Simulations of III-V MOSFETs

My Hylén

Department of Electrical and Information Technology

Faculty of Engineering
Lund University

2012 10

Produced during the summer and fall of 2012
EIT, Lund University

Supervisor: Erik Lind
Examiner: Lars-Erik Wernersson
Presentation: 18/10 2012

Abstract

Simulations of a planar InGaAs MOSFET (Metal-Oxide-Semiconductor Field-Effect Transistor) are used to investigate high electric fields and the resulting impact ionization and band-to-band tunneling within the device. To increase the breakdown voltage of the device a minimization of the electric field is attempted by implementing different doping profiles in the source and drain regions of the device. By studying the electric fields and breakdown voltages preferable doping profiles are found and the impact of these on transconductance and output resistance are investigated.

To further improve the accuracy of the data a more detailed model is implemented and the results are compared to existing data for the real devices.

Acknowledgements

I would like to acknowledge and extend my gratitude to my supervisor Erik Lind not only for helping me set up the directions for this project but also for giving excellent input on my work in progress.

I would also like to thank Mikael Egard for lending me his data on the devices used in this project.

Finally I would like to acknowledge Mikael Hernrup for the help with editing this report and for his never ending support during this project.

Contents

List of Figures

LIST OF FIGURES

List of Tables

LIST OF TABLES

1

Introduction

During the last decades the field effect transistor technology has improved at a dramatic rate. Most industrially utilized transistor technology is based on silicon transistors and the technology is approaching its limits mostly due to problems with scaling. Today efforts are made to develop MOSFETs (Metal-Oxide-Semiconductor Field-Effect Transistors) based on III-V semiconductor materials. These devices are expected to take the technology further due to enhanced material properties such as higher electron mobility (1).

There is an interest in increasing the breakdown voltage of these devices in order to increase the output power. The breakdown voltage is the level of source-drain voltage that when applied results in a major increase in current that deviate strongly from the desired stable saturation current. For at typical MOSFET this extreme increase in source-drain current is due to avalanche and/or Zener breakdown of the drain diode and the effect effectively places a limit on the maximum operating voltage for a device. Therefore, efforts are made to minimize these effects.

Avalanche breakdown is caused by impact ionization, which happens when electrons that are accelerated by a high electric field gains enough energy to break an electron-hole pair upon collision with an atom in the lattice. Both the electron and the hole from the pair can further accelerate in the electric field and gain energy to result in another impact ionization event, creating another electron-hole pair. These numerous events of impact ionization is called avalanche multiplication and it is the main cause for avalanche breakdown (2).

The mechanism behind Zener breakdown is called Zener tunneling or band-to-band tunneling. If the electric field is very high the conduction band and the valence band are bent to the point where an electron can tunnel from the valence band to the conduction

band, creating an increase in both electron and hole current (2). This means that the electric field is high enough for it to break the covalent bond between an electron and an atom. The tunneling and impact ionization processes are shown in figure 1.1.

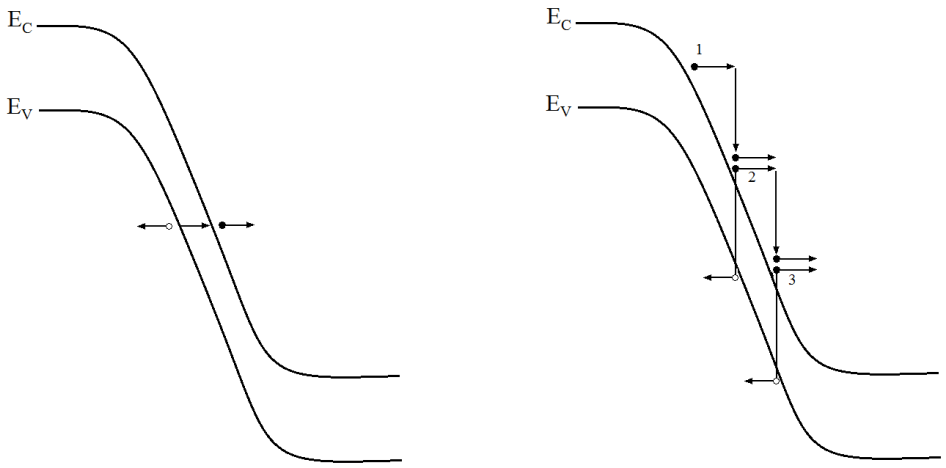


Figure 1.1: Band-to-band tunneling and impact ionization - Left: Energy band diagram for band-to-band tunneling. Tunneling creates an electron-hole pair that are separated by the electric field. Right: Energy band diagram for impact ionization creating the avalanche process. An electron gains energy and creates an electron-hole pair. This electron then gains energy to create another electron-hole pair.

Since avalanche breakdown and band-to-band tunneling is mainly affected by the electric field this is the main factor to consider when trying to increase the breakdown voltage. Several methods have been suggested to achieve this, among them are different buried layers in the MOS structure (3)(4) and modifications in the doping profile (5).

Another problem with these III-V MOSFETs is that new problems and effects appear when we are using new materials and pushing the limitations of scaling. Some of these effects are difficult to characterize and the parameters are hard to measure. To further investigate these devices simulations are made that try to incorporate the physical characteristics of the devices. This will allow us to study parameters that are not possible to measure in a real device, which will give insight in to how these devices operate and help optimize them. Since it is an massive task to incorporate all physical characteristics into a simulation decisions must be made as to what physical models to use and which simplifications can be made.

1.1 Aims and outline

The intent of this project is to simulate and characterize a planar III-V MOSFET based on InGaAs gate late process manufactured by the Nanoelectronics group at LTH (1).

The first part of the project is an investigation of possible methods to increase the breakdown voltage of a simple planar MOSFET model by studying the electric fields within the device. The device features heavily doped source and drain regions to ensure ohmic contacts, and a lightly doped channel. The design of the doping profile in the junction can have a large impact on the electric fields in these devices (2). The aim is to investigate which doping profiles are preferable in that they decrease the maximum electric field but also retain a acceptable level of other important parameters such as on-resistance and transconductance.

These simulations eliminate some of the the difficulties in defining and measuring the breakdown voltage in a real device and the possibility that the device is destroyed when biased in this region (6).

The second part of this project is to simulate the device with a more advanced model that takes non-local effects into account, enabling measurements that give more accurate results. Simulating this device gives a simple way of investigating many different possible structures, and thereby limiting the time-consuming lab work that would otherwise be required. Simulations of this type can present information about characteristics that cannot be measured experimentally and give insight in why some problems occur.

2

Models and methods of simulation

Atlas is a software that enables simulations of semiconductor devices, using an array of tools to insert parameters and present a wide range of data. In this project a platform is used to insert commands, defining the structure and composition of the device, as well as setting up the boundaries and physical models for the calculations. Section 2.1 presents the physics that the simulations are based on, section 2.2 reviews how physical parameters and specifications are set and section 2.3 defines some of the parameters that are used to characterize these devices.

2.1 Electron transport physics in Atlas

The physical models that are used in Atlas are based on at set of fundamental equations that are solved within the device parameters. These equations consist of Poissons Equation, continuity equations and transport equations, and they describe the electrostatic potential and how the electrons and hole densities change throughout the structure.

2.1.1 Poissons equation

Poissons equation relates the electrostatic potential to the space charge density, shown in equation 2.1 where Ψ is the electrostatic potential, ε is the local permittivity and ρ is the local space charge density. The electric field E is obtained from the gradient of the potential as in equation 2.2.

$$div(\varepsilon \bigtriangledown \Psi) = -\rho \tag{2.1}$$

$$E = - \bigtriangledown \Psi \tag{2.2}$$

2.1.2 Carrier continuity equations

The continuity equations defines that the amount of electrons or holes within one local region can only change by the amount that passes in or out or by generation and recombination. This is described by equation 2.3 for electrons and 2.4 for holes. J_n and J_p are electron and hole current densities, n and p are electron and hole concentration, G and R are generation and recombination rates for the carries, and q is the charge of an electron.

$$\frac{\delta n}{\delta t} = \frac{1}{q} div J_n + G_n + R_n \tag{2.3}$$

$$\frac{\delta p}{\delta t} = -\frac{1}{q} div J_p + G_p + R_p \tag{2.4}$$

2.1.3 Transport equations

Equation 2.1, 2.2, 2.3 and 2.4 set up the framework for device simulations in Atlas. To further define the currents J_n and J_p and the generation and recombination rates G and R other equations are required, based on what model is used for the particular device.

