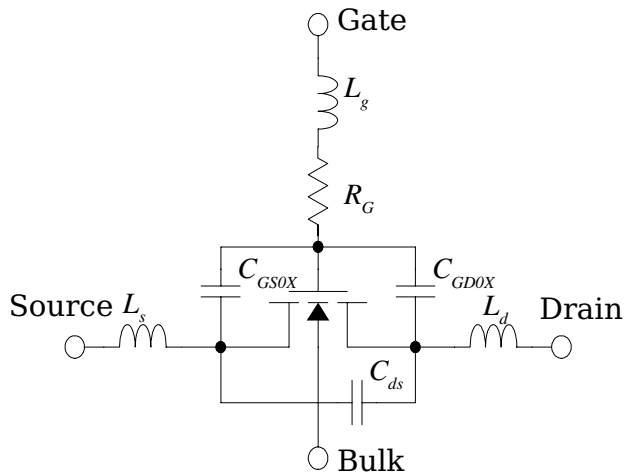
Figure 25. Typical vendor approach for RF modeling. The BSIM3 model is used as a core along with extrinsic parasitic circuitry.

3. General Methods Used in This Thesis

3.1 Device Characterization Procedure

Before any DC or AC simulations can be made with a MOSFET model the parameters defining the technology at hand have to be known. The procedure to obtain parameters for the models is called extraction and is quite time consuming. Many of the methods used in this thesis are explained thoroughly in [77]. First, accurate DC measurements have to be performed to characterize the drain current behavior from weak to strong inversion and the saturation region of operation. This also sets the active charge model, after which the characterization of AC parameters can be performed. Special tailored RF characterization structures have to be processed along with conventional RF devices.

3.1.1 DC Measurements and Extraction

DC measurements were performed using an Agilent 4156A semiconductor parameter analyzer, which was controlled by a PC running either an APLAC circuit simulator [78] or the HP Vee measurement program [79]. Devices were measured straight on-wafer with probe heads on a Cascade RF probe station. Drain current, conductance and transconductance were measured both in drain voltage and gate voltage sweeps. Thus, the six curves measured were:

$$I_{ds} - V_{ds} \qquad I_{ds} - V_{gs}$$

$$g_{ds} - V_{ds} \qquad g_{ds} - V_{gs}$$

$$g_m - V_{ds} \qquad g_m - V_{gs}$$

Due to the derivative problems described in Chapter 2.1, both conductances had to be measured to enable accurate parameter extraction. In the drain voltage sweep the voltages typically varied from 0 V up to 5 V. With lower drain voltages below 1.2 or 1.5 V the step was kept short, 0.1 V, to improve the linear region accuracy of parameter extraction, whereas the higher voltages swept with 0.2 or 0.5 V steps. A typical V_{ds} measurement sweep is shown in Fig. 26 for a long and wide 20 μm x 0.5 μm device. The drain conductance was determined by measuring two additional current points with a higher and lower offset in

drain voltage than the basic current points. Typically, the offset voltages were ±0.04 V. A similar offset to the gate voltage was used to get two additional current points for transconductance calculation. In both cases the derivatives were calculated from the three current points using either linear regression or a polynomial fit.
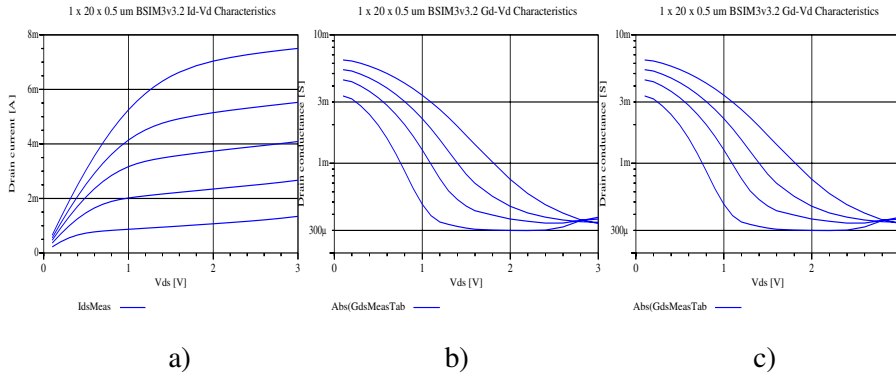
Figure 26. Example of the measured a) drain current, b) transconductance and c) drain conductance in a V$_{ds}$ sweep for a Peregrine 20 μm x 0.5 μm NMOS device.

A large set of devices had to be measured to attain a scalable MOSFET model. A typical set of device geometries required in model extraction routines is in Fig. 27, having a total of 10 devices. First the long and wide 20 μm x 5 μm device is used for extracting the basic parameters of most models, or BSIM3 in this study. These parameters include, for instance, the threshold voltage and mobility and their back gate bias dependence. In order to extract the geometry dependent parameters, a lot of devices are required to have a constant width or length, while varying their length and width, respectively. In Fig. 27 the constant value for both cases is 20 μm, as with the basic long and wide device in the corner. These extractions were made with direct extraction routines for which BSIM3 has more than 20 steps [64] which, after general purpose optimizations, were used to refine the model accuracy and its scalability. The data of all DC extraction devices were used simultaneously as goals for parameter fitting. The extraction process is iterative in such a way that one has to keep changing between parameters affecting the I_{ds}–V_{ds} curves and parameters affecting the I_{ds}–V_{gs} curves. Care must be taken not to fall into numerical minima which are typical of modern complex MOS models.
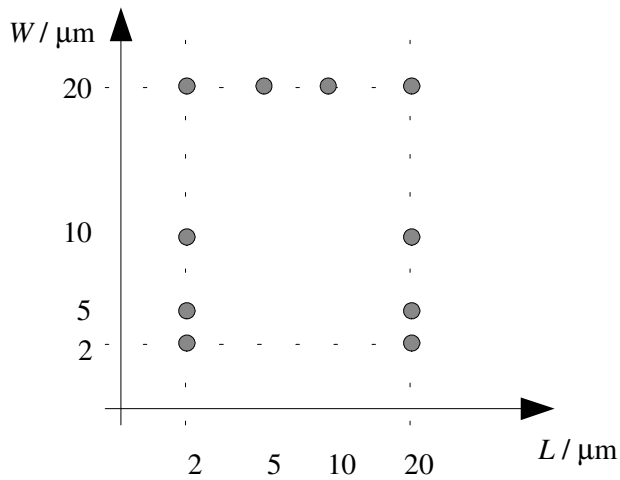
Figure 27. The DC extraction set devices used in Peregrine 0.5 μm UTSi CMOS process.

The measurement accuracy of the Agilent semiconductor parameter analyzer is 1 fA according to the documentation, but the practical measurement accuracy seemed to be in the area of 100 fA. This was the level of current noise in narrow and long channel devices when in deep subthreshold operation. Although short integration times were used in the measurements of larger currents, the measurement uncertainty seems negligible and the largest error in DC extraction is due to the process variation between devices. The measurement conditions are dominated by the systematic errors that have to be taken into account in choosing the data of measured devices.

3.1.2 AC Measurements and Extraction

AC measurements consist of *S* parameter measurements of NMOS devices as two ports straight on-wafer (in Fig. 28), again using a Cascade RF probe station with ground-signal-ground (GSG) probes with a 150 μm pitch and an Agilent 8510B or a Rohde & Schwarz ZVM network analyzer. The frequency was swept from 45 or 200 MHz up to 20 GHz logarithmically, with 101 or 201 points. The device gate is connected to the input port (on the left in Fig. 28) and the drain is connected to the output (on the right in Fig. 28), whereas the source and bulk nodes were grounded. Thus, no bulk bias effects were studied in the AC behavior of the devices. The analyzer input power level to the gate was kept

very low so as not to cause any nonlinear effects to occur in the MOSFET. Usually -40 dBm was used but -30 dBm also had to be applied in the Rohde & Schwarz analyzer due to difficulties in calibration. The DC bias voltages were connected through bias-T units to coaxial cables leading to the RF probes. A conventional short-open-load-through (SOLT) calibration method [80] was applied using a Cascade GSG standard substrate.
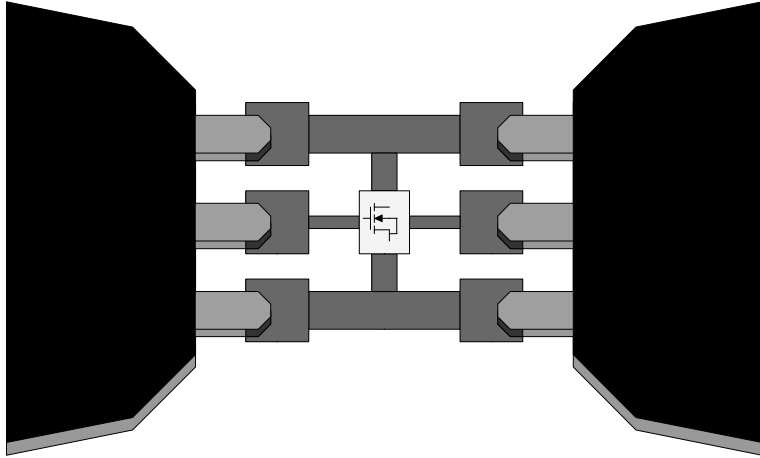
Figure 28. On-wafer measurement of an RF transistor layout with RF ground-signal-ground probes at both ports.

A mere SOLT calibration is not enough to extract the intrinsic device parameters of a MOSFET. The reference plane is shifted to the probe tips and all of the parasitics of the metalizations and pads will be included in the measurement. The procedure of removing the pad and metalization parasitics and moving the reference plane to the edge of the device under test (DUT) itself is called de-embedding [103]. In this study a conventional open and short de-embedding technique was used where an additional open and short layout have to be measured. The open layout in Fig. 29 a) is a structure similar to the one including the DUT itself, but the DUT is missing. In the short structure all the metalizations have been shorted in the middle, as in Fig. 29 b). A two-step correction method is then used, in which the *S* parameters are first converted into *Y* parameters and then the actual transistor *Y* parameters are obtained. An equivalent circuit diagram used for the two-step correction method is shown in Fig. 30. The Y_P parameters in Fig. 30 represent the parallel parasitics, whereas the Z_L parameters represent the series parasitics. It can be shown [103] with

simple mathematics that the actual *Y* parameters of the transistor can be obtained from

$$Y_{Trans} = \cfrac{1}{\cfrac{1}{Y_{DUT} - Y_{open}} - \cfrac{1}{Y_{short} - Y_{open}}} \qquad (35)$$

Y_{DUT} is the measured uncorrected *Y* parameters of the transistor of interest. The open structure measurement directly gives the Y_{open} parameters, determined solely by the Y_P parameters of Fig. 30. Measuring the short structure gives all the Y_P and Z_L component values of Fig. 30, thus requiring a subtraction with Y_{open} to achieve the Z_L component values. With Eq. (35) the parallel parasitics are removed from the uncorrected Y_{DUT} parameters, after which the series parasitics, the Z_L values, are subtracted after a change from admittance to impedance.

a)   b)
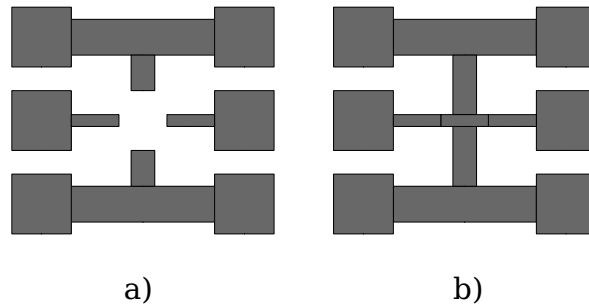
Figure 29. a) open and b) short measurement structure to characterize the pad and metalization parasitics in the two-step de-embedding procedure.

The RF device layouts were large compared to the DC measurement layouts in order to achieve lower impedance levels that are more accurate to measure. Typical RF device layouts were used with many parallel fingers. The advantage of dividing devices into many parallel fingers of the same size is that the gate resistance is divided by the square of the number of parallel devices when comparing a MOSFET with an equal total width. Still, the input capacitance is increased just slightly due to the metalization parasitics.
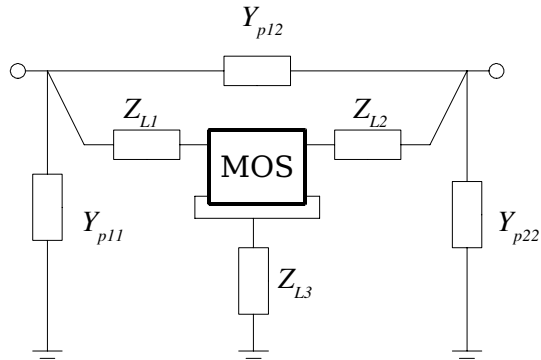
Figure 30. Equivalent circuit of metalization and pad parasitics used in de-embedding.

With the extraction of the DC model most of the AC properties have been set. The zero-bias capacitances, input resistance and bulk-junction diode parameters have to be extracted. Special measurement structures are usually made to maximize the different zero-bias capacitances and bulk-junction diode capacitances for more accurate measurement. These parameters can also be fitted by collecting a large set of *S* parameters from different-sized devices at many bias points. One conventional approach is to use the specially designed measurement structures for the direct extraction of zero-bias and bulk-junction capacitances, after which the *S* parameter data is fitted using general purpose optimization. This was also my approach.

In this study the AC extraction devices were different for different technologies. The technologies used in this thesis are VTT 0.6 µm BiCMOS, AMS 0.35 CMOS, Peregrine 0.5 µm UTSi SOI CMOS and Honeywell 0.35 µm PD SOI CMOS. In the September 1997 run of the VTT 0.6 µm BiCMOS process the AC extraction devices were:

4 x 12.5 µm x 0.8 µm	4 x 12.5 µm x 0.6 µm
4 x 25 µm x 0.8 µm	4 x 25 µm x 0.6 µm
1 x 50 µm x 0.6 µm	1 x 50 µm x 0.8 µm
1 x 50 µm x 1.0 µm	1 x 100 µm x 0.6 µm
1 x 100 µm x 0.8 µm	1 x 100 µm x 1.0 µm

In the January 1999 run of the VTT 0.6 CMOS technology the AC extraction devices were:

8 x 6.25 µm x 0.6 µm (W_T = 50 µm)   16 x 3.12 µm x 1.0 µm (W_T = 50 µm)

8 x 12.5 µm x 0.8 µm (W_T = 50 µm)   16 x 3.12 µm x 1.2 µm (W_T = 50 µm)

16 x 6.25 µm x 0.6 µm (W_T = 100 µm)   16 x 6.25 µm x 1.0 µm (W_T = 100 µm)

4 x 25 µm x 0.8 µm (W_T = 100 µm)   8 x 12.5 µm x 1.2 µm (W_T = 100 µm)

8 x 6.25 µm x 0.6 µm (W_T = 200 µm)   4 x 50 µm x 1.0 µm (W_T = 200 µm)

16 x 12.5 µm x 0.8 µm (W_T = 200 µm)   8 x 25 µm x 1.2 µm (W_T = 200 µm)

In the Peregrine 0.5 µm UTSi SOI CMOS the AC extraction devices were:

8 x 5 µm x 0.5 µm (W_T = 40 µm)   3 x 18 µm x 0.5 µm (W_T = 54 µm)

4 x 10 µm x 0.5 µm (W_T = 40 µm)   20 x 6.8 µm x 0.5 µm (W_T = 136 µm)

4 x 10 µm x 1.0 µm (W_T = 40 µm)   40 x 6.8 µm x 0.5 µm (W_T = 272 µm)

8 x 10 µm x 1.0 µm (W_T = 80 µm)   60 x 6.8 µm x 0.5 µm (W_T = 408 µm)

4 x 20 µm x 0.5 µm (W_T = 80 µm)   161 x 6.8 µm x 0.5 µm (W_T = 1094 µm)

In the Honeywell 0.35 µm PD SOI CMOS the AC extraction devices were:

6 x 10 µm x 0.30 µm   48 x 10 µm x 0.35 µm

6 x 10 µm x 0.35 µm   36 x 10 µm x 0.50 µm

6 x 10 µm x 0.40 µm   36 x 10 µm x 0.80 µm

4 x 10 µm x 0.35 µm   24 x 10 µm x 0.35 µm

4 x 10 µm x 0.50 µm   24 x 10 µm x 0.50 µm

2 x 10 µm x 0.35 µm   12 x 10 µm x 0.35 µm

AC extraction accuracy is much more sensitive to measurement uncertainties than DC characterization. The error sources are the measurement accuracy of the network analyzer, calibration errors and de-embedding errors. Most errors result from the inaccurate calibration of the network analyzer. These calibration errors are seen as systematic errors in AC extraction and thus cannot be removed. The random error of the measurements could be minimized by averaging a lot of measurement data from the same devices. However, this has not been done due to the already large amount of fitting data. Instead, in this work the random errors are averaged out by using a very large number of different devices at different bias points as AC extraction goals simultaneously. The number of Touchstone™ S2P-files in one extraction was 30 to 60, with 101 to 201 points. The number of frequency data points in one extraction was thus from about 3 000 to 12 000 points. When these numbers are multiplied by four, one gets the number of S parameter points in the extraction.

General-purpose optimization leads to some numerical problems when determining the weight of each point. All S parameters must have the same weights, and it is easily defined that different S parameter magnitude weights are equal. However, ensuring that, for instance, the S_{11} magnitude has a similar weight between different Touchstone™ files is not that straightforward. Without any weighing, the priority between different devices or bias points would be unequal. The same weight is given for all frequency points in one S parameter curve of one Touchstone™ file. The weight is calculated, for instance in the case of S_{11} magnitude, by finding the maximum and minimum values of every S_{11} magnitude of different Touchstone™ files. These maximum and minimum values are stored in an array, after which the highest value of these is stored, for instance in *Smax*. The weights of each S_{11} in different bias points are calculated by

$$\frac{Smax}{Max(Min(|S_{11}|), Max(|S_{11}|))} \tag{36}$$

In this way the maximum S_{11} magnitude of all will get a weight of one, while the others will have a weight above one. *Smax* is divided by the larger absolute value of the minimum or maximum value.

There are many formats in which the *S* parameter data can be used in extraction. I used a real-imaginary format in parameter extraction and a magnitude(dB)-phase format in *S* parameter fitting-error calculations. The latter is more commonly used and intuitive for comparison, while the former format gives the imaginary part more even weight in optimization than the magnitude-phase format.

3.2 Extraction of Small Signal Equivalent Circuit Component Values

The MOSFET AC model can also be characterized by studying the small signal component values of a MOSFET shown in the simplified equivalent circuit of Fig. 11. Many articles on different extraction methods have been published [87] –[97].

3.2.1 Simple Approach

The simplest approach is straightforward, assuming one has the even more simplified equivalent circuit of Fig. 31.
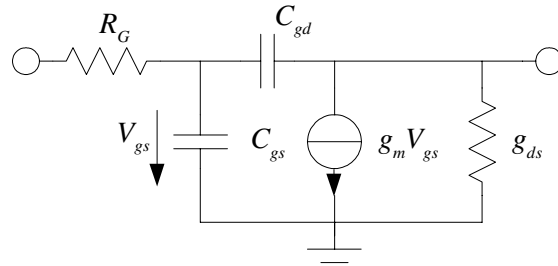
Figure 31 Simple MOS equivalent circuit for AC parameter extraction.

The *Y* parameters of the simple circuit of Fig. 31 can easily be determined as.

$$Y_{11}(\omega) = \frac{1}{R_G + \dfrac{1}{j\omega C_{in}}} \approx \omega^2 C_{in}^2 R_G + j\omega C_{in} \tag{37}$$

$$Y_{12}(\omega) = \frac{-j\omega C_{gd}}{1 + j\omega C_{in} R_G} \approx -j\omega C_{gd} - \omega^2 R_G C_{in} C_{gd} \tag{38}$$

$$Y_{21}(\omega) = \frac{g_m - j\omega C_{gd}}{1 + j\omega C_{in} R_G} \approx g_m - \omega^2 C_{gd} C_{in} R_G - j\omega(C_{gd} + g_m C_{in} R_G) \tag{39}$$

$$Y_{22}(\omega) = g_{ds} - \omega^2 R_G{}^2 C_{in} C_{gd} g_m \\ + j\omega[C_{gd}(1 + g_m R_G) + C_{ds} - \omega^2 C_{gd}{}^2 R_G{}^2 C_{in}] \tag{40}$$

In these equation, C_{in} is the sum of C_{gd} and C_{gs}. From Eq. (37) the input resistance can easily be extracted as the real part of the inverted Y_{11}:

$$R_{in} = \Re\{\frac{1}{Y_{11}}\} \tag{41}$$

Another approach is to calculate like

$$R_{in} = \frac{\Re\{Y_{11}\}}{\Im\{Y_{11}\}^2} \tag{42}$$

This approach has been reported in [81].

Input capacitance is in turn calculated from the imaginary part of the inverted Y_{11}:

$$C_{in} = -\frac{1}{\omega\Im\{\frac{1}{Y_{11}}\}} \tag{43}$$

Or again it can be calculated differently by

$$C_{in} = \frac{\Im\{Y_{11}\}}{\omega} \tag{44}$$

The gate-to-drain capacitance is calculated from Y_{12} of Eq. (38) as

$$C_{gd} = -\frac{\Im\{Y_{12}\}}{\omega} \tag{45}$$

The gate-to-source capacitance is then the subtraction of Eq. (44) and Eq. (45).

$$C_{gs} = C_{in} - C_{gd} = \frac{\Im\{Y_{11}\} + \Im\{Y_{12}\}}{\omega} \tag{46}$$

The transconductance is approximately the real part of Eq. (39)

$$g_m \approx \Re\{Y_{21}\} \tag{47}$$

Eq. (47) is quite accurate at low frequencies.

The drain conductance can be estimated from the real part of Y_{22} from the low frequencies as:

$$g_{ds} \approx \Re\{Y_{22}\} \tag{48}$$

Output capacitance which does not have any real component counterpart is often useful in MOSFET high-frequency characterization. It can be determined as

$$C_{out} = \frac{\Im\{Y_{22}\}}{\omega} \tag{49}$$

The drain-source capacitance of Fig. 31 can be calculated by subtracting C_{gd} from Eq. (49).

$$C_{ds} \approx C_{out} - C_{gd} = \frac{\Im\{Y_{22}\} + \Im\{Y_{12}\}}{\omega} \tag{50}$$

It should be noted, however, that C_{ds} capacitance does not represent any real component, as an intrinsic capacitance between drain and source does not exist. At low frequencies the drain bulk junction capacitance does load the drain as the bulk real impedance is relatively large. However, parasitic capacitances usually do exist between drain and source due to the metalization layer overlapping, and the C_{ds} value can be used for estimation of those, as well as the bulk junction capacitance C_{bd}.

This simple approach for extracting the small signal parameters of MOSFETs does not take into account the non-reciprocity of the gate-to-drain capacitance [82]. If one has to ensure non-reciprocity, the Y_{21} Eq. (39) should be rewritten with C_{dg} instead of C_{gd}, and with $C_{dg} + C_{gs}$ instead of C_{in}. The feedforward

capacitance C_{dg} is usually much larger than the feedback capacitance C_{gd} and it can be extracted by

$$C_{dg} = -\frac{\Im\{Y_{21}\}}{\omega} \tag{51}$$

3.2.2 More Complicated Approaches

Usually, the input or gate resistance is extracted in the cut-off or linear regions using the following expression [90]

$$R_G = \left| \frac{\Re\{Y_{12}\}}{\Im\{Y_{12}\}\Im\{Y_{11}\}} \right| \tag{52}$$

or at any bias the expression is even more complex [92]

$$R_G = \frac{\Re\{Y_{11}\}}{\Im\{Y_{11}\}^2} - R_S\left(1 - 2\frac{\Im\{Y_{12}\}}{\Im\{Y_{11}\}}\right) - (R_S + R_D)\frac{\Im\{Y_{12}\}^2}{\Im\{Y_{11}\}^2} \tag{53}$$

This method requires the values of R_D and R_S to be known.

A "newer" approach [44] for extracting the gate resistance takes into account the input capacitance behavior of short-channel devices

$$R_G = \frac{\Re\{Y_{11}\}}{\Im\{Y_{11}\}^2 \omega C_{in}} \tag{54}$$

In this approach the input capacitance has to be extracted from the imaginary part of Y_{11}. Possibly, the author of [44] just reinvented the approach described in Eq. (42) [81].

Another interesting approach for the extraction of the small signal parameters was reported in [89]. The series resistances R_G, R_S and R_D are first de-embedded from the Z parameters before calculation of the intrinsic model parameters, like C_{gd}, C_{gs}, g_m and g_{ds}. The series resistances themselves are first extracted from the Z parameter real parts using zero-bias conditions for the MOSFET. In another publication [96], it is stated that this approach avoids the possible underestimation of extracted device capacitance by as much as 20%.

4. Effects of RF Measurement Accuracy on Parameter Extraction

The effect of *S* parameter measurement errors resulting from vector network analyzer uncertainties on RF MOSFET parameter extraction were analyzed [84] –[86]. The uncertainty effects on the MOSFET small signal equivalent circuit were studied. The lower uncertainty specification of a high-end network analyzer were used as the basis for the analysis.

In our study of Peregrine devices at RF, we found strange behavior that was very hard to explain. The input impedance real part was very high at lower frequencies, decreasing rapidly with frequency.

Although a lot of research on RF MOSFET characterization has been done [87] –[96], very little has been done on the error analysis of AC extraction. Previous work on FET small signal parameter uncertainties resulting from vector network analyzer uncertainties have concentrated on MESFET and HEMT devices [98]–[101].

In ref. [98], HEMT small signal parameter extraction accuracy is studied experimentally by comparing on-wafer measurements with microstrip measurements, paying a lot of attention to the different calibration methods used in both cases. However, the details about the model parameter uncertainty derivation are not given.

In ref. [99], condition numbers expressing the sensitivity of the computed MESFET small-signal model to *S* parameter measurement uncertainties are derived. Circuit element sensitivities have been calculated analytically using mathematical software capable of symbolic calculations. However, these equations are not presented, as each element has 64 partial derivatives resulting from the complex conversion between Cartesian and polar form *S* parameters and the conversion between *S* and *Y* parameters. Numerical results have been shown for a set of devices, but the authors do not specify whether the data is for one frequency point or an average in a frequency sweep, nor has any bias dependence been considered. It is hard to make general conclusions about the uncertainty of FET parameter extraction.

The most comprehensive study and the one most similar to my work is ref. [101], which uses sensitivity analysis in uncertainty analysis. The parameter variances of R_i, C_{gs}, R_j, C_{gd}, g_{ds}, C_{ds}, g_m and τ of the intrinsic FET small-signal model (Fig. 32) are calculated as a function of S parameter variances.

$$\sigma_x{}^2 = (\boldsymbol{K}_S^x)^2 \sigma_S{}^2 \tag{55}$$

Here, x refers to the model parameter at hand and their variance matrix is

$$\sigma_x{}^2 = [\sigma_{x1}{}^2 ... \sigma_{xn}{}^2]^T \tag{56}$$

and the S parameter variance is

$$\sigma_S{}^2 = [\sigma_{|S_{11}|}{}^2 \sigma_{\theta_{11}}{}^2 ... \sigma_{|S_{22}|}{}^2 \sigma_{\theta_{22}}{}^2]^T \tag{57}$$

The model parameter sensitivities with respect to S parameters are defined with the matrix:

$$\boldsymbol{K}_S^x = \begin{bmatrix} K_{|S_{11}|}{}^{x_1} & \cdots & K_{\theta_{22}}{}^{x_1} \\ \vdots & \ddots & \vdots \\ K_{|S_{11}|}{}^{x_n} & \cdots & K_{\theta_{22}}{}^{x_n} \end{bmatrix} \tag{58}$$

$\boldsymbol{K}^2$ in equation (55) denotes taking the square of each individual element in $\boldsymbol{K}$. θ_{kl} refers to the phase of S parameter S_{kl} in equations (57) and (58). In (57), it is assumed that the S parameter deviations are normal-distributed having a zero mean and being uncorrelated. The small signal parameters are calculated from the Y parameters, thus requiring Y to S parameter conversions to be made. For example, the relative uncertainty values of the small signal parameters are given for an OMMIC GaAs HEMT device at a single optimum extraction frequency, as well as for a weighted wide band extraction. The largest relative errors are with the R_i and R_j parameters, being 380% and 73% respectively. The uncertainty of τ is also in the range of tens of percent, but other parameters are much below 10%. The g_m and C_{gd} errors are only in the range of 2%. Although overall the study in [101] is very comprehensive, the bias or geometry dependence of the determined uncertainties has not been studied.
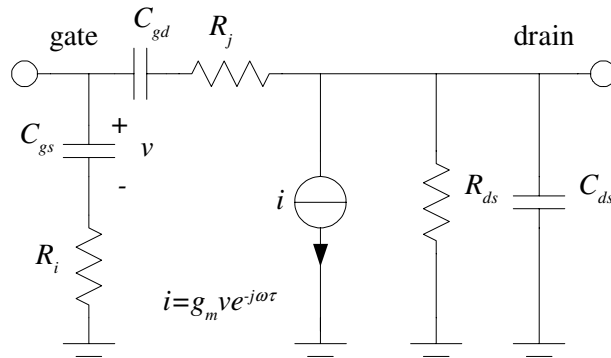
gate C_{gd} R_j drain

C_{gs} + v -

R_i

i

$i=g_m v e^{-j\omega\tau}$

R_{ds} C_{ds}

Figure 32. High frequency intrinsic FET small-signal model.

Previous results of MESFET or HEMT transistors are not directly applicable to MOS devices due to the differences in the equivalent circuit. I have studied the input and output impedance errors as a function of typical MOSFET impedances. This has not been done in previous publications as far as I am aware. This gives the designer a guideline to design devices with optimal impedance levels for accurate parameter extraction, which is not possible with the results of [98]–[101]. Transconductance and feedback capacitance uncertainties have also been analysed for the MOSFET case. Devices for small power applications (e.g. LNA and mixers) are assumed, for which the impedance levels are higher than with power devices.

4.1 Basic Approach of Uncertainty Analysis

The measurement error effect of the network analyzer uncertainties [102] is studied by calculating the total differential error of the parameters of interest. This approach assumes linear expansion in terms of the *S* paremeter uncertainties which is justified by a numerical and measured worst case example in Chapter 4.2. Uncertainty specifications of magnitude and phase are divided (in [102]) into four frequency regions, of which three are used in this study: 0.045–2, 2–8 and 8–20 GHz. No interpolation of data between the discontinuities or transition regions has been done. This results in over-optimistic results just below the transition frequencies 2 and 8 GHz (seen later, especially in Fig. 40 and 58). Only lower uncertainty specifications have been considered.

Usually, the uncertainty specifications are defined as magnitude and phase uncertainties. To calculate the total differential error the small signal parameter equations determined from S parameters have to be defined as a function of S parameter magnitudes and phases.

The relative total differential error is calculated for input resistance, R_{in}; input capacitance, C_{in}; transconductance, g_m; feedback capacitance, C_{gd}; output capacitance, C_{out}; and output resistance, R_{out}. Actually, the latter two of the small signal parameters are not directly present in a conventional MOSFET equivalent circuit but can be used for extracting the desired parameters, depending on the complexity level of the equivalent circuit. Two different approaches were taken for different small signal parameters. Input and output-related model parameter relative errors are calculated theoretically as a function of their absolute values in a simple equivalent circuit. The input impedance measurement errors are studied as a function of different R_{in} and C_{in} combinations as

$$\Delta X_{in} = f\left(R_{in}, C_{in}, \Delta S_{11m}, \Delta \theta_{11}\right) \tag{59}$$

Where X_{in} is either R_{in} or C_{in}. ΔS_{11m} and $\Delta \theta_{11}$ are the S_{11} magnitude and phase uncertainties respectively. The output impedance errors are analogously studied as

$$\Delta X_{out} = f\left(R_{out}, C_{out}, \Delta S_{22m}, \Delta \theta_{22}\right) \tag{60}$$

For the transconductance and feedback capacitance, the more traditional approach was taken by calculating error limits to a specific measurement. This required simplifications to be able to present the error equations in detail. These kinds of simplification have not been used in publications previously.

