Simulating Vertical GaN FinFETs in Sentaurus

ALEXANDER SIMKO
MASTER´S THESIS
DEPARTMENT OF ELECTRICAL AND INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING | LTH | LUND UNIVERSITY

Simulating Vertical GaN FinFETs in Sentaurus

Alexander Simko

Department of Electrical and Information Technology
Lund University

Supervisor: Erik Lind

Examiner: Daniel Sjöberg

June 11, 2024

© 2024
Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

In this thesis a simulation of a vertical GaN FinFET is created. The mobility models are first calibrated against known measurements of a real device. The model is then used to investigate the effects of scaling the device. It is shown that a 1 µm drift layer with a 1 µm fin could achieve a breakdown voltage of 300 V and for a 5 µm drift layer a breakdown voltage of 700 V can be reached. The results also point to that a drift layer doping is too high for a 5 µm and 8 µm drift layer and too low for a 1 µm drift layer.

Populärvetenskaplig sammanfatning

I dag finns transistorer över allt. De kan fylla olika funktioner och därför används också olika material och olika design. För att undersöka hur material och design påverkar prestandan är simuleringar ett viktigt verktyg, vilket denna mastersuppsats ämnar att undersöka.

I en modern telefon finns flera miljarder transistorer. De är oftast några nanometer stora och gjorda i kisel. Dessa är anpassade för att vara så små som möjligt så att de blir snabba och strömsnåla. Det finns dock andra användningsområden för transistorer. Ett exempel på detta är för kraftelektronik. Det handlar då om att styra energiflödet till någon annan enhet. Det kan till exempel vara en laddare till telefonen eller styrenheten till motorn i en elbil. Oavsätt vilket så ställer detta lite andra krav på transistorn. För det första måste dessa ofta tåla höga strömmar och spänningar. Det går därför inte längre att bygga dem hur små som hellst för då skulle de höga spänningarna förstöra komponenten. Man måste också räkna med att transistorn blir varm då det oftast går mycket ström genom den vilket gör att även små förluster ger en större värmeutveckling. Detta gör att kisel inte längre är ett så lockande material att bygga transistorerna av, för dessa ändamål, då det varken tål så höga temperaturer eller klarar att stänga av höga spänningar. Ett material som i dessa aspekter är bättre än kisel är galliumnitrid, eller GaN.

Något annat som är viktigt att ta hänsyn till när man bygger en transistor är den fysiska designen. Oftast när man bygger en transistor vill man att den ska ta upp så lite area som möjligt. Detta möjligör att man kan placera många transistorer brevid varandra för att till exempel kunna leda mer ström. Som ovan nämnt kan man inte göra effekttransistorer för små för då klarar de inte längre att stänga av höga spänningar. För att därför minska arean som transistorn tar upp hjälper det att bygga en vertikal transistor som är lång nog för att kunna blockera all spänning, men som tar upp så liten area som möjligt. För att förbättra prestandan ytterligare kan en fena byggas in i transistorn. Denna fena gör att "gaten", som är den terminal som styr huruvida transistorn är på eller inte, har ett större inflytande på transistorn. Utan fenan blir det mycket svårare att stänga av transistorn vilket inte är önskvärt. Denna typ av transistor, en så kallad vertikal GaN FinFET, är ett nytt designförslag som håller på att testas av ett fåtal forskningsgrupper i världen. Eftersom det är en ny design finns det många frågor som måste besvaras. För att besvara dessa skulle man kunna bygga många transistorer med t.ex. olika dimensioner och testa vilken som fungerar bäst. Detta tar dock långt tid och kostar mycket pengar om man vill

undersöka många designer. Ett bättre sätt att undersöka det på är att simulera transistorn. Just detta har undersökts i denna uppsats. Först byggdes en simulering av en befintlig transistor för att på så sätt kunna kalibrera modelen mot en riktig enhet. Därefter kunde effekterna av ändrade dimensioner undersökas.

Acknowledgements

I would like to extend my gratitude to my supervisor Erik Lind, with whom I have had many helpful and interesting conversations over the course of the project. Our spontaneous meetings were always much appreciated. I would also like to thank my examiner, Daniel Sjöberg.