The simplest model of charge transport that is used in Atlas is the Drift-Diffusion Transport Model. This model is derived from the Boltzmann transport equation which is a simplification of the Fermi-Dirac statistics. Further information about Boltzmann statistics and how the approximations in this model are made can be found in the Atlas user manual (7). The drift-diffusion equations are defined in equation 2.5 and 2.6 where μ_n and μ_p are the electron and hole mobilities and D is the diffusion coefficient.

$$J_n = qn\mu_n E_n + qD_n \bigtriangledown n \tag{2.5}$$

$$J_p = qp\mu_p E_p - qD_p \bigtriangledown p \tag{2.6}$$

The approximations in this model are made with the assumption that the Einstein relationship holds, relating the diffusion coefficient D to the mobility. Using the Boltzmann approximation this can be expressed by equation 2.7. An analogous expression is used for holes.

$$D_n = \frac{kT_L}{q}\mu_n \tag{2.7}$$

This conventional drift-diffusion model is a local model which does not account for non-local effects such as velocity overshoot and the energy dependence of impact ionization rates, which means that it can over or underestimate the current levels . A more advanced model for current densities can be acquired if the current densities are not only related to the carrier concentrations and the electrostatic potential but also the carrier energy. This is called the Energy Balance Transport Model and it introduces T as a an independent variable for carrier temperature in a energy balance equation and associated equations to describe the flux S which is the flow of energy from the carriers to the lattice. The model consists of the following equations:

$$divS_n = \frac{1}{q}J_nE - W_n - \frac{3k}{2}\frac{\delta}{\delta t}(\lambda_n^* nT_n) \tag{2.8}$$

$$J_n = qD_n \bigtriangledown n - q\mu_n n \bigtriangledown \Psi + qnD_n^T \bigtriangledown T_n \tag{2.9}$$

$$S_n = -K_n \bigtriangledown T_n - \frac{k\delta}{q}J_nT_n \tag{2.10}$$

and for holes:

$$divS_p = \frac{1}{q}J_pE - W_p - \frac{3k}{2}\frac{\delta}{\delta t}(\lambda_p^* pT_p) \tag{2.11}$$

$$J_p = -qD_p \bigtriangledown p - q\mu_p p \bigtriangledown \Psi + qpD_p^T \bigtriangledown T_p \tag{2.12}$$

$$S_p = -K_p \bigtriangledown T_p - \frac{k\delta_p}{q}J_pT_p \tag{2.13}$$

K_n and K_p are thermal conductivities of electrons and holes which also have an energy dependance (7). The current density also depends on W which is the energy density loss rate. This parameter is important since it describes the energy exchange for the carriers with the surrounding structure. This includes carrier heating by increasing lattice temperature as well as recombination and generation. How the energy density loss rate depends on temperature can be altered in Atlas by changing the electron and hole energy relaxation time, which is the time constant for the energy exchange. This relationship is described in equation 2.14 for electrons and there is an analogous expression for holes. $TAUREL.EL$ is the energy relaxation time and this value can

be set in Atlas to modify how the energy of the lattice affects the carriers and vice versa. The corresponding parameter for holes is $TAUREL.HO$. These parameters can be derived for these devices through Monte Carlo simulations, but are in this project set to a reasonable standard value.

$$W_n = \frac{3}{2} n \frac{k(T_n - T_L)}{TAUREL.EL} \lambda_n + \frac{3}{2} k T_n \lambda R_{SRH} + E_g(G_n - R_n) \tag{2.14}$$

The energy balance transport model also relates the mobility to the carrier energy. This is done by calculating an effective electric field by solving equation 2.15 and 2.16 for $E_{eff,n}$ and $E_{eff,p}$.

$$q\mu_n(E_{eff,n})E_{eff,n}^2 = \frac{3}{2} \frac{k(T_n - T_L)}{TAUMOB.EL} \tag{2.15}$$

$$q\mu_p(E_{eff,p})E_{eff,p}^2 = \frac{3}{2} \frac{k(T_n - T_L)}{TAUMOB.HO} \tag{2.16}$$

The value of $TAUMOB.EL$ and $TAUMOB.HO$ can be set set to change how the effective electric field is affected by the carrier energy. These effective electric fields are used in some of the mobility models in Atlas which will be discussed in section 2.2.2. The value of $TAUREL.EL$, $TAUREL.HO$, $TAUMOB.EL$ and $TAUMOB.HO$ are shown in table 2.1.

Table 2.1: Electron transport parameters - The values of electron transport parameters for both holes and electrons.

Parameter	Value
$TAUREL.EL$	$0.5 \cdot 10^{-12}$ (s)
$TAUREL.HO$	$0.5 \cdot 10^{-12}$ (s)
$TAUMOB.EL$	$0.5 \cdot 10^{-12}$ (s)
$TAUMOB.HO$	$0.5 \cdot 10^{-12}$ (s)

2.2 Program specifications and models

2.2.1 Structure specification

The first parameters in the simulations specifies the mesh. The mesh is defined by a number of vertical and horizontal lines that create grid points where the solutions to the systems of nonlinear differential equations are calculated. The statements in the

code define what spacing is set between the lines and thereby the density of grid points at different locations in the structure.

A fine mesh with a large number of grid points is required to resolve all significant features of the solution but will also result in longer calculation times. Therefore, a coarse mesh that minimizes the number of grid points is desired. This is solved by using different dense grids at different parts of the structure. We set a coarse grid to start with to ensure that the simulation time is short but set a high amount of nodes locally in regions where we expect large variations in relevant parameters such as the potential, electron and hole concentration, and electric field.

Another part of the structure specification defines the electrodes. The electrodes are defined as a region of the structure that contains the grid points where the voltages will be set during simulations. In this device three electrodes are defined in the structure; the source and drain electrodes that are set to be ohmic and the gate electrode that is set with a workfunction.

The final part of the structure specification is the doping levels. There are several different ways to define doping in Atlas, depending on what doping profile is desired. The simplest way is to define one constant doping level for each region which results in abrupt transitions between regions. Another way is to use a linear doping between two values in a defined region. There is also the possibility to use a doping profiles in the form of the complementary error function or to use a Gaussian doping profile.

2.2.2 Material and model specification

In the material specification different material properties can be set, such as composition or mobility. Atlas provides default values for most materials, and in this project some of the material properties will modified to change the behavior the device.

The next step is to define what models are to be used in the simulations. The different models describe physical characteristics such as carrier statistics, mobility, recombination, impact ionization and tunneling. Each of these has a set of different models depending on the accuracy and variability that is required from the model. Describing all the models is beyond the scope of this report, and only the models used in this project will be stated and detailed. For a more detailed explanation of the different models the interested reader can study the Atlas user manual (7).

The first thing to consider is what type of carrier statistics to use. The two most common models are Fermi-Dirac statistics and Boltzmann statistics, where Boltzmann statistics is a simplification of the Fermi-Dirac statistics that is valid only when the energy difference between the fermilevel and the conduction band is much smaller than

kT. This method makes calculations much simpler. Fermi-Dirac statistics are necessary to accurately predict carrier concentrations in devices with heavily doped regions, and in this project Fermi-Dirac statistics will be used to account for the high doping levels in the source and drain regions.

Mobility modeling is divided into two parts: low field behavior and high field behavior. At low fields the velocity v of the carriers is linearly proportional to the electric field E, as seen in equation 2.17, and the mobility μ has a constant low field value that is dependent on phonon and impurity scattering.

$$v = \mu \cdot E \tag{2.17}$$

At higher fields the mobility starts to decline which causes the carrier velocity to saturate due to an increase in scattering against the lattice. Since the carrier velocity is a product of the mobility and the electric field in the direction of the current flow the electric field will at some point no longer increase the carrier velocity. This is called velocity saturation. In Atlas the constant low field mobility can be set in the materials specification and then a model is chosen to simulate the transition to the velocity saturation regime. In this project a simple parallel electric field dependence mobility model was used. This model handles this effect by calculating an effective field dependent mobility using equation 2.18 and 2.19 where μ_n and μ_p are the effective mobilities, E is the parallel electric field and μ_{n0} and μ_{p0} are the low field mobilities. The parameters $BETAN$ and $BETAP$ are set to 1 in this project.

$$\mu_n(E) = \mu_{n0} \left[\frac{1}{1 + (\frac{\mu_{n0} E}{VSATN})^{BETAN}} \right]^{\frac{1}{BETAN}} \tag{2.18}$$

$$\mu_p(E) = \mu_{p0} \left[\frac{1}{1 + (\frac{\mu_{p0} E}{VSATP})^{BETAP}} \right]^{\frac{1}{BETAP}} \tag{2.19}$$

When the energy balance equation is implemented an effective electric field is calculated as described in section 2.1.3. This is used instead of the parallel electric field described above.