Network analyzer calibration errors were not considered and the de-embedding errors of layout parasitics [103] were assumed to be very small. In reality, inaccurate calibration is probably the largest source of error, whereas the de-embedding error is significant too. However, this study searches for theoretical minimum errors when all conditions are met for perfect measurements with a high-end commercial network analyzer.

4.2 MOSFET Input Impedance Uncertainty

$Z_{in} \longrightarrow$ R_{in} C_{in}

Calculate S parameters

$S_{11m}(R_{in}, C_{in})$ and $\theta_{11}(R_{in}, C_{in})$

Take errors ΔS_{11m} and $\Delta\theta_{11}$ into account

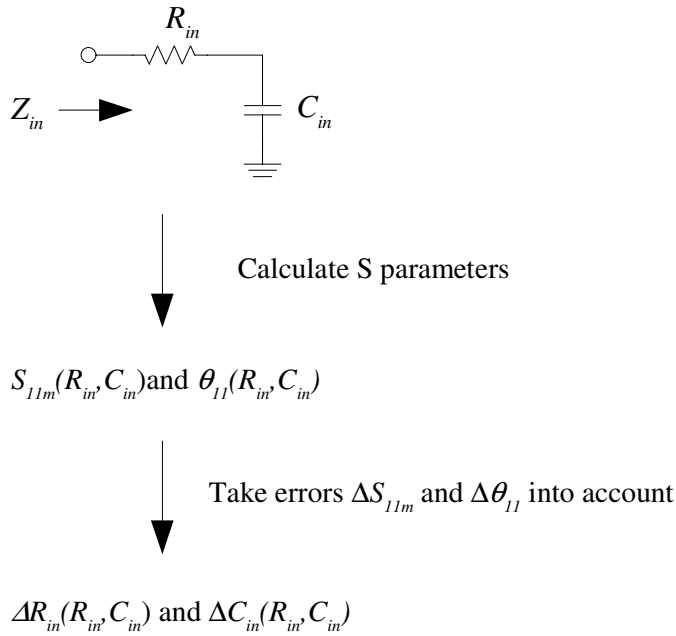$\Delta R_{in}(R_{in}, C_{in})$ and $\Delta C_{in}(R_{in}, C_{in})$

Figure 33. Flowgraph of the input impedance error analysis including the simple input impedance approximation for a MOSFET used in this study.

A graphical presentation of my input impedance sensitivity analysis principle is shown in Fig. 33. The input impedance of a MOSFET is simplified in this study as a simple series connection between a resistor and capacitor, as shown in upper part of Fig. 33. For different R_{in} and C_{in} values, the respective S_{11} parameter is calculated. After adding the magnitude and phase errors of S_{11} the new S_{11} is used to calculate the errors of R_{in} and C_{in}.

This R_{in} and C_{in} series connection is a very accurate approximation at lower frequencies (below 4 GHz) for typical small signal MOSFETs, especially in the cut-off region. At frequencies as high as 20 GHz the error is not more than 10–15 percent for typical devices. This error is caused mostly by the bulk resistance effect becoming more apparent at higher frequencies. Another cause of error in this simple model is the lack of poles caused by the parasitic series resistances.

For this (Fig. 33) port impedance, the S_{11} is easily defined. From it in turn, the impedance can be derived with respect to the S_{11} magnitude and phase as

$$Z_{in} = \frac{1 - S_{11m}{}^2 + j\,2S_{11m}\sin\theta_{11}}{1 + S_{11m}{}^2 - 2S_{11m}\cos\theta_{11}} Z_0 \tag{61}$$

where Z_0 is the reference impedance level. The input resistance seen at the gate is thus the real part of Eq. (61)

$$R_{in} = \Re\{\,Z_{in}\,\} = \frac{1 - S_{11m}{}^2}{1 + S_{11m}{}^2 - 2S_{11m}\cos\theta_{11}} Z_0 \tag{62}$$

and the input capacitance can be calculated simply by

$$C_{in} = \frac{1}{\omega\,|\Im\{\,Z_{in}\,\}|} \tag{63}$$

We calculate the errors using total differential error by taking partial derivatives of R_{in} and C_{in} as

$$\Delta R_{in} = \left|\frac{\partial R_{in}}{\partial S_{11m}}\right| \Delta S_{11m} + \left|\frac{\partial R_{in}}{\partial \theta_{11}}\right| \Delta \theta_{11} = \tag{64}$$
$$\left| 2Z_0 \frac{(2S_{11m} - S_{11m}{}^2 cos\theta_{11} - \cos\theta_{11})}{(1 + S_{11m}{}^2 - 2S_{11m}\cos\theta_{11})} \right| \Delta S_{11m}$$
$$+ \left| 2Z_0 \frac{((1 - S_{11m}{}^2) S_{11m} \sin\theta_{11})}{(1 + S_{11m}{}^2 - 2S_{11m}\cos\theta_{11})} \right| \Delta \theta_{11}$$

and

$$\Delta C_{in} = \left|\frac{\partial C_{in}}{\partial S_{11m}}\right| \Delta S_{11m} + \left|\frac{\partial C_{in}}{\partial \theta_{11}}\right| \Delta \theta_{11} = \tag{65}$$
$$\left| \frac{2S_{11m}{}^3 - S_{11m}(2\cos\theta_{11} - 1) + 2S_{11m}\cos\theta_{11} - 1}{2S_{11m} Z_0 \omega \sin\theta_{11}} \right| \Delta S_{11m}$$
$$+ \left| \frac{S_{11m}}{Z_0 \omega} - \frac{(1 + S_{11m}{}^2 - S_{11m}\cos\theta_{11})}{2S_{11m} Z_0 \omega (\sin\theta_{11})^2} \cos\theta_{11} \right| \Delta \theta_{11}$$

where $S_{11m} = |S_{11}|$, and ΔS_{11m} and $\Delta\theta_{11}$ are the absolute values of S_{11} magnitude and phase uncertainties respectively. The S_{11} magnitude uncertainty of high-end network analyzers is 0.01–0.02 units at typical MOSFET impedances, whereas the phase uncertainty is approximately 1° [102]. Plotting the relative (Fig. 34

71

and Fig. 35) error of input resistance and capacitance as a function of the MOSFET input impedance at 500 MHz, the dramatic effect of S parameter measurement uncertainties is evident. Another set of examples at 2 GHz are calculated in Fig. 36 and Fig. 37. The calculations suggests that a conventional RF MOS input resistance is easily erroneous by hundreds of percent. It is also very dependent on the measured input impedance as well as the frequency. This kind of study, where the extraction accuracy is studied as a function of input resistance and capacitance absolute values, has not been done before. The uncertainty due to the network analyzer has been studied for compound semiconductor FETs, and they suggest similar inaccuracies for intrinsic input resistance extraction as is achieved for R_{in} in this study. According to the example of reference [101], the uncertainty of the intrinsic input resistance between the gate and source is 380% at 40 GHz, when R_i is 0.4 Ω and C_{in} is about 150 fF. Calculating the respective total differential error using my method with these parameter values gives a relative error of 424%. It is in the same range as that of reference [101], although the calculated resistance errors are from somewhat different circuit topologies. These impedances can also be studied on the Smith chart as is done in Fig. 38 at 500 MHz. Curves showing the 7%, 10% and 50% extraction uncertainty of R_{in} are plotted. Considering typical MOSFET input impedances at 500 MHz, it is rather impossible to reach within the 10% "circle".

The capacitance error is much smaller, being in the range of tens of percent, decreasing when the input capacitance increases. In both cases the error is approximately proportional to ~1/f to some extent. This can also be seen at 2 GHz as the error levels are much lower due to the lower impedance level of the MOSFET input. The 2 GHz error does not decrease quite down to ¼ compared to the 500 MHz case, as the RF network analyzer measurement uncertainty is increased. C_{gs} extraction accuracy in the example case of reference [101] suggest a 3.6% relative error at 18 GHz when the C_{gs} value is 136 fF. C_{in} extraction accuracy corresponds to this value and is by my total differential error method 7.7%. Although the C_{gs} extraction uncertainty of [101] takes more S parameter uncertainties into account, it is smaller than the error achieved by my method. This should not be possible, as my method takes only S_{11} measurement uncertainties into account. Again the impedances where the C_{in} can be extracted with 2%, 5% and 10% accuracy can be plotted on the Smith chart as is done in Fig. 39 at 500 MHz.

These calculations suggest that the input capacitance and resistance extraction should not be made at too low frequencies, which is also shown by Fig. 40. R_{in} extraction should possibly be done at frequencies above 5 GHz, as the errors due to the simplified input equivalent circuit of Fig. 33 are considerably smaller compared to the errors caused by measurement uncertainty. This extraction approach has been observed in [90] and [87], possibly for practical reasons. The low frequency accuracy of extracted R_G achieved in [89] must be pure luck. In reference [101], it is found that the optimum extraction frequency of the example HEMT R_i is 40 GHz.

For C_{in} the accuracy does not seem to improve a lot further above frequencies of 500 MHz. The frequency dependence of both input capacitance and resistance is shown in Fig. 40 for a device with an input resistance of 30 Ω and an input capacitance of 500 fF.
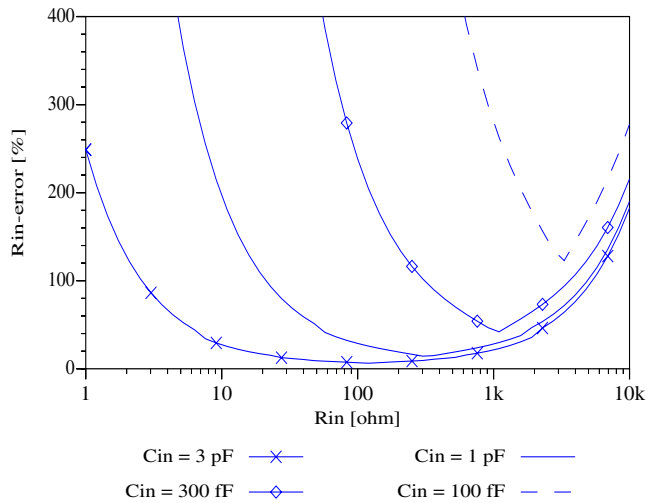

Figure 34. Relative input resistance error as a function of input impedance at 500 MHz.
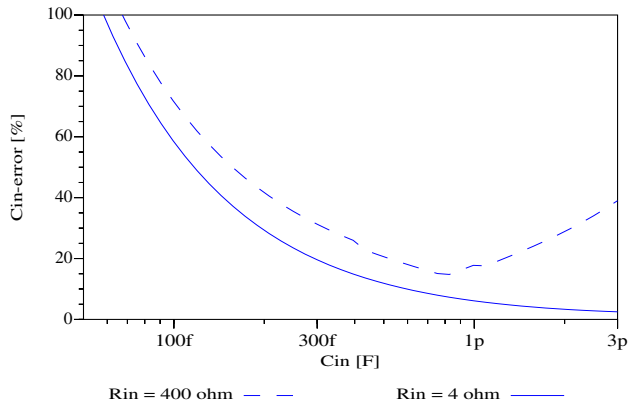
Figure 35. Relative input capacitance error as a function of input impedance at 500 MHz.
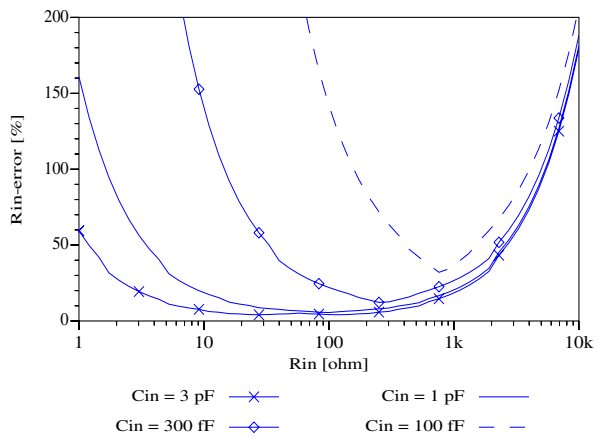

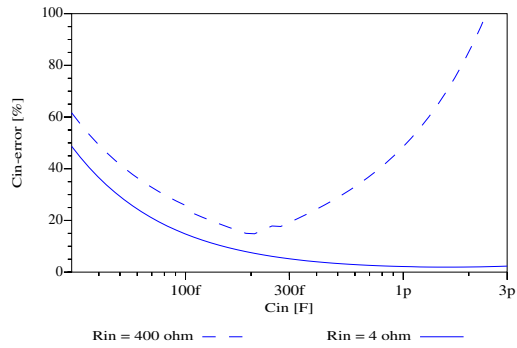*Figure 36. Relative input resistance error as a function of input impedance at 2 GHz.*

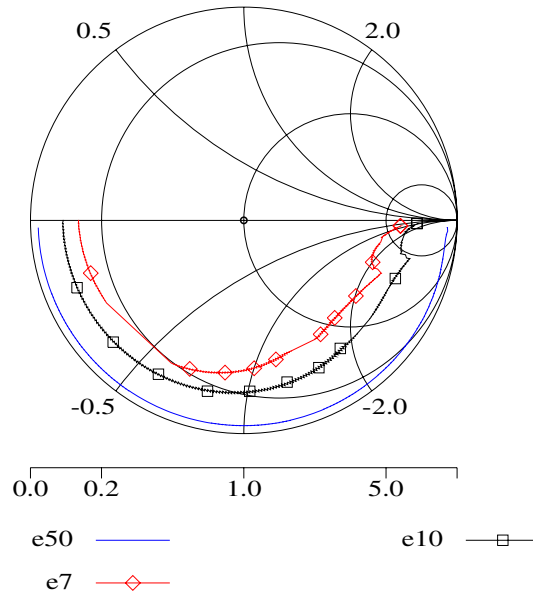*Figure 37. Relative input capacitance error as a function of input impedance at*

Figure 38. Input resistance extraction accuracy as a function of input impedance at 500 MHz. Impedance curves are shown where the input capacitance can be extracted with 7%, 10% and 50% accuracy.
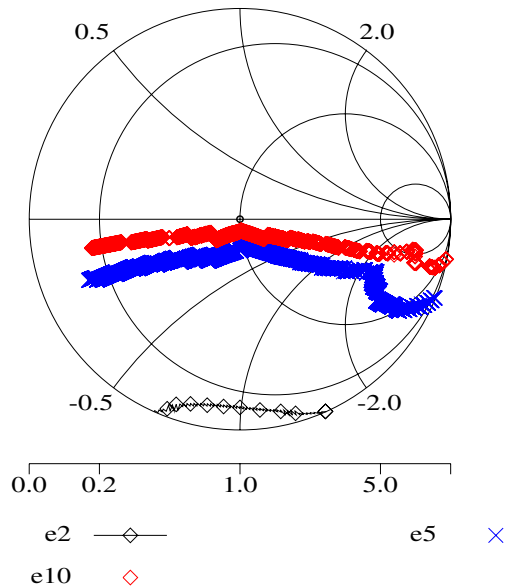
Figure 39. Input capacitance extraction accuracy as a function of input impedance at 500 MHz. Impedance curves are shown where the input capacitance can be extracted with 2%, 5% and 10% accuracy.

Although this study is an approximation, in the case of a MOSFET input impedance it is accurately applied to the series connection of a resistor and capacitor shown in Fig. 33. However, this approximation causes errors that are negligible compared to the uncertainties of the measurement errors.

An example of the measured gate resistance of a VTT 200 x 0.6 μm bulk NMOS device (divided into 16 fingers) is shown in Fig. 41 a). The device is biased in the linear region of operation. The gate resistance has been calculated both with Eq. (41) and Eq. (42), represented by Calc1 and Calc2 markers respectively in the plots. It can be seen that the variation of the extracted values is rather large
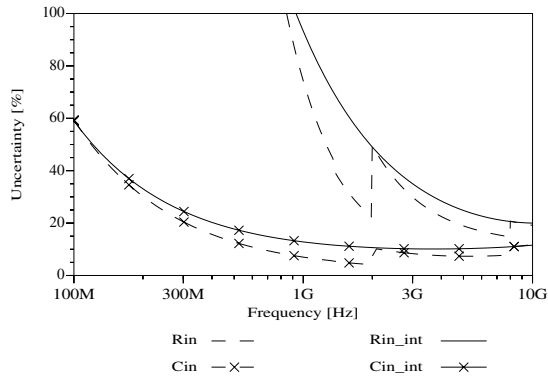
Figure 40. Relative R_{in} and C_{in} errors as a function of frequency for a MOSFET input with R_{in} 30 Ω and C_{in} 500fF. The dashed lines are calculated lines with no interpolation taken into account whereas the solid lines are interpolated error curves.

at lower frequencies (below 2 GHz). This is due to the measurement uncertainty. At higher frequencies Eq. (42) results start to deviate a lot from the other calculation from Eq. (41). At higher frequencies the R_{in} value from Eq. (42) is even doubled compared to the low frequency value. Qualitatively the behavior calculated from Eq. (41) is more correct. According to Fig. 34, the R_{in} extraction uncertainty at 500 MHz for this device with a 230 fF input capacitance is well over 400%. Yet the low frequency value differs only about 100% from the measured high frequency value, as well as the modeled value that can be assumed to be the more correct value. At 2 GHz the uncertainty should still be over 100% according to Fig. 36, but in Fig. 41 a) the 2 GHz value differs only about 10% from the high frequency value. The measurement seems to have succeeded much better than the worst-case uncertainty specifications would have suggested.

Another example in Fig. 41 b) shows the extracted gate resistance of a Peregrine 20 x 6.8 x 0.5 µm NMOS device, also in the linear region of operation, using both R_{in} equations again. The small signal model behavior is also presented in both plots. In this example both equations behave very similarly, having a large input resistance value at lower frequencies. If the high frequency R_{in} value of the device is correct, the total differential error at 500 MHz is out of scale in Fig. 34

when we know that the device capacitance is about 160 fF. Calculating the theoretical total differential error for a device with R_{in} 10 Ω and 160 fF, it is found to be over 6000%. The relative R_{in} error in Fig. 41 b) is over 1000% at 500 MHz, which is well below the theoretical 6000%. At 2 GHz, the theoretical total differential error should be from 200–300 % according to Fig. 36. At 2 GHz, the extracted R_{in} value is about 30 Ω, suggesting a 300 % error. This is again within the measurement uncertainty resulting from network analyzer. Thus, it seems that the strange R_{in} behavior of Fig. 41 b) is not a new device phenomenon but a mere measurement error.
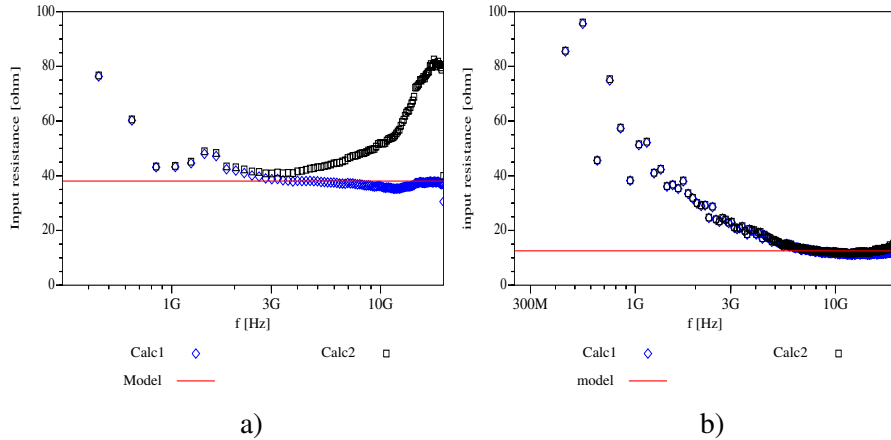

Figure 41. Measured/extracted input resistance value from a) a VTT 16 x 12.5 μm x 0.6 μm bulk NMOS device b) Peregrine 20 x 6.8 μm x 0.5 μm SOI NMOS device.

From both examples it is quite clear that a rather high frequency is required to extract the input resistance value.

To justify the approach of linear expansion in terms of *S* parameter uncertainties numerical uncertainty values are put directly in to the *S* parameters of a Peregrine 20 x 6.8 x 0.5 μm NMOS device. In Fig. 42 the input resistance has been calculated using Eq. (41). The markers show the extracted input resistance of the device while upper and lower error limits have been included by adding or subtracting *S* parameter errors from the original *S* parameters. The approach

in this plot is the traditional one of calculating the error of the measurement at hand. It seems that at 540 MHz the extracted input resistance value is exactly 100 Ω while the upper error is 264 Ω and lower error is -399 Ω. This enormous error is almost solely caused by the S_{11} magnitude uncertainty which can be understood by examining the numerical S_{11} parameters of this case. At 540 MHz S_{11} magnitude is 0.995 and if we add or subtract an uncertainty of 0.018 we see that the impact close to unity is huge. Below 2 GHz a high-end network analyzer reflection uncertainty is 0.018 [102] and adding that to 0.995 results into to a value of 1.013 which is larger than unity and results in turn into a negative input resistance.
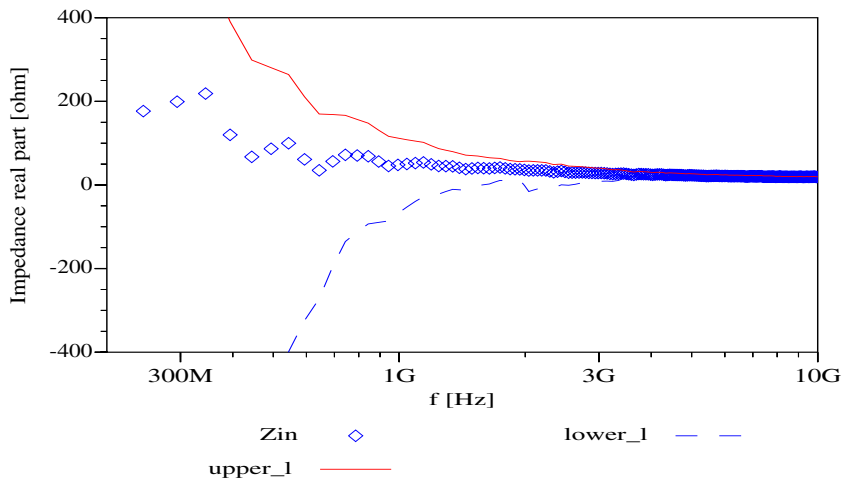

Figure 42. Measured input resistance of a Peregrine 20 x 6.8 μm x 0.5 μm NMOS device along with its upper and lower uncertainty limits. The uncertainty values have been added directly to the S parameters before the extraction of the uncertainty limits.

Experimental input capacitance values are shown in Fig. 43 a) for a VTT 4 x 12.5 μm x 0.6 μm bulk NMOS device calculated with Eq. (43) and Eq. (44). They are referred to as Calc1 and Calc2 in the plots. Up to 2 GHz the values of the different equations seem quite identical, after which the Calc2 curve of (44) decreases more rapidly. The more constant value of the Calc1 curve is qualitatively more correct. With the input resistance value of 80 Ω the error should be less than 70 % at 500 MHz, according to the theoretical total

differential error (in Fig. 35). At 2 GHz the respective uncertainty should be about 20%. There is only a slight variation at low frequencies, except for 300 MHz where the error is about 40 %. The decrease of the input capacitance value for both equations is most likely due to the large input resistance value. Possibly, the series resistances R_G, R_S and R_D should be de-embedded as described in [89], after which the input capacitance should be calculated from the de-embedded Y parameters. In [96] it is stated that this procedure may improve the extraction accuracy of capacitances by as much as 20% at high frequencies. In Fig. 43 a) the extracted C_{in} decrease is steeper than that which is typical, probably due to the very high gate resistance value of the VTT CMOS process.

Another example of an extracted input resistance is shown in Fig. 43 b) for a Peregrine 20 x 6.8 µm x 0.5 µm SOI NMOS in the linear region of operation. The extraction uncertainty with this possible 160 fF input capacitance and 10 Ω gate resistance should be below 40% at 500 MHz, according to Fig. 35. The measurement noise is largest at frequencies below 1 GHz in Fig. 43 b) and it is certainly below the 40% range. At 2 GHz the uncertainty should be very small – about 10–15% according to Fig. 37. At higher frequencies the value decreases by less than 10%, which is the approximate total differential error for that range, too. This is also suggested by the relative C_{in} error in Fig. 40, where it is plotted as a function of frequency. The error does not decrease with increasing frequency after a few GHz as with the case of input resistance. The de-embedding technique of [89] is possibly not useful for this case as the input resistance value does not have a large effect on the input capacitance uncertainty, as depicted by Fig. 35 and Fig. 37.

The study suggests that quite large RF characterization devices should be used for input parameter extraction to achieve lower impedance levels than those typical for MOSFET inputs. Thus, to achieve optimum C_{in} extraction accuracy the device should be as wide and long as possible. Optimum R_{in} extraction requires also as a wide device as possible, but with an optimized length and finger number to get the resistance higher than in typical RF MOS transistors.
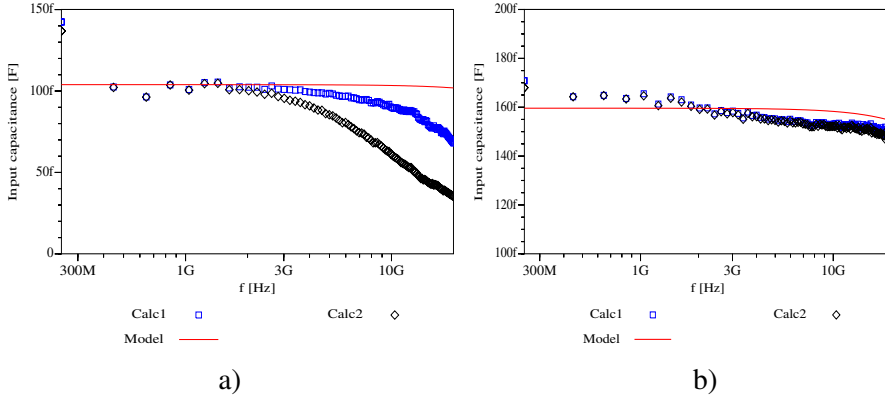
Figure 43. Extracted input capacitance values for a) VTT 4 x 25 μm x 0.6 μm bulk NMOS device and b) Peregrine 20 x 6.8 μm x 0.5 μm SOI NMOS device.

4.3 Transconductance Uncertainty

The transconductance error can be derived from the real part of Y_{21} with Eq. (47) [87]. Other definitions exist but this is the most simple one, resulting in less complicated uncertainty equations. y_{21} was calculated by making the following approximation valid for typical MOSFETs, only at lower frequencies (typically below 2–3 GHz):

$$Y_{21} = \frac{-2\,S_{21}}{Z_0((1+S_{11})(1+S_{22})-S_{12}S_{21})} \approx \frac{-2S_{21}}{Z_0(1+S_{11})(1+S_{22})} \tag{66}$$

Z_0 is the reference impedance level; usually and in this case 50Ω. This approximation works only when S_{12} is very small. For very large RF MOSFETs the approximation in Eq. (66) is valid only up to a few hundred megahertz. The transconductance extraction should be done at very low frequencies (<<0.5 GHz). Due to the approximation, the total differential error could be calculated much more simply for g_m without taking the S_{12} error into account. The transconductance error equation becomes

$$\Delta g_m = \left|\frac{\partial g_m}{\partial S_{11m}}\right|\Delta S_{11m} + \left|\frac{\partial g_m}{\partial \theta_{11}}\right|\Delta \theta_{11} + \left|\frac{\partial g_m}{\partial S_{22m}}\right|\Delta S_{22m} + \left|\frac{\partial g_m}{\partial \theta_{22}}\right|\Delta \theta_{22}$$
$$+ \left|\frac{\partial g_m}{\partial S_{21m}}\right|\Delta S_{21m} + \left|\frac{\partial g_m}{\partial \theta_{21}}\right|\Delta \theta_{21} \tag{67}$$

81

Of course, Eq. (66) has to be defined as a function of S parameter magnitudes (linear) and phases before the partial differentiation. Already, this equation is somewhat complex:

$$g_m = \Re\{ Y_{21} \} = -2S_{21m} \frac{(\cos\theta_{21a} + \sin\theta_{21b})}{Z_0(a^2 + b^2)} \tag{68}$$

Here again the subscript "m" refers to the magnitude value in S_{21m}. In Eq. (68) a and b are functions of S_{11} and S_{22} as

$$a = 1 + S_{11m}\cos\theta_{11} + S_{22m}\cos\theta_{22} + S_{11m}S_{22m}\cos(\theta_{11}+\theta_{22}) \tag{69}$$

and

$$b = S_{11m}\sin\theta_{11} + S_{22m}\sin\theta_{22} + S_{11m}S_{22m}\sin(\theta_{11}+\theta_{22}) \tag{70}$$

Thus, the approximation in Eq. (66) is needed only for practical reasons, as accurate derivation leads to answers extending over 10 pages. All error equations of my approach are presented in Appendix A.

The transconductance plot of a 20 x 6.8 μm x 0.5 μm NMOS device in the saturation region of operation is shown in Fig. 44 as an example. Transconductance and its error have been calculated from experimental data, from which the layout parasitics have been de-embedded [103]. The relative errors at low frequencies are quite low, 6–7%, suggesting that g_m extraction could be performed with decent accuracy. Another example of a 60 x 6.8 μm x 0.5 μm device is shown in Fig. 45. The error is of the same magnitude as in the previous example. A similar increase in the error at 2 GHz as in Fig. 44 is due to the higher frequency uncertainty region of the network analyzer uncertainty specifications. This discontinuity has not been interpolated. In reference [101] the transconductance definition is very similar, and in the example case the relative uncertainty was 2.2%, which is smaller than the smallest total differential error of Fig. 44 and Fig. 45. The reason for this difference is possibly the different uncertainty definition or the different device and operation point of a specific measurement. The g_m uncertainty magnitude of [101] and my work are of the same order.

The g_m error resulting from the approximation can be demonstrated for the 20 x 6.8 μm x 0.5 μm device in Fig. 46.