Table of Contents

Abstract i

Populärvetenskaplig sammanfattning iii

Acknowledgements v

1 Introduction 1
- 1.1 Aim and Objectives . . . 1

2 Theory 3
- 2.1 Semiconductor Physics . . . 3
- 2.2 Device Structure . . . 10

3 Simulation 11
- 3.1 Standalone Fin Simulation . . . 11
- 3.2 Full Device Simulations . . . 13
- 3.3 Scaling . . . 15

4 Discussion 19

Bibliography 21

A Scaling Data 23
- A.1 $1\,\mu m$ drift layer . . . 23
- A.2 $5\,\mu m$ drift layer . . . 25

Chapter 1

Introduction

Silicon has been the dominating material of choice for transistors for a long time. However, for power devices other materials, specifically wide band gap materials such as Gallium Nitride (GaN) or Silicon Carbide (SiC), show superior qualities such as higher switching speeds, higher breakdown voltages and better temperature resistance [1]. Most GaN devices are lateral devices resulting in larger chip area and worse heat dissipation. This can be solved by instead building fully vertical devices. This can however lead to problems in creating a normally OFF device, but by narrowing the channel into a fin the electrostatic control can be improved to the point where a normally OFF behaviour is achieved. There are however still many uncertainties with such a device, e.g. the effects of scaling and breakdown mechanisms [2]. Therefore to further the understanding of such devices a computer simulation would be helpful.

1.1 Aim and Objectives

The aim of this thesis is to create a simulation of a fully vertical GaN FinFET and improve the understanding of such devices. This goal can be split into multiple objectives:

- Create a simulation using relevant models to actually capture the physics of the device.
- Calibrate the models against previously measured results.
- With the calibrated model investigate the effect of device scaling on the device characteristics.

Chapter 2

Theory

2.1 Semiconductor Physics

A semiconductor is a material with characteristics in-between a conductor and an insulator. There are several different semiconductors but most relevant in this thesis is Gallium Nitride, or GaN, which is a crystal consisting of Gallium and Nitrogen atoms.

2.1.1 Band Gap

A single atom has discrete energy levels that its electrons can occupy. If two atoms are brought close together the two atoms will start “sharing” the electrons and each energy level will split into two sub levels. If many more atoms are brought together to form a crystal then this will continue and the discrete energy levels will turn into a continuous band that can be occupied by the electrons which is illustrated in fig. 2.1. Between these bands there will be a gap where no electrons can be. The highest energy band that is filled while the crystal is in its ground state is called the valence band and first empty band is called the conduction band. The gap in energy between these two is called the band gap. This is the energy required to excite an electron from the valence band to the conduction band. Since the valence band is filled with electrons they are unable to move. However if an electron is excited into the conduction band, it will be able to move as the conduction band is empty. At room temperature some electrons will gain enough energy to jump the gap and start moving in the conduction band. This can further be helped by introducing impurities with more electrons into the crystal. In GaN, some gallium atoms can be replaced with silicon atoms. Since silicon has one more electron than gallium the extra electron will be in the conduction band and therefore free to move. [3]

2.1.2 Mobility

Mobility describes how easily charge carriers can move through a semiconductor. It is related to the resistivity of the semiconductor by:

$$\rho = \frac{1}{q\mu n} \tag{2.1}$$

where μ is the mobility, n is the carrier concentration and ρ is the resistivity. The mobility depends on many parameters such as temperature, doping level, electric field and surface

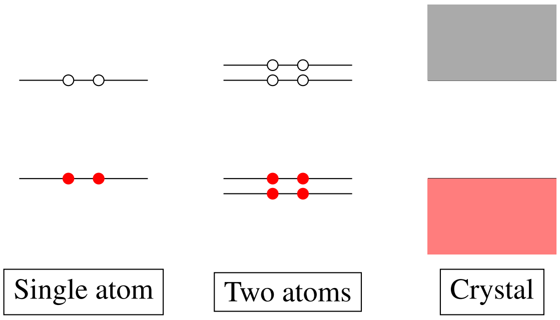


Figure 2.1: Schematic illustration of discrete energy levels transitioning to energy bands.

scattering. The mobility can also be related to the average speed of the charges inside the semiconductor at a given electric field. This relation is linear and is given by:

$$v_D = \mu E \tag{2.2}$$

with v_D being the average drift velocity and E the electric field in the semiconductor [4].

Multiple models for the mobility can be used and combined using Matthiessen's rule:

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \dots \tag{2.3}$$

By modeling different mobility dependances separately they can then be combined using eq. (2.3).