In a semiconductor there will always be generation and recombination due to transitions in the bandgap, impact ionization and tunneling events. In this project the transition-process was modeled using the Shockley-Read-Hall (SRH) recombination where generation/recombination occur in the presence of a trap or defect in the bandgap. This theory was derived by Shockley, Read and Hall and is described further in the Atlas user manual (7).

The impact ionization process in Atlas is described by equation 2.20 where G is the local generation rtate of electron-hole pairs, $\alpha_{n,p}$ are ionization coefficients for electrons and holes and $J_{n,p}$ are the current densities.

$$G = \alpha_n |J|_n + \alpha_p |J|_p \tag{2.20}$$

Impact ionization models are only implemented in the first part of this project and a local electric field model called Selberherr's Impact Ionization Model is used. This model determines the ionization coefficients α_n and α_p using equation 2.21 and 2.22.

$$\alpha_n = ANexp\left[-\left(\frac{BN}{E}\right)^{BETAN}\right] \tag{2.21}$$

$$\alpha_p = APexp\left[-\left(\frac{BP}{E}\right)^{BETAP}\right] \tag{2.22}$$

Here, E is the electric field in the direction of current flow and AN, AP, BN, BP, $BETAN$ and $BETAP$ are user-definable parameters.

The band-to-band tunneling that is implemented in the first part of this project is using equation 2.23 to define the tunneling generation rate G_{BBT}.

$$G_{BBT} = D \cdot BB.A \cdot E^{BB.GAMMA} exp\left(-\frac{BB.B}{E}\right) \tag{2.23}$$

E is the magnitude of the electric field, D is a statistical factor and $BB.A$, $BB.B$ and $BB.GAMMA$ is used-definable parameters. In this implementation these parameters are set to standard values shown in table 2.2 and can also be found in the Atlas User Manual (7).

2.2.3 Numerical methods and solutions

To solve the differential equations in section 2.1 there are three different numerical methods that solve the equations in different manners. The one called Gummel solves the equations by keeping all variables but one constant, solving each unknown in turn and repeating the process until a stable solution is found, called a decoupled solution technique. Another technique, called Newton solves the equations fully coupled, which means that the complete system of unknowns are solved together. The block method will solve some equations fully coupled and some decoupled. Different solution techniques are used in different calculations and can affect the calculation speed depending on what model was chosen. Some models, like the ones implementing impact ionization,

Table 2.2: Impact ionization and band-to-band tunneling parameters - Impact ionization and band-to-band tunneling parameters used in equation 2.21, 2.22 and 2.23.

Parameter	Value
AN	$1 \cdot 10^2$ cm^{-1}
AP	$1 \cdot 10^4$ cm^{-1}
BN	$1 \cdot 10^6$ V/cm
BP	$1 \cdot 10^6$ V/cm
$BETAN$	1
$BETAP$	1
$BB.A$	$5 \cdot 10^{18}$ cm^{-1}V^{-2}s^{-1}
$BB.B$	$3 \cdot 10^7$ V/cm
$BB.GAMMA$	2

often will not converge when using the simple Gummel method and Newton or Block is required instead.

The simplest way of obtaining solutions in atlas is to do DC calculations. The voltage on each electrode is set and Atlas calculates currents and other characteristics in each grid point. The voltage is then ramped by changing the values on the electrodes and output and transfer characteristics can be obtained. The numerical methods use the previous solution as a base for guessing starting values on the next one. Therefore, if Atlas can not find a stable solution it is probably due to a large change in voltage in the electrodes, and this is solved by taking a smaller voltage step. This method is repeated until a stable solution is acquired.

2.3 Figures of merit

To characterize the devices simulated in this project a set of figures of merit are deduced. These are described in the following sections.

2.3.1 On-resistance, R_{on}

The on-resistance of the device is the resistance between source and drain when the device is in the linear region. This is calculated by taking the inverse of the slope of the output characteristics in the linear region. The value of the on-resistance presented in this thesis is the average of the on-resistance between a drain voltage of 0 to 0.1 V.

2.3.2 Transconductance, g_m

The transconductance of a field effect transistor is the ratio the drain current changes when the gate voltage is changed, using a constant drain voltage. This relationship is shown in equation 2.24.

$$g_m = \frac{\delta I_D}{\delta V_G} | V_D \tag{2.24}$$

In this thesis the values presented are calculated by taking the derivative of the transfer characterisitcs at $V_G = 1$ V and $V_D = 1$ V.

3

Results and discussion

Section 3.1 details the results from the measurements of the electric field and breakdown voltages. The maximum electric field is obtained for different types of doping and compared to on-resistance and transconductance. Linear doping is selected and it's effect on impact ionization and band-to-band tunneling is reviewed. Finally validation of the model is attempted by extracting the s-parameters and comparing to existing data and by examining the electron and hole concentrations throughout the structure.

Section 3.2 describes the result from the implementation of the energy balance model, which is an attempt to make a more accurate model of the device than the initial one used in this work. The simulated data is fitted to the experimental data using velocity saturation and mobility and the results shows a good fit.

3.1 Impact of doping on transistor breakdown

The structure of the modeled device is shown in figure 3.1. The device has a gate length of 200 nm and a 5 nm thick HfO_2 as oxide with a permittivity of 15. Below the 10 nm thick InGaAs channel is an InAlAs back barrier with a larger bandgap which traps the carriers in the channel, and thereafter is the bulk which consists of InP. No contact resistance is added to the model to make calculations easier. The gate contact has a workfunction of 4.5 eV. The doping level at the source and drain contact regions is $6 \cdot 10^{19}$ cm^{-3} and the doping level in the channel is $1 \cdot 10^{16}$ cm^{-3}, and it is this abrupt transition between doping levels that is to be altered.

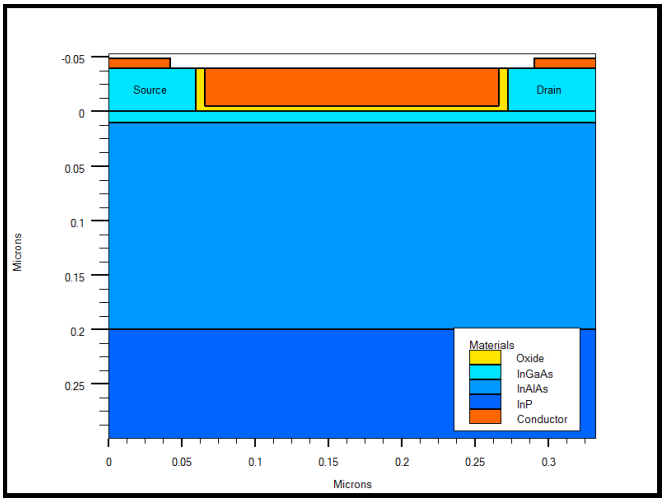


Figure 3.1: Structure of modeled MOSFET - The structure of the modeled MOSFET

3.1.1 Measurements of maximum electric field

To establish how the electric field behaves within the device a simulation was made, and the resulting electric field is shown in figure 3.2. The electric field is largest in the source and drain regions where the abrupt transition between doping levels is present, and this is where the measurements were made. The field was measured at a cutline placed in the drain region, also depicted in figure 3.2.