The bias dependence of g_m absolute and relative errors is shown in Fig. 47 for a 60 x 6.8 μm x 0.5 μm NMOS device. The *S* parameters have been measured at 21 bias points and the transconductance has been calculated at 45 MHz, where the error is lowest. V_{ds} voltage is swept from 0.1 to 2.5 V as the gate voltage is a constant 0.8 V. It can be seen that the absolute error is lowest in the linear region, whereas the relative error is highest. Going in to the saturation region above 0.5 V decreases the relative error down to 1.6% from 4.5%, which it was at $V_{ds} = 0.1$ V. The error of the absolute transconductance value increases with the drain bias from about 200 μS to 640 μS. The absolute error increase begins after 1 V, along with the transconductance increase. This is high above the threshold voltage, which is about 0.5 V. The largest error source is contributed by the S_{21} magnitude error, which is relatively large in the linear region where the S_{21} absolute value is very small. Thus, the extraction should be done above well above the saturation voltage.

gm and its error limits, 20 x 6.8 um x 0.5 um NMOS
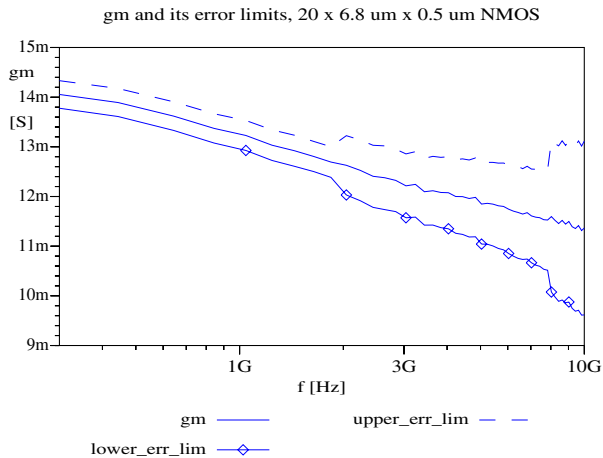
Figure 44. Transconductance of a 20 x 6.8 μm x 0.5 μm NMOS device showing the calculated upper and lower error limits at lower frequencies. The device is in the transition region between the linear and saturation regions of operation, having $V_{gs} = 1.2$ V and $V_{ds} = 1.8$ V.

gm and its error limits, 60 x 6.8 um x 0.5 um NMOS
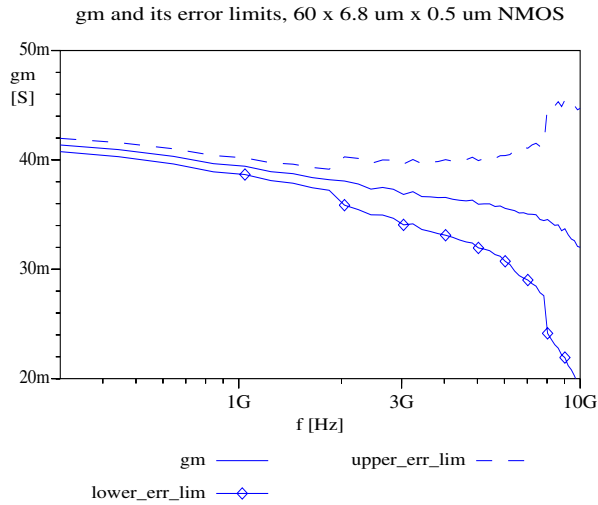
Figure 45. Transconductance of a 60 x 6.8 μm x 0.5 μm NMOS device showing also the calculated upper and lower error limits at lower frequencies. The device is in the saturation region of operation, having $V_{gs} = 0.8$ V and $V_{ds} = 2.6$ V.
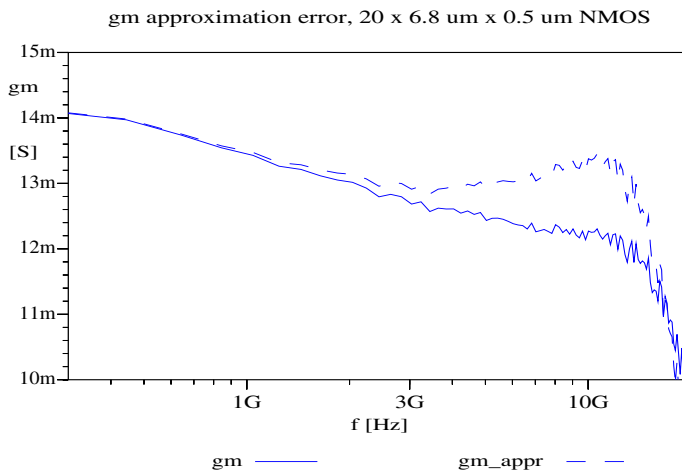
gm approximation error, 20 x 6.8 um x 0.5 um NMOS

Figure 46. Accurate and approximated (by (66)) transconductance of a 20 x 6.8 μm x 0.5 μm NMOS device. The error is apparent above 3 GHz.
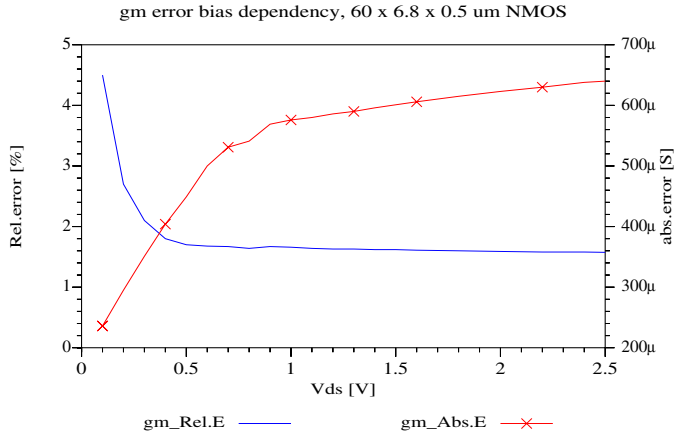
Figure 47. Bias dependence of g_m relative and absolute errors for a 60 x 6.8 μm x 0.5 μm NMOS.

Studying with different sizes of devices suggests that the g_m error is not very dependent on device size or bias point as long as the device is in the right operational region. The error is then approximately in the 6–10% range at frequencies below 500 MHz.

In previous publications no attention has been put on the bias and geometry dependence of the g_m uncertainty.

4.4 Feedback Capacitance Uncertainty

Analogous to the transconductance the feedback capacitance can be derived from the Y_{12} parameter from its most common definition [87]:

$$C_{gd} = \frac{-\Im\{ Y_{12} \}}{\omega}$$

(71)

Analogous to the transconductance case, the Y_{12} parameter can be determined from the S_{12} parameter with

$$Y_{12} = \frac{-2S_{12}}{Z_0((1+S_{11})(1+S_{22})-S_{12}S_{21})} \approx \frac{-2S_{12}}{Z_0(1+S_{11})(1+S_{22})} \qquad (72)$$

As with Y_{21}, the Y_{12} expression was simplified for practical reasons. The C_{gd} expression could more easily be described as a function of S parameter magnitudes and phases, and additionally the reduction in partial derivation was substantial. The C_{gd} equation can be defined by:

$$C_{gd} = -2S_{12m} \frac{(\sin\theta_{12a} - \cos\theta_{12b})}{Z_0\omega(a^2+b^2)} \qquad (73)$$

Here a and b are defined by equations (69) and (70).

Then we get the C_{gd} total differential error as

$$\Delta C_{gd} = \left|\frac{\partial C_{gd}}{\partial S_{11m}}\right|\Delta S_{11m} + \left|\frac{\partial C_{gd}}{\partial\theta_{11}}\right|\Delta\theta_{11} + \left|\frac{\partial C_{gd}}{\partial S_{22m}}\right|\Delta S_{22m} + \left|\frac{\partial C_{gd}}{\partial\theta_{22}}\right|\Delta\theta_{22}$$
$$+ \left|\frac{\partial C_{gd}}{\partial S_{12m}}\right|\Delta S_{12m} + \left|\frac{\partial C_{gd}}{\partial\theta_{12}}\right|\Delta\theta_{12} \qquad (74)$$

A plot of C_{gd} error is shown in Fig. 48 for a 20 x 6.8 µm x 0.5 µm of the device of Fig. 44 in the linear region, with $V_{ds} = 0$ V and $V_{gs} = 0.8$ V. In contrast to the transconductance case, the relative error does not seem to be lowest at low frequencies, but highest. The error decreases rapidly with increasing frequency. Mostly this seem to be the result of S_{12} magnitude uncertainty. At low frequencies the magnitude level is very small, causing more relative uncertainty to the S_{12} magnitude measurement, also reported in [104]. At lower frequencies below 240 MHz the error is more than 100%, decreasing up to 5–6 GHz. At 4 GHz the error seem to be approximately 10% when no interpolation has been taken into account. At higher frequencies (from 5–8 GHz) the error seem to be quite constant, but this cannot be considered an accurate result above 6 GHz in theory due to the approximation done in y_{12} Eq. (72). Practically, the result is quite accurate as at zero drain bias S_{21} is very low, not causing an error in the approximated C_{gd} Eq. (73). The optimal extraction frequency seems to be from 3–5 GHz, with the upper frequency limit determined by the approximation in (72). The reason for the decreasing error with increasing frequency is most probably due to the decreasing impedance of the C_{gd} capacitance, thus becoming an easier challenge for the measurement equipment.
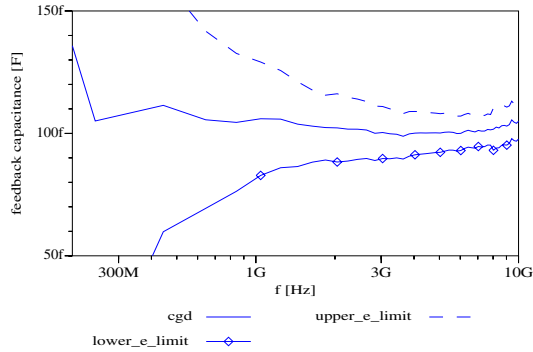
Figure 48. Extracted C_{gd} along with its upper and lower error limits for a 20 x 6.8 µm x 0.5 µm NMOS device.

For the larger 60 x 6.8 µm x 0.5 µm NMOS device in Fig. 49, the error seems to be of the same order as with the smaller 20 x 6.8µm x 0.5µm device of Fig. 48, which has a 23% relative error at 1 GHz. The bias point is chosen again in the deep linear region, with $V_{ds} = 0.1$ V and $V_{gs} = 0.8$ V to minimize the error caused by the approximation in (72). The relative error of the larger device is about 24% at 1 GHz, whereas the difference between approximated and accurate C_{gd} is only 5%. The behavior is qualitatively similar to the error behavior of the smaller device, and the minimum error is reached at a lower frequency. The error is above 100% at lower frequencies (below 200 MHz), decreasing rapidly with increasing frequency. Again, frequencies above 2 GHz cannot be considered due to the approximation done in Eq. (72). Unfortunately, this approximation is very geometry and bias-dependent, as S_{21} and S_{12} rapidly increase in value when the device size is increased and when the bias rises. The optimum C_{gd} extraction frequency is thus not found in a reliable way. Considering the low frequency errors, it seems that C_{gd} extraction is not very sensitive to transistor size.
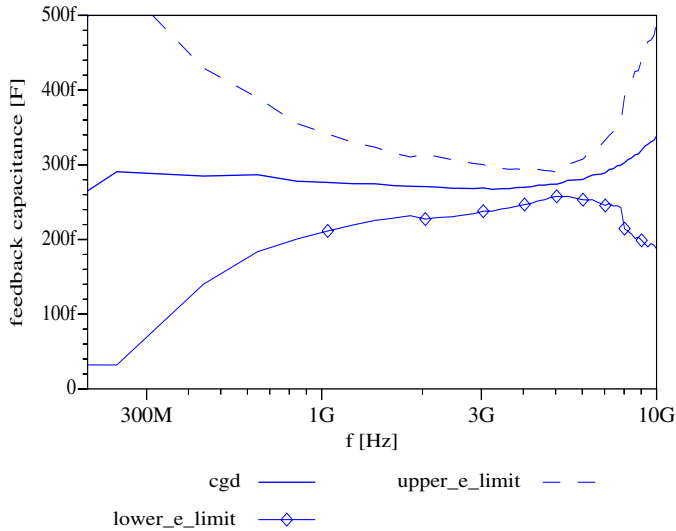
Figure 49. Extracted C_{gd} along with its upper and lower error limits for a 60 x 6.8μm x 0.5μm NMOS device.

In the example case of [101] the FET C_{gd} extraction uncertainty was very small, only 1.6% at the optimum extraction frequency of 6.3 GHz.

The bias dependence of C_{gd} extraction is demonstrated with the 60 x 6.8 μm x 0.5 μm device in Fig. 50, having a 0.8V gate voltage sweeping the drain voltage from 0.1 to 2.5 volts. Here both relative and absolute errors are best in the low drain voltage region. The relative error is 32% at 0.1 V drain voltage, decreasing down to 8% in the saturation region. The absolute C_{gd} error is approximately 90 fF at 0.1 V, decreasing down to 10 fF in the saturation region. Both relative and absolute errors decrease rapidly as the drain voltage is increased until the threshold voltage is reached. After that the change in error magnitudes is very small. In feedback capacitance extraction the largest error contribution is due to the S_{12} magnitude uncertainty. The C_{gd} error decrease with increasing drain voltage is surprising, as the C_{gd} capacitance value decreases when moving from the linear to the saturation region. It would be straightforward to expect that a larger absolute capacitance value would be easier to measure, but there seem to be other uncertainty mechanisms. The S_{12} magnitude uncertainty is the largest source of C_{gd} extraction error, both in the saturation and the linear regions. The

S_{11} and S_{22} uncertainties are very small, except for the S_{22} phase uncertainty in the linear region, which is still much smaller than the S_{12} magnitude error contribution.

Cgd error bias dependency
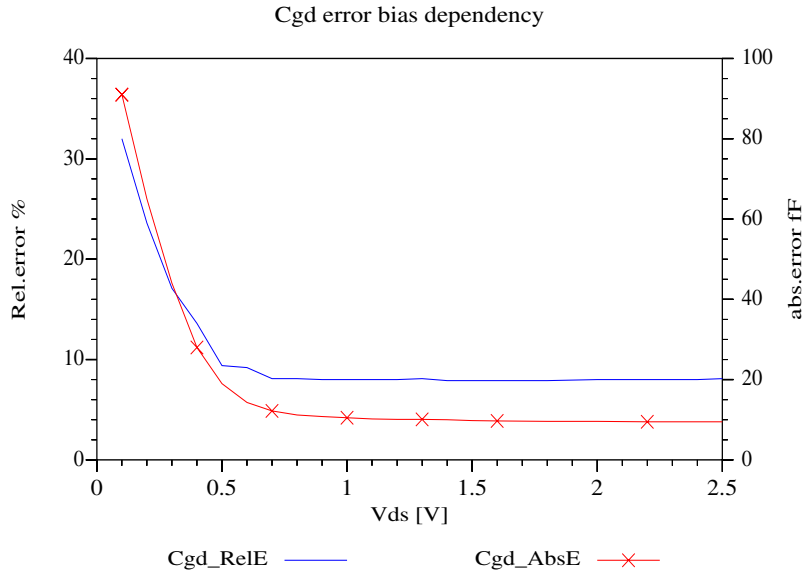
Figure 50. Bias dependence of the C_{gd} extraction of a 60 x 6.8 μm x 0.5 μm device. V_{gs} is 0.8 V while V_{ds} is swept.

It seems that the optimum C_{gd} extraction bias points are at the high drain voltages in the saturation region. If the C_{gd} bias dependence is to be extracted the large relative error should be taken into account in the linear region.

4.5 Uncertainty of Output Impedance
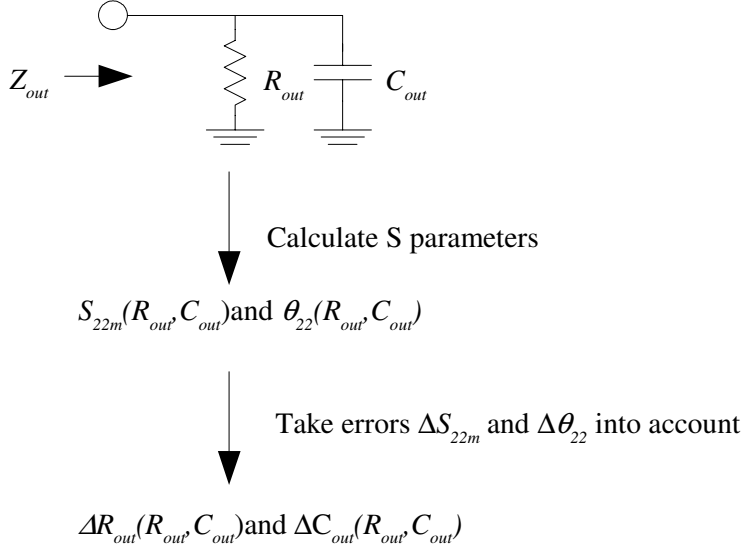
Figure 51. Flowgraph of output impedance error analysis including the simple output impedance approximation for a MOSFET used in this study.

As in the case of input, the output impedance is simplified for error analysis purposes and approximated as a parallel resistor and capacitor connection, as shown in Fig. 51. In contrast to the input impedance, the R_{out} and C_{out} of the output case do not represent any real small signal components but could rather be used as measurement data for the extraction of g_{ds}, C_{ds}, C_{bd} and even R_B. Now the output admittance calculation as a function of S_{22} magnitude and phase gives:

$$Y_{out} = \frac{1 - S_{22m}{}^2 - j2S_{22m}\sin\theta_{22}}{1 + S_{22m}{}^2 + 2S_{22m}\cos\theta_{22}} \cdot \frac{1}{Z_0} \tag{75}$$

The output resistance and capacitance can be defined as [87]

$$R_{out} = \frac{1}{\Re\{\,Y_{out}\,\}} \tag{76}$$

and

$$C_{out} = \frac{\left| \Im\{ Y_{out} \} \right|}{\omega} \qquad (77)$$

Again, solving the total differential error requires partial differential for both Eq. (76) and (77), resulting in

$$\Delta R_{out} = \left| \frac{\partial R_{out}}{\partial S_{22m}} \right| \Delta S_{22m} + \left| \frac{\partial R_{out}}{\partial \theta_{22}} \right| \Delta \theta_{22} = \qquad (78)$$

$$\left| 2Z_0 \frac{(2S_{22m} + \cos \theta_{22}(1+S_{22m}))}{(1-S_{22m}{}^2)^2} \right| \Delta S_{22m} + \left| -2Z_0 \frac{S_{22m} \sin \theta_{22}}{1-S_{22m}{}^2} \right| \Delta \theta_{22}$$

and

$$\Delta C_{out} = \left| \frac{\partial C_{out}}{\partial S_{22m}} \right| \Delta S_{22m} + \left| \frac{\partial C_{out}}{\partial \theta_{22}} \right| \Delta \theta_{22} = \qquad (79)$$

$$\left| \frac{\sin \theta_{22}(1-S_{22m}{}^2 + 2S_{22m}(\cos \theta_{22} + \sin \theta_{22}))}{\omega Z_0 (1+S_{22m}{}^2 + 2S_{22m} \cos \theta_{22})^2} \right| \Delta S_{22m}$$

$$+ \left| \frac{S_{22m} \cos \theta_{22}(1+S_{22m}{}^2) + 2S_{22m}{}^2}{\omega Z_0 (1+S_{22m}{}^2 + 2S_{22m} \cos \theta_{22})^2} \right| \Delta \theta_{22}$$

The relative R_{out} error is plotted in Fig. 52 at 500 MHz as a function of absolute output resistance value with two different output capacitance, with values of 30 fF and 1 pF. The calculation suggests that the relative output resistance error is quite small when close to the 50 Ω reference impedance level, but at the more typical output resistances of saturated MOSFET at the 1 kΩ level, the relative error is already about 20%. The resistance error seems to be quite independent of capacitance value. At higher frequencies the error increases due to increasing network analyzer uncertainty, as shown in Fig. 58 for an output impedance of 1 kΩ and 500 fF. Fig. 53 shows the output capacitance inaccuracy with different R_{out} and C_{out} combinations at 500 MHz. It seems that the error decreases with increasing R_{out} and C_{out} values. In the output case it seems that it is much harder to get reliable MOSFET capacitance values than resistance values. In Fig. 54 the relative R_{out} is plotted at 2 GHz as a function of absolute output resistance value with two different output capacitance values. The difference between this and the 500 MHz error is not very large at lower output resistance values, but at 1 kΩ or higher the difference between the 30 fF and 1 pF curves has increased. At higher frequencies it seems that the parallel capacitance makes the measurement

of a high output resistance harder. At 2 GHz in Fig. 55 the output capacitance measurement seems to be more accurate than at 500 MHz. This result is confirmed in Fig. 58, where the capacitance error decreases at higher frequencies as a function of frequency.

Both output resistance and capacitance extraction errors can also be plotted on the Smith chart as a function of the output impedance. Fig. 56 show the output impedance points where the output resistance extraction results in relative errors of 1%, 2% and 10% at 500 MHz. Similar presentation for the output capacitance is shown in Fig. 57 at 500 MHz. In Fig. 57 the output impedance points are shown where the output capacitance extraction results in realtive errors of 2%, 5% and 20%.

The discontinuities in the Smith chart "curves" in Fig. 38, Fig. 39 , Fig. 56 and Fig. 57 could have been avoided by interpolating the Network analyzer uncertainty specifications as a function of power level.

Figure 52. Calculated total differential error of the output resistance as a function of output impedance at 500 MHz.

Figure 53. Calculated total differential error of the output capacitance as a function of output impedance at 500 MHz.

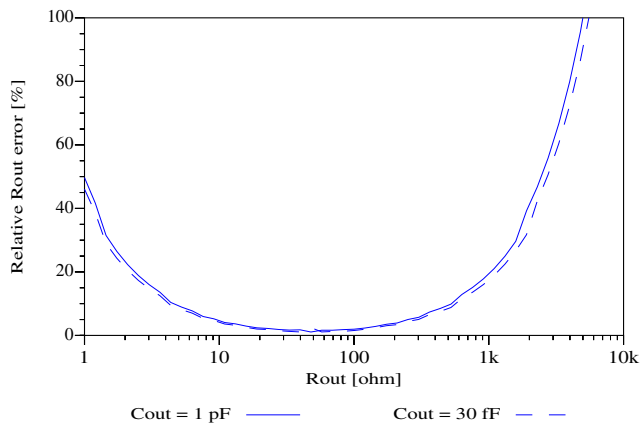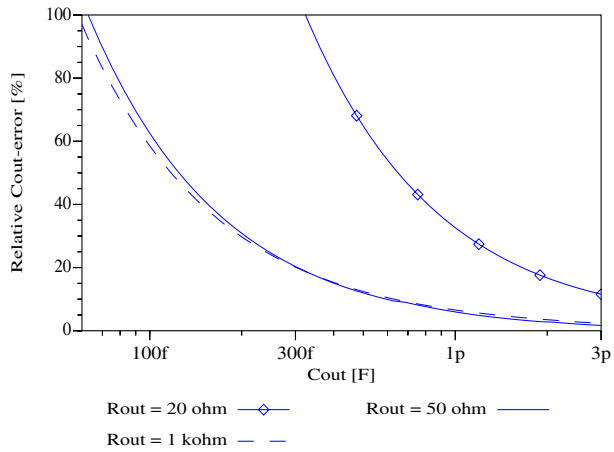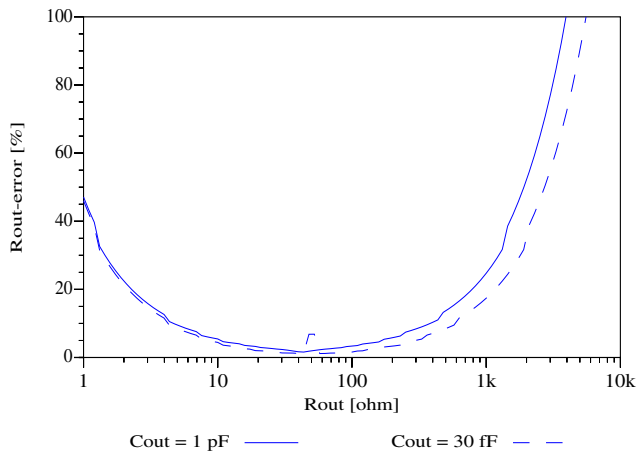*Figure 54. Calculated total differential error of the output resistance as a function of output impedance at 2 GHz.*
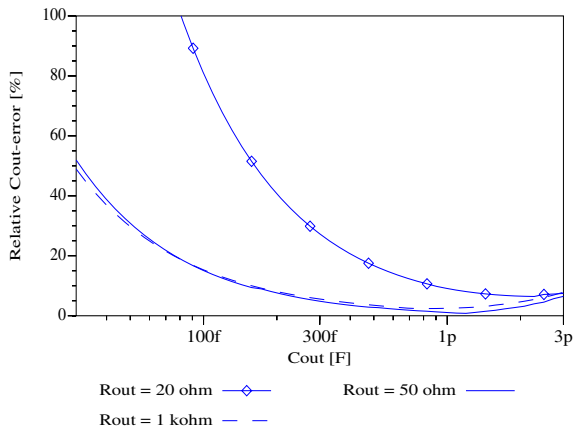
Figure 55. Calculated total differential error of the output capacitance as a function of output impedance at 2 GHz.
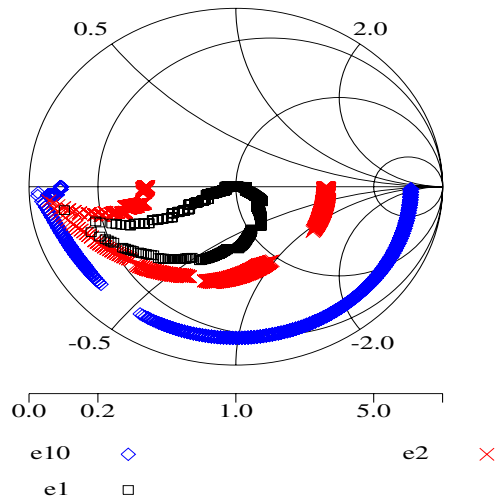
Figure 56. Output resistance extraction accuracy as a function of output impedance at 500 MHz. Impedance curves are shown where the output resistance can be extracted with 1%, 2% and 10% accuracy.
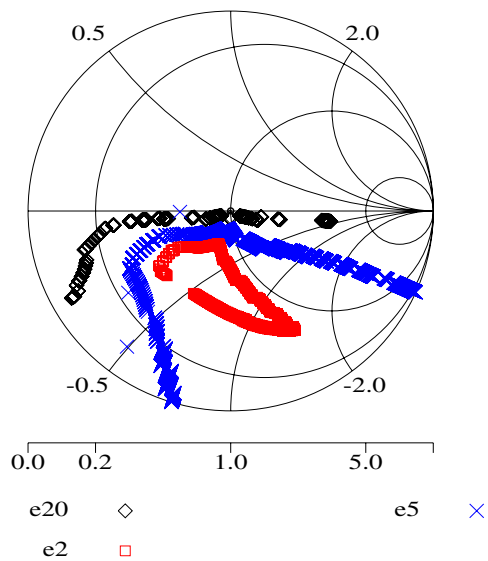
Figure 57. Output capacitance extraction accuracy as a function of output impedance at 500 MHz. Impedance curves are shown where the output capacitance can be extracted with 2%, 5% and 20% accuracy.
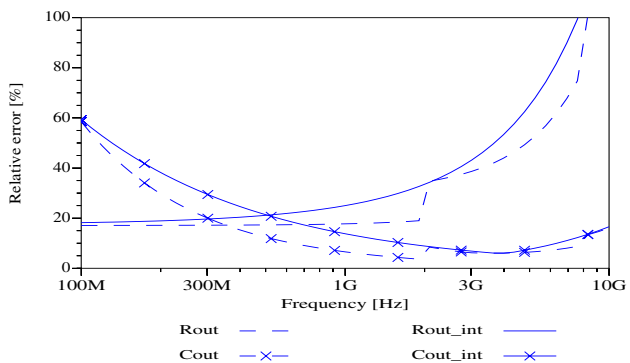
Figure 58. Relative R_{out} and C_{out} errors as a function of frequency for a MOSFET output with R_{out} of 1kΩ and C_{out} of 500 fF. The dashed lines are lines with no interpolation taken into account, whereas the solid lines are the interpolated curves.

95

The minimum g_{ds} uncertainty of the example case of [101] is 5.6%. With the same R_{out} value of 200 Ω and C_{out} value of 26 fF, the total differential error of my method is 3.4%. These numbers are of the same order of magnitude.

4.6 Conclusion of Measurement Uncertainty Study

The results of this measurement error study are especially applicable to AC parameter extraction. For instance, the measurement errors of the different small signal component values can be used as guidelines for optimization goals in parameter extraction. It is not very reasonable to put too much weight on input resistance extraction, as the measured fitting goal can be very erroneous. More emphasis should be put on the other parameters.

A very useful result of this study is that the measured accuracy of input and output-related small signal parameters can be used as guidelines for the design of the optimum AC device extraction set, as well as the bias region. For instance, with an approximate knowledge of the polysilicon sheet resistance and theoretical input capacitance value, it is possible to design devices with the smallest errors for input resistance and capacitance extraction. The resulting scalable MOS model can be used for other geometries too. In the output case the guidelines are quite straightforward. According to the results, the output impedance can be roughly extracted with an error of less than 20% when R_{out} is between 2 Ω and 1kΩ. The most accurate extraction results are achieved when R_{out} is close to 50 Ω. Practically, these output resistance levels are achieved with quite typical wide RF devices, depending on the bias region of interest. In the case of the output capacitance, it seems that the larger the device, the better, as the larger absolute value of the capacitance is easier to measure. The real value of the output resistance should be 50 Ω or larger for the most accurate C_{out} extraction. Thus, C_{out} should be extracted in the linear region of operation if possible.

5. Model Accuracy of Absorbed Parasitic Series Resistances

In BSIM3–5, MOS Model 9 and 11 and BSIM3SOI the parasitic series resistances of drain and source can be absorbed into the current description, reducing the amount of nodes. This simplifies the matrix calculations of the circuit simulator, but at the possible expense of loss in AC simulation accuracy. To my knowledge this has not been analyzed before for high-frequency cases. I have studied the high-frequency response inaccuracy that this simplification can introduce. I considered the differences by studying simplified equivalent circuits in input and output impedances, gain and backward gain, or the S_{12} parameter. Numerical examples have been presented for every case. The substrate effect has been totally neglected to keep the analysis simple. Clear differences can be pointed out with a very simple MOSFET equivalent circuit.

5.1 Input Impedance with and without Absorbed Series Resistances

The effect of absorbed series resistances are most easily considered in the cut-off region of operation, where both transconductance and drain conductance are zero. The equivalent circuit for the input admittance calculation of the conventional model case is simple and presented in Fig. 59. As the channel charge is equally divided between the drain and source, and as the series resistances are practically equal, the equivalent circuit can be simplified even further. The resulting circuit is a series connection of the gate resistance, input capacitance (which is the sum of C_{gd} and C_{gs}) and half of the one side series resistance. The input admittance can be determined as

$$Y_{in}(s) = \frac{sC_{in}}{1 + sC_{in}(R_G + \frac{R_S}{2})} \qquad (80)$$
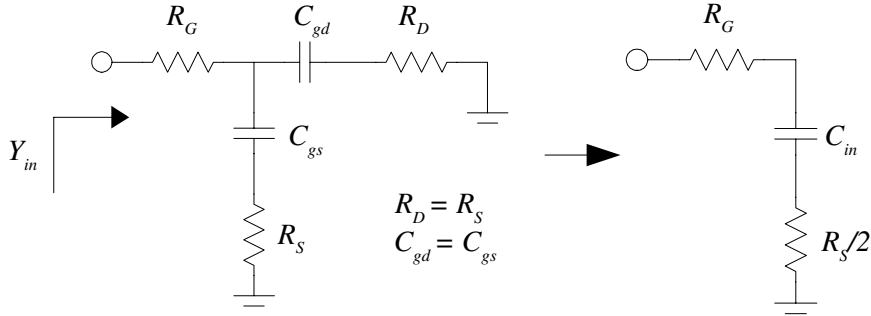
Figure 59. Input admittance calculation of a conventional equivalent circuit in the cut-off region.

The equivalent circuit in the case of absorbed series resistances is simply the series connection of R_G and C_{in}. As series resistances R_S and R_D would end up in series with the drain conductance which is zero, the branch including them disappears. The input admittance is determined for the absorbed series resistance case as

$$Y_{in}(s) = \frac{sC_{in}}{1 + sC_{in}R_G} \tag{81}$$

Comparing (80) and (81) we can see that the number of poles and zeros are the same but the value of the pole is different. Both also have a zero at $s = 0$ and a pole at infinity. The quantitative difference can be seen with the example of a small RF MOSFET of 200 μm x 0.35 μm divided into 20 parallel devices. Typical R_G would be around 2 Ω and R_S also 4 Ω, while input capacitance would be

$$C_{in} = WLC_{ox} = \frac{WL\epsilon_r\epsilon_0}{t_{ox}} = \frac{200\,\mu m \cdot 0.35\,\mu m \cdot 3.9 \cdot 8.85 \cdot 10^{-12}\frac{F}{m}}{8\text{nm}} = 300\text{fF} \tag{82}$$

With these values the poles in (80) and (81) would be 133 GHz and 265 GHz, respectively. The difference is large, but both equations are equally accurate for typical state-of-the-art CMOS technologies where f_t values are somewhat less than the pole frequencies.

