Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 681

Design and Modeling of High-Frequency LDMOS Transistors

BY

LARS VESTLING

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2002

Dissertation for the Degree of Doctor of Philosophy in Solid State Electronics presented at Uppsala University in 2002

Abstract

Vestling, L., 2002. Design and Modeling of High-Frequency LDMOS Transistors. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 681. 50 pp. Uppsala. ISBN 91-554-5210-8.

The lateral double-diffused MOS (LDMOS) transistor has traditionally been a high-voltage device used in switching applications. The use as a high-frequency device has become more important lately since the LDMOS offers an low cost solution for telecommunication applications. An important property of the LDMOS concept is that it can be manufactured in virtually the same process used in standard CMOS production. It only requires one extra process step, which is easily implemented. The other important aspect that gives the LDMOS the good high-frequency performance is that the channel length is a process parameter and not a lithography parameter.

This thesis investigates the LDMOS transistor primarily from two aspects. The first is the high-voltage performance. For a high-voltage device the most important parameter is the breakdown voltage. The second most important parameter is the on-resistance that has the property of being in contradiction of the breakdown voltage and usually trade-offs are made to achieve acceptable performance. In the thesis several methods to improve the breakdown voltage/on-resistance relation are presented.

The other part covers the high-frequency behavior of the LDMOS transistor. High-frequency characterization has been made to gain valuable information for the fundamental understanding of the physical mechanisms inside the transistor. A large part of the thesis covers modeling and parameter extraction of the devices. A new general method for parameter extraction of small-signal equivalent circuit models is presented, which has the appealing properties of not needing any approximation during the extraction which is common with other techniques.

Lars Vestling, Solid State Electronics, The Ångström Laboratory, Uppsala University, Box 534, SE-751 21 Uppsala, Sweden

ISSN 1104-232X
ISBN 91-554-5210-8

Printed in Sweden by Reklam & Katalogtryck, Uppsala 2002

Till dom jag tycker om

Table of Contents

List of Papers

[I] "A Novel High-Frequency High-Voltage LDMOS Transistor using an Extended Gate RESURF Technology"
Lars Vestling, Bengt Edholm, Jörgen Olsson, Stefan Tiensuu and Anders Söderbärg,
Proceedings of IEEE ISPSD'97, pp. 45–48, 1997

[II] "A CMOS Compatible Power MOSFET for Low Voltage GHz Operation"
Lars Vestling, Lars Bengtsson, and Jörgen Olsson,
Proceedings of 29th EuMC, vol. 2, pp. 21–24, 1999

[III] "Drift Region Optimization of Lateral RESURF Devices"
Lars Vestling, Jörgen Olsson and Klas-Håkan Eklund,
submitted to *Solid-State Electronics*, 2001

[IV] "Analysis and Design of a Low-Voltage High-Frequency LDMOS Transistor"
Lars Vestling, Johan Ankarcrona and Jörgen Olsson,
submitted to *IEEE Transactions on Electron Devices*, 2001

[V] "High Voltage LDMOS Transistors for Increasing RF Power Density and Gain"
Niklas Rorsman, Jörgen Olsson, Christian Fager, Lars Vestling, Johan Ankarcrona, Herbert Zirath, and Klas-Håkan Eklund,
Proceedings of GHz2001, 2001

[VI] "A General Small-Signal Series Impedances Extraction Technique"
Lars Vestling and Johan Ankarcrona,
submitted to *IEEE Microwave and Wireless Components Letters*, 2001.

[VII] "An On-Wafer De-embedding Method for Silicon Transistors at Microwave Frequencies"
Lars Bengtsson, Mikael Garcia, Lars Vestling and Jörgen Olsson,
manuscript.

Acknowledgments

Thanks...

- Jörgen, for being the best supervisor you can have.
- Herman, for always having an answer and for always taking a beer on Fridays, what would the department be without you.
- our financial support: NUTEK (Swedish National Board for Industrial and Technical Development) and SSF (Swedish Foundation for Strategic Research).
- the cleanroom staff at Mitel Semiconductor (who recently changed name to Zarlink Semiconductor) for processing the wafers in the *GHz Power Transistors* project.
- Klas-Håkan, for keeping the project running all these years, and a lot of other things as well.
- the Chalmers Crew, Lars, Christian, Niklas, Herbert, for the all things about microwaves you taught us.
- the Co-authors, for making this possible.
- Sören, for not kicking me out.
- Mamma, Pappa and Brorsan for obvious reasons. You're the best.
- Mats, who supervised me when I did my thesis work, I still remember things you taught me.
- Stefan and Bengt, who tried hard to recruit me to the device group and succeeded.
- Anders, for lots of coaching on, and off, the scientific path.
- Matts, for being *the* room mate and a very good friend.

- Uli, Markus, Tony, Johan, I like this team.
- the Åre team, beer and skiing is a great combo.
- Elektroniken, the best department on the first floor. And especially Marianne for keeping it together.
- Karin, for all the reasons you can think of, I love you.
- ...
- And a special thanks to Johan and Andreas for the Star Wars drawings. May the force be with you.

"Hur svårt kan det vara"
Per Vestling

Chapter 1

Introduction

A new era in the modern world started 1947 when Bardeen, Brattain and Shockley at Bell Labs invented the transistor. The next milestone came 1958 when Jack Kilby at Texas Instruments made the first integrated circuit, which recently gave him the Nobel Prize. Since then the evolution in the semiconductor area have been accelerating steadily. This can be illustrated by *Moore's Law* [1] that says that the number of devices per area will double every 18 months. Moore's Law was stated in 1965 and it still holds almost 4 decades later. Moore's Law is also valid on other aspects in the semiconductor business such as the cost per device will half every 18 months.

There have been a variety of transistors types used over the years [2–5], which all have their pros and cons. The most commonly used device type is the field effect transistor, FET, that is used in all kinds of different applications. The MOS (metal-oxide-semiconductor) transistor is an example of an FET device. The other main type of device is the bipolar transistor, which has a basically different operation. The MOSFET is today the most dominating device. The complementary MOS (CMOS) devices are used in all digital logic circuits, such as microprocessors and computers. This makes the silicon MOS transistor the most scrutinized semiconductor device today. This thesis will not cover the CMOS technology, although it needs more research than ever [6].

The CMOS technology is easily scaled to improve performance but this also implies lower supply voltages. For high-voltage MOS transistors the channel region is separated from the voltage supporting part and thus higher breakdown voltages are achieved. There are several different types of high-voltage MOSFET's and an interesting candidate is the lateral double-diffused MOS transistor, the *LDMOS*.

Apart from the high-voltage support the LDMOS transistor can achieve very good high-frequency performance due to that the channel length, that

also sets the high-speed properties, can easily be made very short. An LDMOS transistor can be made using, in the simplest form, one extra process step in addition to the standard CMOS process. The process step is an implantation that forms the channel, which is further explained later.

This thesis focus on the LDMOS transistor and especially on the design of the drift region, which controls the high-voltage operation, and modeling and characterization of the high-frequency operation.

1.1 High-Voltage Transistors

The main use for high-voltage devices are rectifying and switching. For these purposes there are two major criteria for the devices that have to be fulfilled. The purpose of a switching device is to be open or closed, compare to a mechanical switch. If it is open there is no electrical contact between the two terminals of the switch and thus no current can pass through the device. The other state is closed and then the current should flow without losses. A solid-state switching device has the same purpose as the mechanical but the on- and off-state are not as well defined as for the mechanical switch. When the switch is open there will be leakage current through the device and eventually when the voltage is to high breakdown will occur and a large current will pass through the device. When it is closed it will have resistive losses through the semiconductor material. The first criterion that has to be fulfilled is that the breakdown voltage, V_{br}, is high enough for the application it should be used in. Then the resistive losses in on-state should be minimized. The parameter that represent this is called on-resistance, R_{on}. Also the leakage currents should be sufficiently small when the device is off.

The choice of on-resistance and breakdown voltage is often a trade-off since designing for high breakdown also degrades the on-resistance and vice versa. For an idealized device the relation between the specific on-resistance, $R_{on,sp}$ and the parallel-plane breakdown voltage $V_{br,pp}$ can be expressed as, [5]

$$R_{on,sp} = \frac{4V_{br,pp}^2}{\varepsilon_s \mu \mathcal{E}_c^3} \tag{1.1}$$

where ε_s is the dielectric constant for the material, μ is the mobility and $\mathcal{E}_c$ is the critical electric field of the material used. The specific on-resistance is the on-resistance normalized to the area of the device, which makes it possible to compare devices of different sizes. The denominator in (1.1) is called the *Baliga's figure of merit* (BFOM $= \varepsilon_s \mu \mathcal{E}_c^3$) [5]. Since all parameters in BFOM are material dependent this can be used to compare different materials with respect to their usefulness as a high-voltage device, table 1.1 shows BFOM for some common materials.

Material	Silicon	GaAs	4H-SiC	GaN	Diamond
BFOM	1	21	280	1300	11000

Table 1.1: Summary of the Baliga's figure of merit for some materials.

This takes us to the choice of semiconducting material that should be used. Starting from silicon which is the most accessible material today it can be seen that the high-voltage performance is not that good. Another possible material is gallium-arsenide, GaAs, but it has it main use in high-frequency applications. An interesting material which has been a major research object lately is silicon-carbide, SiC. A big problem with SiC is the possibility to manufacture defect free substrates of useful size. Another competitive material that is coming up is gallium-nitride, GaN, which have excellent material properties [7]. The problem so far has been the substrates, which has led to that the only important application has been light-emitting diodes (LED). The most superior material for high-voltage applications is diamond [8, 9]. Since diamond is the hardest material on earth, and very inert, the processing becomes very difficult, but there are attempts to overcome these obstacles and actually manufacture devices in diamond.

A more accessible use of diamond is to manufacture a SOI (silicon-on-insulator) substrate using diamond as the insulator. This would benefit the superior thermal conductivity of diamond compared to silicon oxide which introduces a problem due to its poor thermal conductivity [10]

Apart from the material choice when manufacturing high-voltage devices there are several different types of devices which all have specific properties. In the beginning thyristors were the primarily choice for a high-voltage device. The thyristor is a bipolar type of device that has the properties of large breakdown voltage and high conducting current. Later the GTO (gate turn-off thyristor) was used, which could be switched on and off by a gate voltage and thus the switching losses were reduced compared to the conventional thyristor [11]. Also conventional bipolar devices are used for high-voltage applications. Of the field effect transistors the vertical double-diffused MOS transistor (VDMOS) has been the primarily choice.

The power MOS devices have poor on-resistance for higher voltages. This has led to the development of the insulated gate bipolar transistor (IGBT) which combines the bipolar transistor with field effect transistor. The manufacturing is the same as for the DMOS but the drain contact is of opposite type. In the beginning the IGBT were called IGFET (insulated gate field effect transistor) but later it became evident that this was really a bipolar transistor [12]. The IGBT is today the dominating device for switching device in the region 300-1500 V [5].

All the above mentioned devices are in some respect vertical devices. With vertical means that the current flows from a contact on top of the wafer through the wafer down to a contact on the back side. This makes it possible to use all the silicon in the wafer as the voltage supporting region. The substrate is usually a high-resistivity material to withstand high voltages. It is not uncommon that one device is made on the whole wafer which allows the devices to handle very high currents.

For medium high voltages the losses though the substrate will become significant, and there are integration problems using the back contact. It would be beneficial to have all contacts on the same side. This takes us to another type of devices, the lateral devices. Lateral means that the current flows laterally in the wafer parallel to the surface. The major breakthrough for lateral device came with the *reduced surface field*, RESURF, concept [13] that made it possible to have lateral high-voltage devices on a normal substrate. The RESURF concept only cover the drift region, the part that should sustain the high voltage, and can therefore be used for different kinds of lateral devices. The first presented device was a lateral pn-diode, and it has since then been used in lateral bipolar and field effect devices. An interesting device is the lateral DMOS, LDMOS, that has shown very good performance using a very simple process that is suitable for integration with low-voltage logic.

1.2 High-Frequency Transistors

Traditionally telecommunication applications have used GaAs devices. GaAs has a higher saturation velocity than silicon which gives better high-frequency performance. The fact is that for really high electric fields the saturation velocity for GaAs decreases and becomes comparable to the saturation velocities for silicon.

For silicon devices, the bipolar transistor has been the most used type for high-frequency application. A reason for this is that since the active part of the device, the base region, is controlled by process parameters, i.e. diffusion, it has been relatively easy to achieve the narrow bases that are needed for high-frequency operation [14].