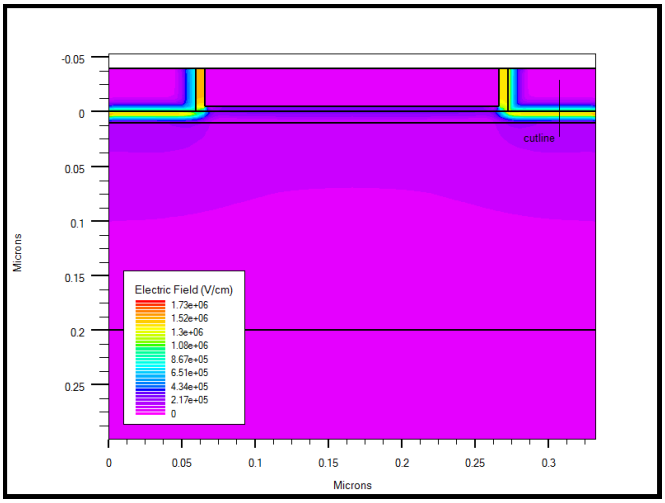
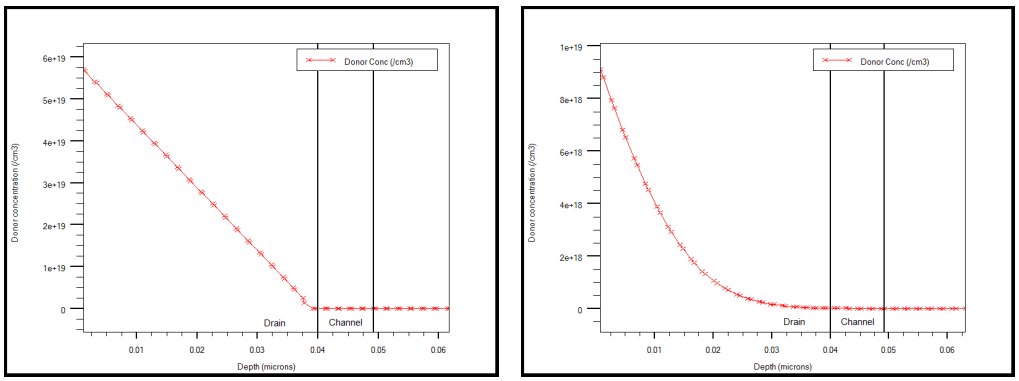


Figure 3.2: Electric field in MOSFET - Electric field in the structure displayed in figure 3.1 with V_G set to 0 V and V_D set to 1 V. A cutline used to measure the electric field in the drain region is also displayed.

From the beginning two different types of doping profiles where tested; a linear profile from the source and drain contact regions to the channel and a complementary error function profile with the top value at the source and drain contact and the bottom value in the channel region. These profiles are shown in figure 3.3.



(a) Linear doping profile

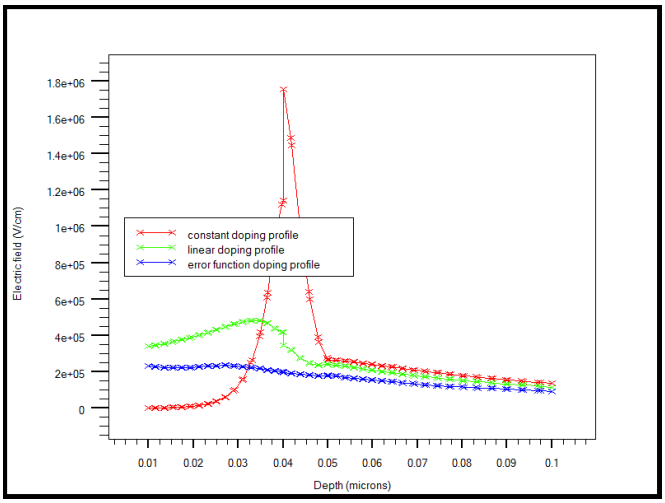
(b) Complementary error function doping profile

Figure 3.3: Doping profiles - Two different doping profiles used in simulations, shown on a linear scale.

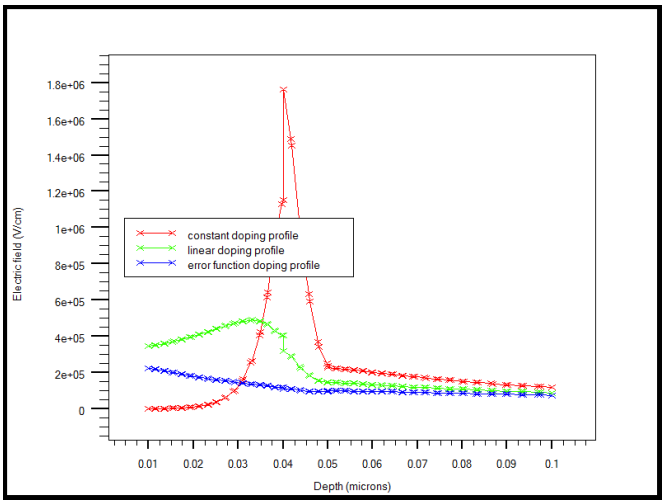
Figure 3.4 show the electric field as a function of depth in the structure along the cutline for the different doping profiles and for different gate voltages. These figures show that there is a benefit in changing the doping profiles in order to decrease the electric field and that the effect is mainly in the transition from the drain/source region and the channel. The two graded doping profiles are spreading the electric field into the source/drain region as opposed to the abrupt doping profile where the electric field peaks at the junction and thereafter declines to a small value. Depending on the level of the field this could have a substantial effect on the amount of impact ionization and band-to-band tunneling in these regions. If this decrease in electric field goes below the level required for impact ionization and tunneling to occur these events would happen at a higher drain voltage where the field is higher, which would in turn increase the breakdown voltage. If the electric field does not decrease enough for these events to stop it could still make the effects less frequent, but the spreading of the electric field could still cancel this effect by increasing the area in which these event can occur.

The two figures also show that the gate voltage does not affect the field in a significant way and the following measurements will be made at a gate voltage of 1 V. This is made with the assumption that the cutline placed in the drain region accurately describes the electric field in the whole drain region. Looking at figure 3.2 we can see that

the cutline is placed some distance from the active region which can affect the results. Placing the cutline closer to the gate might have given better results, which can also be seen later when the impact ionization rates are studied. To investigate the doping profiles further the maximum electric field E_{max} in the cutline, the output resistance R_{on} and the transconductance g_m was extracted from the model and are shown in table 3.1.



(a) $V_G = 0$ V



(b) $V_G = 1$ V

Figure 3.4: Electric fields in cutline - Electric fields for three different doping profiles and two different gate voltages.

The results from table 3.1 show an increase of the output resistance, which is prob-

Table 3.1: Parameters from simulations using different doping profiles - Maximum electric field E_{max} extracted in cutline at $V_G = 1$ V and $V_D = 1$ V, output resistance R_{on} extracted at $V_D = 0.1$ V and transconductance extracted at $V_G = 1$ V.

Parameter	Const. doping profile	Lin. doping profile	Erf. doping profile
E_{max} (V/cm)	$17.6 \cdot 10^5$	$4.90 \cdot 10^5$	$2.24 \cdot 10^5$
R_{on} (Ω)	0.27	0.46	0.91
g_m (S)	2.21	1.09	0.65

ably due to the reduction of carriers in the source and drain regions that the decreased doping leads to. This reduction of carriers will decrease the conductivity of the material and thereby increase the on-resistance which is not desirable, since a higher on-resistance will require a higher drain voltage to reach the saturation region.

The results in table 3.1 also shows a degradation in the transconductance, meaning that the device can not be controlled as easily. This means that a higher gate voltage will be required to get the same current levels as with the constant doping. Both doping profiles shows promising results but the error function doping profile shows a to large degradation in transconductance and output resistance in comparison to the improvement in electric field and it will not be further investigated. The linear doping profile shows almost as high improvement as the error function doping profile but with a smaller degradation of both output resistance and transconductance and will therefore be further characterized in the next section. The linear doping profile is also a reasonable choice since the manufacturing process is based on building the devices from the bottom up by epitaxy.

3.1.2 Effects of linear doping profiles on impact ionization and band-to-band tunneling

The device used in section 3.1.1 was supplemented with models for impact ionization and band to band tunneling. In the previous measurements the linearly graded doping profile was graded from a doping concentration of $6 \cdot 10^{19}$ cm^{-3} in the source and drain regions to $1 \cdot 10^{16}$ cm^{-3} at the edge of the channel region as seen in figure 3.3. A set of different linear doping profiles was used for these additional measurements, graded from $6 \cdot 10^{19}$ cm^{-3} in the source and drain regions to $1 \cdot 10^{16}$ cm^{-3}, $1 \cdot 10^{17}$ cm^{-3}, $1 \cdot 10^{18}$ cm^{-3} and $1 \cdot 10^{19}$ cm^{-3} at the edge of the channel region. In figure 3.6 the impact generation rate is shown, indicating that the impact ionization is occurring mainly in the drain region of the device. With this in mind another doping profile was created,

where only the drain region is graded and the source region is not altered. This could reduce the degradation of the output resistance and transconductance. The resulting output characteristics from the device with different alterations to the doping profile is shown in figure 3.5. The figure reveals that there are some problems with modeling the transfer from low field mobility to high field mobility.