98

5.2 Simple HF Gain with and without Absorbed Series Resistances

The effect of absorbed series resistances on high-frequency gain can be studied by comparing the gain of the conventional simplified equivalent circuit in Fig. 60 with a the absorbed case in Fig. 61. Here the load is a capacitor, C_L. Again, in the conventional equivalent circuit the source node has been removed using series-series feedback, as was done in Chapter 2.4. Also, to keep the analysis simple we assume that in the saturation region of operation the effect of C_{gd} is negligible, and it is removed. The accurate voltage gain for this circuit is

$$A(s) = -\frac{g_m''(\frac{1}{g_{ds}'} + R_S)}{(1 + sC_{gs}'(R_G + R_S))[1 + sC_L(R_S + R_D + \frac{1}{g_{ds}'})]}$$

(83)

It can be approximated further that C_{gs}' is C_{gs}. The notations in (83) are the same as those in Chapter 2.4. Equation (83) has three poles and one zero at infinity. One pole is hidden in transconductance, g_m'', equation (9) resulting from R_S and C_{gs}'

For the absorbed case in Fig. 61 there is no series resistance with C_{gs}, as the R_S effect is solely in the DC description of the model. The modified transconductance g_m' does not have the pole formed by R_S and C_{gs}, as in g_m''. The accurate voltage gain of this absorbed case is
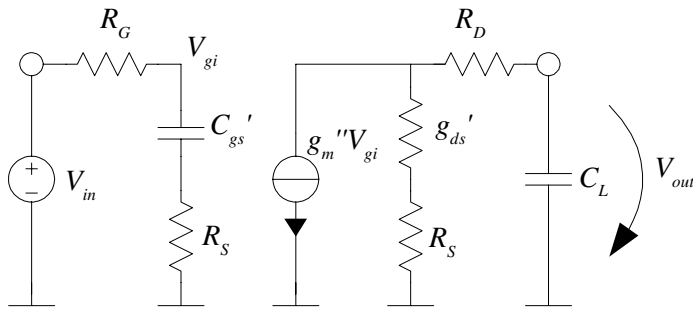

Figure 60. Simplified equivalent circuit for AC gain calculation with the conventional R_S/R_D description.

$$A(s)=-\frac{g_m{}'(\dfrac{1}{g_{ds}{}'}+R_S+R_D)}{(1+sC_{gs}{}'R_G)[1+sC_L(R_S+R_D+\dfrac{1}{g_{ds}{}'})]} \tag{84}$$

This case has two poles and one zero at infinity. The dominant pole is identical for both circuits. What is interesting is that the DC gain is different.
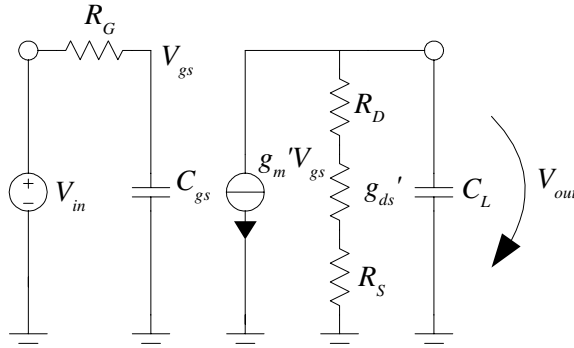
Figure 61. Equivalent circuit for gain calculations for the case with absorbed series resistances.

If we compare the differences between equations (83) and (84) quantitatively with the numerical example of a 20 x 10 µm x 0.35 µm MOS device and a 200 fF load capacitance, we see the difference between the poles. For such a MOSFET, typical realistic values in the saturation region of operation could be $R_G = 2\ \Omega$, $R_S = 4\ \Omega$, $g_{ds} = 200\ \Omega$, $g_m = 10$ mS, $C_{gs} = 150$ fF and $C_L = 150$ fF. The poles of (83) are 255 GHz, 170 GHz and 5.1 GHz, with a DC gain of 2.04. The poles of (84) are 510 GHz and 5.1 GHz, with a slightly different DC gain of 2.08. The dominating pole at 5.1 GHz is due to the load capacitance making equations (83) and (84) behave quite similarly. The difference in this case is seen at very high frequencies.

5.3 Isolation with and without Absorbed Series Resistances

The isolation can be considered as the inverted backward voltage gain from output to input. If we study the cut-off region of operation, the situation is the most simple for analysis. The equivalent circuit of Fig. 62 is again simplified by

not taking the back gate effects into account. The substrate node is totally neglected. Let us first consider the conventional case without absorbed series resistances.
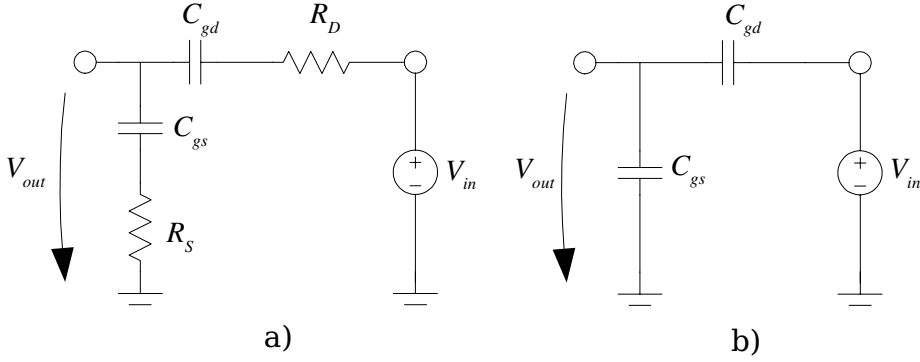
Figure 62. Simplified equivalent circuit of a MOSFET in cut-off for backward gain calculation for a) the conventional case, and for b) the absorbed case.

The output voltage seen at the gate, V_{outb}, is simply a voltage division. From this the backward gain or inverted isolation can be calculated as

$$\frac{V_{outb}}{V_{inb}} = \frac{1 + sC_{gs}R_S}{1 + \dfrac{C_{gs}}{C_{gd}} + sC_{gs}(R_S + R_D)} \tag{85}$$

We see that the isolation has a low frequency value determined by the capacitance ratio, and a frequency dependence determined by the zero caused by C_{gs} and the series resistances.

For the case with series resistances absorbed into the drain current model, the resistors are in series with drain conductance, g_{ds}, which is zero. The resistances practically disappear from the equivalent circuit, which is a mere capacitance connection in Fig. 62 b). The voltage division is completely determined by the capacitance ratio:

$$\frac{V_{outb}}{V_{inb}} = \frac{1}{1 + \dfrac{C_{gs}}{C_{gd}}} \tag{86}$$

We can see that the conventional model has some frequency dependence in the cut-off region, whereas the case with absorbed series resistances does not have any frequency dependence. If we compare the two cases with the numerical example of a 200 µm x 0.35 µm device divided into 20 fingers, both the C_{gs} and C_{gd} capacitances are about 150 fF, and again R_S is taken as 4 Ω. Both the zero and the pole in Eq. (85) are at 265 GHz.

5.4 Output Impedance with and without Absorbed Series Resistances

If we define output impedance as the inversion of Y-parameter y_{22} as

$$Z_{out}(s) = \frac{1}{y_{22}(s)} \tag{87}$$

and if we use a simple equivalent circuit with a zero gate resistance we can calculate Z_{out} from the equivalent circuit of Fig. 63. As the gate is connected to ground, the transconductance current is zero and only the C_{gd} capacitance is seen in parallel with the series connected R_S and g_{ds}' components. In a real transistor circuit the Z_{out} should also include the effect of the source impedance, Z_0, but it is left out in this study for simplicity. Here the drain conductance and transconductance are defined with equations (7) and (8) respectively to simplify the equivalent circuit, as was done in Chapter 2.4. Thus the output impedance can be calculated as

$$Z_{out}(s) = R_D + \frac{R_S + \frac{1}{g_{ds}'}}{1 + sC_{gd}\left(R_S + \frac{1}{g_{ds}'}\right)} \tag{88}$$

The other case is where the series resistances have been absorbed into the drain current description. The equivalent circuit is defined as that of Fig. 64. Practically, the output impedance analysis sees a capacitance in parallel with the series connected R_S, R_D and g_{ds}'. If we again simplify the gate resistance to zero
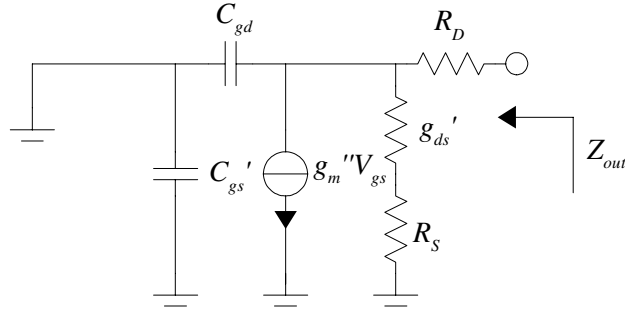
Figure 63. Conventional MOSFET equivalent circuit for simplified output impedance calculation.

the current source of Fig. 64 is zero as the gate to source voltage, V_{gs}, is zero. The output impedance can be calculated as

$$Z_{out} = \frac{R_S + R_D + \dfrac{1}{g_{ds}'}}{1 + s C_{gd}\left(R_S + R_D + \dfrac{1}{g_{ds}'}\right)} \tag{89}$$

If we compare equations (88) and (89), we see that they have the same DC value when $s = 0$ but at high frequencies their behavior is different. (89) has a zero at infinity, whereas (88) does not. Both have one pole, but the poles are slightly different. If we have a somewhat realistic small MOSFET device, for instance with a size of 20 µm x 0.35 µm, the series resistances are around 20 Ω each, with a C_{gd} of 16 fF and a $1/g_{ds}'$ which is for example 100 Ω in linear region of operation, we get a pole of 83 GHz for equation (88) and 71 GHz for (89). The difference is rather small in the saturation region of operation as the drain conductance is very small, resulting in a very small relative effect caused by R_D and R_S.

Obviously, the small difference in output impedance poles between the two modeling approaches causes a negligible difference at small frequencies, as can be seen from the Z_{out} magnitude plot of (88) and (89) in Fig. 65.
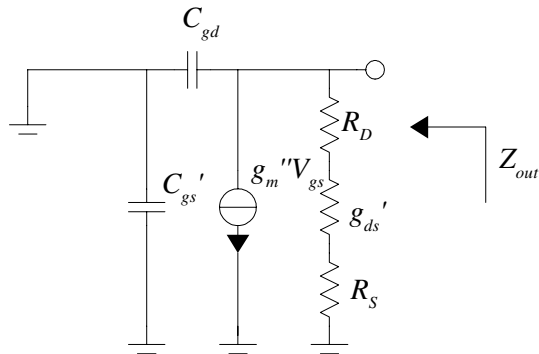
Figure 64. Equivalent circuit of the MOS model with absorbed parasitic series resistances.
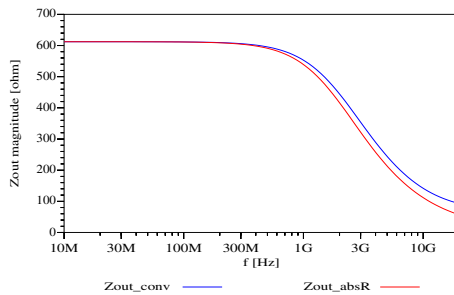
Figure 65. Output impedance magnitude difference between the conventional modeling approach (88) and the absorbed case (89).

5.5 Empirical Results of Distributing R_D/R_S

Comparisons between simulated MOS behavior with absorbed R_D/R_S resistance, a conventional MOS model and measurements have been made. The study has been done with both the Peregrine 0.5 μm UTSi SOI CMOS technology, and with the VTT 0.8 μm CMOS technology. One device will be considered from both technologies. With an accurately characterized DC model, the AC model

was defined as a small signal model as that of Fig. 66 for the Peregrine case. The AC accuracy of the reflection coefficients, as well as the input and output capacitances and resistances, are considered. The transconductance, S_{12} and S_{21} parameters are also studied.
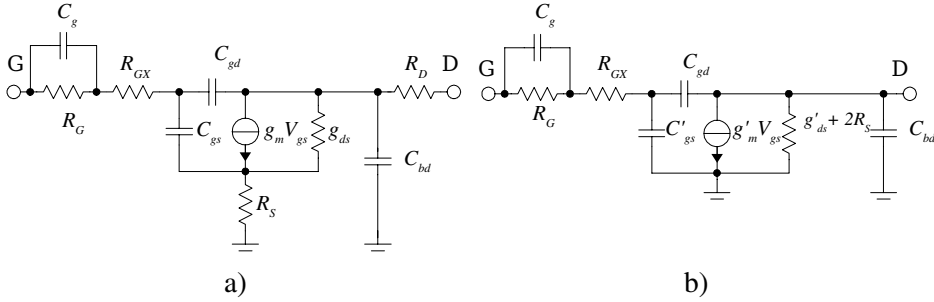
Figure 66. Simple AC equivalent circuit for a) the conventional Peregrine model and b) the model with absorbed series resistances.

The first case is a Peregrine SOI device having a geometry of 20 x 6.8 µm x 0.5 µm. The small signal component values of the device are shown in Table 1. A special input equivalent circuit at the gate was used to improve the abnormal input resistance behavior by adding a C_g capacitance in parallel with the gate resistance. An additional R_{GX} resistance was put in series at the gate. This is a very similar approach to the Peregrine vendor model.

The input reflection seen in Fig. 67 a) presents how the input resistance difference shown in Fig. 67 b) affects the high frequency behavior. Here, input resistance is determined from simulations and measurements by the simple equation

$$R_{in} = \Re \frac{1}{y_{11}}$$

(90)

At low frequencies the difference is not seen in Fig. 67 a). The model with absorbed series resistances has a worse match in input reflection. This is seen across the whole bandwidth from 300 MHz to 20 GHz. We can estimate the approximate high-frequency value of the input resistances. Small modifications to (80) (replacing R_G with R_{GX}) have to be made to account for the gate equivalent circuits depicted in Fig. 66. Using the input resistance defined by

(90) results in an approximated input resistance at high frequencies for the conventional model of

$$R_{in}(f=f_{high})=R_{GX}+\frac{R_S}{2}$$

<div align="right">(91)</div>

Table 1. Small signal model parameters of the studied models in Peregrine technology of a 20 x 6.8 µm x 0.5 µm SOI NMOS device.

Parameter	*Value*
R_G	30 Ω
R_{GX}	2.2 Ω
C_g	1.9 pF
C_{gs}	120 fF
C_{gd}	45 fF
C_{bd}	31 fF
g_m	14.8 mS
R_S / R_D	6.4 Ω
R_{ds} or $1/g_{ds}$	890 Ω

For the model with absorbed series resistances the input resistance is solely determined by R_{GX} at high frequencies. This can easily be observed in Fig. 67 b), where the input resistance value is approximately 7 Ω at 20 GHz for the conventional model and only 3 Ω, which is almost the same as R_{GX} from Table 1, for the absorbed case. There is a constant difference of over $R_S/2$ across the whole bandwidth. In cut-off the difference would accurately be $R_S/2$, according to the comparison in Chapter 5.1, but practically the difference is slightly larger in active device operation, as C_{gd} and C_{gs} are not equal as assumed in Fig. 59 for the cut-off region, thus giving a slightly different weight for the R_S and R_D "parallel" connection.

It should be noted that the input resistance could independently be fitted accurately for both modeling approaches, but the total behavior would be distorted for the absorbed case. The output match, as well as S_{12} and S_{21}, are also affected by the series resistances. All of these characteristics should be fitted simultaneously to achieve an accurate MOS model.
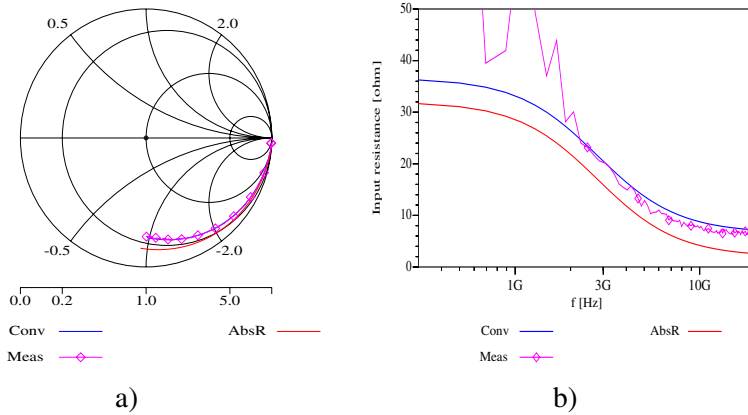
Figure 67. a) Input reflection and b) input resistance for the conventional MOS model, with the absorbed case and for the measurement of a 20 x 6.8 μm x 0.5 μm SOI NMOS device.

The S_{12} and S_{21} magnitude and phase fits are shown in Fig. 68. The differences in S_{12} are not noticeable, and in the S_{21} case the fit differences are very small between the two different models. Clearly, the measured curve is rather different to the models and this is due to an inaccurate model. Although the high-frequency gain seems to be different according to equations (83) and (84) in the two different series resistance modeling approaches, the difference is rather small in this case. This is partly due to the S parameter definition where the source and load impedances have to be defined.

When comparing the transconductance defined by the real part of y_{21}, as in equation (47), we get the situation of Fig. 69. The difference between the two simulated curves is smaller than their difference to the measurement. However, above 10 GHz the difference between the two simulated curves increases in the favor of the conventional model. The higher the frequency, the less accurate the absorbed resistance model is. Also, it seems that there is a zero at a few GHz,

with the absorbed resistance model not present in the measured device nor in the conventional MOS model.
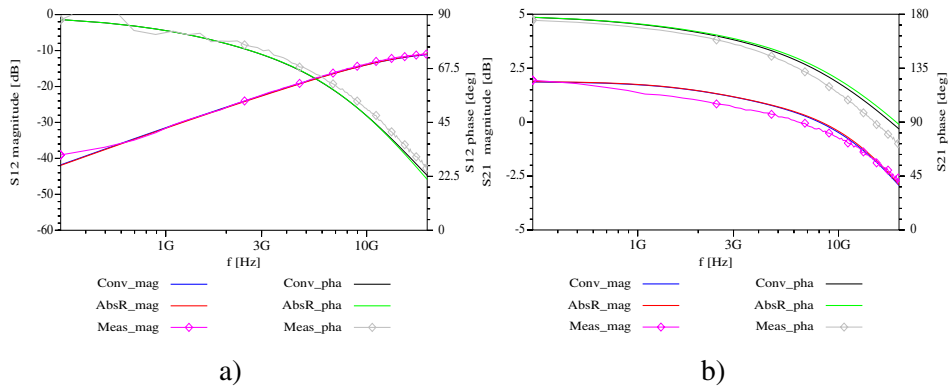
Figure 68. S_{12} and S_{21} magnitude and phase fits for the two different series resistance modeling approaches and measurements of a 20 x 6.8 μm x 0.5 μm SOI NMOS device.
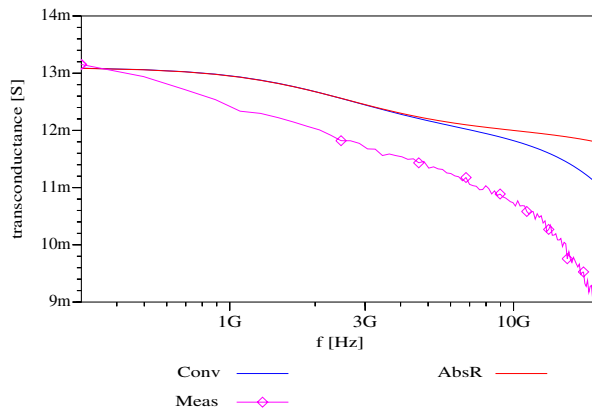
Figure 69. Transconductance comparison of the two series resistance modeling approaches, as well as a measured curve of a 20 x 6.8 μm x 0.5 μm SOI NMOS device.

The output reflection behavior appears to be more accurate with the conventional case of modeling the series resistances, as seen in Fig. 70 a). In Fig. 70 b) we can see that one reason for this is the output resistance behavior,

which is different at high frequencies. Here, the output resistance is defined simply by equation (89). The difference at very low frequencies is negligible as expected by equations (88) and (89). Also, the output capacitance differs between the two models, causing the better S_{22} fit with the conventional series resistance description. This can be seen in Fig. 71, where both input and output capacitances have been shown. The capacitances have been calculated from y-parameters with equations (44) and (77).

a)   b)
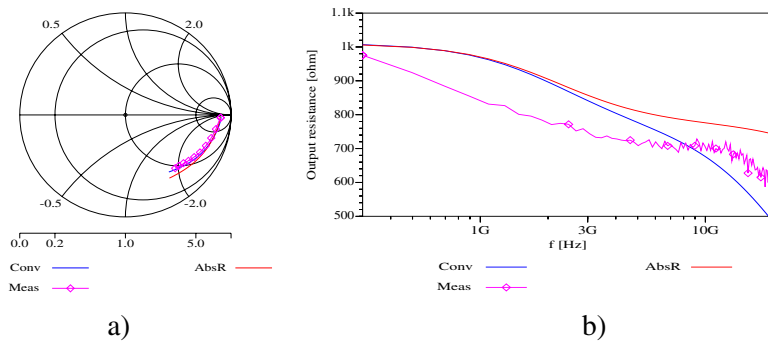
Figure 70. a) Output reflection and b) output resistance models compared to a measured result of a 20 x 6.8 μm x 0.5 μm SOI NMOS device.
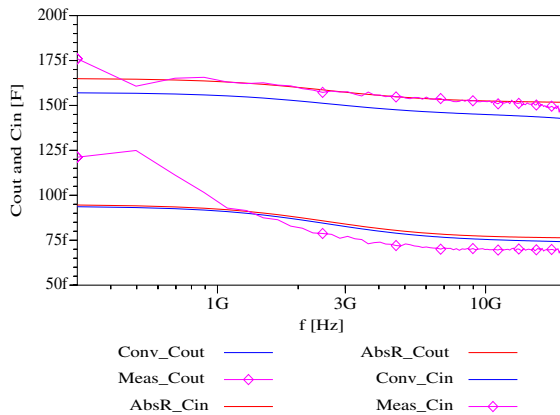
Figure 71. Input and output capacitances of a conventional and absorbed series resistance description. The measured curve for a 20 x 6.8 μm x 0.5 μm SOI NMOS device is also presented. The input capacitance curves are above the output capacitance curves.

Empirical studies of the effect of absorbing the series resistances for bulk CMOS devices were done using the more complicated small signal equivalent circuits of Fig. 72. The bulk junction diode capacitances are included, as well as one single substrate resistance. No modified input resistance networks are required.

a)   b)
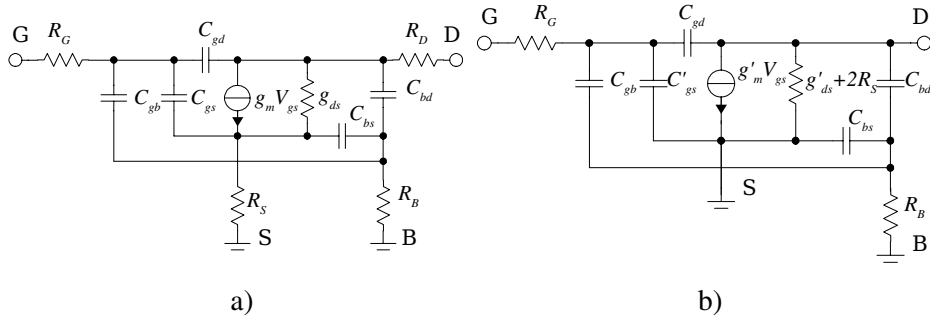
Figure 72. Simple AC equivalent circuit for a) the conventional VTT bulk NMOS model and b) the model with absorbed series resistances.

The small signal equivalent circuit component values are presented in Table 2 for an example device of 100 μm x 0.6 μm consisting of 16 parallel devices. The parameters also describe the bulk NMOS device to be in deep saturation, resulting in a small drain conductance or large drain-source resistance. The drain conductance is much smaller, $1/4k\Omega$ in the bulk CMOS case compared to the SOI device. This is the result of a much smaller I_{ds}–V_{ds} curve slope in the bulk device case. Also, here the input resistance shown in Fig. 73 b) is calculated by (90). The difference, caused by series resistance absorption, again causes the $R_S/2$ difference between the two modeling approaches across the whole bandwidth. This results in a more accurate input match, shown by Fig. 73 a) for the conventional model. R_S is about 10 Ω in this case. The input resistance error is rather large at the lower frequencies, jumping above and below the probable correct value. This is a normal result of measurement uncertainty, as explained in Chapter 4.2.

Table 2. Small signal model parameters of the studied models in VTT technology of a 100 μm x 0.6 μm bulk NMOS device divided into 16 parallel fingers.

Parameter	Value
R_G	16 Ω
R_B	51 Ω
C_{gb}	4 fF
C_{gs}	130 fF
C_{gd}	36 fF
C_{bd}	62 fF
C_{bs}	21 fF
g_m	7.6 mS
R_S / R_D	10.4 Ω
R_{ds} or $1/g_{ds}$	3980 Ω

a)   b)
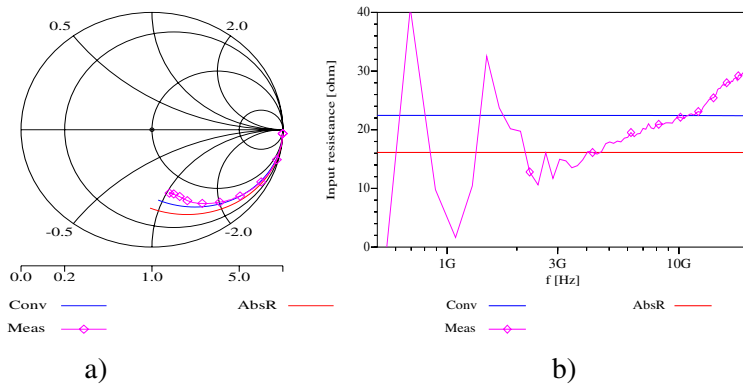
Figure 73. a) Input reflection and b) input resistance for the conventional MOS model, with the absorbed case and for the measurement of a 100 μm x 0.6 μm bulk NMOS device divided into 16 parallel fingers.

The S_{12} and S_{21} fits are shown in Fig. 74 a) and b). As in the case of the SOI example, the differences between the two resistance modeling approaches seem to be negligible. Although the models do not fit the experimental curves perfectly, the difference with the simulated curves can hardly be seen.

The transconductance defined by the real part of Y_{21} (as in equation (47)) is

a)   b)
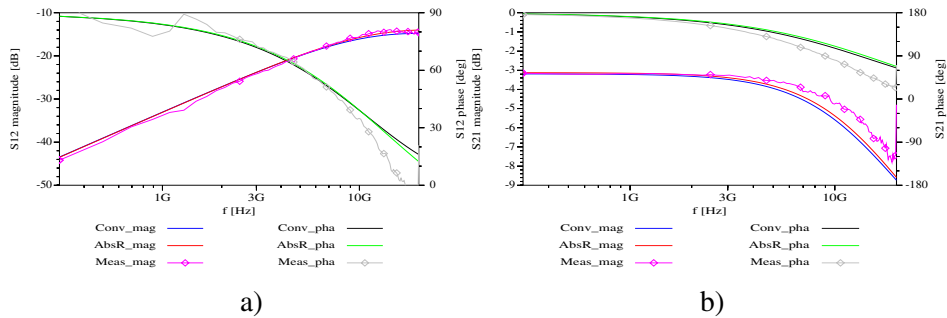
Figure 74. S_{12} and S_{21} fits for the two different series resistance modeling approaches, along with the measurements of a 100 μm x 0.6 μm bulk NMOS device divided into 16 parallel fingers.

shown in Fig. 75. Again, the difference between the two simulated curves is smaller than their difference compared to the measurement. Still, the transconductance fit of the conventional MOS model is more accurate at higher frequencies, similarly to the SOI device case. The benefit of the conventional series resistance description seems to be clear.

In Fig. 76 a) the output impedance match is shown for the 100 μm x 0.6 μm bulk NMOS device, along with the two different modeling approaches. As in the case of the SOI device, the conventional model has a much better fit than the model with absorbed series resistances. This is partly due to the better output resistance match shown in Fig. 76 b), again defined by equation (89). Below 3 GHz the measurement errors of the output resistance seem to be large, and at higher frequencies the variations decrease.
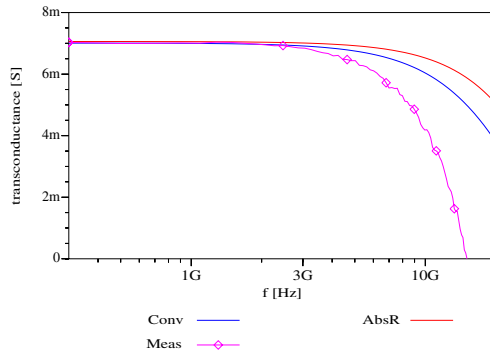
Figure 75. Transconductance comparison of the two series resistance modeling approaches as well as a measured curve of a 100 μm x 0.6 μm bulk NMOS device.

a)   b)
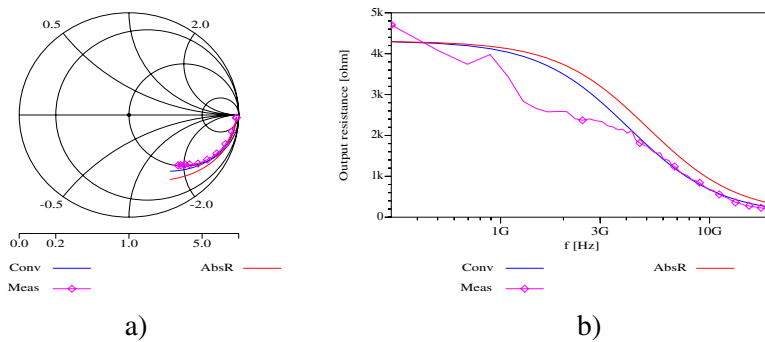
Figure 76. a) Output match, S₂₂, and b) output resistance of the two series resistance models compared to the experimental data of a 100 μm x 0.6 μm bulk NMOS device.

The input and output capacitances in Fig. suggest here as well that the fit of the conventional series resistance model is better. In the input capacitance there is a constant level difference between the two approaches, but the high-frequency behavior is also more accurate for the conventional model. The computational capacitance as a function of frequency decreases more rapidly for the conventional case. Similar differences can also be seen with the output capacitance, but the difference is rather small between the two approaches. In the output capacitance fits there is practically very little difference below 10 GHz.

Overall it seems that the trade-off of absorbing the series resistances into the drain current description results in poor AC characteristics across the whole frequency band, especially for the real parts of input and output impedances. Below 10 GHz the output capacitance and gain can be simulated with quite good accuracy.
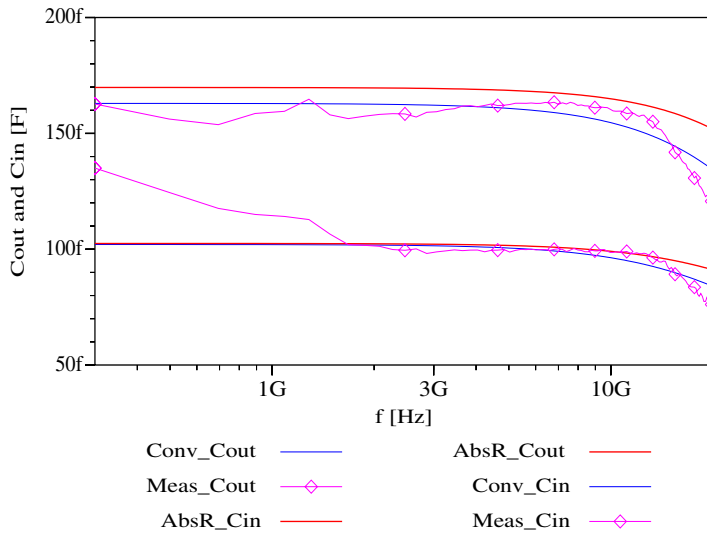
Figure 77. Input and output capacitances of the conventional and absorbed series resistance description compared to the measurement results of a 100 μm x 0.6 μm bulk NMOS device. The input capacitance curves are above the output capacitance curves.