Another established high-frequency device is the silicon germanium heterojunction bipolar transistor, SiGe-HBT, which has good high-frequency properties [15]. A SiGe device is not entirely made in SiGe. The manufacturing starts with a silicon wafer and then SiGe is deposited in the active region of the device, i.e. the base region in a SiGe bipolar device. The amount of germanium and silicon is not 50-50 as can be interpreted from SiGe. The relation is $Si_{1-x}Ge_x$ where x is somewhere between 10% and 25%. The germanium profile can also be graded.

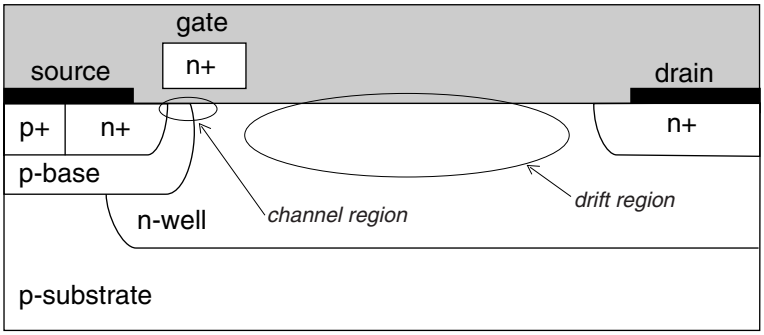


Figure 1.1: Schematic of an LDMOS transistor.

The most interesting high-frequency devices today are made in MOS-technology. The fastest processors in a computer has an operation frequency of several GHz and the production volume is very large. This is the driving technology for the main part of the semiconductor industry. Today gate lengths below 50 nm [16, 17], gate delays at 1 ps [16] and current gain cut-off frequencies above 100 GHz [18] have been presented.

1.3 The LDMOS Transistor

The LDMOS transistor has been around for many years and has primarily been used in switching applications. In the early 70's came the first publications that demonstrated the LDMOS in microwave operation [19]. An f_{MAX} of 10 GHz was presented and the main reason for this very impressive result was the fact that the channel length is a process parameter and not a lithography parameter.

Since then the use has increased steadily every year. Several manufacturers of devices for telecommunication have LDMOS transistors in their product line [20–22]. The advantages include properties like linearity and negative temperature coefficients. The possibility to incorporate the LDMOS in a standard CMOS process is also of great interest.

The structure of an LDMOS transistor is shown in figure 1.1. There are two distinct parts of the transistor that have different functionality, the channel region and the drift region. A conventional MOSFET has these two parts in the same place which makes it difficult to design the parts independently. The first part is the channel region, which is the actual transistor. The channel is formed by first implanting the p-base which is self-aligned to the poly gate edge. The p-base is then diffused under the gate and sets the drain edge

of the channel. Then the source contact implantation is done, also self-aligned to the same gate edge. The source implant is activated but the diffusion is much smaller than the p-base. This approach to form the channel makes it easy to have a short channel without having lithography with the same resolution, it only requires good control over the process. The transistors in this work have a channel length of 0.3 μm [23] in a 1.5 μm process and the gate lengths varies from 0.75 μm to 4 μm. The benefit of a short channel is that a high transconductance can be achieved which gives good high-frequency properties. Another advantage with a short channel is that the transistor always works in velocity saturation which makes the transconductance independent of the gate voltage and thus a good linearity.

The other important part is the drift region which extends from the drain end of the channel and to the drain contact. The drift region should support the applied voltage and protect the channel region from high voltage. The drift region design utilizes the RESURF concept, which is explained later. In a first approximation the breakdown voltage and also the on-resistance are linearly dependent on the drift region length for a well designed device.

A desirable property of the LDMOS transistor, unlike bipolar devices, is the negative temperature coefficient which is useful for high-voltage device since it does not require any external protection circuits.

Chapter 2

High-Voltage Operation

High-voltage operation using semiconductor devices is usually a trade-off between the on-state losses, determined by the on-resistance, and the maximum voltage that can be applied with full functionality of the device, the breakdown voltage. Traditionally a high breakdown voltage is achieved by using high-resistivity substrates that are almost semi-insulating, but the obvious problem is then that the resistance through the device is high when the device is in on-state.

In 1979 the RESURF concept was introduced and since then it has become a guideline for designing lateral high-voltage devices. In this chapter the RESURF concept is explained and improvements that can be made to further stretch the limits.

2.1 Breakdown Mechanisms

The breakdown in a semiconducting device occurs when the applied voltage becomes too high and a large current starts to flow through the device. Within the semiconducting material two types of breakdown are possible, *zener* and *avalanche* breakdown [4].

The zener breakdown is a tunneling breakdown and it occurs over a highly doped pn-junction. When both sides of the junction are highly doped the n-side conduction band will cross the p-side valence band at a relatively low reversed bias. Then there is a possibility for the electrons to tunnel from the p-side to the n-side and thus create a reverse current. For the zener breakdown to occur the doping profile over the junction should be sharp and the doping levels high, otherwise the avalanche breakdown will occur before the zener breakdown. The zener effect is used in zener-diodes which are used as voltage

references. The requirement for a zener breakdown was that both sides of the junction should be highly doped and the zener breakdown therefore occurs at low voltages. This gives us a hint that for a high-voltage device, were at least on side of the junction is low doped, the zener breakdown will not occur and here another type of breakdown occurs.

The other kind of breakdown in the semiconductor material is the avalanche breakdown. For a parallel-plane reverse biased pn-junction the voltage is primarily distributed over the depletion region. When the free carrier (electrons and holes) travel through the depletion region they gain energy from the accelerating electric field. They also lose energy when they collide with atoms in the lattice. If a carrier have gained enough energy and collide with an atom it will ionize the atom. Ionization simply means that the atom will lose one charge. The result of the collision is thus that the free carrier has created a new electron-hole pair. Then the new carriers accelerates by the electric field and eventually will be swept out of the depletion region and contributes to the leakage current. If the electric field is high enough something else happens. The first carrier created two extra carriers. Then the new carriers are accelerated and each of them eventually collide again and the 3 carriers becomes 9. The 9 carrier triples again and again and so on. This is called avalanche breakdown. This is the normal way a high-voltage device breaks down. This kind of breakdown is normally non-destructive but it may happen that the dissipated power gets high enough to destroy the device Appendix C shows how the breakdown voltage can be calculated for a parallel-plane diode.

As if this was not enough, the breakdown can also occur outside the semiconducting material. Take the gate oxide in a MOS transistor for example. If the potential over the oxide gets too high the oxide will not stand the high electric field and a current will pass through the oxide. This is called dielectric breakdown. Silicon-dioxide, SiO_2, for example has a dielectric strength that is 30 times higher than that of silicon.

2.2 The RESURF Concept

Traditionally high-voltage devices have been made on thick high-resistivity substrates. High-resistivity substrates give good high-voltage duration but the on-state losses are large. In 1979, Appels and Vaes presented the RESURF concept [13] which totally changed the way to design a lateral high-voltage device.

The concept of RESURF can be used on any lateral device since it only covers the drift region (the region that supports the high voltage). To illustrate the concept its use in a diode structure is explained. Consider the structure shown in figure 2.1a. The structure can be separated in two parts

each having a pn-junction, there is a vertical p^-/n^--junction and a lateral p^+/n^--junction. Considering the junctions separately the p^+/n^--junction has the lowest breakdown voltage. For a parallel plane junction the breakdown is proportional to $N_E^{-3/4}$ where $N_E = N_A N_D/(N_A + N_D)$ is the effective doping. If $N_D^{n^-} = N_A^{p^-} \ll N_A^{p^+}$, then $N_E \approx N_D$ for the p^+/n^--junction and $N_E = N_D/2$ for the p^-/n^--junction.

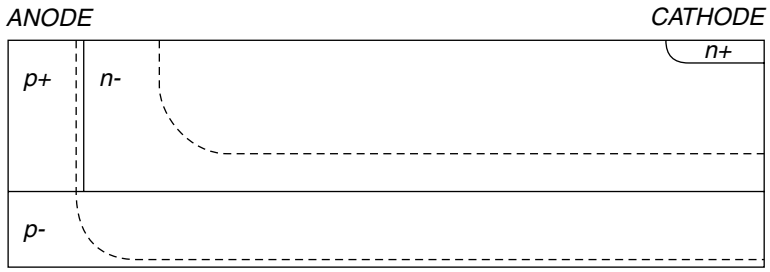
For a thinner n^--layer the depletion of the vertical p^-/n^--junction will interact with the depletion region from the lateral p^+/n^--junction as shown in figure 2.1b. The substrate will then deplete the p^+/n^--junction area from below and thus a larger depletion region is achieved which lowers the maximum electric field at the same applied voltage. Further increasing the applied voltage as in figure 2.1c will increase the depletion region all the way to the cathode contact and thus will the field be uniformly distribute all the way from the anode to the cathode. The breakdown then occurs when the p^-/n^--junction reaches its maximum voltage.

For the fully utilizing of the RESURF concept there are criteria that have to be fulfilled. The total charge in the n^--layer should be 10^{12} cm^{-2}, which is the same as $N_{epi}d_{epi}$. It has later been shown that the optimum charge is not a fixed number but it depends on the substrate and the epi-thickness [24].

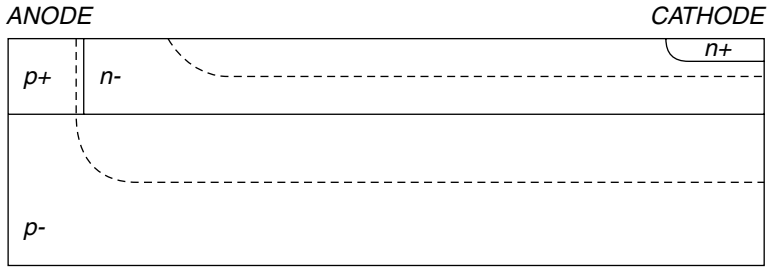
2.3 Advanced RESURF Concepts

The RESURF introduced a major breakthrough in the design of lateral high-voltage devices. As explained in the previous section the concept has some limitations. The on-resistance is primarily set by the desired breakdown voltage and can not be improved using the original principle. There are however several methods to expand the concept for improving the on-resistance [25, 26]. Recently after the first publication on the RESURF concept a double RESURF device was presented by the same authors [27]. The double RESURF means that the drift region is depleted from two directions. The implementation of the double RESURF is an extra p-doped layer in the surface of the drift region. This p-doped layer will deplete the drift region from above in conjunction with the substrate that depletes from below. This enables us to increase the total n-well charge to maintain charge balance and the on-resistance can be reduced by 50%.

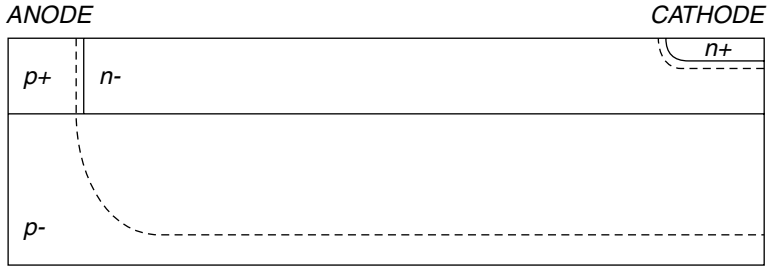
Another important design issue is the use of field plates. The use of field plates is not a RESURF specific technique but it is often used for these devices. A field plate has the functionality of controlling the electric field by putting metal plates (from the normal metalization) over critical regions in the structure. The metal plates can be connected to the appropriate voltage and are usually used to shield junctions. Figure 2.2 shows an example



(a)



(b)



(c)

Figure 2.1: Diode structures to explain the RESURF concept. (a) Diode with a thick epi-layer. (b) Thinner epi-layer at the same applied voltage as (a). (c) Same diode as (b) at a higher applied voltage.

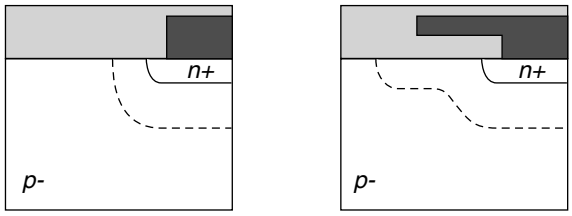


Figure 2.2: Example of the use of field plates to protect critical regions. (a) a device without a field plate and (b) a device with a field plate.

of a contact with a field plate and the same contact without the field plate. When applying voltage the electric field will be highest at the n^+-contact junction and the breakdown voltage will be determined by the junction curvature. By introducing the field plate the depletion region will be changed as in figure 2.2b. This will increase the voltage that can be applied on this junction and thus a higher breakdown voltage.