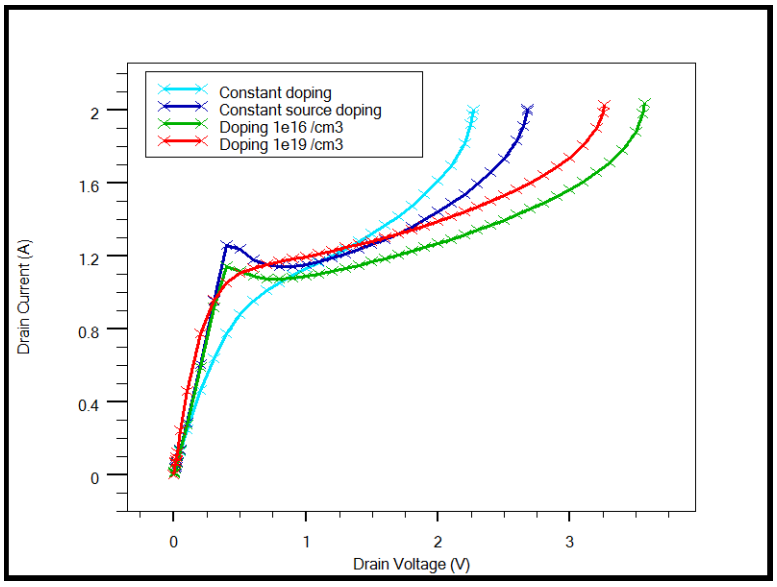


Figure 3.5: Output characteristics - Output characteristics for simulated device with models for impact ionization and band-to-band tunneling that causes breakdown to occur as seen by the large increase in current.

The maximum field for the different doping profiles was derived in the same way as in section 3.1.1. The breakdown voltage was derived by stopping the calculations at a given current level where the solutions did no longer diverge, indicating that breakdown was attained, and extracting the drain voltage at that point. The maximum electric field E_{max} and the breakdown voltage V_{BD} is shown in table 3.2

Table 3.2 show that there is only a small difference in maximum electric field in the drain region between the different cases of graded junctions. This is expected because on a linear scale they will be almost identical. This result is also mirrored in the breakdown voltage. The breakdown voltage in the transistor that had linear grading only in the drain region does not increase as much as in the other devices. This could be an effect of some impact ionization and band-to-band tunneling present in the source region. To further evaluate these results the on-resistance and transconductance was extracted and is shown in table 3.3.

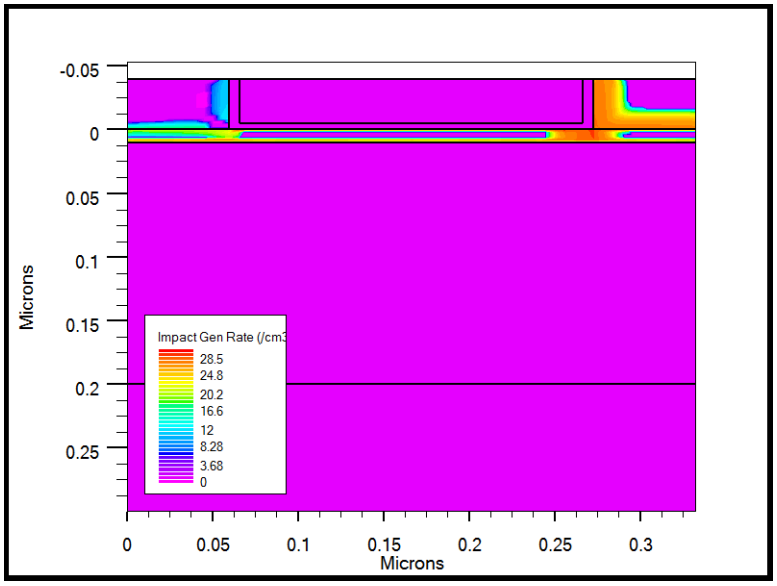


Figure 3.6: Impact generation rate - Impact generation rate in simulated device with $V_G = 0.6$ V and $V_D = 1$ V.

Table 3.2: Maximum electric field and breakdown voltage - Maximum electric field E_{max} extracted at $V_G = 0.6$ V and $V_D = 1$ V and breakdown voltage V_{BD} extracted at $V_G = 0.6$ V.

Doping profile	E_{max} (V/cm)	V_{BD} (V)
Constant doping	$17.5 \cdot 10^5$	2.27
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{16}$ cm^{-3}	$4.91 \cdot 10^5$	3.56
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{17}$ cm^{-3}	$4.90 \cdot 10^5$	3.55
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{18}$ cm^{-3}	$4.82 \cdot 10^5$	3.51
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{19}$ cm^{-3}	$5.81 \cdot 10^5$	3.25
Drain graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{16}$ cm^{-3}	$4.91 \cdot 10^5$	2.67

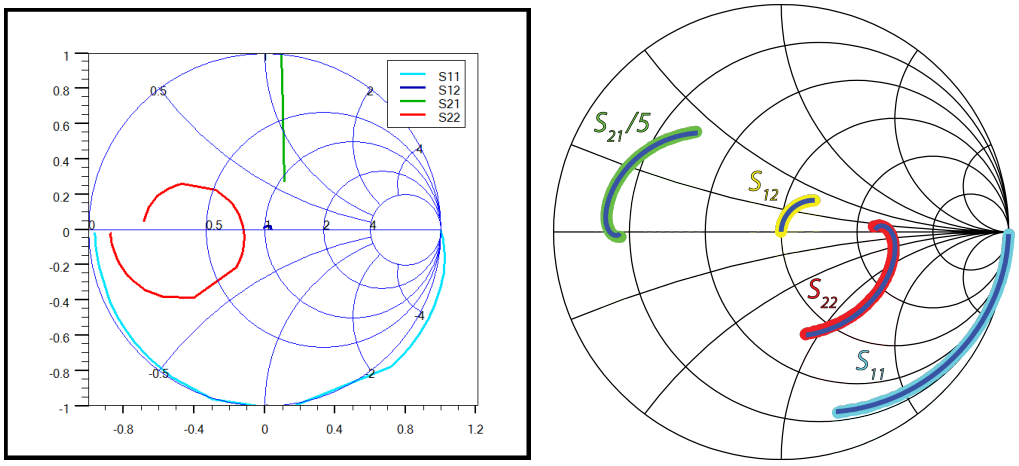
The results in table 3.3 show a degradation in both on-resistance and transconductance as expected. An interesting thing about this data is that it shows that grading only the drain region in the device makes sure that the on-resistance and transconductance does not degrade as much as in the other cases but still manages to increase the breakdown voltage. The increased transconductance in this case will also increase the current in the device and thereby increasing the impact ionization, leading to a lower breakdown voltage. The results are still promising with a 17% increase in breakdown voltage compared to a 11% decrease in transconductance for the case where only the

Table 3.3: On-resistance and transconductance - On-resistance R_{on} extracted at $V_D = 0.1$ V and transconductance g_m extracted at $V_G = 1.0$ V.

Doping profile	R_{on} (Ω)	g_m (S)
Constant doping	0.24	2.36
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{16}$ cm^{-3}	0.37	1.73
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{17}$ cm^{-3}	0.37	1.73
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{18}$ cm^{-3}	0.35	1.75
Graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{19}$ cm^{-3}	0.24	2.07
Drain graded $6 \cdot 10^{19}$ cm^{-3} to $1 \cdot 10^{16}$ cm^{-3}	0.26	2.11

drain region is graded.

To investigate how accurate this model is and make sure that these results are relevant the s-parameters for the device was extracted. These parameters plotted in a Smith chart is shown in figure 3.7 together with the s-parameters measured on a real device.



(a) Simulated S-parameters **(b)** Measured S-parameters

Figure 3.7: S-parameters - Smith-charts for both simulated and measured S-parameters from 360 MHz to 70 GHz on a device with 55 nm gate length.

These s-parameters show that the model has the same basic behavior as the real devices and that the impedance, such as parasitic capacitance and on-resistance, have the same pattern in the smith chart even though it does not have the same exact values. The only parameter that does not have the same behavior is S_{21} because the value of

the measured parameter has been decreased 5 times and it can not be viewed in the simulated chart.

In figure 3.8 the hole concentration for the simulated device is shown. There is a hole concentration below the gate which comes from the impact ionization and band-to-band tunneling events. This effect is present even is only one of the models (impact ionization and band-to-band tunneling) is applied which shows that the two models are actually working and that these effects is infact the reason for the increase in current leading to breakdown.