6. Input Impedance Accuracy of Models

6.1 Input Capacitance Accuracy

Input capacitance accuracy, or in practice the active charge model accuracy of the BSIM3v3.2 model, was evaluated using Peregrine UTSi CMOS technology and VTT CMOS technology. A lot of S parameter measurements were performed at many bias points to extract the input capacitance bias dependence. Such comparisons have not been made at RF before, and most of the active charge model publications have concentrated on studying the MOSFET charge model DC dependence. In Fig. 78 a) the input capacitance fit for a 60 x 6.8 μm x 0.5 μm UTSi SOS NMOS device is shown. The V_{ds} voltage is swept with 5 different gate voltages: V_{gs} = 0.8, 1.3, 1.8, 2.3 and 3.0 V. The theoretical maximum input capacitance at zero bias determined by the oxide capacitance is approximately 580 fF for this case, and it is seems correct for the simulation. That is when it has been calculated as:

$$C_{in} = W_{eff} L_{eff} C_{OX} + 2 CGD0 W_{eff} \qquad (92)$$

Here, *CGD0* is the gate-drain zero bias capacitance. According to Fig. 78 a), the measurement is not that accurately fitted. Practically, the simulated value is underestimated in all bias regions and the simulated capacitance change, when the device transits from the linear to the saturation region, seems too small and too steep compared to measured behavior. Here one should note that the relative input capacitance error is approximately 5% from Fig. 37 for this size device when carefully extracted in the optimum frequency range, which is around 2 GHz. However, the difference at low frequencies is well above 10%, which suggests that the extraction error is larger in this measurement case than the theoretical minimum uncertainty from Fig. 37. In Fig. 78 b) the gate-drain capacitance for the same device is shown, and it seems that the C_{gd} fit accuracy is somewhat better compared to C_{in}. The zero-bias level is quite accurate, as is the capacitance difference between the linear and saturation regions. As the intrinsic C_{gs} change is of the same order as the C_{gd} change, it seems that the active charge model of BSIM3 is capable of describing the capacitance behavior of a MOSFET. The input capacitance change should mostly depend on the C_{gd} change when the drain bias is varied, suggesting that the measured input capacitance in Fig. 78 a) suffers from extraction uncertainty.
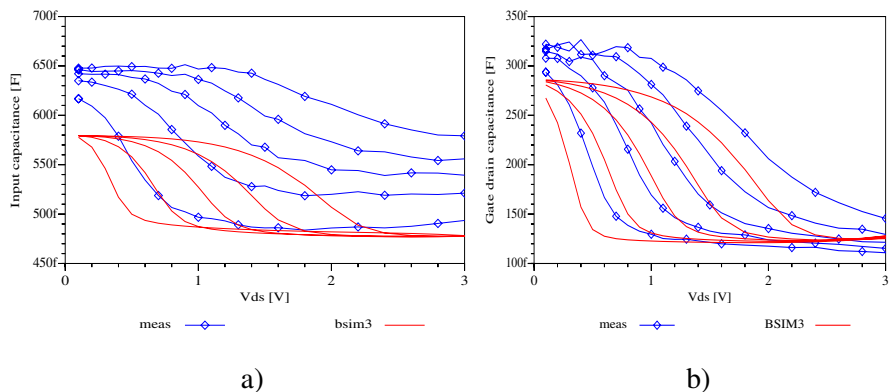
Figure 78. Measured and simulated BSIM3 a) input capacitance and b) gate-drain capacitance of a 60 x 6.8 µm x 0.5 µm NMOS SOS device in a drain voltage sweep.

Another input capacitance example is of an 8 x 10 µm x 1.0 µm UTSi CMOS device in Fig. 79 a), along with the C_{gd} fit in b). The simulated input capacitance with a low drain bias is quite accurately fitted, whereas the charge model inaccuracy increases when the device operates in the saturation region. This seems especially true for lower gate biases. According to theory [4], the input capacitance should have a constant value in deep saturation. The measurement gives different saturation capacitance results for different gate voltages. This can be the result of an extraction error caused by the different input resistance, which is a result of its transconductance dependence [4], studied in the next subchapter. The error is, at worst, almost 20%, although the minimum uncertainty should be closer to 10% at 2 GHz according to Fig. 37. Again, the theoretical low-frequency input capacitance value calculated by (92) is accurately determined by BSIM3, which is about 237 fF in this case. As with the first example, the C_{gd} fit in Fig. 79 b) is somewhat accurate. The capacitance difference between zero-bias and saturation described by BSIM3 quite accurately resembles the measured difference. The zero drain bias capacitance value is also quite accurate. It seems that the result in Chapter 4.4 of a more accurate feedback capacitance extraction than input capacitance extraction is correct.

116
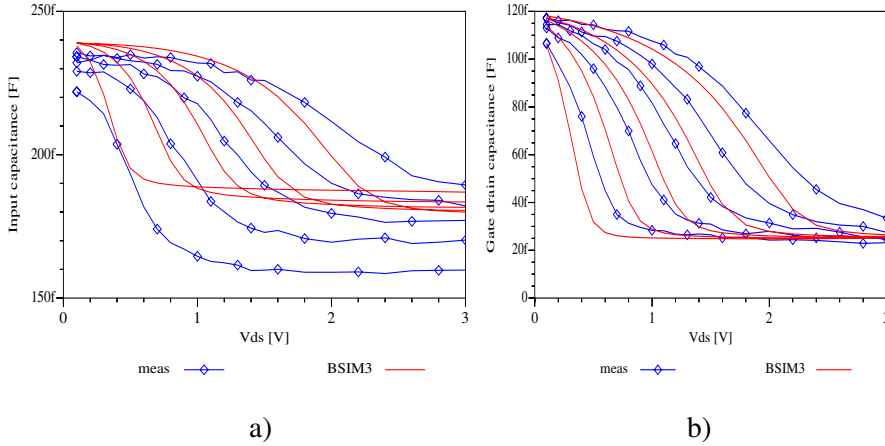
Figure 79. Measured and simulated BSIM3 a) input capacitance and b) gate-drain capacitance of an 8 x 10 µm x 1.0 µm NMOS SOS device in a drain voltage sweep.

A rough input capacitance evaluation was also done for bulk CMOS devices, presented in Fig. 80 for a 16 x 6.25 µm x 0.6 µm VTTB6 NMOS device. Here there is only one gate voltage, 1.2 V, with 9 drain voltage bias points. As the lowest drain voltage is 0.5 V, it can only be estimated that it seems that the measured and simulated input capacitance have a somewhat equal zero-bias value. According to equation (92) and extracted zero-bias capacitance values, the calculated maximum input capacitance is 212 fF. This corresponds well with the simulated result. Generally, the input capacitance fit is not that good. The capacitance value transition from the linear to the saturation region occurs at different drain voltages and the capacitance derivative seems different. However, as the input resistance value is about 40 Ω the error is 15–20%, according to Fig. 37. The fit error in Fig. 80 is slightly smaller, suggesting the difference may just be due to measurement uncertainty.
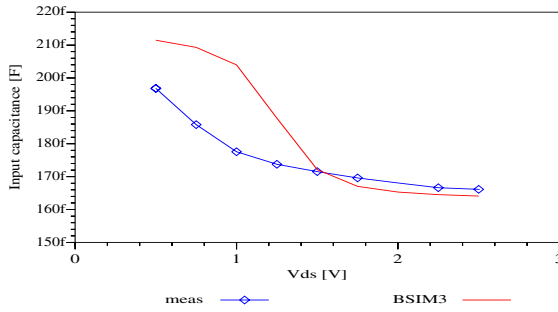
Figure 80. Measured and simulated BSIM3 input capacitance of a 16 x 6.25 μm x 0.6 μm NMOS bulk device in a drain voltage sweep.

In all input capacitance fits, the zero drain bias value seem to be quite well estimated, as does the capacitance difference between the linear and saturation regions. However, the capacitance bias dependence is not that well fitted. The place of the maximum derivative of the capacitance and of the derivative value itself seem to be inaccurate. These values are set by the transistor DC model and the active charge model. If the DC extraction is accurate and the active charge model behavior is inaccurate, there is little one can do for accurate AC extraction.

6.2 Input Resistance Bias Dependence

According to [37], the MOSFET intrinsic input resistance seen at the internal gate electrode is bias dependent in saturation and strong inversion by

$$R_{in,i} = \frac{1}{8g_m}$$

(93)

This resistance is due to the non-quasi-static effect of the inversion charge. According to an older but more theoretical study in [4], the input resistance in saturation is of the form

$$R_{in,i} = \frac{1}{5g_m}$$

(94)

118

In another publication [26] by the BSIM3 model developers, the intrinsic input resistance is suggested to be of the form

$$R_{in,i} = \frac{R_{st} R_{ed}}{\gamma (R_{st} + R_{ed})} \qquad (95)$$

R_{st} is the quasi-static channel resistance, which is basically V_{ds}/I_d, and R_{ed} is the excess-diffusion channel resistance, which is bias independent but scales with geometry. γ is a fitting parameter. Verification of the model was done with both device simulations and measurements. The latter gave unrealistically accurate R_{in} extraction results, as the extraction was performed at quite a low frequency. Other measured verifications of the bias dependence of the intrinsic input resistance have not been done, except for some data in [105]. In that reference there is only one plot of R_{in} versus g_m^{-1} with varying channel lengths, but there is no explanation of how the R_{in} extraction has been performed – the extraction frequency and the bias region – and what kind of test layouts were used.

This dependence was verified by extracting the input resistance of Peregrine UTSi CMOS devices as a function of drain voltage with five gate voltage steps. In Fig. 81 the input resistance has been extracted from 115 Touchstone™ files between 9 and 10 GHz for two different devices. In Fig. 81 a) the transistor dimensions are 4 x 10 µm x 1 µm, and in b) 8 x 10 µm x 1 µm. Along with the measured input resistance, the calculated intrinsic input resistances are plotted as well. They have been calculated with (93) added with an "extrinsic" offset value. In both cases the calculated resistance value fit the measured one with some accuracy when the transistor transits into the saturation region of operation. In Fig. 81 b) the device is precisely twice as large as the one in Fig. 81 a). The extrinsic offset in Fig. 81 a) is 24 Ω, whereas it is 13 Ω in Fig. 81 b). This result suggests there really is a bias dependence in the input resistance, as according to worst-case uncertainty specifications the relative error of extracted input resistance is below 20%. This can be approximated from Fig. 40.

Studying the input behavior with larger RF devices did not show the bias dependence, as the extracted resistance was in the same range as the error caused by measurement uncertainties. An example input resistance plot of a 60
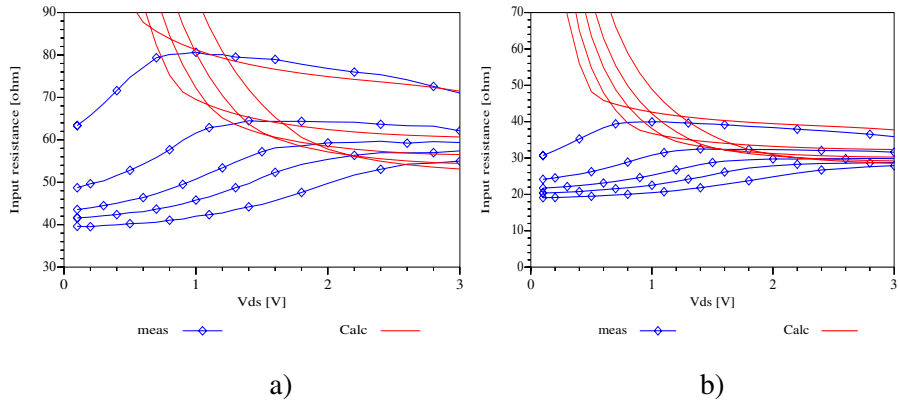
Figure 81. Input resistance bias dependence of a) 4 x 10 μm x 1 μm b) 8 x 10 μm x 1 μm UTSi CMOS transistors.

x 6.8 μm x 0.5 μm transistor is shown in Fig. 82, along with the calculated resistance having a 0.5 Ω offset due to the parasitic series resistance. In the saturation region the transconductance is rather large, resulting in very small resistance values when inverted. Thus it seems to be impossible to characterize the bias dependence of the intrinsic input resistance of large RF devices.
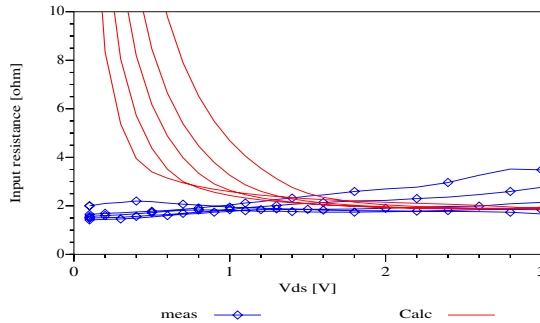
Figure 82. Extracted and calculated input resistance of a 60 x 6.8 μm x 0.5 μm NMOS transistor.

According to these measurement results, it seems that the intrinsic input resistance is bias dependent in SOI devices as well. This study compared with the previous work with bulk CMOS devices [4], [37], [105] prove the phenomenon is similar in SOI technology.

6.3 Bulk and SOI CMOS Comparison

From a circuit designer's point of view, the bulk and SOI CMOS input behavior is quite identical. The gate capacitance is formed in a similar way, being practically equal between devices of the same geometry in SOI and bulk technology. There are slightly different capacitances in SOI and bulk technologies that load the gate. Still, the differences seem so small that they are not measurable by standard RF device layouts. Specifically tailored layouts are required for this purpose.

7. Bulk Effect on Model Accuracy

7.1 CMOS

The bulk effect on RF behavior was studied empirically by comparing scalable models with different devices at different bias points simultaneously. A scalable BSIM3 DC model was extracted and the core MOSFET model was used for the intrinsic behavior; the extrinsic model described the series resistances, the bulk junction diodes along with their capacitances and the substrate resistance network as shown in Fig. 83. The drain and source series resistance and the bulk junction capacitances are described precisely according to the BSIM3v3 model, and the reason for their external modeling was just to reach the internal drain, source and bulk nodes of the BSIM3 model and to be able to build the substrate network at the internal bulk node.

In this study, different substrate resistance models and their simplifications have been evaluated. Fig. 83 shows the different topologies chosen for comparison. The case without any substrate network in Fig. 83 a) was chosen for reference only. However, it is still common in many analog models and vendor models to neglect the substrate network totally. Case b) is the basic implementation in an Aplac circuit simulator and c) is a suggestion presented [25]. Topology d) is a simplification of Fig. 83 e), which has been suggested in [106], and it is implemented in the BSIM4 model. The last topology in Fig. 83 f) has been suggested in [90].

In [30] and [31], three different substrate network topologies are compared at a single bias point with only two devices. The three substrate models were Fig. 83 b), c) and e) without resistor R_B. The substrate model with a single bulk resistor was considered to be the writers' "new" model. Substrate resistance network values are extracted from y_{22}, in contrast to my approach of general optimization, having all S parameters as goals simultaneously. Thus my study better takes into account the bulk resistance effect on the input, output and transadmittance parameters. In addition to this, I used at least ten different-sized devices in at least three different bias points to ensure the study validity with different geometry and operating point conditions.

As the geometry and bias dependence of substrate resistances in a conventional interdigital RF layout is not very well defined, the layouts of the test devices were designed in such a way as to have a well-known geometry dependence. A method of determining the single substrate resistance geometry dependence of typical RF devices has been suggested [45]. However, the model scalability was verified with only three devices. Another more interesting approach to determine scalable substrate resistances for a multifinger device with body ties on both sides of the layout is presented in [46]. The approach used in my study was to use specific layouts that are dominated by the width dependence, decreasing the amount of unknown factors. The substrate contacts were designed to run along the width of the device. Thus, the layout of Fig. 84 was chosen to result in a dominating width dependence. This simple approach removes the need to verify the accuracy of the geometry dependences developed in [45] and [46]. The example layout of Fig. 84 is a 50 µm wide and 0.8 µm long NMOS, divided into 8 parallel devices. Unfortunately, a drawback of the layout is an additional parasitic capacitance in parallel with the gate resistance, as the second metalization was chosen to run along the gate polysilicons. All of the substrate resistances were scaled according to

$$R_X = \frac{RX}{W_{eff}\, n_f} \tag{96}$$

where RX is the respective model parameter. W_{eff} is the effective total width and n_f is number of fingers or parallel devices. Bias independence was assumed although the size of the depletion region probably affects the substrate conductivity. According to [45] and [46] the bias dependence is rather small justifying this study. A worst case substrate resistance change of 6% [45] is possibly smaller than the extraction error due to measurement uncertainty.
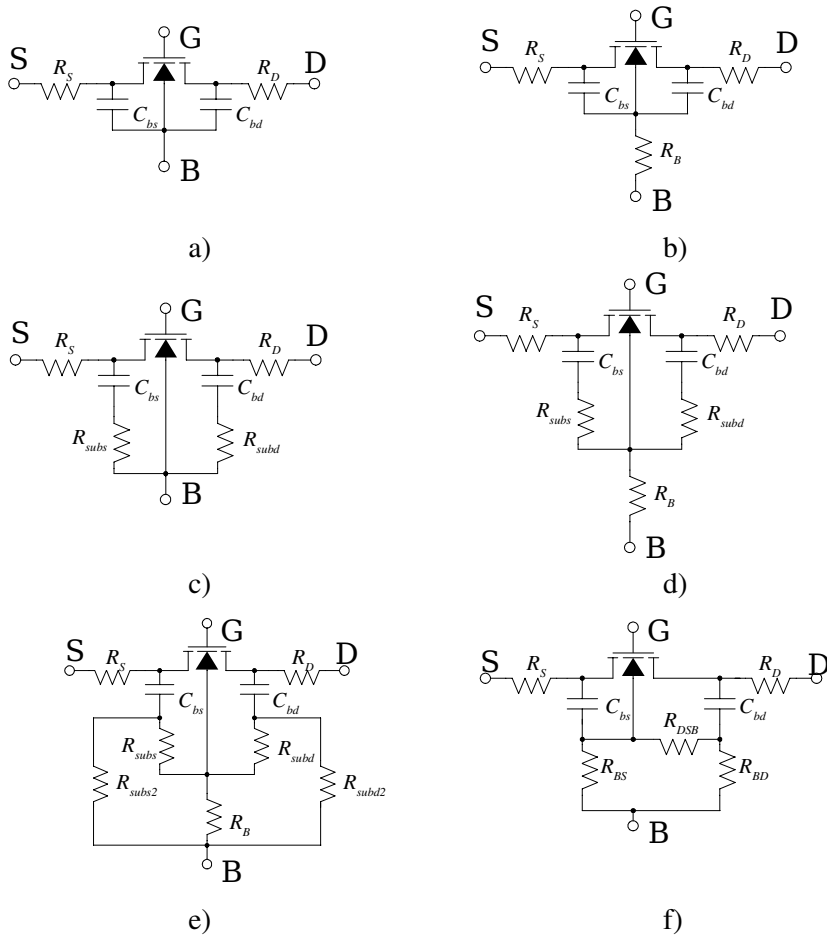
*Figure 83. Equivalent circuits used for substrate resistance model comparison.
a) No substrate resistances. b) The basic Aplac implementation. c) An
improvement proposal from* [25]. *d) Combination of b) and c) or a
simplification of e). e) An improved proposal published in* [106]. *f) A topology
suggested in* [90].
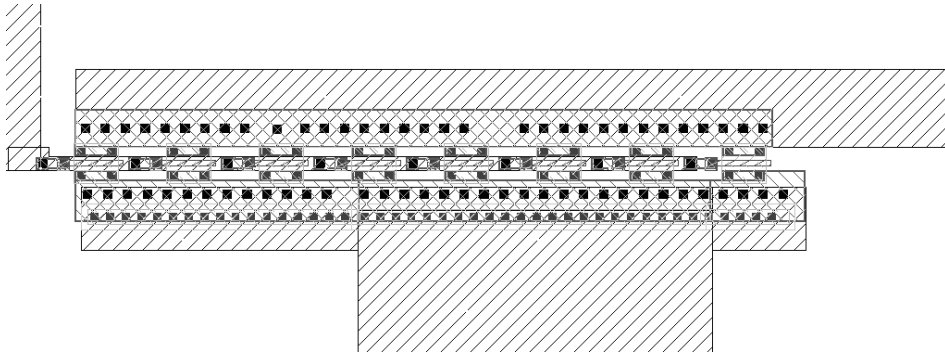
Figure 84. Example layout of a 50 µm x 0.8 µm device with 8 parallel devices designed specifically for bulk resistance characterization.

The comparisons have been made at a frequency range of 300 MHz up to 10 GHz. The relative *S* parameter errors are shown in Table 3 for the substrate resistance network topologies of Fig. 83. AC data is from the January 1997 VTT CMOS process run, but similar results were achieved fitting the January 1999 data as well. The differences between the substrate networks are quite obvious, especially when studying the S_{22} fit errors. This is due to the huge improvements in the output resistance and capacitance fits using the substrate resistance models. Thus the a) circuit not using any substrate resistance models, gives the poorest AC fit. The basic Aplac implementation has quite a good fit with the single resistance compared to the c) option with two resistances. The c) configuration did, however, have the best output reflection coefficient, but with a worse S_{12} fit than the b) circuit. The reason for the bad S_{12} magnitude fit may be due to optimization falling to a local minimum. As the DC model seldom provides accurate low frequency values, the AC optimization may end up over-tuning some parameters to compensate for the error.

The best substrate network approach seems to be the d) circuit, which combines the benefits of the b) and c) models. The d) and e) topologies have practically the same fit errors, as the e) topology was at its optimum when the R_{subd2} and R_{subs2} resistances were very large. The f) topology AC behavior is much worse compared to e), although it is only a somewhat simplified version of e). A symmetrical model is more physical and gives better AC behavior, but f) is good for simplified modeling if the trade-off between model simplicity, simulation speed and accuracy must be considered. It should be noted, however, that using

125

one single resistance, as in case b), results in the best improvement that can be achieved. More accurate AC measurements may also have given other results for topology e). According to [31], the c) model produces the worst fit. This result is consistent with my study. The single resistor network was as good as the simplified e) topology, according to [31], which again is somewhat consistent with my results. No real numerical values were given to be able to compare the fits of the different substrate networks.

Table 3. Relative S *parameter fit errors using different substrate resistance topologies.*

% / error	a)	b)	c)	d) and e)	f)
S_{11} mag (dB)	33.1	32.8	33.3	32.7	33.2
S_{11} phase	12.3	13.4	12.3	12.9	13.5
S_{12} mag (dB)	11.4	12.0	22.1	14.3	13.0
S_{12} phase	44.0	23.8	35.2	23.5	27.4
S_{21} mag (dB)	11.4	10.9	11.7	10.3	10.4
S_{21} phase	9.7	8.5	8.2	8.0	7.8
S_{22} mag (dB)	43.2	17.0	12.4	14.2	15.2
S_{22} phase	31.0	17.4	15.3	15.5	16.0
Tot. relative	**25.7**	**16.9**	**18.8**	**16.4**	**17.1**

Clearly, the largest effect of substrate network is seen at the output, and typical examples of *S* parameter fits are shown in Fig. 85 for the simple circuits of Fig. 83 a) and d). The device is a 4 x 25 µm x 0.8 µm NMOS transistor in the saturation region of operation. All four *S* parameters are shown, and in every plot there is the reference simulation without any substrate resistance network and the circuit of Fig. 83 d) that gave the best scalable fit. These two simulations are compared to the measured curves. The fit is not perfect, which is due to the scalability requirement. A single small signal equivalent circuit and a single device are easily characterized accurately, but this is not typically the case when trying to fit many devices at many bias points. Both circuits suffer from the bad DC accuracy, leading to the small inaccuracy at the low frequency

value of the S_{22} magnitude. The mere 0.2 dB inaccuracy at lower frequencies may even be due to a 50% conductance inaccuracy, explained in Chapter 1.2. DC model inaccuracy is also the reason for the S_{21} magnitude difference compared to the measurement at the low frequencies. The S_{12} accuracy in Fig. 85 b) is worse than typical MOS RF layouts and their models, and the reason for this is probably the less optimum layout for RF applications. There are not many parallel devices – only four in this case – as required for the smaller RC constant at the gate. The gate resistance is thus very large. Nevertheless, the improvement is seen both in magnitude, and especially in the phase of S_{12}. The S_{11} fit in Fig. 85 a) is also worse than typical MOS RF models, probably due to the less optimal gate structure. The advantage of using a substrate network is hardly seen. The most dramatic improvement achieved by using the more complicated substrate resistance network is seen in the S_{22} fit in Fig. 85 d). In the S_{21} fit in Fig. 85 c) the difference is somewhat noticeable, but smaller than in the S_{22} case.

In addition to the output impedance fit, improvement by using a substrate resistance network can also be seen in the output impedance real part. In Fig. 86 the output resistance of a 4 x 25 µm x 0.8 µm is shown when modeled with and without a substrate resistance network. The latter model uses the most accurate substrate resistance network of Fig. 83 d). Both simulation models suffer from an inaccurate DC value or low frequency R_{out} value, but the model with the substrate resistance network has the dominating pole at a lower frequency. A similar result can be seen in Fig. 87 for a smaller 4 x 12.5 µm x 0.8 µm NMOS device, where the difference is even more clearly in the favor of the model with substrate resistances.
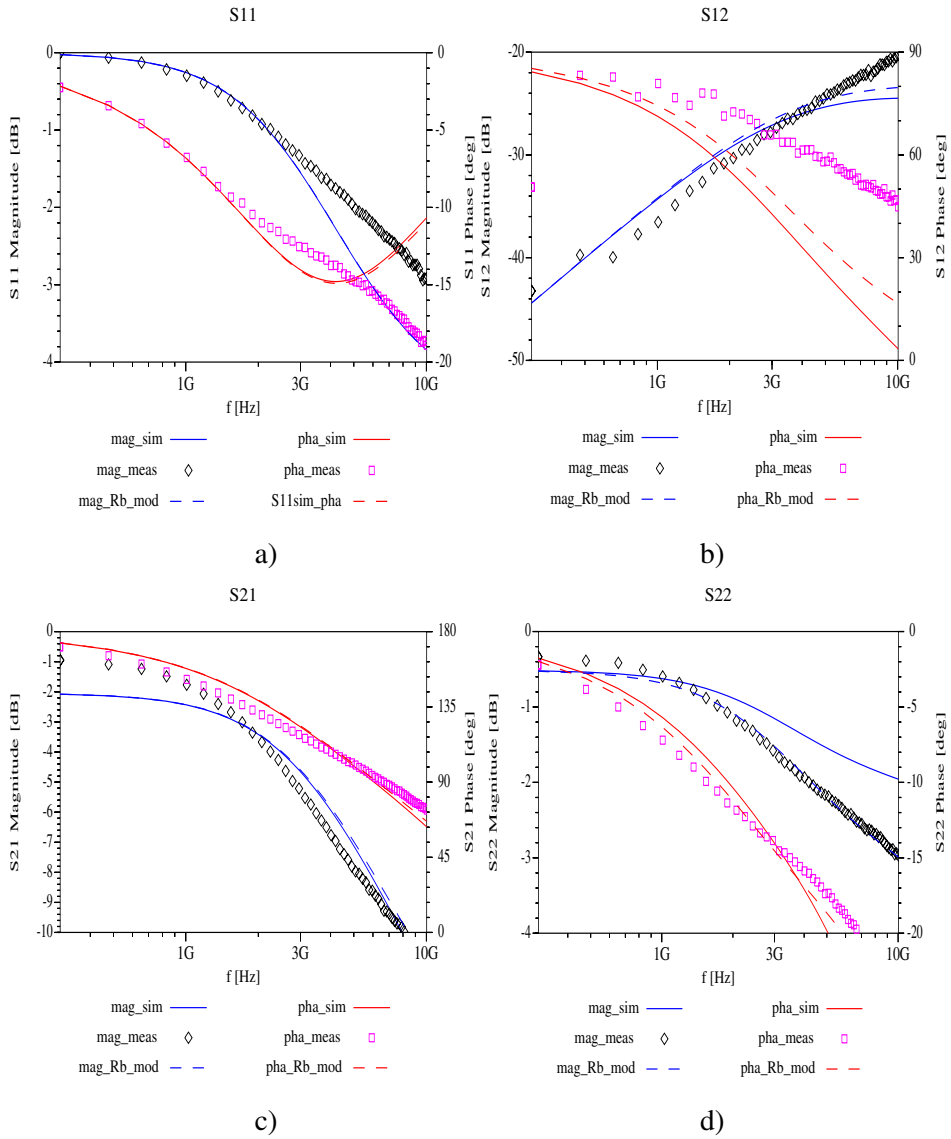
Figure 85. a) S_{11}, b) S_{12}, c) S_{21} and d) S_{22} magnitude and phase fits of a 4 x 25 μm x 0.8 μm NMOS device simulated with a simple model without any substrate resistors and with the equivalent circuit of Fig. 83 d).
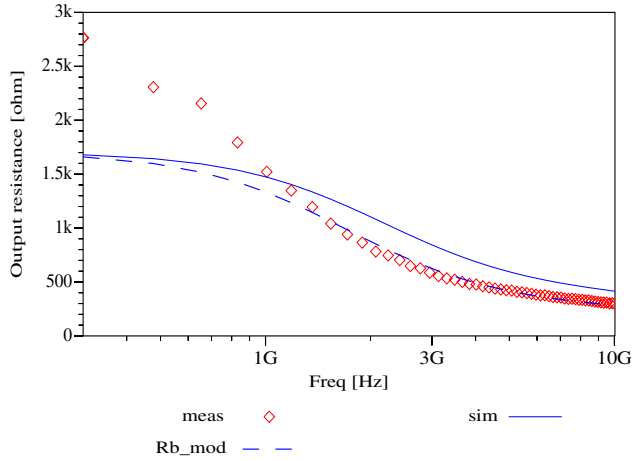
Figure 86. Output resistance of a 4 x 25 μm x 0.8 μm NMOS device in deep saturation modeled with and without a substrate resistance network.
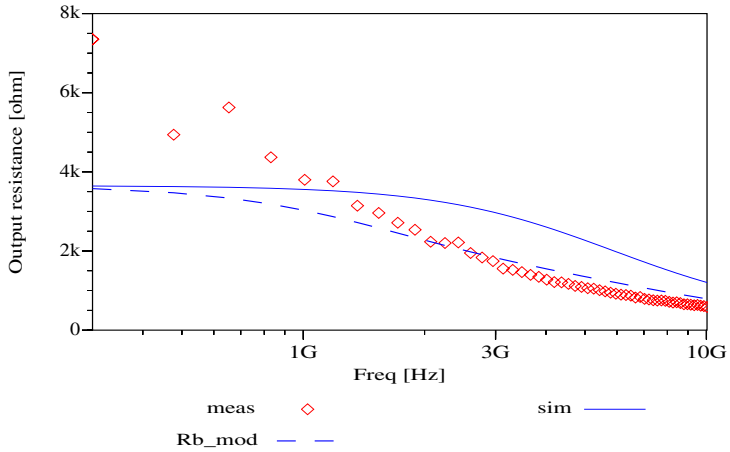
Figure 87. Output resistance of a 4 x 12.5 μm x 0.8 μm NMOS device simulated with and without a substrate resistance network. Both model fits are compared to the measurement.