A problem with field plates is that high electric fields are created in the silicon under the edge of the field plate. These weak spots can be removed by extending the field plate with a semi-insulating layer that spreads the field more. The semi-insulating material used is polycrystalline silicon, SIPOS (semi insulating polycrystalline silicon). The SIPOS layer can also be extended to cover the whole drift region and if it is connected in the ends a small current will pass through which will further push the electric field to be more uniform. Bipolar devices with breakdown voltages of 1000 V and 1500 V using SIPOS layer have been presented [28]. Paper I shows a similar concept for a high-voltage high-frequency LDMOS transistor where the poly silicon gate extends all the way to the drain where it is connected to the drain potential. The part of the extended gate that is located over the drift region is intrinsic poly silicon and the edges are highly doped to form a reverse biased pn-diode. This allows for a leakage current from gate to drain which spreads the electric field and thus a higher breakdown voltage. The breakdown voltage is increased from 60 V to 240 V by introducing the extended gate, in the example in paper I.

A more easily implementable approach to reduce the on-resistance for a given structure and breakdown is to use the concept of the *p-top* [29, 30], [paper III]. The use of the p-top is an extension of the double RESURF concept and it is said that the device has dual conducting paths. The im-

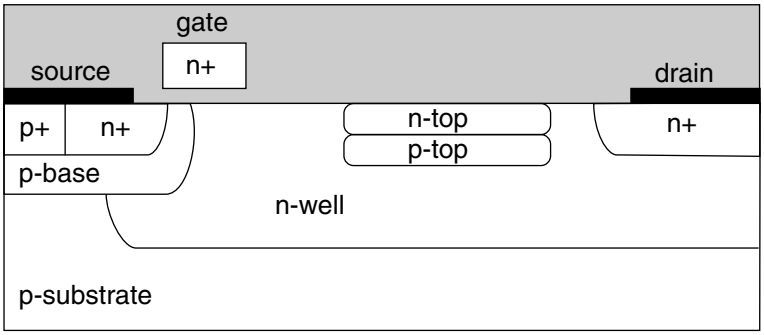


Figure 2.3: A LDMOS transistor with additional p- and n-doped layers, called p-top and n-top, in the drift region.

plementation of the p-top is shown in figure 2.3, which should be compared to the case without the p-top in figure 1.1. In other publications the p-top refers to the surface implant in a double RESURF structure and what I call p-top has been called p-buried but in this thesis p-top refers to the p-top as in figure 2.3. In the figure there is also another layer called *n-top* but more about that later. From the RESURF concept there is an optimal n-well charge for highest breakdown voltage and this sets a limit for the on-resistance. By introducing the p-top the n-well charge can be increased since it is affected by the p-top as well as the substrate. How much the charge can be increased and how much the on-resistance can be reduced is thoroughly studied in paper III.

Above the p-top is a thin n-doped layer which conducts current in the on-state and is quickly depleted in the off-state. To improve the conduction an extra n-doped layer can be implanted above the p-top as shown in figure 2.3 and the layer is called n-top. By adding the n-top the p-top charge has to be increased equally much not to disturb the charge balance. The result is then that the on-resistance can be reduced further with only a small penalty in the breakdown voltage.

The most important design issue about the n/p-top is the total charge in the different layers, but there are other issues. The p-top can be either grounded or floating, which gives some differences. The grounding is done in the third dimension by contacting the p-top to the p-base. In general a floating p-top will give better on-resistance but poorer breakdown voltage and it should probably be grounded in most cases due to a better control over the operation. Another issue is the lateral length and position of the p-top. It should preferably be as long as possible to have the improved effect over the whole drift region. But if the p-top comes to close to the drain contact there will be problems and the device will have much lower breakdown. On the

other side the p-top can not be to closed to the p-base since it will then cut off the conducting path from the channel end to down to the n-well. It also affects the input capacitance if it comes to close to the gate. Then there is the location of the n-top. In the simplest case it is implanted using the same mask as the p-top as in figure 2.3 or it can be implanted without a mask and thus extend all the way from the gate to the drain contact. The most optimal case is probably somewhere in between, but this have not been studied yet. Comparing the case using the p-top mask with the case without the mask shows that without the mask the on-resistance is lowered but the breakdown voltage suffers.

An interesting device that is basically a VDMOS is the CoolMOS^{TM} [31, 32]. A problem with vertical devices is the poor on-resistance but the CoolMOS^{TM} concept offers a five times reduction of R_{on}. The concept can be seen as the double RESURF concept turned 90 degrees. In a cross section of the device there will be colons of n- and p-doped material where the n-doped is the drift region were the current flows. The p-doped colons are used to deplete the n-doped drift region quickly in the blocking state and this allows the n-doping to be one magnitude higher which decreases the on-resistance. The drawback of the concept is the manufacturing of the colons. This is made during an epitaxial deposition process which makes it rather expensive.

The concept with multiple layer in the drift region has also been proposed for lateral devices [33, 34].

Chapter 3

High-Frequency Operation

The advantages for the LDMOS transistor in high-frequency operation are mainly due to an easily achieved short channel. A short channel makes the device operate in velocity saturation which gives a constant transconductance, resulting in good linearity [35]. The use of LDMOS transistors at high-frequencies has mainly been in the area of telecommunication where the frequencies of operation are up to some GHz today, but they are increasing. The standard supply voltage in a base station is 28 V which needs devices with a breakdown voltage around 50-60 V. The output power density of these devices is about 0.5 W/mm at 2 GHz [21]. Normal operation of these devices are class AB, and work on class E has shown that increasing supply voltage gives drastically increases output power [36]. It has also been shown, for dual depletion layer LDMOS devices, that if the supply voltage is increased the efficiency is increased and power densities up to 2 W/mm have been presented at V_D=70 V and 1 GHz [paper V]. Figure 3.1 shows the 20 mm gate width device from the paper.

Other application areas that can utilize these devices are in lightning and heating, using micro-magnetrons. The LDMOS has also been shown to give competitive performance for low-voltage applications [paper II], with supply voltages at 5 V and breakdown voltages at 15-30 V.

When characterizing high-frequency devices there are several different key parameters. There are the basic DC-parameters such as on-resistance, transconductance and breakdown voltage. Then there are the small-signal parameters of which the cut-off frequencies f_T and f_{MAX} are the most common. The current gain cut-off frequency, f_T is the frequency at which the current gain, H_{21}, is 0 dB. The maximum frequency of oscillation, f_{MAX}, is the frequency where the unilateral power gain, U_{pg}, is 0 dB. The unilateral power gain is the maximum gain that can be achieved assuming that the

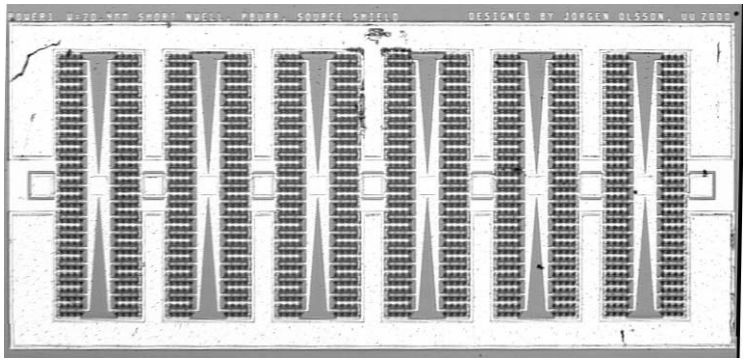


Figure 3.1: A power LDMOS transistors, with a gate width of 20 mm.

source and load impedances are conjugately matched to the in and output impedance of the device. The f_{MAX} is also the frequency when the maximum available/stable gain is 0 dB.

For the large-signal characteristics there are parameters such as the power gain, power density and power added efficiency (PAE) that are important. These parameters are measured using load-pull measurements. Other important parameters include linearity and intermodulation which are very important for e.g. WCDMA (a mobile telecommunication standard).

3.1 High-Frequency Characterization

There are several different methods to characterize high-frequency devices. The methods includes scattering parameter (S-parameter) measurements using a network analyzer. The S-parameter measurement is a small-signal measurement and a main issue is calibration and to accurately de-embed the parasitics. For power measurements the load-pull measurement is a very useful technique, although it is expensive and time consuming.

3.1.1 S-parameters

The most important tool for high-frequency characterization is the network analyzer. The type of network analyzer primarily used in this work is a *vector network analyzer* (VNA). The basic principle is that a power wave is transmitted into the *device under test* (DUT) and the reflected power wave is measured in magnitude and phase. There are network analyzers that only measures the magnitude of the reflected wave and they are called scalar network analyzers. Usually more than one port of the network analyzer is used and then

the transmitted power waves from one port to the other are also measured. For the most common case when two ports are used (2-port measurement) the result is two reflection coefficient and two transmission coefficient. These coefficients are called scattering parameters. The S-parameters are usually represented in matrix form,

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \tag{3.1}$$

where S_{ij} is the measured power on port i when the power wave is applied on port j. Using a VNA for the measurements gives each S-parameters with magnitude and phase. In an ordinary network analyzer the port is terminated with 50 Ω. The S-parameter matrix can be used to calculate different small-signal gains and they can be converted to other two-port representations such as impedance Z, or admittance Y for e.g. parameter extraction.

The first thing to consider when doing S-parameter measurements using a VNA is to properly calibrate the system. All cables and connectors between the VNA and the DUT will affect the measurement and they should be calibrated so that the reference plane is located as close to the DUT as possible. This is done by replacing the DUT with calibration standards with known characteristics. There are numerous different calibration methods that have different complexity and correctness for different frequency ranges. The simplest case is the open calibration, which assumes that the parasitics between the signal and ground are capacitive. Then the DUT is replaced by an open structure which is assumed to reflect the whole signal and transmit zero signal. Measuring the open structure gives a response that represents the cables and connectors. They can then be de-embedded from the measurement on the real device. The open calibration can be used when the frequencies are relatively low, but for measurements in the GHz-range more advanced methods should be used. Other methods includes the SOLT (short-open-load-thru) which uses 4 calibration standards, and the TRL/LRM which uses 3 standards. A modern VNA has several of the calibration methods built-in and the VNA removes the errors automatically from the measurements so that the result is the DUT.

The next problem is if the devices are mounted or not. The device can be mounted in a package and then the measured result includes the packaging parasitics, bond wires between the chip and the package, the metal structure on the chip, etc. This gives the true performance of the package device and since all devices are sold in a package this is relevant. But in research and development the intrinsic behavior of the devices is of great interest. Then the devices are still on the wafer and then on-wafer measurements have to be done. On-wafer measurements are done using wafer probes that are needles that are put on the metal pads on the wafer.

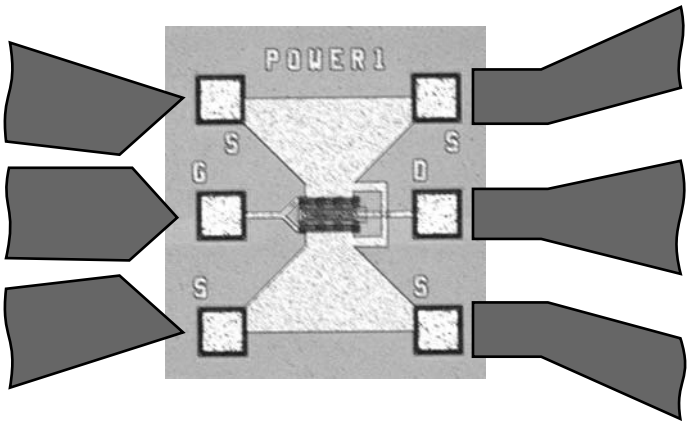


Figure 3.2: A typical pad structure which clearly shows the ground-signal-ground pad configuration.

Figure 3.2 shows a typical structure where the GSG (ground-signal-ground) pads are clearly shown. The drawings outside the picture shows how the probe tips looks like. To the right a Cascade Air Coplanar Probe [37] is shown and to the left a Picoprobe [38] is shown. The typical distances in the figure are the pad size that is 70×70 μm and the pitch (center-to-center distance between the pads) that is 150 μm.

To calibrate the probes a calibration substrate is used. The calibration substrate consists of several metal lines that create the calibration standards needed. This makes it possible to put the reference plane at the probe tips. But the problems does not end here. First, the contact between the probe tip and the wafer will change between different touch downs. Just lifting the probe and lowering it again on the same spot will give a different contact resistance. It is therefore important to probe with the same pressure and skating every time. Skating is the movement of the probe over the metal when the probe tip has landed. Another problem is that the interest is in the intrinsic device and not the surrounding metal pattern. This takes us to the third big issue in high-frequency characterization, de-embedding.