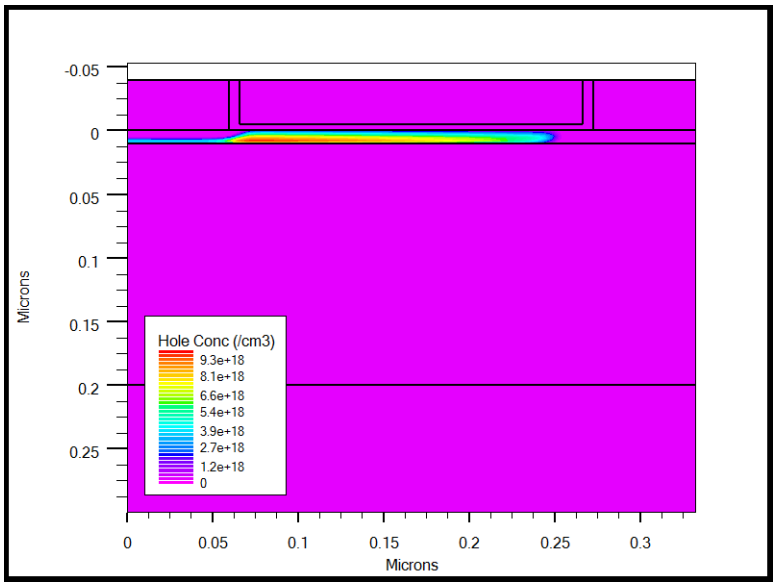


Figure 3.8: Hole concentration - Hole concentration in simulated device with $V_G = 0.6$ V and $V_D = 1$ V.

3.2 Implementation of energy balance model

The simulations in this section are made using the energy balance model. This model is implemented to make more accurate simulations of the devices, and as described in section 2.1.3 this model accounts for non-local effects that can have a large effect on the current levels. Comparing the output characteristics in this project the current levels increases with a factor of around two using a 140 nm gate length, which gives a of how important these models are.

The device in this section is the same as in section 3.1 but has gate lengths of 14, 55 and 140 nm and a 5 nm thick delta doping starting 1 nm below the channel, with

a doping level of $1 \cdot 10^{18}$ cm^{-3}. Contact resistance of 87.5 $\Omega\mu m$ is added to the model at each contact, which is a value that was measured on the real devices. No impact ionization or band-to-band tunneling was modeled in these simulations.

Two following equations is a simple description of the current in a MOSFET, where equation 3.1 describes the linear region and equation 3.2 describes the saturated region.

$$I_D = \frac{W}{L} C_{ox} \mu_n (V_{GS} - V_T) \cdot V_{DS} \tag{3.1}$$

$$I_D = W C_{ox} v_{sat} (V_{GS} - V_T) \tag{3.2}$$

These equations show that a fit to the experimental data can be made by fitting transfer characteristics in the saturated region by altering the velocity saturation and by altering the threshold voltage with the workfunction, and then fitting the output characteristics by altering the low field mobility.

3.2.1 Fitting to the experimental data - transfer characteristics

Figure 3.9, 3.10 and 3.11 show the fit off the transfer characteristics for the different gate lengths. Measurements were performed at a drain voltage of 1 V and gate voltages of -1 V through 1.8 V.

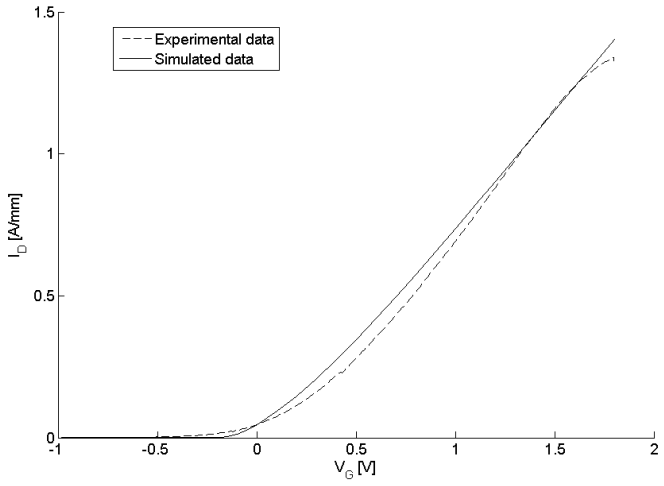


Figure 3.9: Transfer characteristics for 140 nm gate length. - Fitting transfer characteristics to experimental data for the device with 140 nm gate length.

For the 140 and 55 nm gate lengths a good fit is found with a saturation velocity of $8 \cdot 10^6$ cm/s and $6 \cdot 10^6$ cm/s respectively. These values are lower than previous results

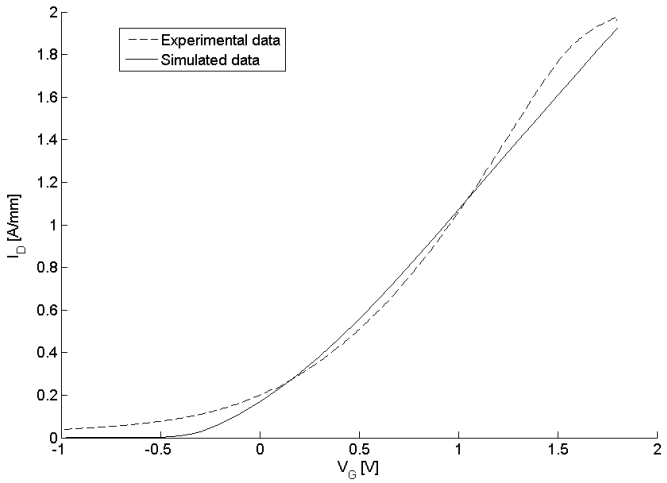


Figure 3.10: Transfer characteristics for 55 nm gate length. - Fitting transfer characteristics to experimental data for the device with 55 nm gate length.

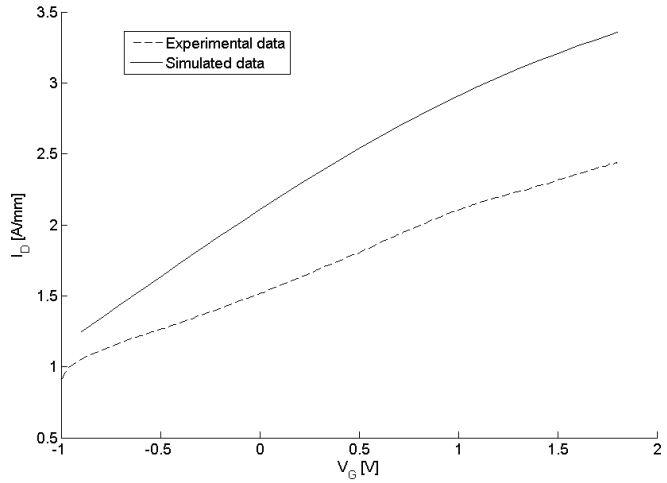


Figure 3.11: Transfer characteristics for 14 nm gate length. - Fitting transfer characteristics to experimental data for the device with 14 nm gate length.

on these devices (8), probably due to the increase in current levels due to non-local effects that were not regarded in that previous work. The fitting of the simulated data is not as good using 55 nm gate length as 140 nm gate length, and the fit of the 14 nm device is not satisfactory using a saturation velocity of $5 \cdot 10^6$ cm/s. This probably depends on a number of different effects, or rather the absence of these effects, since

no impact ionization, band-to-band tunneling or any form of defects were modeled.

3.2.2 Fitting to the experimental data - output characteristics

Figure 3.12, 3.13 and 3.14 show the fit off the output characteristics for the different gate lengths.

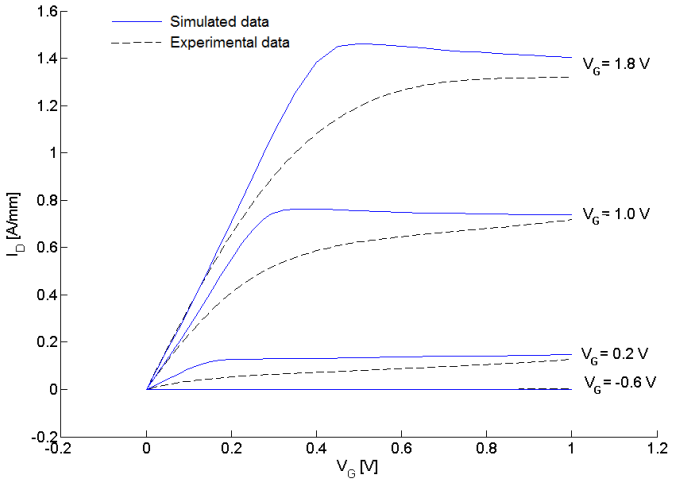


Figure 3.12: Output characteristics for 140 nm gate length. - Fitting output characteristics to experimental data for the device with 140 nm gate length.