The output capacitance fit defined with (77) is also improved. This can be seen for the two previous devices of Fig. 86 and Fig. 87. Their output capacitance fits are shown in Fig. 88 and Fig. 89 with and without a bulk resistance network. The bias point is the same, and again the the simple simulation does not take any substrate resistance into account, whereas the other uses the best substrate equivalent circuit of Fig. 83 d). The fit suffers from some inaccuracy but still shows the advantage of using the substrate resistance model, especially in the case of the larger device. In the smaller device in Fig. 89, the fit is not that much better quantitatively when using the bulk resistance network, but qualitatively the shape of the model resembles the measured curve more closely. An interesting side note in the capacitance plots is that the low-frequency output capacitance value is different depending on the substrate resistances. This is probably exaggerated in these plots due to the less optimal RF device layout. There are only four parallel devices in the whole transistor, resulting in a very high gate resistance.
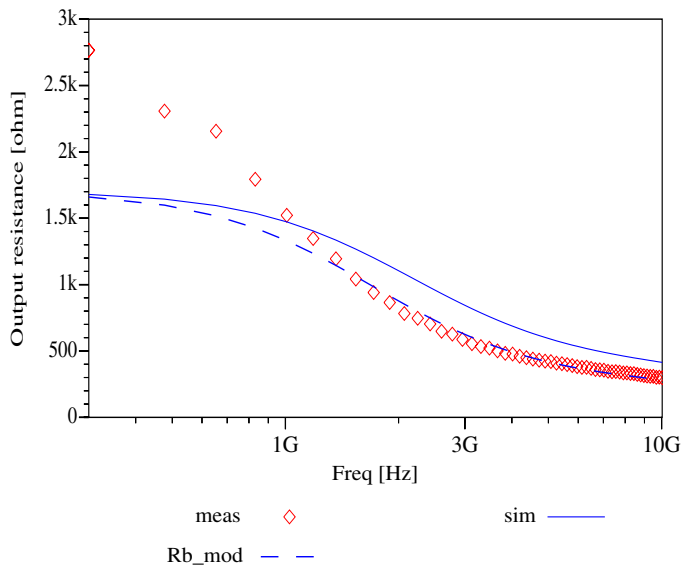
Figure 88. Output capacitance fit of a 4 x 25 μm x 0.8 μm NMOS device with and without the substrate network model.
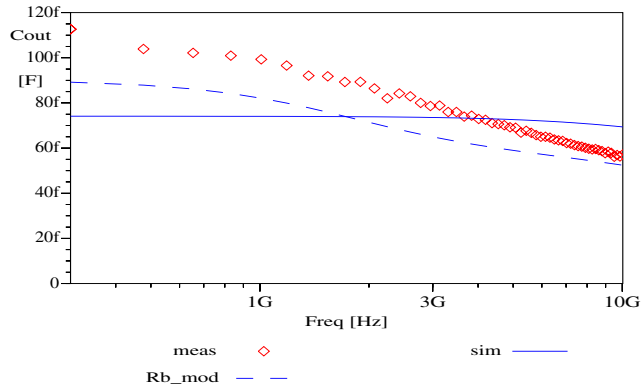
Figure 89. Output capacitance fit of a 4 x 12.5 μm x 0.8 μm NMOS device with and without the substrate network model.

With the simplified output equivalent circuit of Fig. 90, the output impedance real part can be approximated and qualitatively examined. In Fig. 90 the substrate effect is very simply added to the circuit by loading the intrinsic drain node with a series connection of R_B and C_{bd}. This approximation does not accurately take into account the more complex couplings of the substrate node, but gives a qualitative understanding of the output resistance behavior.
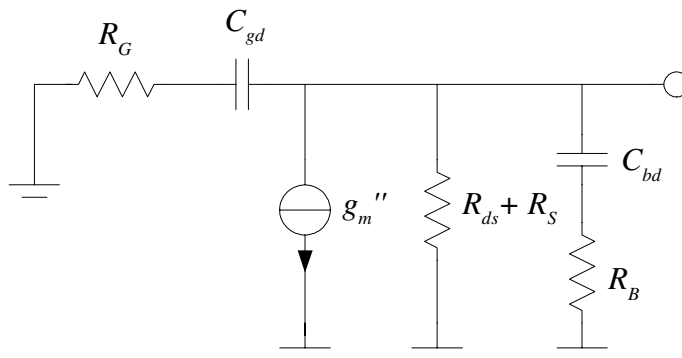
Figure 90. Simplified output impedance model for the qualitative approximation of output resistance when a substrate network is added to the equivalent circuit.

For the simple case without the R_B bulk resistance, the output resistance according to equation (76) would be

$$R_{out,\,Rb=0}(\omega) = \frac{(R_{ds}+R_S)(1+\omega^2 C_{gd}{}^2 R_G)}{(1+\omega^2 C_{gd}{}^2 R_G (R_G + R_{ds}+R_S))} \tag{97}$$

Output resistance for the complete equivalent circuit of Fig. 90 can be determined as:

$$R_{out,\,Rb}(\omega) = \frac{(R_{ds}+R_S)(1+\omega^2 C_{gd}{}^2 R_G{}^2)(1+\omega^2 C_{bd}{}^2 R_B{}^2)}{(1+\omega^2 p_1{}^2)(1+\omega^2 p_2{}^2)} \tag{98}$$

Here, p_1 and p_2 are the poles, which are rather complex, determined from the general solution for second order equation

$$p_{1,2} = \frac{1}{2\pi}\sqrt{\frac{-(-b)\mp\sqrt{b^2-4ac}}{2a}} \tag{99}$$

Here, the a, b and c terms are defined as

$$a = C_{gd}{}^2 C_{bd}{}^2 R_G R_B (R_G R_B + (R_G + R_B)(R_{ds}+R_S)) \tag{100}$$

$$b = C_{gd}{}^2 R_G (R_{ds}+R_S+R_G) + C_{bd}{}^2 R_B (R_{ds}+R_S+R_B) \tag{101}$$

and

$$c = 1 \tag{102}$$

Comparing equations (97) and (98), it is evident that their behavior must be inherently different although the low frequency value is the same. If we consider a numerical example with $R_G = 10\ \Omega$, $R_B = 200\ \Omega$, $R_{ds} + R_S = 350\ \Omega$, $C_{gd} = 200$ fF and $C_{bd} = 320$ fF, we get a pole of 13.3 GHz for the output resistance of equation (97). These values correspond to a typical 40 x 10 µm x 0.6 µm NMOS device in the linear region, which is practical in RF design. Taking the R_B effect into account using equation (98), we get a dominant pole of 1.5 GHz which is almost 12 GHz lower than the case with R_B being zero. Not only are the poles affecting the frequency dependence, but there are zeros as well. Of course, there is a second pole in (98) which is at over 20 GHz. With the example values the zero in (97) is 122 GHz, whereas (98) has an additional zero of 10 GHz. Being

at a much lower frequency, its effect is very important in determining the output resistance AC fit.

Considering again the single bulk resistor network, we can examine the pole frequency as a function of R_B value. Using the typical values of the RF device of the previous example, we can plot the dominant pole and zero as a function of frequency using equation (99) for the pole calculation and the zero of (98). The plot is shown in Fig. 91. It can be seen that the dominant pole decreases rapidly for this device size when the bulk resistance value is increased. The error in the simple equivalent circuit is zero in the case where R_B is zero, but increases rapidly as soon as a small bulk resistance is present.
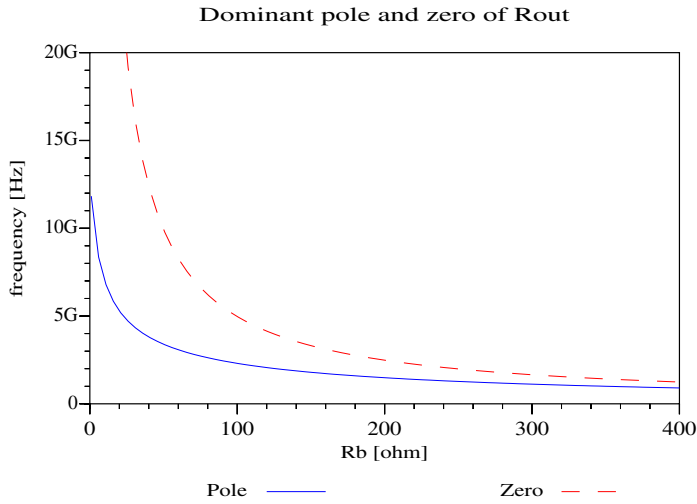
Figure 91. The dominant pole and zero frequency of the output resistance as a function of bulk resistance value.

Similar qualitative approximation of the output capacitance (defined by (77)) behavior can be done. Calculating Y_{out} from Fig. 90 we get

$$Y_{OUT}(s) = \frac{1}{R_{ds}+R_S} + \frac{sC_{gd}}{1+sC_{gd}R_G} + \frac{sC_{bd}}{1+sC_{bd}R_B}$$

(103)

From this the capacitance can be determined as:

$$C_{OUT}(\omega) = \frac{C_{gd}}{1 + \omega^2 C_{gd}^2 R_G^2} + \frac{C_{bd}}{1 + \omega^2 C_{bd}^2 R_B^2} \qquad (104)$$

If R_B is zero, the pole in the C_{bd} term disappears, resulting in the simple model output capacitance of

$$C_{OUT}(\omega) = \frac{C_{gd}}{1 + \omega^2 C_{gd}^2 R_G^2} + C_{bd} \qquad (105)$$

This simple comparison suggests that the output capacitance frequency behavior is affected by the bulk resistance. In this simple case the low frequency value is equal for both equations, being just the sum of C_{gd} and C_{bd}. This is probably a more common case in layouts suitable for RF devices than in the device fits of Fig. 88 and Fig. 89, as their layouts have a very large gate resistance value.

AC accuracy comparison of substrate resistance networks should be done up to frequencies as high as 20 GHz. Although my data were measured up to 20 GHz, the measurement uncertainty deteriorated the practical comparison. This may be due to a less accurate de-embedding or calibration. Although no bias dependence was considered, the simple resistance approach in the substrate networks did have a clear effect on the AC fit accuracy. This might be due to measurement uncertainty or the bias points of the *S* parameter files being mostly in the saturation region of operation.

7.2 PD SOI CMOS

The possible body tie effect of PD SOI technology was studied by making a similar AC fit with the Honeywell process as with the VTT technology. The devices were not designed by us and they were not that practical for a theoretical bulk coupling study. However, the devices had a constant finger size and only the number of parallel devices varied. Due to isolating trenches there were small substrates only beneath the gates, simplifying the situation quite a lot. Thus, it could be considered that there was a constant bulk resistance between the device substrate and the body tie contact.

A scalable AC fit was optimized for the set of extraction devices presented in Chapter 3.1.2, using the BSIM3 and EKV models with an equivalent circuit having only one bulk resistance. Table 4 shows the total AC fit as relative *S*

parameter errors, with and without using a substrate resistance. The largest difference between the two cases is in the S_{12} phase error, similar to the CMOS case in the previous chapter. The magnitude fits of S_{21} and S_{22} also improve, as in the fit comparison of VTT CMOS devices. The total relative error difference is still rather small, as expected, as the SOI technology should not have such a large substrate resistance as bulk technologies. The extracted bulk resistance value for one finger is about 6 kΩ as, according to Honeywell documentation, the body tie resistance should be around 5 kΩ.

Table 4. Substrate resistance effect on the fit in Honeywell PD SOI Technology using BSIM3 model.

% / error	$R_B = 0$	R_B used
S_{11} mag (dB)	14.6	14.8
S_{11} phase	64.8	65.2
S_{12} mag (dB)	11.3	10.8
S_{12} phase	52.6	40.7
S_{21} mag (dB)	104.6	101.6
S_{21} phase	6.66	6.84
S_{22} mag (dB)	44.1	41.3
S_{22} phase	13.9	13.7
Tot. relative	**55.5**	**53.5**

An example fit change of the S_{22} parameter for a 101 x 10 μm x 0.35 μm device is shown in Fig. 92 a) using the BSIM3 model and one single bulk resistor. In this particular fit the substrate resistance has quite a clear effect, improving the model accuracy. Another S_{22} fit example is shown in Fig. 92 b) for a 101 x 10 μm x 0.4 μm with quite similar behavior. In both fit examples the devices have been biased in the saturation region. In both cases the S_{22} magnitude fit is more accurate above 3 GHz. Somehow, the S_{22} phase fit worsens in both cases, although Table 4 suggest the situation should be the contrary.
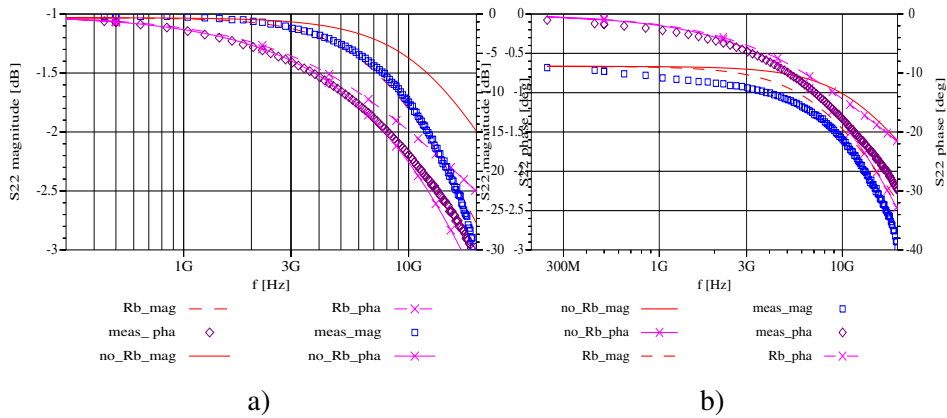
Figure 92. S_{22} fit of a) 101 x 10 μm x 0.35 μm and b) 101 x 10 μm x 0.4 μm Honeywell PD SOI NMOS transistors modeled with and without a substrate resistor.

7.3 SOS Device Characteristics

In the AC fit comparison of Peregrine SOS devices, the substrate resistance was used to see whether it improves the fit or not. In theory, there should be no effect caused by the substrate as the silicon film is only tens of nanometers and the substrate is lossless sapphire. Again, the BSIM3 model was used with a single substrate model.

In Table 5 the relative *S* parameter errors are shown for the two fit cases. One uses the substrate resistor and the other fit uses a model without any substrate resistance. 40 different Touchstone™ files were used as fit goals measured from 10 devices. Although the total relative error difference between the two modeling cases is possibly within measurement accuracy, the difference in the S_{22} magnitude seems to be clear. The difference is not large but visible in many cases, as can be seen in Fig. 93 a) S_{22} fit for a 4 x 20 μm x 0.5 μm device and in Fig. 93 b) S_{22} fit for a 8 x 5 μm x 0.5 μm device. There is almost a 0.2 dB difference at 10 GHz between simulations of the two modeling approaches in the former example in Fig. 93 a). In the latter example, the absolute dB error of

S_{22} is smaller across the whole bandwidth. In both device fits the S_{22} phase is more accurate when the bulk resistance value is zero, although the overall fit comparison in Table 5 suggests the contrary. A couple of S parameter datasets that had a bad DC fit probably distorted the optimization.

Table 5. Peregrine SOS transistor fits using BSIM3 model with a single bulk resistance.

% / error	$R_B = 0$	R_B used
S_{11} mag (dB)	57.2	57.5
S_{11} phase	6.1	6.1
S_{12} mag (dB)	25.6	26.0
S_{12} phase	14.0	14.5
S_{21} mag (dB)	71.1	70.9
S_{21} phase	8.7	9.0
S_{22} mag (dB)	25.5	23.7
S_{22} phase	13.8	13.5
Tot. relative	**27.8**	**27.7**

a)   b)
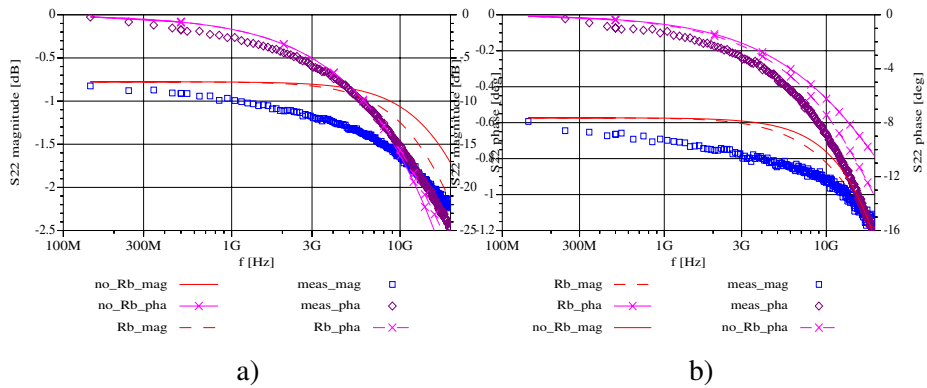
Figure 93. Using and not using a single substrate resistance in MOSFET equivalent circuit of Peregrine sapphire substrate devices for a) a 4 x 20 μm x 0.5 μm NMOS and b) a 8 x 5 μm x 0.5 μm NMOS device in the saturation region of operation.

According to these AC fits, the differences between the two cases seem quite subtle, suggesting that the sapphire substrate is indeed isolating. Yet, the output fit of the MOS model improves when using the substrate resistance, possibly indicating that either the output capacitance measurement accuracy is less than perfect, or that the measurement uncertainty is causing the seemingly unrealistic SOS behavior. The extra freedom gained by the more complex model possibly enables the better fit, although the situation the model is describing should be unphysical. The only real purpose of this SOS device fit comparison using bulk resistance models was to see if the bulk CMOS comparison is valid. According to these comparisons, it seems that in the bulk CMOS case the improvement using substrate resistance networks seems quite clear.

8. Other Improvements to the MOSFET Equivalent Circuit

8.1 Additional Gate Capacitance in Parallel with the Gate Resistance

One possible modification to the equivalent circuit is to add a capacitance parallel to the gate resistance [28], as is shown in Fig. 94. The authors in [28] intended the model to describe the distributed gate behavior of a wide transistor. The C_g capacitance was added to the small signal equivalent [28] circuit by making a transmission line approximation of the transistor gate [108]. Some suggestions to model the polysilicon depletion effect [109] with this equivalent circuit have also been presented. This may not work in that situation either, as the polysilicon depletion effect is not frequency-dependent, but gate bias-dependent. The extra capacitance caused by polysilicon depletion is in series with the oxide capacitance, typically being over ten times larger. This results in a 5–10% decrease in effective gate capacitance.

In [28] the modification of Fig. 94 was used to improve the input resistance model accuracy, as the measured device showed a decrease in input resistance as a function of frequency. Nothing was said of the device size where this C_g model is applicable. A capacitance value of one-fifth of the total input capacitance was suggested, based on theoretical equivalent circuit calculations.
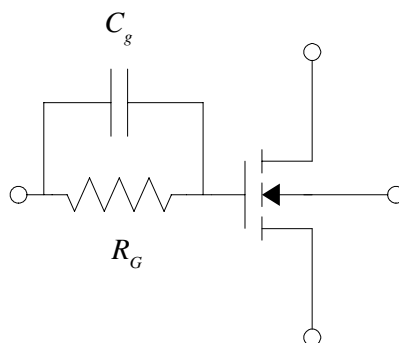
Figure 94. An improvement to the equivalent circuit suggested in [28]. An additional C_g capacitance is added in parallel with R_G.

Two AC fit comparisons were made. The first fit was made using the basic extraction set of the January 1999 run of the VTTB6 technology presented in Chapter 3.1.2. The BSIM3 model was used as the basic model onto which the extrinsic C_g capacitance was built. The other AC fit was made using wide single-finger transistors from the 1997 run of the VTTB6 technology. In order to make the C_g capacitance scale with the device size area and perimeter, terms were used to determine C_g as

$$C_g = [W_f L C_{GA} + 2(W_f + L)C_{GP}]n_f \qquad (106)$$

C_{GA} and C_{GP} are the area and perimeter terms, respectively. W_f is the width of one finger and n_f is the number of parallel fingers.

8.1.1 AC Fit Comparison with Multifinger Devices

In contrast to the previous purely theoretical research reported in [28], no real improvement was achieved using the C_g capacitance, although test devices were used with wide finger widths and small numbers of parallel devices. A scalable fit was trialled, and the total C_g values were varied from a few fF's up to 10 pF. Thus C_g values from $C_{in}/20$ to $20C_{in}$ were trialled. I suspect much of the input resistance behavior is due to measurement inaccuracy, as explained in Chapter 4.2.

24 Touchstone™ data files measured from 12 devices were used as fitting goals. The process was the January 1999 run of the VTTB6 technology. S parameter fits of a basic model implementation compared to a model with the additional C_g capacitance are presented in Table 6. The total relative error is smaller in the model with the additional C_g capacitance. The S_{11} and S_{12} phase fits especially are slightly better . The difference is small and the fit of the C_g model is even worse in the S_{21} and S_{22} case compared to the basic BSIM3 model Aplac implementation.

Table 6. S parameter fit of the basic BSIM3 model compared to the model with an additional C_g capacitance. AC extraction devices were multifinger transistors.

% / error	Basic BSIM3	C_g added to the model
S_{11} mag (dB)	69.5	66.5
S_{11} phase	24.0	20.7
S_{12} mag (dB)	5.1	5.2
S_{12} phase	25.1	20.5
S_{21} mag (dB)	13.7	15.8
S_{21} phase	14.5	14.9
S_{22} mag (dB)	75.5	75.9
S_{22} phase	11.2	11.2
Tot. relative	**29.8**	**28.9**

Most of the improvement achieved by the C_g capacitance is due to the better S_{11} accuracy. This can also be seen in examples of the input resistance fits of the two models. In Fig. 95 a) a 4-finger 200 µm x 1.0 µm NMOS device is fitted both with the basic BSIM3 Aplac implementation and with the C_g model. Fig. 95 b) shows the respective plots of a device with more parallel fingers, the measures being 100 µm x 0.6 µm divided into 16 fingers. First of all, the gate polysilicon resistivity is quite high, being about 70 Ω/square, which results in high input resistances. In Fig. 95 a) the input resistance is about 400 Ω and decreases with increasing frequency. This is, with some inaccuracy, qualitatively modeled correctly with the C_g capacitance, and it is definitively better than the basic BSIM3 implementation fit. According to Fig. 34 and Fig. 36, the input resistance uncertainty in this case is about 50% at 500 MHz and 20% at 2 GHz. The respective uncertainties for the input capacitance according to Fig. 35 and Fig. 37 are about 20% at 500 MHz and 30% at 2 GHz. It seems that the input resistance behavior is real and not just a result of measurement uncertainties. For the smaller device with four times more parallel fingers in

Fig. 95 b), the fit does not practically change between the two models compared. As expected, the input resistance is much smaller in the device with many parallel fingers, being over 20 Ω. Both measurements suffer from the R_{in} extraction uncertainty explained in Chapter 4.2. The fit errors between measurement and simulation are also quite large, possibly due to the extraction uncertainty and possibly due to the challenge of fitting a scalable model and not just a simplified small signal equivalent circuit.
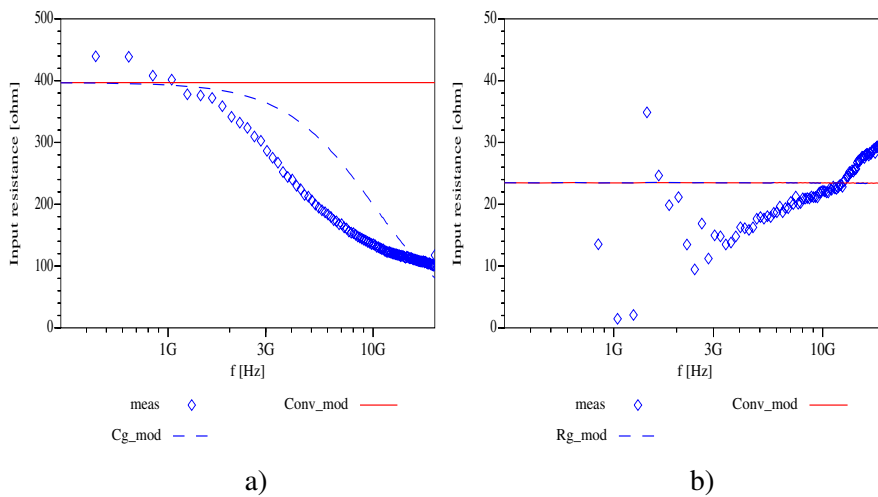
Figure 95. Input resistance fit with and without the C_g capacitance for two device sizes. a) Four finger 4 x 50 µm x 1.0 µm NMOS device and b) a 16 finger 16 x 6.75 µm x 0.6 µm NMOS device.

The input capacitance fits of the two previous devices show a similar behavior where the effect of the C_g capacitance is negligible or smaller in the device with many parallel devices. In the larger 200 µm x 1.0 µm device with only four parallel fingers in Fig. 96 a), the input capacitance decreases as a function of frequency, whereas it stays quite constant in the other case. For the 16-finger device in Fig. 96 b) there is hardly any difference in C_{in}, whether the C_g capacitance is present or not.

a)   b)
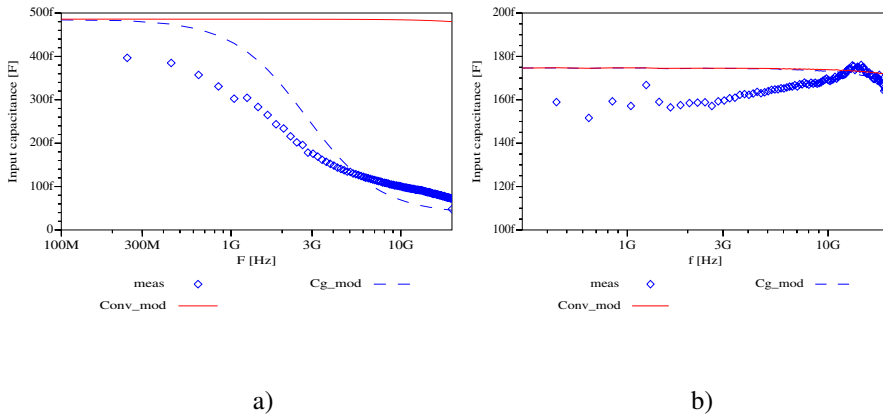
Figure 96. Input capacitance fit with and without the C_g capacitance for two device sizes. a) a four-finger 4 x 50 μm x 1.0 μm NMOS device and b) a 16-finger 16 x 6.75 μm x 0.6 μm NMOS device.

8.1.2 AC Fit Comparison with Single-Finger Devices

The extraction set of single-finger devices included six transistors with two widths: 100 and 50 μm. For both widths there were three channel lengths: 0.6 μm, 0.8 μm and 1.0 μm. Two Touchstone™ files were measured at a two different bias points (in the linear and saturation regions) for all devices. In total there were 12 Touchstone™ datafiles as optimization goals in AC extraction. All devices showed quite strong input resistance, input capacitance and transconductance frequency dependence. A comparison of the AC fits with and without the C_g capacitance is shown in Table 7. The relative S parameter fit errors are shown for both cases and only the S_{22} magnitude has slightly worsened after using the C_g capacitance. All other parameters have improved or been constant. Again, the most dramatic improvement is achieved in the S_{11} fit.

143

Table 7. S parameter fit of the basic BSIM3 model compared to the model with an additional C_g capacitance. AC extraction devices were single finger transistors.

% / error	*Basic BSIM3*	*C_g added to the model*
S_{11} mag (dB)	42.4	27.8
S_{11} phase	22.2	7.7
S_{12} mag (dB)	9.6	9.6
S_{12} phase	46.1	30.9
S_{21} mag (dB)	8.9	6.8
S_{21} phase	9.9	7.5
S_{22} mag (dB)	11.8	13.7
S_{22} phase	9.9	9.5
Tot. relative	**20.1**	**14.2**

A typical example of the input impedance fit improvement can be seen in the input resistance and capacitance fit of Fig. 97 a) for a single-finger 100 μm x 0.8 μm device. Although the input resistance fit improves radically by connecting a C_{in}/5 capacitance in parallel, the approximation is at best below 3 GHz in this case. Above that frequency the model underestimates the input resistance by short circuiting it. There seems to be a 500–600 Ω constant level at 10 GHz and above that cannot be described by the model. Still, the relative improvement is radical compared to the conventional model without any C_g capacitance. The C_g model was verified theoretically only up to 5 GHz in [28], and it seems that the model loses its accuracy even below that frequency in this case. The input capacitance case in Fig. 97 b) suggests that the C_g model describes the distributed gate behavior of the input capacitance quite sufficiently up to 20 GHz. At lower frequencies the errors are relatively large, but this is probably due to the R_{in} extraction uncertainty described in Chapter 4.2.
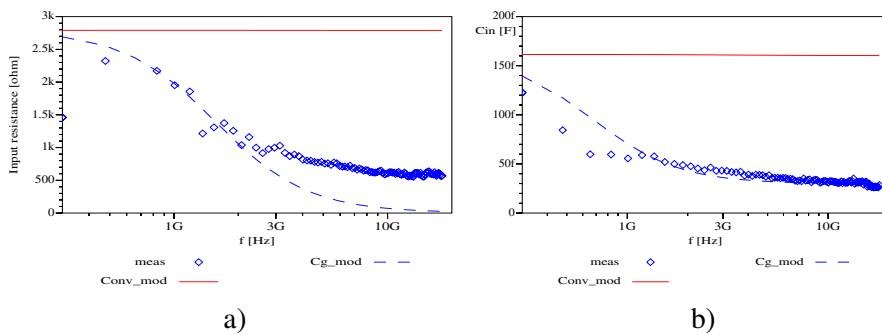
Figure 97. a) Input resistance and b) input capacitance fits of single-finger 100 μm x 0.8 μm MOSFET.

In contrast to the *S* parameter fit of multifinger devices, the S_{21} magnitude error decreases which can be seen by comparing Tables 6 and 7. This is partly due to the improved fit of the real part of the y_{21} parameter. This is presented in Fig. 98 by comparing the transconductance fit of the conventional BSIM3 model to the C_g modified model. The low frequency fit below 3 GHz is clearly improved, whereas the better model at higher frequencies is still questionable. It seems that using the small signal value of $C_{in}/5$ in parallel with R_G does not improve the AC behavior at very high frequencies, but works fine at lower frequencies.
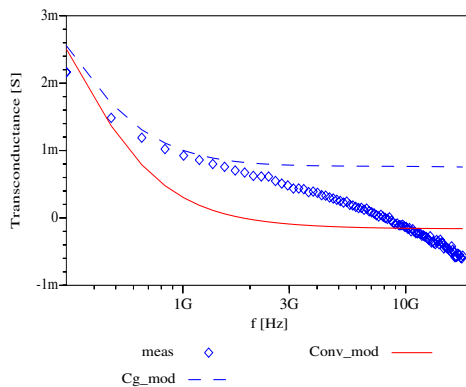
Figure 98. Transconductance fit of a single-finger100 μm x 0.8 μm MOSFET using the conventional BSIM3 model and the model with an additional parallel C_g capacitance.