3.1.2 De-Embedding

De-embedding is how to remove, de-embed, the surrounding parasitics so that only the intrinsic part of the device can be characterized. Numerous ways to de-embed the parasitics with different levels of complexity and accuracy, have been presented [39–41], [paper VII].

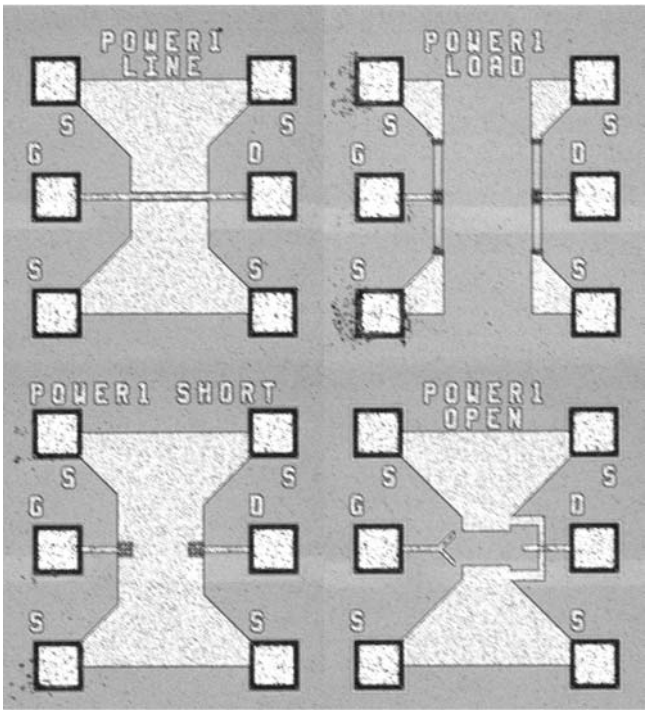


Figure 3.3: The manufactured on-wafer calibration standards used in paper VII.

A straight forward and intuitively simple way is to represent the parasitics with a lumped circuit model. The simplest case is to represent the pad parasitics using a capacitance that should represent the capacitance between the pad metal and the substrate.

There are more advanced lumped representations of the pad parasitics [41]. Then several de-embedding structures are manufactured together with the transistors that each tries to isolate some of the lumped parameters and thus by combining several measurements get a lumped model for the parasitics that can be de-embedded from the transistor measurements.

Another approach is to make the DUT with different widths using the same metal pattern and then approximate the parasitics with the admittance that are achieved when the width is extrapolated to zero [40].

A third method is to use the same calibration routines that are built-in in the VNA [paper VII]. Figure 3.3 shows the calibration standards that are used when measuring a device as in figure 3.2. In the figure are the four standards

used in a SOLT-calibration. In the short standard (S) the device is replaced by a metal sheet that shortcuts the signal pad to the ground pads. The open structure (O) has exactly the same metal pattern as the real device but the intrinsic device is removed. In the load structure (L) the device is replaced by poly silicon resistors that are dimensioned from test pattern data to result in a 50 Ω load between the signal pad and the ground pads. For the thru line (T) the signal pads are connected with a metal strip above the ground plane and the sizes are designed to give a characteristic impedance of the thru line of 50 Ω. The on-wafer calibration standards can also be used for TRL/LRM calibrations. The de-embedding can be done either by using the calibration structures shown in figure 3.3 when calibration the network analyzer or by calibration on the standard ISS (impedance standard substrate) and then do the de-embedding afterward.

3.1.3 Power Measurements

For power measurement the load/source-pull measurement is common. In the measurement setup each port has a variable impedance tuner. The tuner can be either manual or automated. During the measurement the are tuners varied (pulled) and the result will be isoplots of the gain over the Smith chart. This is used to find the optimum loads for gain or efficiency. The results will be gain, output power and efficiency as function of input power.

3.1.4 Impedance Measurements

The S-parameter measurements of transistors are usually made by sweeping the frequency for several bias points. This makes it unnecessary complicated to study the AC-performance as function of applied bias. For example, to study the input capacitance, C_{in}, as a function of drain voltage, would require a s-parameter measurement for each drain voltage since the network analyzer can not handle the DC-bias by itself. Then all these s-parameters measurements should be put together and converted to the Y-parameters that give C_{in} ($Y_{11} = G_{in} + j\omega C_{in}$). A more convenient way is to use some kind of impedance analyzer that measures the impedance (or admittance) while sweeping either the frequency or the applied bias. The impedance analyzer conducts an one-port measurement so transmission parameters can not be measured. From the measurement the result can be interpreted in different ways. It can for example be seen as a resistance in series with a capacitance, a parallel combination of a resistance and an inductance or a reflection coefficient. Appendix E shows some of the possible representations and how to convert the results between them.

3.2 High-Frequency Modeling

In all kind of transistor modeling there are basically 3 different kinds of models [42]. The first is the physical model, which is based on physical phenomena. These models can predict what happens when a physical parameter is changed. This is the preferred way to model transistors but it may led to very complex models that are less suitable for large simulations, e.g. circuits simulations consisting of thousands of devices. The second kind of models is empirical models. These use appropriate mathematical functions to represent the behavior and have several "fitting" parameters with no physical meaning. These models are difficult to use when trying to predict changes in the process etc. They are also often bad in predicting the behavior outside the region of validity. The third kind of models is table based models that are achieved by measuring the device under several different conditions and then the results are put in a table together with appropriate interpolation functions. Many models are combinations of the above mentioned kinds of models. Often a model is physically based but some complex behaviors are represented with empirical functions or uses fitting parameters, these models are sometimes called semi-empirical.

An important kind of model for high-frequency transistors is the small-signal equivalent circuit. Small-signal means that the applied signal is assumed to be small enough that the response is linear, i.e. no higher order harmonics are created. The equivalent circuit consists of lumped elements, e.g. resistances and capacitances, that should represent the device behavior when small signals are applied. A small-signal equivalent circuit for a FET-device is shown in figure 3.4.

The outer parts named, Y_g and Y_a, represents the parasitics that should be de-embedded using an appropriate method. The parts named Z_i are the series impedances that represents the contact resistances and the inductances in the connectors. The remaining part is the *intrinsic* transistor, which represents the actual device.

After the small-signal comes the large signal. The small-signal model is a linear model that is valid for small applied signals. The large signal model should resemble the characteristics from DC up to microwave frequencies. This should be done for a large range of drain and gate biases. Most large signal models are non-linear. Non-linear models are defined as those models that cannot be analyzed using the superposition principle. Another definition of non-linearity is the creation of higher order harmonics, i.e. applying a signal with frequency, f, on the input will create a signal with frequencies, f, $2f$, $3f$, etc. on the output. An example to visualize this is to imagine a large signal, with frequency f, that is applied on the gate. If the signal is large it will vary the drain current from 0 ($V_G < V_T$) to I_{Dsat}. If the amplitude of

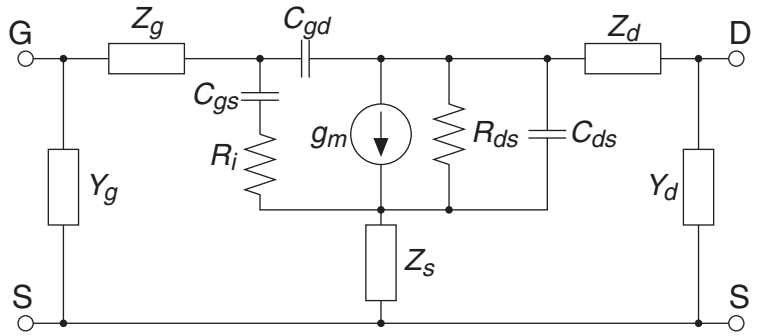


Figure 3.4: Simple small-signal equivalent circuit for a FET-device.

the input signal is large enough the drain current characteristics will resemble a square topped signal. From elementary Fourier analysis it is known that a square wave has harmonics as multiples of the base frequency.

3.2.1 Network Analysis

This section will explain some fundamental concepts useful when dealing with equivalent circuits and parameter extraction. The basis for setting up analytical expressions for the admittance or impedance matrices is also provided.

The theory for setting up an analytical admittance matrix is based on nodal analysis. In the fundamental form every node in a circuit is given a number. The ground node can be included, this is called floating admittance matrix, or not included, then it is called grounded admittance matrix. The total number of nodes is then N which makes the admittance matrix a $N \times N$-matrix. The sum of all elements that are connected to node i should be written in the position for Y_{ii} in the matrix. The elements should be written in the admittance form, e.g. a resistance is written as G or $1/R$ and a capacitance as $j\omega C$. In place Y_{ij} should all elements connected between node i and j be written with a minus sign. This is the case for all passive elements in the circuit (R, L and C).

As an example the circuit in figure 3.5 is used. Applying the above rules for the passive elements in the circuit yields:

$$Y_p = \begin{pmatrix} j\omega(C_{gs}+C_{gd}) & -j\omega C_{gd} \\ -j\omega C_{gd} & G_{ds}+j\omega(C_{ds}+C_{gd}) \end{pmatrix} \tag{3.2}$$

Repeating the general rules yield that connected to node 1 is the capacitances, C_{gs} and C_{gd}, which are added and multiplied by $j\omega$ to get the admittance form and inserted in place Y_{11}. Between node 1 and 2 are just C_{gd}

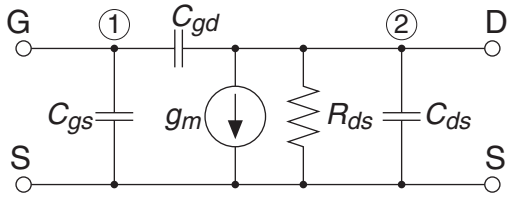


Figure 3.5: Intrinsic small-signal equivalent circuit.

connected and is therefor inserted in both Y_{12} and Y_{21} with a minus sign. A general consequence of the passive Y matrix is that $Y = Y^t$, i.e. $Y_{ij} = Y_{ji}$.

What about g_m then? The transconductance is represented with a voltage-controlled current-source which is an active element. To treat g_m the correlation between the Y-matrix and the nodal voltages and currents must be established.

$$\begin{pmatrix} i_1 \\ i_2 \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{3.3}$$

Since the circuit can be treated as a linear circuit the active and passive part can be treated apart and then be added together in the end (superposition). So for now, remove all passive elements in figure 3.5 so that just g_m remains. The result is that node 1 is floating and node 2 is connected to ground through g_m. One strategy to evaluate the active Y-matrix, Y_a is to change v_1 and v_2 and see how this affects the currents. When v_2 is changed nothing happens. This means that $Y_{12} = Y_{22} = 0$. If v_1 is increased the current through g_m will increase with $g_m v_1$ which means that Y_{21} should be g_m and nothing happens to i_1 so $Y_{11} = 0$. The active Y-matrix is then summarized as (3.4) which yields the total Y-matrix (3.5) for the circuit.

$$Y_a = \begin{pmatrix} 0 & 0 \\ g_m & 0 \end{pmatrix} \tag{3.4}$$

$$Y = Y_a + Y_p = \begin{pmatrix} j\omega(C_{gs} + C_{gd}) & -j\omega C_{gd} \\ g_m - j\omega C_{gd} & G_{ds} + j\omega(C_{ds} + C_{gd}) \end{pmatrix} \tag{3.5}$$

This is the basic methodology for creating an admittance matrix. To get the impedance matrix Z for the same circuits the relation is $Z = Y^{-1}$. The circuit in figure 3.5 is the simplest case of an intrinsic equivalent model. To make it more realistic some series impedances can be added on each terminal as in figure 3.4. The admittance matrix for this circuit will have more than 2 nodes. Then the admittance matrix has to be reduced to a two-port Y-matrix when comparing to measured data which usually are two-port data.

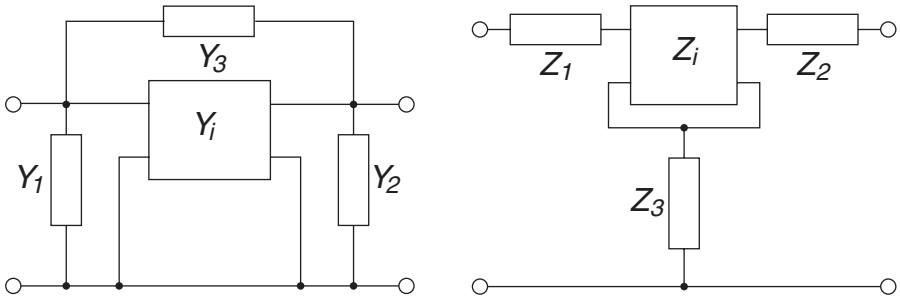


Figure 3.6: Circuit examples to explain the relation between Y and Z parameters.