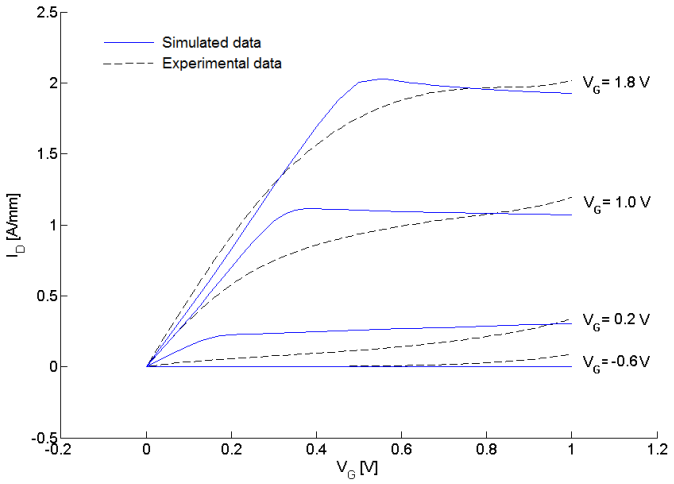


Figure 3.13: Output characteristics for 55 nm gate length. - Fitting output characteristics to experimental data for the device with 55 nm gate length.

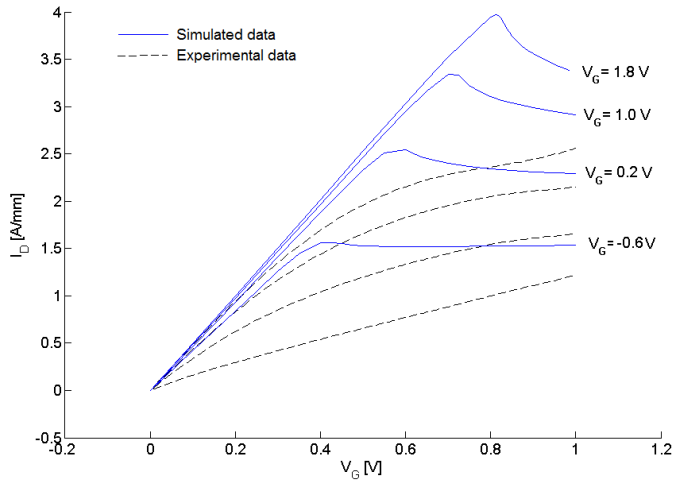


Figure 3.14: Output characteristics for 14 nm gate length. - Fitting output characteristics to experimental data for the device with 14 nm gate length.

For the 140 nm device the simulated on-resistance, that was fitted by altering the low-field mobility, is a good match to the measured results using a mobility of 1200 cm^2/Vs. The 55 nm device the results diverge more from the measured data and for the 14 nm device the results are far from the measured values using mobilities of 1100 cm^2/Vs and 1000 $cm^2/V{\cdot}s$ respectively.

A problem that is apparent when studying the output characteristics is that the modeling of the transfer from low-field mobility to high-field mobility is not accurate to the data. The model that was used is a simple model and a more advanced model could probably solve this problem.

4

Conclusions and future work

In the first part of this project the aim was to study how different doping profiles affected the electric field and breakdown voltage of a device. The result were positive and using a linear doping profile the breakdown voltage was increased. The change in doping profile also degraded the on-resistance and transconductance resulting in a tradeoff between a high breakdown voltage and degradation of other parameters. The devices simulated in this part of the project mostly use ideal parameters and does not include simulations of defects and some other models required to exactly describe the real devices. This leads to a more qualitative than quantitative reasoning; it is safe to say that these alterations of the doping profile will increase the breakdown voltage and degrade other parameters, but there are difficulties determining exactly to what grade.

In the second part of this project the Energy Balance Transport Model was implemented and transistor characteristics was simulated and compared to existing data on these devices. The aim is derive a good model for these devices to give insight into how the devices work. A good fit of the simulated data was achieved with the larger gate lengths. With smaller gate lengths the results deviated from the measured data, and further work needs to be done. This involves implementing more advanced models for mobility and structural defects as well as models for impact ionization and band-to-band tunneling. This model could then also be used to further study the impact of doping profiles on device breakdown voltage.

References

[1] Mats rlelid Karl-Magnus Persson B. Mattias Borg Filip Lenrick Reine Wallenberg Erik Lind Mikael Egard, Lars Ohlsson and Lars-Erik Wernersson. High-frequency performance of self-aligned gate-last surface channel $In_{0.53}Ga_{0.47}As$ mosfet. *IEEE Electron device letters*, 33(3), 2012. 1, 3

[2] S.M Sze. *Semiconductor Devices*. John Wiley Sons, Inc, 2002. 1, 2, 3

[3] Ye Xingning Han Lei and Chen Xingbi. Increasing breakdown voltage of ldmost using buried layer. *Journal Of Electronics*, 20(2), 2003. 2

[4] Ali A. Orouji Mahsa Mehrad. New trench gate power mosfet with high breakdown voltage and reduced on-resistance using a sige zone in drift region. *Current Applied Physics*, 1(12), 2012. 2

[5] Ran Yan-Nima Dehdashti Akhavan Weize Xiong Chi-Woo Lee, Aryan Afzalian and Jean-Pierre Colinge. Drain breakdown voltage in mugfets: Inluence of physical parameters. *IEEE Transactions on electron devices*, 55(12), 2008. 2

[6] Jesus A. del Alamo George Duh Mark H. Somerville, Roxann Blanchard and P. C. Chao. A new gate current extraction technique for measurement of on-state breakdown voltage in hemts. *IEEE Electron device letters*, 19(11), 1998. 3

[7] Silvaco, June 2012. ATLAS User's Manual - device simulation software. 6, 7, 9, 10, 11

[8] Johannes Bengtsson. Iii/v-transistor simulation. Master's thesis, Lund University, 2012. 25