8.1.3 Usability of the Additional C_g Capacitance

In wide multifinger RF devices the C_g capacitance would seem to be useless, as the impedance of the gate resistance is much smaller compared to the C_g capacitance. If we compare the transition frequency where the C_g and R_G absolute impedance values are equal between a conventional RF device with multiple parallel fingers to a single-finger device, we get dramatic differences, as can be seen from the transition frequency equation

$$f_{t,gate} = \frac{n_f{}^2}{2\pi R_G C_g} \tag{107}$$

where n_f is the number of parallel devices or fingers and R_G is the gate resistance in the single-finger case. A quantitative example with a 400 µm x 0.6 µm NMOS as a 1 and 40 finger layout would give transition frequencies of 240 MHz and 380 GHz respectively. A fifth of the input capacitance in this case is typically around 100 fF, and the sheet resistance of the polysilicon gate is typically around 10 Ω/square. It is easily understood that the effect of C_g is negligible in an RF device operating up to tens of GHz. This is why the curves of the conventional BSIM3 model and the C_g added model do not differ a lot in Fig. 95 b) and Fig. 96 b).
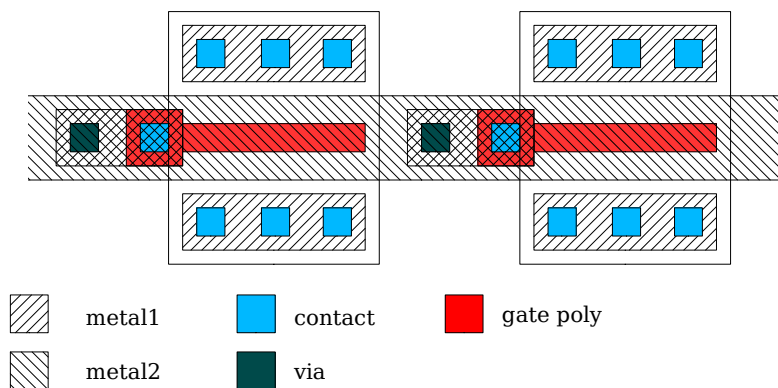
Figure 99. MOS layout causing parasitic capacitance between gate polysilicon and metal 2.

The reason for the out of the ordinary input behavior in this study seems to be partly the result of the distributed gate effect and partly that of the parasitic capacitances between the gate polysilicon and the gate metalization. The situation is depicted in Fig. 99 with the layout type chosen for the scalable substrate resistance study. It is intuitively understood that a parasitic capacitance is formed between the second metal and the gate poly. It is interesting to note that the resulting fit value for C_{GA} of 0.044 fF/µm^2 in the scalable model comparison is the same magnitude of order as the area capacitance value between metal 2 and the gate polysilicon, for which the theoretical value is 0.024 fF/µm^2. These numbers are more than two or three times smaller than the suggestions in [28], which was one-fifth of the MOSFET input capacitance value. According to a theoretical study in [32], the metal-over-gate structure would result in a parallel C_g capacitance with a capacitance value slightly above the metal-to-gate capacitance. It seems that the C_g capacitance has modeled parasitic metal-over-gate couplings in the scalable device set comparison. Comparing the capacitance values, the $C_{in}/5$ value is much larger than parasitic capacitances, and thus should make the former effect at the gate stronger.

Thus, in some cases the model of Fig. 94 could be practical for modeling parasitic couplings. C_g is needed if, for some reason, a MOS device is designed as a single-finger layout or very few parallel devices and a metal runs above the gate and they are connected to the same node, similar to the case in Fig. 84. The metalizations result in parasitic capacitances from the gate, as shown by the cross section in Fig. 100. Not only is the C_g capacitance introduced but also extra overlap capacitances from the gate to drain and source. These extra overlap capacitance values are about one-tenth in the VTTB6 technology compared to the intrinsic overlap capacitances and have quite an important effect on modeling the MOSFET AC behavior.
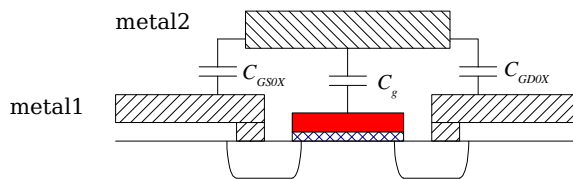
Figure 100. Cross section of MOSFET with parasitic capacitances resulting from metalizations.

Finally, the C_g capacitance in parallel with the gate resistance can only be used in small signal simulation if it is determined as one-fifth of the input capacitance [28]. The C_g capacitance value is voltage-dependent and cannot be used in conventional charge-based device models. Nevertheless, it is possible to use the MOGFET approach [32] in a large signal model, as the capacitance value is not bias-dependent and just models parasitic couplings.

8.2 Distributed Gate Model

The suggestion of distributing the gate is quite an old one [110]–[112]. The idea is to better describe the *RC* delay at the gate by dividing the device model into many parallel devices and by putting gate resistances between them, as in Fig. 101. In this study the wide MOSFET was divided into nine subdevices, as in Fig. 101. The width of all devices from M2 to M8 is the total width divided by eight. M1 and M9 is just half of the size of the other devices being thus

$$W_1 = W_9 = \frac{W_{tot}}{16} \tag{108}$$

8.2.1 AC Fit of the Distributed Gate Model Using Single-Finger Devices

First, an accurate scalable DC model was extracted using the Berkeley SPICE Level 3 model. The same set of AC extraction devices of the September 1997 VTT 0.6 μm CMOS run was used, as in the BSIM3v3 extraction of Chapter 3.1.2 in preliminary AC extraction. After the basic DC and AC extraction, the measurement devices were switched to six single-finger devices of sizes 50 μm x 0.4 μm, 50 μm x 0.6 μm, 50 μm x 0.8 μm, 100 μm x 0.4 μm, 100 μm x 0.6 μm and 100 μm x 0.8 μm. Then the conventional Level 3 AC fit of these single-finger devices was compared to the fit of the distributed-gate model, with only a small optimization of the AC models. Gate resistance values of the distributed model were determined similarly to the device widths, being

$$R_{G1} = R_{G9} = \frac{R_{Gtot}}{16} \tag{109}$$

for the first and last subdevice and the rest were determined as

$$R_{Gi} = \frac{R_{Gtot}}{8} \qquad (110)$$

Here, R_{Gtot} is the total gate resistance of the polysilicon. It should be observed that the effective gate resistance of the conventional MOS model is defined by

$$R_G = \frac{R_{Gtot}}{3} \qquad (111)$$
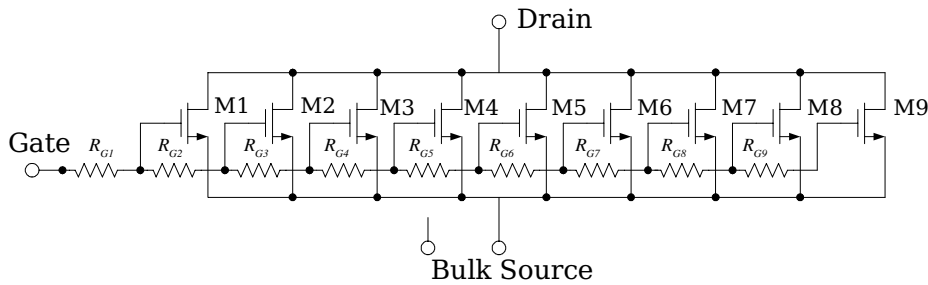
Figure 101. Distributed-gate model to account for the realistic RC delay in a wide device.

The Level 3 model was chosen because its model parameters do not have width dependences, in contrast to the newer-generation models like EKV, BSIM3 and MOS Model 9. In the newer models dividing a large MOS circuit model into smaller subcircuits affects the DC model and results, for instance, in a different threshold voltage and different current and charge characteristics. Practically, the effect is small when considering devices much wider than the minimum gate width allowed by the semiconductor technology. For simplicity, the Level 3 model was used. In the Level 3 model one only needs to ensure that the effective width of the distributed transistor is equal to the width of the original device.

In Table 8 shows the conventional Level 3 S parameter fit of the six single-finger devices compared to the fit of the distributed-gate model. All devices have one Touchstone™ file associated with them, measured in the saturation region of operation. The relative magnitude and phase errors are shown for all S parameters, and it seems that the distribution of the gate model improves dramatically all other S parameter fits, except for the S_{22} fit, where the worsening is very subtle. Dramatic improvements are achieved in the S_{11} and S_{21}

fits. It seems that the proper RC delay model of the input especially improves the S_{11} fit a lot.

Table 8. The S parameter fit comparison of the distributed-gate model to the conventional Level 3 model using single-finger devices.

% / error	*Non distributed gate model*	*Distributed gate model*
S_{11} mag (dB)	50.0	28.6
S_{11} phase	39.4	18.5
S_{12} mag (dB)	15.0	9.5
S_{12} phase	82.8	69.5
S_{21} mag (dB)	14.3	2.4
S_{21} phase	13.9	4.7
S_{22} mag (dB)	37.4	39.7
S_{22} phase	26.9	29.0
Tot. relative	**35.0**	**25.2**

If we study the input resistance difference between the conventional lumped gate resistance model and the distributed-gate resistance model of a 50 µm x 0.8 µm MOS device, we see the dramatic improvement in Fig. 103 a). The input resistance values are such that the decrease is unlikely to be a result of measurement uncertainty. It seems that the distribution of the gate resistance is able to describe this decreasing input resistance behavior. The transconductance behavior in Fig. 103 b) is also quite different from the respective curves of a conventional RF MOS device. Still, the distributed gate model seems to describe this behavior as well. Quite similar results are obtained with the other single-finger devices, as can be seen for a 100 µm x 1.0 µm device in Fig. 102.

Figure 102. a) Input resistance and b) transconductance fit comparison of the lumped-gate resistance model to the distributed-gate resistance model. The device is a single-finger device with a total width of 100 μm and a length of 1.0 μm.
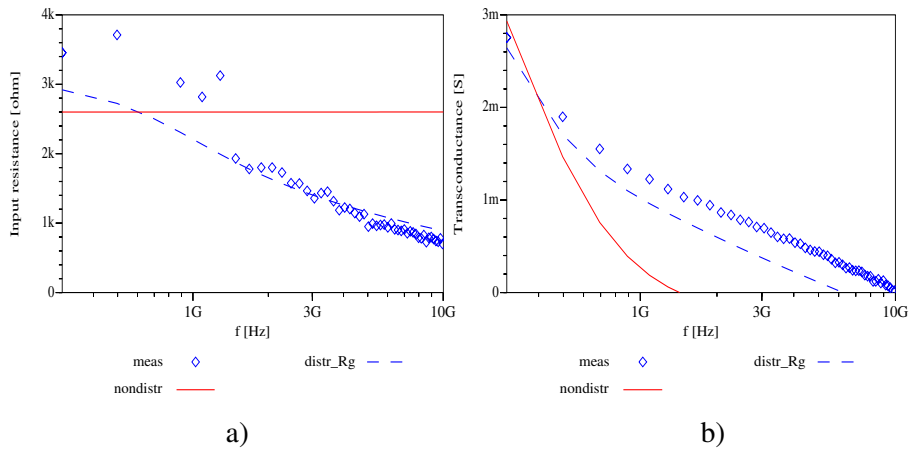

Figure 103. a) Input resistance and b) transconductance fit comparison of the lumped gate resistance model to the distributed gate resistance model. The device is a single finger device with a total width of 50 μm and a length of 0.8 μm.

8.2.2 AC Fit of the Distributed-Gate Model Using Four-Finger Devices

If the distributed-gate model is trialled with the device sizes of the basic AC extraction set used for this process run, the difference is smaller, as can be seen in Table 9. The device sizes used for this comparison fit were 50 µm x 0.8 µm, 50 µm x 0.6 µm, 100 µm x 0.8 µm and 100 µm x 0.6 µm, all having four parallel fingers. The bias points are chosen in the saturation region of operation and a total of six different Touchstone™ files are used for fit comparison. The simulation models consisted of 36 parallel devices, where each four-finger distributed gate had nine parallel devices. Practically, the equivalent circuit had four circuits of Fig. 101 in parallel. The improvements are rather subtle in all S parameter fits, except for the S_{12} fit where the fit is slightly worse compared to the lumped-resistance model. The total relative error decreases just 2% compared to the almost 10% improvement in single-finger devices. The 50 µm-wide devices especially have a very small change in AC behavior when switched from the lumped-resistance model to the distributed-gate model. Again, this can easily be seen from the input resistance and transconductance fits in Fig. 104. The transconductance behavior is similar to conventional multifinger RF devices, but still the distributed-gate model is slightly better in accuracy than the lumped-gate-resistance model. The difference of the model accuracy improvement compared to the two previous transconductance examples of single-finger devices seems to be rather small. Just four parallel devices deteriorate the advantage of the distributed-gate model; thus it is probably not useful at all for multifinger RF devices. This is especially true when considering the simulation efficiency with a lot of complex MOS devices.

Table 9. The S parameter fit comparison of the distributed-gate model to the conventional Level 3 model using four-finger devices.

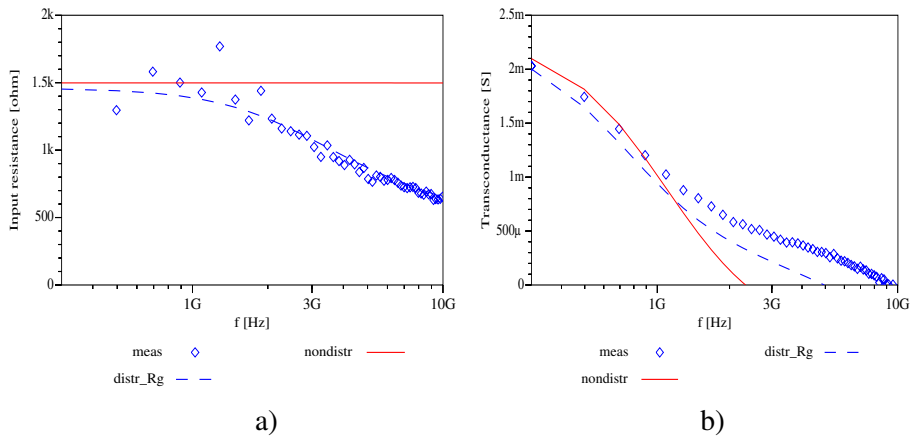| *% / error*	*Non distributed gate model*	*Distributed gate model*
S_{11} mag (dB)	49.9	44.9
S_{11} phase	59.0	51.5
S_{12} mag (dB)	15.0	16.5
S_{12} phase	35.0	33.8
S_{21} mag (dB)	7.9	7.0
S_{21} phase	5.6	13.6
S_{22} mag (dB)	64.3	63.5
S_{22} phase	20.7	20.0
Tot. relative	**32.2**	**30.1**

a)   b)
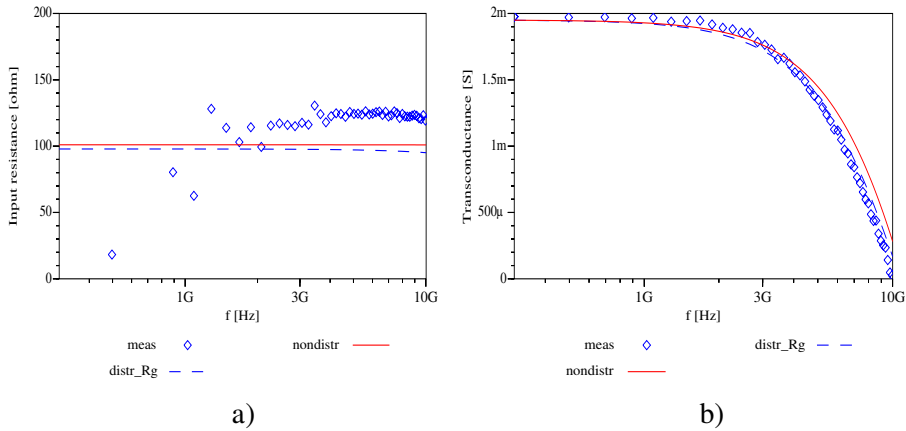
Figure 104. a) Input resistance and b) transconductance fit comparison of the lumped-gate-resistance model to the distributed-gate-resistance model. The device is a four-finger device with a total width of 50 μm and a length of 0.8 μm.

The comparison results suggest that distributing the gate does improve the AC accuracy of the MOS models. According to these results, this method seems to be more accurate than modeling the distributed gate effect by the parallel C_g capacitance. By comparing the input resistance and the transconductance fits achieved with these different approaches, it seems that the C_g capacitance approximation fails to describe the behavior above 3 to 5 GHz. Qualitatively, the latter approach of distributing the gate with many parallel devices having gates series connected seems to model a single-finger device transistor at RF, although the quantitative fit was not perfect. Yet, there seems to be no advantage to using either of these methods for conventional RF devices, only for wide devices having only a few parallel fingers. In any case, such layouts are impractical at high frequencies due to the high gate resistance.

9. Discussion

Ideally, the circuit designer would not need to have an in-depth knowledge of the circuit models he or she is using, but could use it almost as a black box. With some vendor's analog models this ideal can be reached to some extent, but, practically, the designer has to know the basic differences between analog and digital MOS device models. Digital models seldom have critical RF parasitics taken into account, like the gate resistance or the bulk resistance network. The results of this study can be used as guidelines for some approximate values of the gate resistance as well as the bulk resistances.

A critical step in model characterization is the extraction procedure with accurate DC and AC measurements. No matter which model is used, the DC extraction has to be made with the utmost care to make a basis for AC parameter extraction. Most of the low-frequency AC characteristics are determined by the DC model and, basically, AC parameter extraction is used to fit the high-frequency properties. In this step accurate measurements are required, as is a good knowledge of the accuracy of the different small signal parameters used in extraction. Even with previous publications of compound semiconductor FETs, a feeling of small signal parameter extraction uncertainties can be achieved. This knowledge can be used in AC parameter extraction as weights for optimization goals. With my study results the model developer can even search for optimized device layouts to extract the input and output-related small signal parameters. Practically, the results give very good guidelines for input resistance and capacitance extraction layouts. However, this requires at least a slight idea of the polysilicon sheet resistance to be able to approximate the wanted gate resistance value. Usually, the value of polysilicon sheet resistance is known and this method is straightforward. In the case of the output resistance and capacitance, this study does not give such useful guidelines for device layout design. Although the most accurate input resistance and capacitance ranges are shown by my study, the parameter uncertainties are quite independent to each others' absolute values. The extracted output capacitance uncertainty is not greatly affected by the absolute value of output resistance, unless it is well below 50 Ω. Instead, it seems that the absolute capacitance value should be as large as possible to increase the relative measurement accuracy, which is intuitive. Output resistance uncertainty is not very dependent on the absolute output capacitance value. The only practical guidelines for output resistance and

capacitance extraction given by my study are that the device should be as large as possible to get as large a drain conductance and output capacitance as possible. This lowers the MOS output impedance to levels where it is less challenging to measure. My approach to calculating the uncertainty of extracted transconductance and feedback capacitance is only matched by Christian Fager et al. [101] among previous publications, but the equations achieved by my method can be simply presented directly as a function of S parameters. No additional Y-to-S parameter conversion equations are needed in between. The transconductance and feedback capacitance extraction of my method suffers from the approximation carried out in the Y-to-S parameter conversion. The analysis could be developed further by not doing this approximation, but this would result in extremely lengthy equations so complex that they lose their information value.

The possibility of absorbing the drain and source parasitic resistances into the drain current equations does simplify the equivalent circuit and is an accurate approach at low frequencies. However, my study suggests that the absorbed model fails at higher frequencies. This is proved by calculating simplified cases. The comparison to the conventional modeling approach is not thus a general answer but covers MOS characteristics that are quite similar in the general case. The absorption of drain and source resistance effect into the drain current equation is quite a good approximation in many cases when the drain and source resistance value is small, i.e. when applying the model to a wide multifinger RF device. The simplified model becomes inaccurate, especially when trying to get all speed out of the technology of interest. For instance, if a ring oscillator is designed with quite small devices to optimize speed, the parasitic resistance values are relatively large, deteriorating the accuracy of the simple model. In such a case, from the circuit designer's point of view, there seems no other choice but to define the outer resistors by him or herself if the resistances have been absorbed into the drain current equation. This is quite a straightforward process if the designer has the model documentation and access to the model parameter list.

The input impedance study suggests that the BSIM3 model is capable of describing bias-dependent capacitance behavior when S parameter data was used to extract the capacitances. Also, the transconductance-dependent input resistance characteristic was empirically confirmed, but this is something that

has never been taken into account in vendor models. From the designer's point of view, the input impedance study suggests that there is practically no difference when designing with SOI or bulk CMOS. This is different to the case of output impedance, where SOI has a considerably smaller capacitance value. When switching from bulk to SOI CMOS, the RF designer must notice the differences in output matching, whereas the input behavior is quite identical to bulk CMOS.

Previous studies concerning the substrate network effect on MOSFET characteristics have focused on a narrow field of equivalent circuits, and the extraction has been performed with a very limited number of devices. Also, the extraction algorithm has mainly put weight on the output admittance fit. I tried to make the comparison with many different-size devices using a general AC parameter optimization algorithm with S parameters as goals. Different technologies were compared as well. I evaluated many different substrate network equivalent circuits and found out that just a single bulk resistor network is enough to get the most dramatic improvement up to 10 GHz. Putting more resistors into the equivalent circuit is basically just fine-tuning the result. Doing this same evaluation at frequencies above 10 GHz would be important to accurately cover future circuit frequencies. CMOS is already commercially used up to 5 GHz, and future mass markets may reach frequencies over 60 GHz in a period of ten years. This puts more challenges on the substrate models, and it may be possible that a mere resistive substrate network is not sufficient as the dielectric substrate impedance turns reactive at a few tens of GHz. Possibly, it will be found at higher frequencies that the BSIM4 type substrate network is much better than a one or three resistor approach. However, our measurement data was not accurate enough above ten 10 GHz for substrate network study, although a high-end network analyzer was used. If the circuit designer needs to improve the AC characteristics of a digital MOS vendor model, even the numerical results of this study can be used as approximations. Putting a single resistor at the bulk is a considerable improvement to the case of no substrate network at all. Even a few hundred per cent error in the bulk resistance value seems to be better than no substrate network.

Distribution of the gate resistance is an interesting way of improving the model accuracy at high frequencies. However, the two approaches studied in this thesis are practically not applied by the circuit designer. The other approach, of

distributing the gate by defining many parallel MOS devices and putting the gate resistances in series, requires the optimizing of DC model parameters. The characteristics of many models depend on geometry, and dividing a channel width into parallel devices scales the current and charge behavior nonlinearly. Thus, it can be a heavy workload for a designer not familiar with parameter extraction. The other approach of defining an additional gate capacitance in parallel with the gate resistance basically works well for small signal equivalent circuits, but is not directly applicable for large signal analysis. In the special situations were the device is not practical for RF layouts (a wide device with a small number of parallel fingers) the distributed-gate models do improve the simulation accuracy. From the circuit designer's point of view these models are somewhat needless, as conventional RF devices are designed to minimize gate resistance and the devices do not practically suffer from the distributed nature of the gate.

10. Summary

Before the MOS model high-frequency behavior can be analyzed in detail, one must be aware of the AC extraction accuracy. Without this knowledge the transistor AC extraction can lead to unphysical or unrealistic results. A study of the extraction of small signal component values has not been done yet, as the main focus has been on the extraction procedures alone. The results of my study can be used, for example, as a guideline for AC extraction error requirements so as not to fall into numerical traps in optimization. The results of the input and output parameter extractions especially can be used as guidelines for RF test device design to minimize the measurement uncertainty, and thus to improve AC extraction accuracy.

Modern modeling approaches and some suggested improvements were evaluated by theoretical studies and by empirical comparisons. The modern way of absorbing parasitic drain and source series resistances into the drain current description is studied. This technique has been used in, for instance, BSIM3 and its derivatives, as well as in MOS Model 9, to decrease the amount of nodes required by the model description. The results suggest that only the input impedance is quite similar in the two approaches, differing only at very high frequencies. Isolation and gain of the two models behave differently even at DC!

The importance of the substrate network was studied in detail for bulk CMOS, as well as for SOI PD and SOS technologies. In the case of bulk CMOS and PD SOI, the study suggests that the bulk substrate network is crucial for describing output impedance behavior accurately. Both the output impedance real part and the output capacitance behavior better resemble that of the measurements when the substrate network has been included in the model. It seems that good accuracy improvement is achieved with just a single bulk resistor. Additional improvement is achieved by increasing the number of resistors to three. At this used frequency in the 300 MHz to 10 GHz range no further accuracy improvement was achieved by increasing the resistor amount over three.

Two modeling approaches of describing distributed gates were studied. The first was a small signal approach where an additional capacitance was put in parallel with the gate resistance, and the other was a heavy method of connecting MOS

models in parallel and putting the gates in series with small gate resistances in between. Both methods seemed to improve the AC fit of devices with a very small number of parallel devices. According to these results, the small signal method was found to work up to 3 GHz quite accurately, after which it started to behave qualitatively incorrectly. The second modeling method seemed to behave qualitatively correctly up to the maximum frequency used, although there were some inaccuracies at frequencies above 3 GHz. In practice, it was found that both of these models are useless for typical RF MOS layouts. The small signal model cannot be used in large signal analysis and the second approach is heavy and less suitable for modern mainstream models with complicated geometry dependences.

References

1   Doan, C., Emami, S., Niknejad, A., Brodersen, R., "Millimeter-Wave CMOS Design". IEEE Journal of Solid-State Circuits, Vol. 40, No. 1, pp. 144–155, Jan. 2005.

2   Ainger, R., "High performance RF-filter suitable for above IC integration: film bulk-acoustic-resonators (FBAR) on silicon". Proc. IEEE Custom Integrated Circuits Conference, pp. 141–146, 2003,

3   Sung, P.-H., Fang, C.-M., Chang, P.-Z., Chin, Y.-C., Chen, P.-Y., "The Method for Integrating FBAR with Circuitry on CMOS Chip". Proc. IEEE Frequency Control Symposium and Exposition, pp. 562–565, Aug. 2004.

4   Tsividis, Y., *Operation and Modeling of the MOS Transistor.* New York, McGraw-Hill, 1987.

5   Razavi, B., Yan, R-H., Lee, K., "Impact of Distributed Gate Resistance on the Performance of MOS Devices", IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications, Vol. 41, No. 11, pp. 750–754, 1994.

6   Chang, J., Abidi, A., Gaitan, M., "Large Suspended Inductors on Silicon and Their Use in a 2-μm CMOS RF Amplifier", IEEE Electron Device Letters, Vol. 14, No. 5, May. 1993.

7   Basedau, P., Huang, Q., "A 1GHz, 1.5V Monolithic LC Oscillator in 1-μm CMOS", 1994 ESSCIRC, 20th European Solid-State Circuit Conference Proceedings, pp. 172–175, 1994.

8   Rofougaran, M., Rofougaran, A., Ølgaard, C., Abidi, A., "A 900 MHz CMOS RF Power Amplifier with Programmable Output", 1994 Symp. on VLSI Circuits, Digest of Technical papers, pp. 133–134, 1994.

9   Crols, J., Steyaert, M., "A 1.5 GHz Highly Linear CMOS Downconversion Mixer", IEEE Journal of Solid-State Circuits, Vol. 30, No. 7, pp. 736–742, July 1995.

10   Abou-Allam, E., Manku, T., El-Masry, E., "A Digital Tunable Single and Differential-Ended 900MHz Front End RF Amplifier using CMOS Technology", 1995 CCECE/CCGEI, pp. 215–217, 1995.

11   Thamsirianunt, M., Kwasniewski, T., "CMOS VCOs for PLL Frequency Synthesis in GHz Digital Mobile Radio Communications", IEEE 1995 Custom Integrated Circuits Conference, pp. 331–334, 1995.

12   Craninckx, J., Steyaert, M., "A 1.8-GHz CMOS Low-Phase-Noise Voltage-Controlled Oscillator with Prescaler", IEEE Journal of Solid-State Circuits, Vol. 30, No. 12, pp. 1474–1482, Dec. 1995.

13   Crols, J., Steyaert, M., "A Single-Chip 900 MHz CMOS Receiver Front-End with High Performance Low-IF Topology", IEEE Journal of Solid-State Circuits, Vol. 30, No. 12, pp. 1483–1492, Dec. 1995.

14   Rofougaran, A., Rael, J., Rofougaran, M., Abidi, A., "A 900MHz CMOS LC-Oscillator with Quadrature Outputs", 1996 IEEE International Solid-State Circuits Conference, pp. 392–393, 1996.

15   Abou-Allam, E., Manku, T., El-Masry, E., "Q-enhanced 1.9 GHz tuned CMOS RF amplifier", Electronics Letters, Vol. 32, No. 5, pp. 448–449, Feb. 1996.

16   Razavi, B., "A Study of Phase Noise in CMOS Oscillators", IEEE Journal of Solid-State Circuits, Vol. 31, No. 3, pp. 331–343, March 1996.

17   Sullivan, P., Xavier, B., Costa, D., Ku, W., "Silicon MOSFET Distributed Amplifier", Electronics Letters, Vol. 32, No. 12, pp. 1106–1108, June 1996.

18   Ho, Y., Biyani, M., Colvin, J., Smithhisler, C., O, K., "3V Low Noise Amplifier Implemented Using a 0.8μm CMOS Process with Three Metal Layers for 900MHz operation", Electronics Letters, Vol. 32, No. 13, pp. 1191–1193, June 1996.

19   Rofougaran, A., Chang, J., Rofougaran, M., Abidi, A., "A 1 GHz CMOS RF Front-End IC for a Direct-Conversion Wireless Receiver", IEEE Journal of Solid-State Circuits, Vol. 31, No. 7, pp. 880–889, July 1996.

20   Karanicolas, A., "A 2.7-V 900 MHz CMOS LNA and Mixer", IEEE Journal of Solid-State Circuits, Vol. 31, No. 12, pp. 1939–1944, Dec. 1996.

21   Kinget, P., Steyaert, M., "A 1-GHz CMOS Up-Conversion Mixer", IEEE Journal of Solid-State Circuits, Vol. 32, No. 3, pp. 370–376, March 1997.

22   Shaeffer, D., Lee, T., "A 1.5-V, 1.5-GHz CMOS Low Noise Amplifier", IEEE Journal of Solid-State Circuits, Vol. 32, No. 5, pp. 745-759, May 1997.

23   Saijets, J., Andersson, M., Åberg, M., "Comparison of Four CMOS Models at Radio Frequencies", 1998 Norchip Conference, pp. 87–92, Nov. 1998.

24   Saijets, J., Åberg, M., "RF CMOS Modeling", IEEE Circuits and Devices, Vol. 15, No. 2, pp. 7–11, 1999.

25   Liu, W., Gharpurey, R., Chang, M., Erdogan, U., Aggarwal, R., Mattia, J., "R.F. MOSFET Modeling Accounting for Distributed Substrate and Channel Resistances with Emphasis on the BSIM3v3 SPICE Model", 1997 IEDM, IEEE Electron Devices Meeting, Tech. Digest, pp. 309–312, 1997.

26   Jin, X., Ou, J-J., Chen, C-H., Liu, W., Deen, J., Gray, P., Hu, C., "An Effective Gate Resistance Model for CMOS RF and Noise Modeling", 1998 IEDM, IEEE Electron Devices Meeting, Tech. Digest, pp. 961–964, 1998.

27   Scholten, A., Tiemeijer, L., de Vreede, P., Klaassen, D., "A Large Signal Non-Quasi-Static MOS Model for RF Circuit Simulation", 1999 IEDM, IEEE Electron Devices Meeting, Tech. Digest, pp. 163–166, 1999.

28   Tin, S., Osman, A., Mayaram, K., "Comments on "A Small-Signal MOSFET Model for Radio Frequency IC Applications"", IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems, Vol. 17, No. 4, pp. 372–374, April 1998.

29   Tin, S., Osman, A., Mayaram, K., "BSIM3 MOSFET Model Accuracy for RF Circuit Simulation", RAWCON'98 Proceedings, pp. 351–354, 1998.

30   Tin, S., Mayaram, K., "Substrate Network Modeling for CMOS RF Circuit Simulation", IEEE 1999 Custom Integrated Circuits Conference, pp. 583–586, 1999.