Another way to do this that is also good for the fundamental understanding about de-embedding is to successively build up the matrix. Consider the circuit shown in figure 3.6a. The black box named Y_i is an arbitrary two-port network represented with its Y-parameters. The total admittance, Y_e, for the circuit is

$$Y_e = \begin{pmatrix} Y_{i,11} + Y_1 + Y_3 & Y_{i,12} - Y_3 \\ Y_{i,21} - Y_3 & Y_{i,22} + Y_2 + Y_3 \end{pmatrix} \tag{3.6}$$

The admittances $Y_{1,2,3}$ can apparently be separated from Y_i easily. The Y_e matrix can thus be expressed as $Y_e = Y_i + Y_{123}$ where Y_{123} is given by:

$$Y_{123} = \begin{pmatrix} Y_1 + Y_3 & -Y_3 \\ -Y_3 & +Y_2 + Y_3 \end{pmatrix} \tag{3.7}$$

If the analytical expression for Y_i is known and the circuit topology shows a Π-structure, such in figure 3.6, it is easy to add the elements that corresponds to $Y_{1,2,3}$. This can also be seen as a parallel circuit with Y_i and Y_{123} were the result is the sum of the admittances. For the circuit in figure 3.6b the Z-matrix can be expressed in a similar way.

$$Z_e = \begin{pmatrix} Z_1 + Z_3 & Z_3 \\ Z_3 & Z_2 + Z_3 \end{pmatrix} + Z_i \tag{3.8}$$

The other view of this is that the impedances are in series were they can be added.

This can be used to successively build up a Y or Z-matrix. This is also a fundamental technique when embedding or de-embedding parasitics, or series resistances from measurements [43]. Take for example the circuit in figure 3.4. The intrinsic device Y_i is expressed similar to (3.2). Then the series Z-impedances are seen to correspond to figure 3.6b and the Z-matrix is Z_e (3.9). The pad admittances Y_g and Y_d can be summarized in the admittance matrix Y_a (3.10). The total Y-matrix is then given by (3.11).

$$Z_s = \begin{pmatrix} Z_g + Z_s & Z_s \\ Z_s & Z_d + Z_s \end{pmatrix} \tag{3.9}$$

$$Y_a = \begin{pmatrix} Y_g & 0 \\ 0 & Y_d \end{pmatrix} \tag{3.10}$$

$$Y_m = Y_a + [Z_s + Y_i^{-1}]^{-1} \tag{3.11}$$

If the measured data, Y_m, are known, the pad parasitics, Y_a, are known and the series impedances, Z_e, are known the intrinsic Y_i can be calculated as

$$Y_i = [(Y_m - Y_a)^{-1} - Z_s]^{-1} \tag{3.12}$$

3.3 Parameter Extraction

When using small-signal equivalent circuits it is important to extract the parameter values in a correct and efficient way. A method that almost always works, but seldom gives reliable values is optimization. Optimization is what it says, using a starting approximation an optimizer tries to find the solution that minimizes the error between the measured data and the model. The optimizer usually finds the nearest local minimum and then stops. This is usually not the global minimum. Most simulation tools include some kind of optimizer.

A more reliable method is to use some kind of direct extraction [43]. The important thing is not to rely on a good starting approximation. Assume for example that the measured data should be applied on the equivalent circuit shown in figure 3.5. Here the series impedances are assumed to be de-embedded in a previous step. The admittance matrix for this circuit is shown in (3.13).

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} j\omega(C_{gs} + C_{gd}) & -j\omega C_{gd} \\ g_{m0}e^{-j\omega\tau} - j\omega C_{gd} & G_{ds} + j\omega(C_{ds} + C_{gd}) \end{bmatrix} \tag{3.13}$$

By plotting $-\Im(Y_{12})/\omega$ versus the frequency the model predicts that the plot should show a straight horizontal line at the value of C_{gd}. If the plot does not show a constant value the explanation can be that the measured data is incorrect or that the model is wrong. But if the error is small an approximate value of C_{gd} can be extracted. The same can be done for the rest of the parameters using e.g. equations (3.14-3.19).

$$C_{gd} = -\frac{\Im(y_{12})}{\omega} \tag{3.14}$$

$$C_{gs} = \frac{\Im(y_{11}+y_{12})}{\omega} \tag{3.15}$$

$$C_{ds} = \frac{\Im(y_{12}+y_{22})}{\omega} \tag{3.16}$$

$$G_{ds} = \Re(y_{22}) \tag{3.17}$$

$$g_{m0} = |y_{21}-y_{12}| \tag{3.18}$$

$$\tau = -\frac{1}{\omega}\arctan\left(\frac{\Im(y_{21}-y_{12})}{\Re(y_{21}-y_{12})}\right) \tag{3.19}$$

This is an example of an overdetermined equation system. There are 6 unknown and there are 8 independent measured parameters. Actually there are two measured parameters that are not used and they are expected to be zero in the model, $\Re(y_{11})$ and $\Re(y_{12})$. By examining them another indication of the validity of the model is revealed. There are other ways to extract the parameters, consider for example (3.17). What happens if $\Re(y_{12}) \neq 0$, then y_{12} has a resistive part which may be modeled by putting a resistance in series with C_{gd}, this is not an improbable case. This resistive part will also be seen in y_{22} which implies that a more correct value of G_{ds} is achieved by $\Re(y_{22}+y_{12})$

3.3.1 Series Impedances

As been seen in the previous section for the intrinsic part of the transistor it is fairly easy to extract the parameters directly if the equivalent circuit is not to complicated. A more challenging task is to extract the series impedance which consists of, in the general case, a resistance in series with an inductance on each terminal. In some cases some of the series elements can be neglected without degrading the model.

A simple method to extract the series resistance [44] assumes that the intrinsic part of the transistor is purely capacitive when the transistor if off, i.e. all applied voltages are zero, and that the series impedances are purely resistive. The method also assumes that the series resistances are bias inde-

pendent. By measuring the Z-parameters when the device is off the following relations are achieved:

$$R_s = Z_{12} = Z_{21} \tag{3.20}$$
$$R_g = Z_{11} - R_s \tag{3.21}$$
$$R_d = Z_{22} - R_s \tag{3.22}$$

This method has sometimes been called the coldFET method but one can question its validity on MOSFET's. The coldFET method was developed for MESFET's which has a fundamental difference compared to a MOSFET. The gate is dielectrically isolated from the drain and source and this is not the case for a MESFET. For the original coldFET method the MESFET is biased a V_D=0 and V_G greater than the barrier height. Then some assumptions about the intrinsic part is done which leads to the series impedances [45, 46].

The above discussed method has some major drawbacks such as the assumption that the inductances are neglected, the series resistances are assumed to be bias independent and the assumption that the intrinsic part is capacitive in off-state. The next step would be to put back the inductances and try to extract the parameters when the device is on which is the case when the device is used. By assuming the equivalent circuit similar to the one shown in figure 3.7 and putting up the expressions for the Z-parameters the following is achieved by looking at the real and imaginary parts separately [47],

$$\Re(Z_{12}) = R_s + \frac{A_s}{\omega^2 + B} \tag{3.23}$$
$$\Re(Z_{22} - Z_{12}) = R_d + \frac{A_d}{\omega^2 + B} \tag{3.24}$$
$$\Re(Z_{11} - Z_{12}) = R_g + \frac{A_g}{\omega^2 + B} \tag{3.25}$$
$$\frac{\Im(Z_{12})}{\omega} = L_s - \frac{E_s}{\omega^2 + B} \tag{3.26}$$
$$\frac{\Im(Z_{22} - Z_{12})}{\omega} = L_d - \frac{E_d}{\omega^2 + B} \tag{3.27}$$
$$\frac{\Im(Z_{11} - Z_{12})}{\omega} = L_g - \frac{E_g}{\omega^2 + B} - \frac{F_g}{\omega^2(\omega^2 + B)} \tag{3.28}$$

where the constants A_x, B, E_x and F_x are functions of the intrinsic parameters and independent of the frequency. The series resistances and inductances are then extracted by doing a curve fit to the above equations. A sanity check can be made by investigating the different B's that are achieved from every

expressions and they should be the same. For (3.28) the curve fit can be done at higher frequencies when the third term on the right hand side can be neglected.

To visualize the above extraction technique the series resistances and inductances can be extracted by manipulation (3.23- 3.28) mathematically [40]. By for instance combining (3.23) and the corresponding expression using $\Re(Z_{21})$ which is the same with a different A_g the following is achieved,

$$\Re(Z_{12}) = \left(1 - \frac{A_{12}}{A_{21}}\right) R_s + \frac{A_{12}}{A_{21}} \Re(Z_{21}) \tag{3.29}$$

then by plotting $\Re(Z_{12})$ versus $\Re(Z_{21})$ the result should be a straight line with the intersect $(1 - A_{12}/A_{21})R_s$ and the slope A_{12}/A_{21} and then can R_s easily be evaluated.

The previous methods, for extracting the series resistances and inductances, make assumptions and approximations that seems to work in some cases. In the general case it may not be possible to make approximations without losing accuracy. In paper VI, a different approach on how to extract the series impedances is proposed. The method does not introduce any approximations or assumptions. The method does, as all methods, need a good model. By measuring the S-parameters and after the necessary de-embedding of the pad parasitics a set of Z-parameters are achieved. For the circuit in figure 3.7 for example the analytical expression for the different Z-parameters can be evaluated. The expressions can be rewritten to the form:

$$Z_{ij} = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_n s^n}{1 + b_1 s + b_2 s^2 + \cdots + b_m s^m} \tag{3.30}$$

where s is the complex frequency ($s = j\omega$). The parameters a and b can be extracted by fitting the measured Z-parameters to the analytical expressions. The fit is done using the least square method which does not require any starting values of the parameters. Each of the a and b coefficients can be expressed as a function of the circuit parameters and, for the example in figure 3.7 and most MOS models, the series inductances can be evaluated as the ratio between a_n and b_m, i.e.

$$L_g + L_s = a_n^{11}/b_m \tag{3.31}$$

$$L_s = a_n^{12}/b_m \tag{3.32}$$

$$L_s = a_n^{21}/b_m \tag{3.33}$$

$$L_d + L_s = a_n^{22}/b_m \tag{3.34}$$

where a_n^{ij} is the highest order coefficient for the numerator in Z_{ij} and b_m is the highest order coefficient for the denominator for any Z_{ij} (all different Z_{ij} have

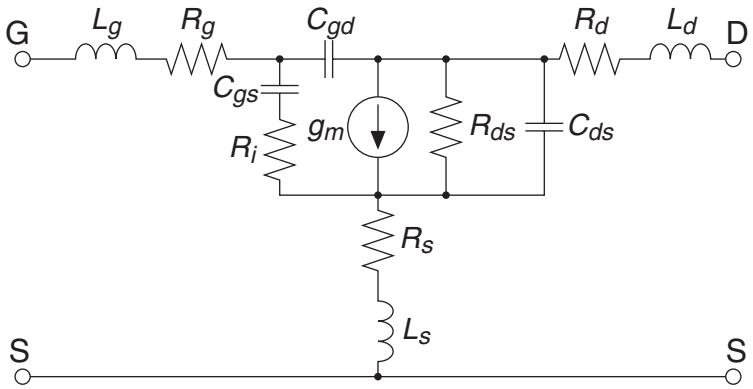


Figure 3.7: Example of a small-signal equivalent circuit.

the same denominator). The series resistances can be extracted in a similar way. This methodology can also be used to extract the intrinsic parameters.

3.3.2 Example of an Extraction

An example of how to extract the small-signal equivalent circuit model parameters for a low-voltage LDMOS transistor is shown together with some considerations that should be taken into account when doing the parameter extraction. The equivalent circuit that should be considered is shown in figure 3.7.