Appendix A

Source code - Energy balance

```
go atlas
Title    MOSFET Id-Vd

#
# SILVACO International 1996
#
#
# SECTION 1: Mesh input
#

mesh width=21.6
set Lg=0.014
set contactXSize=0.06
set contactYSize=-0.04

set chTh=0.010
set oxTh=0.005
set oxWallTh=0.006
set xSize=2*$contactXSize+2*$oxWallTh+$Lg

set workFunction=3

x.mesh loc=0.0   spac=0.005
x.mesh loc=$contactXSize   spac=0.005
```

```
x.mesh loc=$contactXSize+$oxWallTh  spac=0.001
x.mesh loc=$contactXSize+$oxWallTh+$Lg  spac=0.001
x.mesh loc=$contactXSize+2*$oxWallTh+$Lg  spac=0.001
x.mesh loc=$xSize   spac=0.005

y.mesh loc=$contactYSize  spac=0.002
y.mesh loc=-$oxTh   spac=0.002
y.mesh loc=0.0   spac=0.002
y.mesh loc=$chTh+0.006   spac=0.002
y.mesh loc=$chTh+0.19   spac=0.01
y.mesh loc=$chTh+0.29   spac=0.1

#
# SECTION 2: Structure Specification
#
region num=1  oxide y.min=$contactYSize y.max=0.00 x.min=$contactXSize _
x.max=$xSize-$contactXSize
region num=2  material=InGaAs y.min=0.00 y.max=$chTh x.composition=0.47
region num=3  material=InAlAs y.min=$chTh y.max=$chTh+0.19 x.composition=0.48
region num=4  material=InP y.min=$chTh+0.19 y.max=$chTh+0.49
region num=5  material=InGaAs y.min=$contactYSize y.max=0.0 _
x.min=0.00 x.max=$contactXSize x.composition=0.4
region num=6  material=InGaAs y.min=$contactYSize y.max=0.0 _
x.min=$xSize-$contactXSize x.max=$xSize x.composition=0.4

electrode       num=1  name=source y.min=$contactYSize y.max=$contactYSize _
x.min=0.0 x.max=$contactXSize-0.01
electrode       num=2  name=gate   y.min=$contactYSize y.max=-$oxTh  _
x.min=$contactXSize+$oxWallTh x.max=$xSize-$contactXSize-$oxWallTh _
electrode       num=3  name=drain  y.min=$contactYSize y.max=$contactYSize _
x.min=$xSize-$contactXSize+0.01  x.max=$xSize

#Contact buffer area doping#
doping     uniform y.min=$contactYSize y.max=0.0 x.min=0 _
x.max=$contactXSize n.type conc=6e19
doping     uniform y.min=$contactYSize y.max=0.0 x.min=$xSize-$contactXSize _
```

```
x.max=$xSize n.type conc=6e19

#Structure doping#
doping     uniform y.min=0 y.max=$chTh+0.49 x.min=0.0 x.max=$xSize n.type _
conc=1e16
doping     uniform y.min=$chTh+0.001 y.max=$chTh+0.006 n.type conc=1e18
doping     uniform y.min=$chTh y.max=$chTh+0.19 p.type conc=5e16

#intdefects NTA=0 NGA=0 NGD=0 NTD=3e13 WTD=0.4 S.I

# SECTION 3: Material Models

material material=InGaAs mun=700 mup=1000 taurel.el=0.5e-12 taumob.el=0.5e-12 _
taurel.ho=0.5e-12 taumob.ho=0.5e-12 vsat=4e6
material material=InAlAs mun=350 mup=35 taurel.el=1e-12 taumob.el=1e-12 _
taurel.ho=1e-12 taumob.ho=1e-12

model fldmob srh print hcte
model material=InGaAs print evsatmod=1

contact name=source resistance=87.5
contact name=gate   workfun=$workFunction
contact name=drain  resistance=87.5

#
# SECTION 4: Id Vg
#

method newton  maxtrap=20 carriers=2
output con.band val.band e.velocity

solve init
save outf=init00.str

solve vstep=-0.1 vfinal=-1.0 name=gate
solve vdrain=0 vstep=0.01 vfinal=1 name=drain
```

```
log outf=Id_Vg_Vd10.log
solve vstep=0.1 vfinal=1.8 name=gate
log off

tonyplot Id_Vg_Vd10.log

#
# SECTION 5: Id Vd
#

method newton  maxtrap=20 carriers=2
output con.band val.band e.velocity

solve init
save outf=init00.str

solve vstep=-0.1 vfinal=-0.6 name=gate
log outf=Id_Vd_Vg-06.log
solve vdrain=0 vstep=0.1 vfinal=1 name=drain
log off

solve vdrain=0

solve vstep=0.1 vfinal=0.2 name=gate
log outf=Id_Vd_Vg02.log
solve vdrain=0 vstep=0.1 vfinal=1 name=drain
log off

solve vdrain=0

solve vstep=0.1 vfinal=1.0 name=gate
log outf=Id_Vd_Vg10.log
solve vdrain=0 vstep=0.1 vfinal=1 name=drain
log off
```

```
solve vdrain=0

solve vstep=0.1 vfinal=1.8 name=gate
log outf=Id_Vd_Vg18.log
solve vdrain=0 vstep=0.1 vfinal=1 name=drain
log off

save outf=init10.str

tonyplot Id_Vd_Vg-06.log -overlay Id_Vd_Vg02.log Id_Vd_Vg10.log _
Id_Vd_Vg18.log -set IdVd.set
tonyplot init10.str

quit
```

Appendix B

Source code - Planar linear doping

```
go atlas
Title    MOSFET Id-Vd

#
# SILVACO International 1996
#
#
# SECTION 1: Mesh input
#

mesh width=1000
set Lg=0.20
set contactXSize=0.06
set contactYSize=-0.04

set chTh=0.010
set oxTh=0.005
set oxWallTh=0.006
set xSize=2*$contactXSize+2*$oxWallTh+$Lg

set DopingTop=6e19
set DopingBottom=1e16
```

```
set munChannel=3000
set vsatChannel=2.0e7
set workFunction=4.5

x.mesh loc=0.0   spac=0.005
x.mesh loc=$contactXSize   spac=0.005
x.mesh loc=$contactXSize+$oxWallTh  spac=0.001
x.mesh loc=$contactXSize+$oxWallTh+$Lg  spac=0.001
x.mesh loc=$contactXSize+2*$oxWallTh+$Lg  spac=0.001
x.mesh loc=$xSize   spac=0.005

y.mesh loc=$contactYSize  spac=0.002
y.mesh loc=-$oxTh   spac=0.002
y.mesh loc=0.0   spac=0.002
y.mesh loc=$chTh+0.006   spac=0.002
y.mesh loc=$chTh+0.19   spac=0.01
y.mesh loc=$chTh+0.29   spac=0.1


#
# SECTION 2: Structure Specification
#
region num=1  oxide y.min=$contactYSize y.max=0.00 x.min=$contactXSize _
x.max=$xSize-$contactXSize
region num=2  material=InGaAs y.min=0.00 y.max=$chTh x.composition=0.47
region num=3  material=InAlAs y.min=$chTh y.max=$chTh+0.19 x.composition=0.48
region num=4  material=InP y.min=$chTh+0.19 y.max=$chTh+0.49
### Linear doping profile ###
region num=5  material=InGaAs y.min=$contactYSize y.max=0.0 x.min=0.00 _
x.max=$contactXSize x.composition=0.4 ND.TOP=$DopingTop ND.BOTTOM=$DopingBottom
region num=6  material=InGaAs y.min=$contactYSize y.max=0.0 _
x.min=$xSize-$contactXSize x.max=$xSize x.composition=0.4 _
ND.TOP=$DopingTop ND.BOTTOM=$DopingBottom

electrode       num=1  name=source y.min=$contactYSize y.max=$contactYSize  _
```

```
x.min=0.0 x.max=$contactXSize-0.01
electrode       num=2  name=gate   y.min=$contactYSize y.max=-$oxTh  _
x.min=$contactXSize+$oxWallTh x.max=$xSize-$contactXSize-$oxWallTh
electrode       num=3  name=drain  y.min=$contactYSize y.max=$contactYSize  _
x.min=$xSize-$contactXSize+0.01  x.max=$xSize

#Structure doping#
doping     uniform y.min=0 y.max=$chTh+0.49 x.min=0.0 x.max=$xSize _
n.type conc=1e16
doping     uniform y.min=$chTh y.max=$chTh+0.19 p.type conc=5e16

#intdefects NTA=0 NGA=0 NGD=0 NTD=3e13 WTD=0.4 S.I

#
# SECTION 3: Material Models
#
material material=InGaAs mun=$munChannel mup=350 vsat=$vsatChannel _
taun=1e-8 taup=1e-8
material material=InAlAs mun=100 mup=35 vsat=1e6 taun=1e-12 taup=1e-12
material region=1 permittivity=15

model fermi srh fldmob temp=300

impact gradqfl selb region=2 AN1=1.0E2 BN1=1.0E6 AP1=1.0E2 BP1=1.0E6 AN2=1.0E2 _
BN2=1.0E6 AP2=1.0E2 BP2=1.0E6 BETAN=1 BETAP=1
impact new selb region=5 AN1=1.0E2 BN1=1.0E6 AP1=1.0E2 BP1=1.0E6 AN2=1.0E2 _
BN2=1.0E6 AP2=1.0E2 BP2=1.0E6 BETAN=1 BETAP=1
impact new selb region=6 AN1=1.0E2 BN1=1.0E6 AP1=1.0E2 BP1=1.0E6 AN2=1.0E2 _
BN2=1.0E6 AP2=1.0E2 BP2=1.0E6 BETAN=1 BETAP=1


#contact name=source resistance=87.5
contact name=gate   workfun=$workFunction
#contact name=drain  resistance=87.5

#
```

```
# SECTION 4: Breakdown calculation
#

method newton itlim=25 trap  maxtrap=20  carriers=2
output con.band val.band

solve init
save outf=init00.str

solve vstep=0.1 vfinal=0.6 name=gate
log outf=V_bd_vg10_1e16.log
solve vdrain=0 vstep=0.1 vfinal=10 name=drain compl=2 cname=drain
log off
save outf=impact.str

tonyplot impact.str -set hole_conc.set
tonyplot V_bd_vg10_1e16.log -set V_bd.set

extract init infile="V_bd_vg10_1e16.log"
extract init infile="V_bd_vg10_1e16.log"
extract name="Vbd" x.val from curve(v."drain",i."drain") where y.val=2

quit
```