31   Tin, S., Osman, A., Mayaram, K., Hu, C., "A Simple Subcircuit Extension of the BSIM3v3 Model for CMOS RF Design", IEEE Journal of Solid-State Circuits, Vol. 35, No. 4, pp. 612–624, April 2000.

32   MacEachern, L., Manku, T., "Metal-Over-Gate MOSFET Modeling for Radio Frequency IC Applications", IEEE Int. Symposium on Circuits and Systems, ISCAS 2000, pp. 164–167, May 2000.

33   Enz, C., "MOS Transistor Modeling for RF Integrated Circuit Design", IEEE 2000 Custom Integrated Circuits Conference, pp. 189–196, 2000.

34   Cheng, Y., Matloubian, M., "On the High-Frequency Characteristics of Substrate Resistance in RF MOSFETs", IEEE Electron Device Letters, Vol. 21, No. 12, pp. 604–606, Dec. 2000.

35   Kim, C., Park, J-W., Yu, H., Cho, H., "Gate Layout and Bonding Pad Structure of a RF n-MOSFET for Low Noise Performance", IEEE Electron Device Letters, Vol. 21, No. 12, pp. 607–609, Dec. 2000.

36   Hjelmgren, H., Litwin, A., "Small-Signal Substrate Resistance Effect in RF CMOS Identified Through Device Simulations", IEEE Trans. on Electron Devices, Vol. 48, No. 2, pp. 397–399, Feb. 2001.

37   Enz, C., "An MOS Transistor Model for RF IC Design Valid in All Regions of Operation", IEEE Trans. on Microwave Theory And Techniques, Vol. 50, No. 1, pp. 342–359, Jan. 2002.

38   Lin, Y-S., Lu, S-S., Lee, T-H., Liang, H-B., "Characterization and Modeling of Small-Signal Substrate Resistance Effect in RF CMOS", 2002 IEEE Radio Freq. Integrated Circuit Symposium, pp. 315–318, 2002.

39   Saijets, J., Andersson, M., Åberg, M., "A Comparative Study of Various MOSFET Models at Radio Frequencies", Analog Integrated Circuits and Signal Processing, Vol. 33, No. 1, pp. 5–17, Oct. 2002.

40   Je, M., Shin, H., "Gate Bias Dependence of the Substrate Signal Coupling Effect in RF MOSFETs", IEEE Electron Device Letters, Vol. 24, No. 3, March 2003.

41   Lin, Y-S., Tu, H-Y., Chiu, H-W., Lu, S-S., "Characterization and Modeling of Size Effect on the Performance of 0.10 μm RF MOSFETs for SOC Applications", 2003 IEEE Radio Freq. Integrated Circuits Symposium, pp. 543–546, 2003.

42   Lin, Y-S., Lee, T-H., Liang, H-B., Lu, S-S., "Characterization and Modeling of 100 nm RF Generic CMOS and 500 nm RF Power CMOS", IEEE 2003, pp. 105–108, 2003.

43 Yang, M., Wang, Y., Yeh, T., Ho, P., Chia, Y., Young, K., "Characterization and Model of 4-Terminal RF CMOS With Bulk Effect", Proc. IEEE 2004 Int. Conference on Microelectronic Test Structures, Vol. 17, pp. 189–193, March 2004.

44 Lin, H., Chung, S., Wong, S., Huang, G., "An Accurate RF CMOS Gate Resistance Model Compatible with HSPICE", Proc. IEEE 2004 Int. Conference on Microelectronic Test Structures, Vol. 17, pp. 227–230, March 2004.

45 Chang, R., Yang, M-T., Ho, P., Wang, Y-J., Chia, Y-T., Liew, B-K., Yue, P., Wong, S., "Modeling and Optimization of Substrate Resistance for RF-CMOS", IEEE Trans. on Electron Devices, Vol. 51, No. 3, pp. 421–426, March 2004.

46 Cheng, Y., Matloubian, M., "Parameter Extraction of Accurate and Scaleable Substrate Resistance Components in RF MOSFETs", IEEE Electron Device Letters, Vol. 23, No. 4, pp. 221–223, April 2002.

47 Lee, T-Y., Cheng, Y., "High-Frequency Characterization and Modeling of Distortion Behavior of MOSFETs for RF IC Design", IEEE Journal of Solid-State Circuits, Vol. 39, No. 9, pp. 1407–1414, Sep. 2004.

48 Yang, M., Ho, P., Lin, C., Yeh, T., Wang, Y., Voinigescu, S., Tazlauanu, M., Chia, Y., Young,K., "BSIM4 High-Frequency Model Verification For 0.13 μm RF-CMOS Technology", 2004 IEEE MTT-S Digest, pp. 1049–1052, 2004.

49 Bayrajtariglu, B., Camilleri, N., Lambert, S., "Microwave Performance of n-p-n and p-n-p AlGaAs/GaAs heterojunction bipolar transistors", IEEE Trans. on Microwave Theory and Techniques, Vol. 36, pp. 1869–1873, Dec. 1988.

50 Tsividis, Y., Suyama, K., "MOSFET Modeling for Analog Circuit CAD: Problems and Prospects", IEEE Journal of Solid-State Circuits, Vol. 29, No. 3, pp. 210–216, March 1994.

51 Sah, C., "Characteristics of the Metal-Oxide-Semiconductor Transistor", IEEE Trans. on Electron Devices, Vol. ED-11, pp. 324–345, 1964.

52 Ihantola, H., Moll, J., "Design Theory of a Surface Field-Effect Transistors", Sol. St. Elec., Vol. 7, pp. 423–430, 1964.

53   Foty, D., *MOSFET Modeling with SPICE: Principles and Practice*, Prentice Hall, 1997.

54   Kang, S., Choi, B., Kim, B., "Linearity Analysis of CMOS for RF Applications", IEEE Trans. on Microwave Theory and Techniques, Vol. 51, No. 3, pp. 972–977, March 2003.

55   Langevelde, R., Klaassen, F., "Accurate Drain Conductance Modeling for Distortion Analysis in MOSFETs", IEEE IEDM Tech. Digest, pp. 313–317, 1997.

56   Langevelde, R., Tiemeijer, L., Havens, R., Knitel, M., Roes, R., Woerlee, P., Klaassen, D., "RF-Distortion in Deep-Submicron CMOS Technologies", IEEE IEDM, pp. 807–810, 2000.

57   Xi, X., Cao, K., Jin, X., Wan, H., Chan, M., Hu, C., "Distortion Simulation of 90nm nMOSFET for RF Applications", Solid-State and IC Tech. Proceedings, Vol. 1, pp. 247–250, 2001.

58   Lee, H., Chun, K-S., Yi, J-H., Lee, J., Park, Y., Min, H., "Harmonic Distortion due to Narrow Width Effects in Deep sub-micron SOI-CMOS Device for analog-RF applications", 2002 IEEE Int. SOI Conference, pp. 83–85, 2002.

59   Choi, C-H., Yu, Z., Dutton, R., "Impact of Poly-Gate Depletion on MOS RF Linearity", IEEE Electron Device Letters, Vol. 24, No. 5, pp. 330–332, May 2003.

60   Bendix, P., Rakers, P., Wagh, P., Lemaitre, L, Grabinski, W., McAndrew, C., Gu, X., Gildenblat, G., "RF Distortion Analysis with Compact MOSFET Models", IEEE 2004 Custom Integrated Circuits Conference, pp. 9–12, 2004.

61   Lee, T.-Y., Cheng, Y., "MOSFET HF distortion behavior and modeling for RF IC design", Proceedings of the IEEE 2003 Custom Integrated Circuits Conference, pp. 87–90, Sept. 2003.

62   Grabinski, W., Vandamme, E., Schreurs, D., Maeder, H., Pilloud, O., Mcandrew, C., "5.5 GHz LSNA MOSFET modeling for RF CMOS circuit design", ARFTG Conference Digest, pp. 39–47, Dec. 2002.

63   Iniquez, B., Picos, R., Kwon, I., Shur, M., Fjeldly, T., Lee, K., "Compact MOSFET modeling for harmonic distortion analysis", Proceedings of

Fifth IEEE Caracas Conference on Devices, Circuits and Systems, Vol. 1, pp. 111–117, Nov. 2004.

64   Cheng, Y. et al., BSIM3v3 Manual, Report, University of California, Berkeley, Dept. of Electrical Engineering and Computer Sciences, 1996.

65   Yang, P., Epler, B., Chatterjee, P., "An Investigation of the Charge Conservation Problem for MOSFET Circuit Simulation", IEEE Journal of Solid-State Circuits, Vol. 18, No. 1, pp. 128–138, Feb. 1983.

66   Velghe, R., Klaassen, D., Klaassen, F., MOS Model 9 level 902, Unclassified Report NLUR 003/94, Philips Research Laboratories, Eindhoven, Netherlands, June 1995.

67   Langevelde, R., MOS Model 11, Level 1100, Unclassified Report NL-UR 2001/813, Philips Electronics, June 2004.

68   Y.C. et al., BSIM4 Manual, Report, University of California, Berkeley, Dept. of Electrical Engineering and Computer Sciences, 2004.

69   Chenming, H. et al., BSIM3SOI v1.3 Manual, University of California, Berkeley, Dept. of Electrical Engineering and Computer Sciences, 1998.

70   Bucher, M., Lallement,C., Enz, C., Theodoloz, F., Krummenacher, F., The EPFL-EKV MOSFET Model Equations for Simulation, Model Version 2.6, Revision II, Report, Electronics Laboratories, Swiss Federal Institute of Technology (EPFL), July 1998.

71   Smit, G., Langevelde, R., Scholten, A., Klaassen, D., Gildenblat, G., Li, X., Wang, H., Wu, W., "PSP 102.0" model documentation, 2006.

72   Gildenblat, G., Wang, H., Chen, T., Gu, X., Cai, X, "SP: an advanced surface-potential-based compact MOSFET model", IEEE Journal on Solid-State circuits, Vol. 39, No. 9, pp. 1394–1406, Sep. 2004.

73   Gildenblat, G., Li, X., Wu, W., Wang, H., Jha, A., Langevelde, R., Smit, G., Scholten, A., Klaassen, D., "PSP: An Advanced Surface-Potential-Based MOSFET Model for Circuit Simulation", IEEE trans. Electron Devices, Vol. 53, No. 9, pp. 1979–1993, Sep. 2006.

74   Wang, H., Chen, T., Gildenblat, G., "Quasi-static and Nonquasi-static compact MOSFET Models Based on Symmetric Linearization of the Bulk

and Inversion Charges," IEEE trans. Electron Devices, Vol. 50, No. 11, pp. 2262–2272, Nov. 2003.

75   Gray, P., Meyer, R., *Analysis and Design of Analog Integrated Circuits*, New York, Wiley, pp. 579–584, 1993.

76   Lu, S-S., Meng, C., Chen, T-W., Chen, H-C., "The Origin of the Kink Phenomenon of Transistor Scattering Parameter S_{22}", IEEE Trans. on Microwave Theory and Techniques, Vol. 49, No. 2, pp. 333–340, 2001.

77   Saijets, J., "NMOS Transistor Modeling at Radio Frequencies", Masters' Thesis, Department of Electrical and Communications Engineering, Helsinki University of Technology, 1998.

78   M. Valtonen, P. Heikkilä, A. Kankkunen, K. Mannersalo, R. Niutanen, P. Stenius, T. Veijola, T. Virtanen, "APLAC -A New Approach to Circuit Simulation by Object-Orientation". Proc. of the 10th European Conference on Circuit Theory and Design, vol. I, pp. 351–360, Copenhagen, July 1991.

79   User Guide: Vee OneLab User's Guide, Agilent Technologies, (2000)

80   Hewlett Packard, Specifying Calibration Standards for the HP 8510 Network Analyzer, Product Note 8510-5a, Feb. 1988.

81   Je, M., Kwon, I., Han, J., Shin, H., Lee, K., "On the Large-Signal CMOS Modeling and Parameter Extraction for RF Applications", Int. Conf. on Simulation of Semiconductor Processes and Devices, pp. 67–70, Sept. 2002.

82   Kwon, I., Je, M., Lee, K., Shin, H., "A New Small Signal Modeling of RF MOSFETs including Charge Conservation Capacitances", ESSCIRC 2000, pp. 296–299, 2000.

83   Pehlke, D., Schroter, M., Burstein, A., Matloubian, M., Chang, M., "High frequency application of MOS compact models and their development for scalable RF model libraries", Proceedings of the IEEE Custom Integrated Circuits Conference, pp. 219–222, May 1998.

84   Saijets, J., Åberg, M., "MOSFET RF Extraction Uncertainties Due to S Parameter Measurement Errors", The 20th Nordic Semiconductor Meeting, Tampere, pp. 48–49, Aug. 2002.

85   Saijets, J., Åberg, M., "An Analytical Approach to Determine MOSFET RF Extraction Uncertainties Due to S Parameter Measurement Errors", 62nd ARFTG Microwave Measurement Conference, Boulder, CO, pp. 99–108, Dec. 2003.

86   Saijets, J., Åberg, M., "MOSFET RF Extraction Uncertainties Due to S Parameter Measurement Errors", Physica Scripta, Vol. T114, pp. 224–247, 2004.

87   Ou, J-J., Jin, X., Ma, I., Hu, C., Gray, P.R., "CMOS RF modeling for GHz communications IC's," Symposium on VLSI Technology, pp. 94–95, 1998.

88   Kondoh, H., "An Accurate FET Modelling from Measured S-parameters", IEEE MTT-S Technical Digest, pp. 377–380, 1986.

89   Lovelace, D., Costa, J., Camilleri, N., "Extracting Small-Signal Model Parameters of Silicon MOSFET Transistors", 1994 IEEE MTT-S Digest, pp. 865–868, 1994.

90   Jen, S., Enz, C., Pehlke, D., Shröter, M., Sheu, B., "Accurate Modeling and Parameter Extraction for MOS Transistors Valid up to 10 GHz", IEEE Trans. on Electron Devices Vol. 46, No. 11, pp. 2217–2226, 1999.

91   Enz, C., Cheng, Y., "MOS Transistor Modeling for RF IC Design", IEEE Trans. on Solid-State Circuits Vol. 35, No. 2, pp. 186–201, 2000.

92   Cheng, Y., Matloubian, M., "High Frequency Characterization of Gate Resistance in RF MOSFETs", IEEE Electron Device Letters, Vol. 22, No. 2, pp. 98–100, 2001.

93   Pengg, F., "Direct Parameter Extraction on RF-CMOS", 2002 IEEE MTT-S Technical Digest, pp. 271–274, 2002.

94   Pengg, F., "Direct Parameter Extraction on RF-CMOS", 2002 IEEE Radio Freq. Integrated Circuits Symposium, pp. 355–358, 2002.

95   Cheng, Y., Chen, C., Matloubian, M., Deen, M., "High-Frequency Small Signal AC and Noise Modeling of MOSFETs for RF IC Design", IEEE Trans. on Electron Devices, Vol. 49, No. 3, pp. 400–408, March 2002.

96   Voinigescu, S., Tazlauanu, M., Ho, P., Yang, M., "Direct Extraction Methodology for Geometry-Scalable RF-CMOS Models", Proc. IEEE

2004 Conference on Microelectronic Test Structures, Vol. 17, pp. 235–240, March 2004.

97   Guo, J., Lin. Y., "A New Lossy Substrate De-Embedding Method for Sub-100 nm RF CMOS Noise Extraction and Modeling", IEEE Trans. on Electron Devices, Vol. 53, No. 2., Feb. 2006.

98   Walters, P., Pollard, R., Richardson, J., Gamand, P., Suchet, P., "On-Wafer Measurement Uncertainty For 3-Terminal Active Millimetre-wave Devices", IEEE GaAs IC Symposium, pp. 55–58, 1992.

99   King, F., Winson, P., Snider, A., Dunleavy, L., Levinson, D., "Math Methods in Transistor Modeling: Condition Numbers for Parameter Extraction", IEEE Trans. Microwave Theory and Techniques, Vol. 46, No. 9, pp. 1313–1314, 1998.

100   Choubani, F., Béjaoui, W., Bouallegue, A., "Condition Number for Reliable Microwave Transistor Modeling", Wiley International Journal of RF and Microwave Computer-Aided Engineering, Vol. 14, No. 2, pp. 174–181, 2004.

101   Fager, C., Linnér, P., Pedro, J., "Optimal Parameter Extraction and Uncertainty Estimation in Intrinsic FET Small-Signal Models", IEEE Trans. Microwave Theory and Techniques, Vol. 50, No. 12, pp. 2797–2803, 2002.

102   HP 8510B Documentation, Uncertainty Specifications

103   Koolen, M., Geelen, J., Versleijen, M., "An improved de-embedding technique for on-wafer high-frequency characterization" IEEE 1991 Bipolar Circuits and Technology Meeting, pp. 188–191, 1991.

104   Vidkjaer, J., "Identification of small-signal transistor models", ECCTD'89, proceedings of., pp. 99–103, 1989.

105   Tiemeijer, L., Maaijer, L., Langevelde, R., Scholten, A., Klaassen, D., "RF CMOS Modelling", in Advances in Analog Circuit Design (AACD'99), pp. 129–149, 1999.

106   Tiemeijer, L., Klaassen, D., "Geometry Scaling of the Substrate Loss of RF MOSFETs", ESSDERC 98, Tech. Dig, pp. 480–483, 1998.

107　Mahalingan, U., Rustagi, S., Samudra, G., "Direct Extraction of Substrate Network Parameters for RF MOSFET Modeling Using a Simple Test Structure", IEEE Electron Device Letters, Vol. 27, No. 2, pp. 130–132, Feb. 2006.

108　Abou-Allam, E., Manku, T., "A Small-Signal MOSFET Model for Radio Frequency IC Applications", IEEE Trans. Computer Aided Design, Vol. 16, No. 5, pp. 437–447, 1997.

109　Ricco, B., Versari, R., Esseni, D., "Characterization of Polysilicon-Gate Depletion in MOS Structures", IEEE Electron Device Letters, Vol. 17, No. 3, pp. 103–105, 1996.

110　Kim, L-S., Dutton, W., "Modeling of the distributed gate effect in MOSFET's", IEEE Trans. Computer Aided Design, Vol. 8, pp. 1365–1367, 1989.

111　Razaui, B., Yan, R-H., Lee, K., "Impact of distributed gate resistance on the performance of MOS devices," IEEE Trans. Circuits Syst. 1, Vol. 41, pp. 750–754, 1994.

112　"HSPICE User's Manual: Elements and Device Models, Volume II", Meta-Software Inc., pp. 16-103–16-113, 1996.

Appendix A: Complete Uncertainty Equations

Transconductance related error equations:

Transconductance is repeated here from Eq. (68) and can be written approximatively as

$$g_m = \Re\{ Y_{21} \} = -2S_{21m} \frac{(\cos\theta_{21a} + \sin\theta_{21b})}{Z_0(a^2 + b^2)} \tag{112}$$

where a and b are defined with Eq. (69) and Eq. (70) respectively.

The total differential error of transconductance is calculated as

$$\Delta g_m = \left|\frac{\partial g_m}{\partial S_{11m}}\right|\Delta S_{11m} + \left|\frac{\partial g_m}{\partial\theta_{11}}\right|\Delta\theta_{11} + \left|\frac{\partial g_m}{\partial S_{22m}}\right|\Delta S_{22m} + \left|\frac{\partial g_m}{\partial\theta_{22}}\right|\Delta\theta_{22} \tag{113}$$
$$+ \left|\frac{\partial g_m}{\partial S_{21m}}\right|\Delta S_{21m} + \left|\frac{\partial g_m}{\partial\theta_{21}}\right|\Delta\theta_{21}$$

The partial differentials are the following:

$$\frac{\partial\Re\{ Y_{21} \}}{\partial S_{11m}} = -2\frac{S_{21m}}{Z_0} \tag{114}$$
$$\left[\frac{\cos\theta_{21}((\cos\theta_{11} + S_{22m}\cos(\theta_x)) + \sin\theta_{21}(\sin\theta_{11} + S_{22m}\cos(\theta_x)))}{a^2 + b^2} \right.$$
$$\left. -2\frac{\cos\theta_{21}a + \sin\theta_{21}b}{(a^2 + b^2)^2}(S_{11m} + \cos\theta_{11})(1 + S_{22m}{}^2 + 2S_{22m}\cos\theta_{22}) \right]$$

here

$$\theta_x = \theta_{11} + \theta_{22} \tag{115}$$

$$\frac{\partial \Re\{Y_{21}\}}{\partial \theta_{11}} = -2 \frac{S_{21m}}{Z_0} S_{11m}$$

$$\cdot [\, \frac{-\cos\theta_{21}(\sin\theta_{11}+S_{22m}\sin\theta_x)+\sin\theta_{21}(\cos\theta_{11}+S_{22m}\cos\theta_x)}{a^2+b^2}$$

$$+2\frac{\cos\theta_{21}a+\sin\theta_{21}b}{(a^2+b^2)^2}\left(\sin\theta_{11}(1+S_{22m}{}^2)+2S_{22m}\cos\theta_{22}\sin\theta_{11}\right)\,]$$

(116)

$$\frac{\partial \Re\{Y_{21}\}}{\partial S_{22m}} = -2 \frac{S_{21m}}{Z_0}$$

$$[\, \frac{\cos\theta_{21}(\cos\theta_{22}+S_{11m}\cos\theta_x)+\sin\theta_{21}(\sin\theta_{22}+S_{11m}\sin\theta_x)}{a^2+b^2}$$

$$-2\frac{\cos\theta_{21}a+\sin\theta_{21}b}{(a^2+b^2)^2}(S_{22m}+\cos\theta_{22})(1+S_{11m}{}^2+2S_{11m}\cos\theta_{11})\,]$$

(117)

$$\frac{\partial \Re\{Y_{21}\}}{\partial \theta_{22}} = -2 \frac{S_{21m}}{Z_0} S_{22m}$$

$$\cdot [\, \frac{-\cos\theta_{21}(\sin\theta_{22}+S_{11m}\sin\theta_x)+\sin\theta_{21}(\cos\theta_{22}+S_{11m}\cos\theta_x)}{a^2+b^2}$$

$$+2\frac{\cos\theta_{21}a+\sin\theta_{21}b}{(a^2+b^2)^2}\left(\sin\theta_{22}(1+S_{11m}{}^2)+2S_{11m}\cos\theta_{11}\sin\theta_{22}\right)\,]$$

(118)

$$\frac{\partial \Re\{Y_{21}\}}{\partial S_{21m}} = \frac{-2(\cos\theta_{21}a-\sin\theta_{21}b)}{a^2+b^2}$$

(119)

$$\frac{\partial \Re\{Y_{21}\}}{\partial \theta_{21}} = \frac{2S_{21m}(\sin\theta_{21}a+\cos\theta_{21}b)}{a^2+b^2}$$

(120)

Feedback capacitance-related error equations:

Transconductance is repeated here from Eq. (68) and can be written approximatively as

$$C_{gd} = -2S_{12m}\frac{(\sin\theta_{12a}-\cos\theta_{12b})}{Z_0\omega(a^2+b^2)}$$

(121)

where a and b are defined with Eq. (69) and Eq. (70) respectively.

The total differential error of feedback capacitance is calculated as

$$\Delta C_{gd} = \left|\frac{\partial C_{gd}}{\partial S_{11m}}\right| \Delta S_{11m} + \left|\frac{\partial C_{gd}}{\partial \theta_{11}}\right| \Delta \theta_{11} + \left|\frac{\partial C_{gd}}{\partial S_{22m}}\right| \Delta S_{22m} + \left|\frac{\partial C_{gd}}{\partial \theta_{22}}\right| \Delta \theta_{22}$$
$$+ \left|\frac{\partial C_{gd}}{\partial S_{12m}}\right| \Delta S_{12m} + \left|\frac{\partial C_{gd}}{\partial \theta_{12}}\right| \Delta \theta_{12}$$

(122)

As a function of Y_{12} this can be expressed as

$$\Delta C_{gd} = -\frac{1}{\omega}\left(\left|\frac{\partial \Im\{Y_{12}\}}{\partial S_{11m}}\right| \Delta S_{11m} + \left|\frac{\partial \Im\{Y_{12}\}}{\partial \theta_{11}}\right| \Delta \theta_{11} + \left|\frac{\partial \Im\{Y_{12}\}}{\partial S_{22m}}\right| \Delta S_{22m} \right.$$
$$\left. + \left|\frac{\partial \Im\{Y_{12}\}}{\partial \theta_{22}}\right| \Delta \theta_{22} + \left|\frac{\partial \Im\{Y_{12}\}}{\partial S_{12m}}\right| \Delta S_{12m} + \left|\frac{\partial \Im\{Y_{12}\}}{\partial \theta_{12}}\right| \Delta \theta_{12} \right)$$

(123)

The partial differentials are the following:

$$\frac{\partial \Im\{Y_{12}\}}{\partial S_{11m}} = -2\frac{S_{12m}}{Z_0}$$

(124)

$$[\ \frac{\sin\theta_{12}(\cos\theta_{11} + S_{22m}\cos\theta_x) + \cos\theta_{21}(\sin\theta_{11} + S_{22m}\cos\theta_x)}{a^2 + b^2}$$
$$-2\frac{\sin\theta_{12}a + \cos\theta_{12}b}{(a^2 + b^2)^2}(S_{11m} + \cos\theta_{11})(1 + S_{22m}^2 + 2S_{22m}\cos\theta_{22})\ ]$$

$$\frac{\partial \Re\{Y_{12}\}}{\partial \theta_{11}} = -2\frac{S_{12m}}{Z_0}S_{11m}$$

(125)

$$\cdot[\ \frac{-\sin\theta_{12}(\sin\theta_{11} + S_{22m}\sin\theta_x) + \cos\theta_{12}(\cos\theta_{11} + S_{22m}\cos\theta_x)}{a^2 + b^2}$$
$$+2\frac{\sin\theta_{12}a + \cos\theta_{12}b}{(a^2 + b^2)^2}\left(\sin\theta_{11}(1 + S_{22m}^2) + 2S_{22m}\cos\theta_{22}\sin\theta_{11}\right)\ ]$$

$$\frac{\partial \Im\{Y_{12}\}}{\partial S_{22m}} = -2\frac{S_{12m}}{Z_0}$$

(126)

$$[\ \frac{\sin\theta_{12}(\cos\theta_{22} + S_{11m}\cos\theta_x) + \cos\theta_{21}(\sin\theta_{22} + S_{11m}\cos\theta_x)}{a^2 + b^2}$$
$$-2\frac{\sin\theta_{12}a + \cos\theta_{12}b}{(a^2 + b^2)^2}(S_{22m} + \cos\theta_{22})(1 + S_{11m}^2 + 2S_{11m}\cos\theta_{11})\ ]$$

$$\frac{\partial \Re\{Y_{12}\}}{\partial S_{12m}} = \frac{-2(\sin\theta_{12}a + \cos\theta_{12}b)}{a^2 + b^2}$$

(127)

$$\frac{\partial \Re\{Y_{12}\}}{\partial \theta_{12}} = \frac{-2S_{12m}(\cos\theta_{12}\,a - \sin\theta_{12}\,b)}{a^2 + b^2} \tag{128}$$

Author(s)
Saijets, Jan

Title

MOSFET RF Characterization Using Bulk and SOI CMOS Technologies

Abstract

MOSFET radio-frequency characterization and modeling is studied, both with SOI CMOS and bulk CMOS technologies. The network analyzer measurement uncertainties are studied, as is their effect on the small signal parameter extraction of MOS devices. These results can be used as guidelines for designing MOS RF characterization layouts with as small an AC extraction error as possible. The results can also be used in RF model extraction as criteria for required optimization accuracy.

Modifications to the digital CMOS model equivalent circuit are studied to achieve better RF behavior for the MOS model. The benefit of absorbing the drain and source parasitic series resistances into the current description is evaluated. It seems that correct high-frequency behavior is not possible to describe using this technique. The series resistances need to be defined extrinsically. Different bulk network alternatives were evaluated using scalable device models up to 10 GHz. Accurate output impedance behavior of the model requires a bulk resistance network. It seems that good accuracy improvement is achieved with just a single bulk resistor. Additional improvement is achieved by increasing the number of resistors to three. At this used frequency range no further accuracy improvement was achieved by increasing the resistor amount over three. Two modeling approaches describing the distributed gate behavior are also studied with different MOS transistor layouts. Both approaches improve the RF characteristics to some extent but with limited device geometry. Both distributed gate models describe well the high frequency device behavior of devices not commonly used at radio frequencies.

Tekijä(t)
Saijets, Jan

Nimeke

MOSFETien radiotaajuuskarakterisointi bulk- ja SOI CMOS -teknologioita käyttäen

Tiivistelmä

MOSFETin radiotaajuuskarakterisointia ja mallitusta tarkastellaan sekä SOI CMOS että bulk CMOS -teknologioilla. Piirianalysaattorien mittausepävarmuutta tarkastellaan ja niiden vaikutusta MOS-transistorin piensignaaliparametrien ekstraktointiin. Näitä tuloksia voidaan käyttää ohjenuorana RF MOS- karakterisointiin käytettävien piirikuvioiden suunnittelussa, kun halutaan AC-ekstraktoinnin virhe mahdollisimman pieneksi. Tuloksia voidaan käyttää myös RF-mallin ekstraktoinnissa halutun optimointitarkkuuden kriteerinä.

Digitaalisen CMOS-mallin vastinpiirimuunnelmia on tarkasteltu tarkoituksena saada MOS-mallille paremmat radiotaajuusominaisuudet. Kanavan kanssa sarjassa olevien parasiittisten vastusten vaikutusta on tarkasteltu, kun ne ovat joko erillisinä tai suoraan virtayhtälöön sisällytettyinä. Jälkimmäisen tavan hyötyä on arvioitu. Näyttää siltä, että oikeanlaatuisen suurtaajuuskäyttäytymisen kuvaaminen ei onnistu tällä tekniikalla. Kanavan kanssa sarjassa olevat vastukset on määriteltävä ulkoisiksi. Erilaisia substraattivastinpiirien vaihtoehtoja on arvioitu käyttäen skaalautuvia transistorimalleja 10 GHz:n taajuuteen asti. Tarkan ulostuloimpedanssin kuvaaminen edellyttää transistorille substraattivastinpiiriä. Näyttää siltä, että merkittävään mallin tarkkuuden parantumiseen riittää yksi ainoa substraattivastus. Tarkkuus paranee tästäkin lisättäessä vastusten määrää kolmeen. Käytetyllä taajuusvälillä ei saavutettu mallin tarkkuuden lisäparannusta, kun yritettiin nostaa substraattiverkon vastusten määrää yli kolmen. Erilaisilla MOS-transistorin piirikuvioilla tarkasteltiin myös kahta mallitustapaa, joilla voidaan kuvata jakautuneen hilan käyttäytymistä. Kummatkin lähestymistavat parantavat mallin radiotaajuusominaisuuksia johonkin rajaan asti, mutta vain tietyillä transistorigeometrioilla. Molemmat jakautuneen hilan mallit kuvaavat hyvin sellaisten transistoreiden suurtaajuusominaisuuksia, joita ei yleensä käytetä radiotaajuuksilla.

This work deals with CMOS transistor characterization at radio-frequencies (RF). An accurate transistor model is the basis for circuit simulation and design. Previously MOS transistor models have been less accurate at RF and have prevented the use of cheap CMOS circuitry in the radio parts of mobile terminals. In recent years a lot of research has been made to correct this problem. This thesis work has produced new knowledge and scientific results in the following areas: 1) RF measurement uncertainty effect on the transistor characterization; 2) the input impedance accuracy of MOS models compared to experimental results; 3) the benefit of different modifications to the basic digital CMOS model equivalent circuit. The equivalent circuit modifications includes different approaches to describe MOSFET bulk resistance network, absorption of the series resistances into the current description as well as two approaches of describing the distributed nature of gate polysilicon behavior. Both SOI CMOS and bulk CMOS technologies have been used in this work.