First the series impedances should be extracted. Using the method in paper VI, the first thing to do is to set up the analytical expressions for the Z-parameters and look for the orders of the polynomials in the numerator and denominator. Be aware that some of the equations must be multiplied with s so that all a_0's are non-zero and $b_0 = 1$. The result is shown in (3.35-3.38)

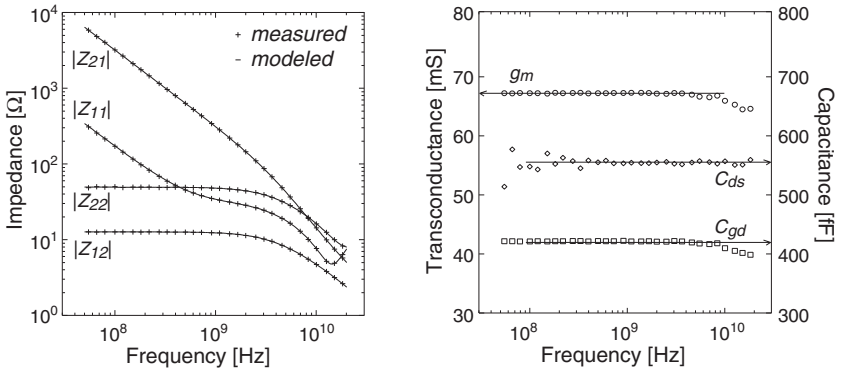
$$sZ_{11} = \frac{a_0 + a_1 s + \cdots + a_4 s^4}{1 + b_1 s + \cdots + b_2 s^2} \tag{3.35}$$

$$Z_{12} = \frac{a_0 + a_1 s + \cdots + a_3 s^3}{1 + b_1 s + \cdots + b_2 s^2} \tag{3.36}$$

$$sZ_{21} = \frac{a_0 + a_1 s + \cdots + a_4 s^4}{1 + b_1 s + \cdots + b_2 s^2} \tag{3.37}$$

$$Z_{22} = \frac{a_0 + a_1 s + \cdots + a_3 s^3}{1 + b_1 s + \cdots + b_2 s^2} \tag{3.38}$$

Then the coefficient are extracted using the method in paper VI. At this



(a) (b)

Figure 3.8: (a) The first least square fit to the measured Z-parameters. (b) The direct extraction of some of the intrinsic parameters.

point it should be pointed out that the above circuit generates 25 different coefficient but there are only 14 unknowns so the equation system is overdetermined. For example, the a_4 term in (3.36) is the same as the a_5 term in (3.37). This could have been taken care of when setting up the equation system but the result then will be the same as replacing them with their mean value. There are other coefficients that should be the same and there are also more complicated relations. The first least square fit for the Z-parameters are shown in figure 3.8. As can be seen the first least square fit easily finds the coefficients that make a nearly perfect fit to the measured data.

The series inductances are then calculated using (3.31-3.34). Then a new set of coefficient are calculated which corresponds to the analytical expression were the inductances are removed. How this is done is also shown in paper VI. From this new set of coefficient the series resistances are extracted in the same way as the inductances. The intrinsic Y-matrix can be calculated using (3.39), where Z_e is the measured Z-matrix.

$$Y_i = [Z_e - Z_s]^{-1} \tag{3.39}$$

$$Z_s = \begin{pmatrix} R_g + R_s + s(L_g + L_s) & R_s + sL_s \\ R_s + sL_s & R_d + R_s + s(L_d + L_s) \end{pmatrix} \tag{3.40}$$

The analytical expression for Y_i look as:

$$Y_i = \begin{pmatrix} j\omega(C_{gs} + C_{gd} + \frac{C_{gc}}{1+j\omega C_{gc}R_i}) & -j\omega C_{gd} \\ \frac{g_m}{1+j\omega C_{gc}R_i} - j\omega C_{gd} & \frac{1}{R_{ds}} + j\omega(C_{gd} + C_{ds}) \end{pmatrix} \tag{3.41}$$

The intrinsic parameters will be extracted using traditional direct extraction methods. For this particular circuit the following expressions will get the intrinsic parameters.

$$C_{gd} = -\frac{\Im(Y_{12})}{\omega} \tag{3.42}$$

$$C_{ds} = \frac{\Im(Y_{12} + Y_{22})}{\omega} \tag{3.43}$$

$$R_{ds} = \frac{1}{\Re(Y_{22})} \tag{3.44}$$

$$g_m = \frac{\Re(Y_{21} - Y_{12})^2 + \Im(Y_{21} - Y_{12})^2}{\Re(Y_{21} - Y_{12})} \tag{3.45}$$

$$C_{gc} = -\frac{g_m \Re(Y_{11} + Y_{12})}{\omega \Im(Y_{21} - Y_{12})} \tag{3.46}$$

$$C_{gs} = \frac{1}{\omega}\left(\Im(Y_{11} + Y_{12}) + \frac{\Re(Y_{11} + Y_{12})\Re(Y_{21} - Y_{12})}{\Im(Y_{21} - Y_{12})}\right) \tag{3.47}$$

$$R_i = -\frac{1}{\omega C_{gc}} \frac{\Im(Y_{21} - Y_{12})}{\Re(Y_{21} - Y_{12})} \tag{3.48}$$

Some of the parameters in (3.42-3.48) are plotted against frequency in figure 3.8 where it is clearly shown that the parameters are constant versus frequency. When all parameters are extracted the circuit is simulated using the extracted parameters. The correspondence between measured data and the simulated are shown in figure 3.9.

3.4 Numerical Simulation

An important tool for fundamental understanding of the mechanisms that controls the operation of semiconductor devices are the TCAD (technology CAD) simulation tools [48–50]. The simulators are fed with a structure where all different regions of material and the doping profiles are defined. The structure is also supplied with a FEM-mesh (finite element method) which should be dense enough to cover all active regions in the structure. Then the boundary conditions are set, such as terminal voltages and ambient temperature. The simulator then solves the basic electro-magnetic differential equations in

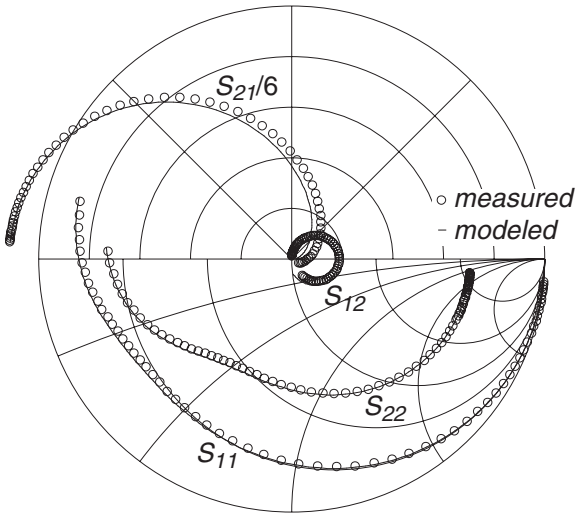


Figure 3.9: The measured S-parameters compared to the S-parameters obtained from the extracted equivalent circuit.

all mesh points and iterates until the solution is reached. When the solution is known all terminal voltages and currents can be seen. The terminal voltages and currents could be measured on a real device as well, but from the simulation the carrier concentrations, the potential distribution and other things can be visualized in every mesh point. This is of valuable help when trying to recognize critical point inside the structure, were for instance the breakdown starts.

It is also possible to perform other simulations than the steady-state simulations. AC-simulations are valuable since it gives the small-signal AC-performance in term of terminal capacitances and conductances. It is also possible to achieve the scattering parameters. An interesting simulation is the transient simulation which gives the possibility to monitor how the current flows and the potential distribution changes over time when applying a step function or an sinusoidal signal.

Another valuable benefit from the numerical simulations is the fact that it is easy to see how process or design changes affects the performance. It might take a day to simulate how a certain doping level affects the breakdown. To do this the traditional way by manufacturing the devices and then measuring the devices might take 3 months if you can squeeze in that into your project.

Numerical simulations have been used extensively in papers I and III.

Chapter 4

Summary

The LDMOS transistor has gained large interest lately, especially as an RF power transistor. Two important properties of an RF power transistor are the high-voltage performance and the high-frequency performance. This thesis discusses how the high-voltage properties can be improved in terms of on-resistance and breakdown voltage. These two properties are usually in contradiction so the answer is not simple. The other part of the thesis cover the high-frequency part. The focus is on high-frequency modeling and parameter extraction.

When discussing high-voltage lateral devices the concept of RESURF always comes up. The RESURF makes it possible to manufacture high-voltage lateral devices on normal substrates. Usually high-voltage devices are made on special high-resistivity material which is useless for normal transistors. The RESURF concept can be used for all kinds of lateral devices since it covers the voltage supporting drift region and therefore the active part can be for example a diode, a bipolar transistor or a MOS transistor. The RESURF has some fundamental limitation in terms of on-resistance and breakdown voltage. In the ideal case the ratio between the breakdown voltage and the on-resistance is constant, so increasing the breakdown voltage will also increase the on-resistance. There have been many improvements on the RESURF concept which decreases the on-resistance at preserved breakdown voltage. This thesis discusses some of them and in the papers two new improvements are proposed.

For the high-frequency characterization the ability to make correct measurements are important. The most important tool for high-frequency characterization is the network analyzer and it is important to calibrate the instrument rigorously at every measurement occasion. Since the VNA measures at high frequencies every piece of cable or connector will be measured as

well as the actual DUT. The transistor itself also contains part that are of less importance such as pad parasitics. These should be de-embedded from the measured data to show the real transistor characteristics. A method for on-wafer de-embedding of the pad parasitics is presented. When using small-signal equivalent circuits it is important to accurately extract the circuit parameters in a correct and consistent way. The direct extraction methods are very important and a new method for extracting the series impedances is presented that has the attractive property of not needing any approximations or assumptions that are common for other methods.

The LDMOS transistors that are manufactured and analyzed during this project have shown performance comparable and sometimes even better than other published power LDMOS transistors.

Based on the results in this thesis one can foresee substantial improvements for the dual-conducting layer LDMOS transistor. Both the breakdown voltage and the current drive capability can be increased, which will increase the RF output power density. The ability to operate at higher supply voltages may very well open up for new microwave applications in the GHz region.

A catchphrase in this project has been *"Exploring the limits for MOS based Microwave Power Transistors"* and the conclusion is that we haven't seen the limit yet.

Appendix A

Summary of Papers

Paper I

A Novel High-Frequency High-Voltage LDMOS Transistor using an Extended Gate RESURF Technology

A novel high-voltage LDMOS transistor using an extended gate is presented. The gate is extended all the way from the channel region to the drain contact were it is connected to the drain. The doping profile in the gate is highly doped n-type over the channel and the extended part is a reversed n^+p-diode which allows a small leakage current from the gate to the drain. The function of the extended drain is first that the small leakage current uniformly distributes the voltage over the drift region and thus a higher breakdown voltage. The second function is that the charge in the gate helps to deplete the drift region and thus can the n-well charge be increased which leads to a lower on-resistance.

The device has a breakdown voltage of 240 V compared to 60 V for the same device without the extended gate. The specific on-resistance is 24 mΩcm^2 and due to the short channel the high-frequency performance is good, f_T=2.8 GHz and f_{MAX}=5.8 GHz.

Paper II

A CMOS Compatible Power MOSFET for Low Voltage GHz Operation

A low-voltage microwave LDMOS transistor manufactured in a standard CMOS process is presented. The influence of the gate length for a fixed channel length has been studied. The study shows that the on-resistance and breakdown voltage is linearly dependent on the gate length. The transconductance is independent of the gate length but the increase of the input ca-

pacitance degrades the cut-off frequencies for longer gates.

The devices have breakdown voltages ranging from 15 to 30 V and for the shortest gate length, f_T=6.6 GHz and f_{MAX}=14 GHz. The PAE is 45% and the output power is 25 dBm for a 2 mm device at 1 GHz.

Paper III

Drift Region Optimization of Lateral RESURF Devices

High-voltage lateral devices commonly uses the RESURF concept for low on-resistance and high breakdown voltage. There are limitations in the RESURF concept and several advanced versions of RESURF have been presented. This paper presents a method for improving the on-resistance while maintaining the breakdown voltage. An extensive study on how to optimize the drift region implantations for optimal performance is done. The results shows that the on-resistance can be reduced a factor 2 by utilizing the proposed method with a dual conducting layer RESURF technique.

Paper IV

Analysis and Design of a Low-Voltage High-Frequency LDMOS Transistor

The output conductance of a transistor has large influence on the f_{MAX}. In this paper, a low-voltage LDMOS transistor has been analyzed in order to improve the output conductance and thus improve f_{MAX}. The analysis includes DC-measurements, impedance measurements, S-parameter measurements and a small-signal analysis is performed.

The analysis resulted in a design change where the p-base/n-well overlap was decreased and thus a lower drain-source capacitance. The measured result was an increased f_{MAX} with 30% and only a slight decrease in f_T.

Paper V

High Voltage LDMOS Transistors for Increasing RF Power Density and Gain

In this paper the dual conducting microwave power LDMOS transistor is presented. Two types of devices are evaluated, with the n-top patterned as the p-top and with the n-top implanted blanket. The focus is on the power performance measured with load-pull measurements. The devices have breakdown voltages around 100 V and 60 V respectively and the power measurements are measured for supply voltages up to 70 V for the device with patterned n-top.

The small-signal f_T and f_{MAX} is 4.5 GHz and 9 GHz. For the 100 V device the power density is above 2 W/mm at V_{DS}=70 V and 1 GHz which is the best ever published result for a power MOSFET (to our knowledge). At

3.2 GHz the power density is above 1 W/mm (V_{DS}=50 V) and 0.6 W/mm at V_{DS}=28 V.

Paper VI

A General Small-Signal Series Impedance Extraction Technique

A new method for extracting the series resistances and inductances is presented. The method is based on the least square method and thus not requires initial guessed values. The method does not include any assumptions or approximations and it should be specified for each specific equivalent circuit used. The method can also be used to extract the parameters of intrinsic circuits. It may also be used for lumped shunt admittances.

Paper VII

An On-Wafer De-embedding Method for Silicon Transistors at Microwave Frequencies

A method for on-wafer de-embedding of the pad parasitics for silicon-based transistors is presented. The method relies on the usual calibration methods that normally are built-in in the network analyzer. For the on-wafer calibration special calibration structures are manufactured in the same process as the transistors. There are four structures, a short, an open, a load and a thru line. The structures makes it possible to perform a SOLT or a TRL/LRM calibration directly on the wafer and thus no other calibrations are made. Another approach is to calibrate using the normal ISS and then do the parasitic de-embedding afterward which may be useful.

Appendix B

Acronyms

A summary of the acronyms used in the text.

BFOM - Baliga's figure of merit [5]

BiCMOS - bipolar complementary metal-oxide-semiconductor

BSIM - Berkeley short channel insulated gate field effect transistor model [51]

CMOS - complementary metal-oxide-semiconductor

DC - direct current

AC - altering current

DMOS - double-diffused metal-oxide-semiconductor

DUT - device under test

FEM - finite element method

FET - field effect transistor

GSG - ground-signal-ground

GTO - gate turn-off thyristor

HBT - hetero-junction bipolar transistor

IGBT - insulated gate bipolar transistor

IGFET - insulated gate field effect transistor

ISS - impedance standard substrate

LDD - lightly doped drain

LDMOS - lateral double-diffused metal-oxide-semiconductor

LED - light-emitting diode

LIGBT - lateral insulated gate bipolar transistor

LRM - line-reflect-match, calibration method

MESFET - metal-semiconductor field effect transistor

MOS - metal-oxide-semiconductor

MOSFET - metal-oxide-semiconductor field effect transistor

PAE - power added efficiency

RESURF - reduced surface field

RF - radio frequency

SIPOS - semi-insulating polycrystalline silicon

SOI - silicon-on-insulator

SOLT - short-open-load-thru, calibration method

TCAD - technology computer aided design

TRL - thru-reflect-line, calibration method

VDMOS - vertical double-diffused metal-oxide-semiconductor

VNA - vector network analyzer

GaAs - gallium arsenide

GaN - gallium nitride

Si - silicon

SiC - silicon carbide

SiGe - silicon germanium

SiO_2 - silicon oxide

Appendix C

Parallel-plane breakdown

To calculate the avalanche breakdown voltage the *impact ionization coefficient for electrons*, α_n, and *holes*, α_p, are important. The coefficients defines the number of electron-hole pairs created by a electron or hole traveling one unit length in the depletion region. The ionization coefficient has been measured and is given by the empirical expression

$$\alpha = ae^{-b/\mathcal{E}} \tag{C.1}$$

with different a and b for electrons and holes [52]. An useful approximation for deriving analytical solutions for the breakdown voltage is [53]

$$\alpha = 1.8 \cdot 10^{-35}\mathcal{E}^7 \tag{C.2}$$

To calculate the avalanche breakdown voltage the condition in which the impact ionization achieves an infinite rate has to be determined. Using the definition of the ionization coefficient, $\alpha_n dx$ is the number of electron-hole pairs created by a electron traveling the distance dx. Then can the total number of generated electron-hole pairs created by a single electron-hole pair generated at distance x by calculated by

$$M(x) = 1 + \int_0^x \alpha_n M(x)dx + \int_x^W \alpha_p M(x)dx \tag{C.3}$$

By solving (C.3) for the condition that avalanche breakdown occurs when $M(x)$ goes to infinity the solution yields that the *ionization integral* should be unity (C.4), [5] .

$$\int_0^W \alpha dx = 1 \tag{C.4}$$

The electric field distribution in the depletion region can be expressed as (C.5) and integrating the electric files yields the potential distribution (C.6).

$$\mathcal{E}(x) = \frac{qN_A}{\varepsilon_s}(W - x) \tag{C.5}$$

$$V(x) = \frac{qN_A}{\varepsilon_s}\left(Wx - \frac{x^2}{2}\right) \tag{C.6}$$

Inserting (C.5) in (C.4) and solving for W yields:

$$\int_0^W 1.8 \cdot 10^{-35} \left(\frac{qN_A}{\varepsilon_s}(W - x)\right)^7 dx = 1 \tag{C.7}$$

$$W_{c,PP} = 28600 \left(\frac{\varepsilon_s}{qN_A}\right)^{7/8} = 2.67 \cdot 10^{10} N_A^{-7/8} \tag{C.8}$$

The avalanche breakdown voltage, $V_{br,PP}$, is then calculated by inserting (C.8) in (C.6) for $W = W_{c,PP}$.

$$V_{br,PP} = 5.34 \cdot 10^{13} N_A^{-3/4} \tag{C.9}$$

The maximum electric field $\mathcal{E}_c$ is achieved by inserting (C.9) in (C.5) at $W = 0$

$$\mathcal{E}_c = 4010\ N_A^{1/8} \tag{C.10}$$

The above calculation were done for an abrupt n^+p junction were the depletion region can be assumed to only extend in the low doped p-side of the junction. This makes the field distribution as a right angled triangle. If the p and n doping levels are comparable the depletion region will extend on both sides of the junction. Then N_A in the above calculation has to be replace by an "effective doping", N_E

$$N_E = \frac{N_A N_D}{N_A + N_D} \tag{C.11}$$

which yields

$$W_{c,PP} = 2.67 \cdot 10^{10}\ N_E^{-7/8} \tag{C.12}$$

$$V_{br,PP} = 5.34 \cdot 10^{13}\ N_E^{-3/4} \tag{C.13}$$

$$\mathcal{E}_{crit} = 4010\ N_E^{1/8} \tag{C.14}$$

The above equations holds as long as the avalanche breakdown occurs before the depletion regions reaches the highly doped contacts. In those cases

the length of the drift region must be taken into account when solving (C.7). This exercise is left to the reader.

Appendix D

Least square fitting

complement to paper VI

The derivation of the matrix equation in [paper VI] is not explained in the paper but will be explained here.

In this example only two z-parameters will be used but when the concept is understood it will be fairly easy to increase the number of equations and coefficients. Assume that the following equations represent your measured data,

$$z_1 = \frac{a_0^1 + a_1^1 s + a_2^1 s^2}{1 - b_1 s - b_2 s^2} \tag{D.1}$$

$$z_2 = \frac{a_0^2 + a_1^2 s}{1 - b_1 s - b_2 s^2} \tag{D.2}$$

where s is the complex frequency. The order of the highest order terms in the numerators and denominators are chosen arbitrary. Then the following error function is applied

$$\begin{aligned}\eta = \sum_{i=1}^{N} \Big[&\left\{ z_{1i}(1 - b_1 s_i - b_2 s_i^2) - (a_0^1 + a_1^1 s_i + a_2^1 s_i^2) \right\}^2 \\ &+ \left\{ z_{2i}(1 - b_1 s_i - b_2 s_i^2) - (a_0^2 + a_1^2 s_i) \right\}^2 \Big]\end{aligned} \tag{D.3}$$

where z_{1i} and z_{2i} are the measured data at the frequencies s_i. The total number of measurement points is N. The error function, η, is differentiated with respect to a and b which gives e.g.

$$\begin{aligned}\frac{\partial \eta}{\partial a_0^1} &= \sum_{i=1}^{N}\left[-2\left\{z_{1i}(1-b_1 s_i - b_2 s_i^2) - (a_0^1 + a_1^1 s_i + a_2^1 s_i^2)\right\}\right] \\ &= -2\sum_{i=1}^{N}\left\{z_{1i} - b_1 s_i z_{1i} - b_2 s_i^2 z_{1i} - a_0^1 - a_1^1 s_i - a_2^1 s_i^2\right\} \qquad \text{(D.4)} \\ &= -2(\Sigma z_{1i} - b_1 \Sigma s_i z_{1i} - b_2 \Sigma s_i^2 z_{1i} - a_0^1 N - a_1^1 \Sigma s_i - a_2^1 \Sigma s_i^2)\end{aligned}$$

When all derivations are made they should be set to 0 to minimize the error. The criteria for a minimum is that the derivate is zero. So for the example above the following relation holds

$$a_0^1 N + a_1^1 \Sigma s_i + a_2^1 \Sigma s_i^2 + a_0^2 \cdot 0 + a_1^2 \cdot 0 + b_1 \Sigma s_i z_{1i} + b_2 \Sigma s_i^2 z_{1i} = \Sigma z_{1i} \qquad \text{(D.5)}$$

On the left hand side of (D.5) there are 7 terms one for each coefficient. When differentiating η for each coefficient gives us 7 equations similar to (D.5). Arranging the equations as above makes it possible to represent the equation system on matrix form:

$$X\mathbf{a} = \mathbf{y} \qquad \text{(D.6)}$$

where

$$\begin{bmatrix} N & \Sigma s_i & \Sigma s_i^2 & 0 & 0 & \Sigma s_i z_{1i} & \Sigma s_i^2 z_{1i} \\ \Sigma s_i & \Sigma s_i^2 & \Sigma s_i^3 & 0 & 0 & \Sigma s_i^2 z_{1i} & \Sigma s_i^3 z_{1i} \\ \Sigma s_i^2 & \Sigma s_i^3 & \Sigma s_i^4 & 0 & 0 & \Sigma s_i^3 z_{1i} & \Sigma s_i^4 z_{1i} \\ 0 & 0 & 0 & N & \Sigma s_i & \Sigma s_i z_{1i} & \Sigma s_i^2 z_{1i} \\ 0 & 0 & 0 & \Sigma s_i & \Sigma s_i^2 & \Sigma s_i^2 z_{1i} & \Sigma s_i^3 z_{1i} \\ \Sigma s_i z_{1i} & \Sigma s_i^2 z_{1i} & \Sigma s_i^3 z_{1i} & \Sigma s_i z_{1i} & \Sigma s_i^2 z_{1i} & \Sigma s_i^2(z_{1i}+z_{2i}) & \Sigma s_i^3(z_{1i}+z_{2i}) \\ \Sigma s_i^2 z_{1i} & \Sigma s_i^3 z_{1i} & \Sigma s_i^4 z_{1i} & \Sigma s_i^2 z_{1i} & \Sigma s_i^3 z_{1i} & \Sigma s_i^3(z_{1i}+z_{2i}) & \Sigma s_i^4(z_{1i}+z_{2i}) \end{bmatrix} = X \qquad \text{(D.7)}$$

$$\mathbf{a} = \begin{bmatrix} a_0^1 \\ a_1^1 \\ a_2^1 \\ a_0^2 \\ a_1^2 \\ b_1 \\ b_2 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} \Sigma z_{1i} \\ \Sigma s_i z_{1i} \\ \Sigma s_i^2 z_{1i} \\ \Sigma z_{2i} \\ \Sigma s_i z_{2i} \\ \Sigma s_i (z_{1i}^2 + z_{2i}^2) \\ \Sigma s_i^2 (z_{1i}^2 + z_{2i}^2) \end{bmatrix} \qquad \text{(D.8)}$$

The matrix equation is then solved by

$$\mathbf{a} = X^{-1}\mathbf{y} \qquad \text{(D.9)}$$

Appendix E

Impedance equivalents

When measuring an impedance the result is basically a magnitude, $|Z|$, and a phase, ϕ, at a given radial frequency $\omega = 2\pi f$. The mathematical representation of Z is shown in (E.1). Rewriting the expression for Z in its real (resistance) and imaginary (reactance) part yields (E.2).

$$Z = |Z|e^{j\phi} \tag{E.1}$$

$$= |Z|\cos\phi + j|Z|\sin\phi = R + jX \tag{E.2}$$

The impedance is often represented in admittance. The admittance, Y, is simply the inverse of the impedance ($Y = 1/Z$). The admittance can be separated in its conductance, G, and susceptance, B, $Y = G + jB$. When measuring the impedance the results can be given in a numerous combinations of resistances, inductances and capacitances. Figure E.1 shows some representations of the impedance.

A common representation is the capacitance and the conductance. This should be interpreted as that the impedance is a capacitance in parallel with a resistance, figure E.1a, were $R_P = 1/G$. The total impedance Z^a is given by:

$$Z^a = \frac{R_P}{1 + j\omega C_P R_P} = \frac{R_P}{1 + \omega^2 C_P^2 R_P^2} - j\frac{\omega C_P R_P^2}{1 + \omega^2 C_P^2 R_P^2} \tag{E.3}$$

Similar equation for the representations in fig. E.1b,c can be derived.

$$Z^b = R_S + \frac{1}{j\omega C_S} \tag{E.4}$$

$$Z^c = \frac{j\omega L_P}{R_P + j\omega L_P} \tag{E.5}$$

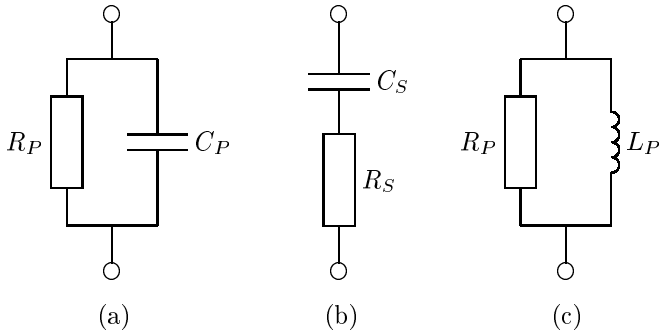


Figure E.1: Some impedance representations.

Setting $Z^a = Z^b$ and solving for R_S, C_S, R_P and C_S gives expressions to convert a the RC-parallel representation to the series representation and vice versa:

$$R_S = \frac{R_P}{1+\omega^2 C_P^2 R_P^2} \quad \text{(E.6)} \qquad R_P = R_S + \frac{1}{\omega^2 C_S^2 R_S} \quad \text{(E.8)}$$

$$C_S = C_P + \frac{1}{\omega^2 C_P R_P^2} \quad \text{(E.7)} \qquad C_P = \frac{C_S}{1+\omega^2 C_S^2 R_S^2} \quad \text{(E.9)}$$

Similar expression for the RL series or parallel representation yields:

$$R_S = \frac{\omega^2 L_P^2 R_P}{R_P^2+\omega^2 L_P^2} \quad \text{(E.10)} \qquad R_P = R_S + \frac{\omega^2 L_S^2}{R_S} \quad \text{(E.12)}$$

$$L_S = \frac{L_P R_P^2}{R_P^2+\omega^2 L_P^2} \quad \text{(E.11)} \qquad L_P = L_S + \frac{R_S^2}{\omega^2 L_S} \quad \text{(E.13)}$$

The correlation between L and C is simply $C = -1/\omega^2 L$. Other useful representations are the reflection coefficient, Γ, which is a scattering parameter, and the Q-factor.

$$\Gamma = \frac{Z - Z_0}{Z + Z_0} \quad \text{(E.14)} \qquad Q = \frac{|X|}{R} \quad \text{(E.15)}$$

References

[1] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, 1965.

[2] A. S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, 1967.

[3] S. M. Sze, *Physics of Semiconductor Devices*, Wiley, 2nd edition, 1981.

[4] Streetman B. G., *Solid State Electronic Devices*, Prentice-Hall, 5th edition, 1999.

[5] B. J. Baliga, *Power Semiconductor Devices*, PWS, 1995.

[6] SIA, "International Technology Roadmap for Semiconductors, 1999 edition," `http://www.itrs.org/`.

[7] R. Borges, "Gallium nitride electronic devices for high-power wireless applications," *RF Design*, pp. 72–82, September, 2001.

[8] K. Shenai, R. S. Scott, and B. J. Baliga, "Optimum Semiconductors for High-Power Electronics," *IEEE Trans. Electron Devices*, vol. 36, no. 9, pp. 1811–1823, 1989.

[9] M. N. Yoder, "Wide Bandgap Semiconductor Materials and Devices," *IEEE Trans. Electron Devices*, vol. 43, no. 10, pp. 1633–1636, 1996.

[10] B. Edholm, L. Vestling, M. Bergh, S. Tiensuu, and A. Söderbärg, "Silicon-On-Diamond MOS-Transistors with Thermally Grown Gate Oxide," *Proc. IEEE SOI Conf.*, pp. 30–31, 1997.

[11] M. Rosling, *Investigations of Interior Electrical Properties in Power Semiconductor Devices*, Ph.D. thesis, Uppsala University, 1992.

[12] E. O. Johnson, "The Insulated-Gate Field-Effect Transistor - A Bipolar Transistor in Disguise," *RCA Review*, vol. 34, pp. 80–94, 1973.

[13] J. A. Appels and H. M. J. Vaes, "High Voltage Thin Layer Devices (RESURF Devices)," *IEDM Tech. Dig.*, pp. 238–241, 1979.

[14] P Ashburn, *Design ans Realization of Bipolar Transistors*, Wiley, 1988.

[15] A. Schüppen, M. Tortschanoff, J. Berntgen, P. Maier, D. Zerrweck, H. von der Ropp, J. Tolonics, and K. Burger, "The Proliferation of Silicon Germanium," *Proc. ESSDERC*, pp. 88–91, 2000.

[16] R. Chau, et al., "30nm Physical Gate Length CMOS Transistors with 1.0 ps n-MOS and 1.7 ps p-MOS Gate Delays," *IEDM Tech. Dig.*, pp. 45–48, 2000.

[17] S.-H. Oh, et al., "50 nm Vertical Replacement-Gate (VRG) pMOSFETs," *IEDM Tech. Dig.*, pp. 65–68, 2000.

[18] Y. Momiyama, T. Hirose, H. Kurata, K. Goto, Y. Watanabe, and T. Sugii, "A 140 GHz ft and 60 GHz fmax DTMOS Integrated with High-Performance SOI Logic Technology," *IEDM Tech. Dig.*, pp. 451–454, 2000.

[19] H. J. Sigg, G. D. Vendelin, T. P. Cauge, and J. Kocsis, "DMOS Transistor for Microwave Applications," *IEEE Trans. Electron Devices*, vol. ED-19, pp. 45–53, 1972.

[20] H. F. F. Jos, "Novel LDMOS Structure for 2 GHz High Power Basestation Application," *Proc. 28th EuMC*, pp. 739–744, 1998.

[21] A. Wood, C. Dragon, and W. Burger, "High Performance Silicon LDMOS Technology for 2GHz RF Power Amplifier Applications," *IEDM Tech. Dig.*, pp. 87–90, 1996.

[22] J. Johansson, N. af Ekenstam, and Q. Chen, "Silicon LDMOS Technology for RF Power Amplifier Applications," *Proc. GigaHertz97*, pp. 20–21, 1997.

[23] R. Valtonen, J. Olsson, and P. De Wolf, "Channel length extraction for DMOS transistors using capacitance-voltage measurements," *IEEE Trans. Electron Devices*, vol. 48, no. 7, pp. 1454–1459, 2001.

[24] Z. Parpia and C. A. T. Salama, "Optimization of RESURF LDMOS Transistors: An Analytical Approach," *IEEE Trans. Electron Devices*, vol. 37, no. 3, pp. 789–796, 1990.

[25] A. W. Ludikhuize, "A Review of RESURF Technology," *Proc. ISPSD*, pp. 11–18, 2000.

[26] A. W. Ludikhuize, "Performance and Innovative Trends in RESURF Technology," *Proc. ESSDERC*, pp. 35–42, 2001.

[27] H. M. J. Vaes and J. A. Appels, "High Voltage, High Current Lateral Devices," *IEDM Tech. Dig.*, pp. 87–90, 1980.

[28] G. Charitat, D. Jaume, A. Peyre-Lavigne, and P. Rossel, "1000 and 1500 Volts Planar Devices using Field Plate and Semi-Resistive Layers: Design and Fabrication," *IEDM Tech. Dig.*, pp. 803–806, 1990.

[29] A. Söderbärg, B. Edholm, J. Olsson, F. Masszi, and K.-H. Eklund, "Integration of a Novel High-Voltage Giga-Hertz DMOS Transistor into a Standard CMOS Process," *IEDM Tech. Dig.*, pp. 975–978, 1995.

[30] D. R. Disney, A. K. Paul, M. Darwish, R. Basecki, and V. Rumennik, "A new 800V lateral MOSFET with dual conduction paths," *Proc. ISPSD*, pp. 399–402, 2001.

[31] G. Deboy, M. März, J.-P. Stengl, H. Strack, J. Tihanyi, and H. Weber, "A new generation of high voltage MOSFETs breaks the limit line of silicon," *IEDM Tech. Dig.*, pp. 683–385, 1998.

[32] G. Lorenz, G. Deboy, A Knapp, and M. März, "CoolMOS^{TM} - a new milestone in high voltage Power MOS," *Proc. ISPSD*, pp. 3–10, 1999.

[33] J. G. Mena and C. A. T. Salama, "High-Voltage Multiple-Resistivity Drift-Region LDMOS," *Solid-State Electronics*, vol. 29, no. 6, pp. 647–656, 1986.

[34] J. Cai, C. Ren, N. Balasubramanian, and J. K. O. Sin, "A Novel High Performance Stacked LDD RF LDMOSFET," *IEEE Electron Device Letters*, vol. 22, no. 5, pp. 236–238, 2001.

[35] J.-J. Bouny, "Advantages of LDMOS in high power linear amplification," *Microwave Engineering Europe*, pp. 37–40, April 1996.

[36] H. Zirath and D. B. Ruthledge, "An LDMOS VHF Class-E Power Amplifier Using a High-Q Novel Variable Inductor," *IEEE Trans. Microwave Theory and Techniques*, vol. 47, no. 12, pp. 2534–2538, 1999.

[37] Cascade Microtech, `http://www.cmicro.com/`.

[38] GGB Industries, `http://www.ggb.com/`.

[39] H. Cho and D. E. Burk, "A Three-Step Method for De-Embedding of High-Frequency S-parameter Measurements," *IEEE Trans. Electron Devices*, vol. 38, no. 6, pp. 1371–1375, 1991.

[40] J.-P. Raskin, R. Gillon, J. Chen, D. Vanhoenacker-Janvier, and J.-P. Colinge, "Accurate SOI MOSFET Characterization at Microwave Frequencies for Device Performance Optimization and Analog Modeling," *IEEE Trans. Electron Devices*, vol. 45, no. 5, pp. 1017–1025, 1998.

[41] T. E. Kolding, "A Four-Step Method for De-Embedding Gigahertz On-Wafer CMOS Measurements," *IEEE Trans. Electron Devices*, vol. 47, no. 4, pp. 734–740, 2000.

[42] Y. Tsividis, *Operation and Modeling of The MOS Transistor*, McGraw-Hill, 2nd edition, 1999.

[43] G. Dambrine, A Cappy, F. Heliodore, and E. Playez, "A New Method for Determining the FET Small-Signal Equivalent Circuit," *IEEE Trans. Microwave Theory and Techniques*, vol. 36, no. 7, pp. 1151–1159, 1988.

[44] D. Lovelace, J. Costa, and N. Camilleri, "Extracting Small-Signal Model Parameters of Silicon MOSFET Transistors," *Proc. IEEE MTT-S*, pp. 865–868, 1994.

[45] R. Anholt and S. Swirhun, "Equivalent-Circuit Parameter Extraction for Cold GaAs MESFET's," *IEDM Tech. Dig.*, pp. 1243–1247, 1991.

[46] J. C. Costa, M. Miller, M. Golio, and G. Norris, "Fast, Accurate, On-Wafer Extraction of Parasitic Resistances and Inductances in GaAs MESFETs and HEMTs," *IEDM Tech. Dig.*, pp. 1011–1014, 1992.

[47] S. Lee, H. K. Yu, C. S. Kim, J. G. Koo, and K. S. Nam, "A Novel Approach to Extracting Small-Signal Model Parameters of Silicon MOSFET's," *IEEE Microwave and Guided Wave Letters*, vol. 7, no. 3, pp. 75–77, 1997.

[48] Silvaco International, `http://www.silvaco.com/`.

[49] Integrated Systems Engineering, `http://www.ise.com/`.

[50] Avant!, `http://www.avanticorp.com/`.

[51] BSIM at Berkeley, `http://www-device.eecs.berkeley.edu/~bsim3/`.

[52] R. van Overstraeten and H. de Man, "Measurement of the Ionization Rates in Diffused Silicon *p-n* Junctions," *Solid-State Electronics*, vol. 13, pp. 585–608, 1970.

[53] W. Fulop, "Calculation of Avalance Breakdown of Silicon *p-n* Junctions," *Solid-State Electronics*, vol. 10, pp. 39–43, 1967.