Doping Dependence

Experimental data shows that mobility decreases with increasing doping. One simple model for this is the *Arora model* given by [5]:

$$\mu_{dop} = \mu_{min} + \frac{\mu_d}{1 + \left(\frac{N_A + N_D}{N_0}\right)^A} \tag{2.4}$$

where N_A and N_D is the doping concentration and μ_{min}, μ_d, N_0 and A are fitting parameters. In order to account for temperature dependence every fitting parameter can be multiplied with a temperature dependant scaling factor, k_i

$$k_i = \left(\frac{T}{300}\right)^{\alpha_i} \tag{2.5}$$

where T is the temperature in Kelvin and α_i is another unique fitting parameter for each parameter mentioned above. Combining eqs. (2.4) and (2.5) then yields:

$$\mu_{dop} = \mu_{min} \cdot \left(\frac{T}{300}\right)^{\alpha_m} + \frac{\mu_d \cdot \left(\frac{T}{300}\right)^{\alpha_d}}{1 + \left(\frac{N_A + N_D}{N_0 \cdot \left(\frac{T}{300}\right)^{\alpha_N}}\right)^{A \cdot \left(\frac{T}{300}\right)^{\alpha_A}}} \tag{2.6}$$

Electric Field Dependence

Given equation 2.2 the drift velocity should have a linear dependence on the applied electric field. This is however only true for low electric field and it has been observed that at high electric fields the velocity saturates and becomes almost constant. To account for this, the mobility has to decrease when the electric field increases [4]. One model that accounts for this is the *Extended Canali Model* which is given by [6]:

$$\mu(E) = \frac{(\alpha+1)\mu_{low}}{\alpha + \left[1 + \left(\frac{(\alpha+1)\mu_{low}E}{v_{sat}}\right)^{\beta}\right]^{1/\beta}} \tag{2.7}$$

with α and β being fitting parameters, μ_{low} the mobility at low electric fields and v_{sat} being the saturation velocity.

Surface Scattering Dependence

Interfaces between semiconductor and oxide introduce defects in the crystal structure. This leads to a mobility degradation close to the surface due to several scattering phenomena. These include surface roughness scattering, acoustic phonon scattering and Coulomb scattering due to charged traps at the interface. Surface roughness and acoustic phonon scattering are both described in the *Enhanced Lombardi Model* and depend strongly on the electric field normal to the interface ($E_{\perp}(x)$). The mobility degradation due to acoustic phonon scattering is given by [7]:

$$\mu_{ac}(x) = \frac{B}{E_{\perp}(x)} + \frac{C\,(N_A + N_D)^{\lambda}}{E_{\perp}^{\frac{1}{3}}(x) \cdot \left(\frac{T}{300}\right)^{k}} \tag{2.8}$$

and the degradation from surface roughness is given by:

$$\mu_{sr}(x) = \left(\frac{E_{\perp}^{2}(x)}{\delta} + \frac{E_{\perp}^{3}(x)}{\eta}\right)^{-1} \tag{2.9}$$

with B, C, λ, k, δ and η being fitting parameters and x being the distance from the interface. An illustration of the x and y direction is shown in fig. 2.2. The mobilities from eqs. (2.8) and (2.9) can then be combined with any bulk mobility (μ_b) using eq. (2.3). However as these effects only occur at the surface they need to be weighted with a distance factor $D(x) = \exp(-x/l_{crit})$. l_{crit} is also a fitting parameter. This gives the final mobility:

$$\frac{1}{\mu} = \frac{1}{\mu_b} + \frac{D(x)}{\mu_{ac}(x)} + \frac{D(x)}{\mu_{sr}(x)} \tag{2.10}$$

where the mobility far from the interface is roughly equal to μ_b and only closer to the interface does μ_{ac} and μ_{sr} effect the mobility.

The mobility degradation from Coulomb scattering, i.e. scattering from trap charges at the interface can be modeled using the following formula [8]:

$$\mu_c(x) = \frac{\mu_1\left(\frac{T}{300}\right)^{k}\left[1 + \left(\frac{n}{c_{trans}\cdot\left(\frac{N_{p,i}}{N_0}\right)^{\eta_1}}\right)^{v}\right]}{\left(\frac{N_{p,i}}{N_0}\right)^{\eta_2}\cdot D(x)\cdot f(E_{\perp}(x))} \tag{2.11}$$

Coulomb scattering only occurs if the field normal to the interface is sufficiently large. Therefore the field dependant function $f(E_\perp)$ is introduced:

$$f(E_\perp(x)) = 1 - \exp\left[-\left(\frac{E_\perp(x)}{E_0}\right)^\gamma\right] \tag{2.12}$$

In eqs. (2.11) and (2.12) μ_1, k, ν, c_{trans}, N_0, η_1, η_2, γ and E_0 are fitting parameters and T is the temperature, n is the electron charge density, $N_{p,i}$ is the positive interface charge density and $D(x)$ is the above mentioned distance factor.

Effective Mobility

As many of the above mentioned mobility models depend on the position it can be hard to measure these in physical devices. Therefore the concept of *effective mobility* is very useful. The effective mobility is given by [4]:

$$\mu_e = \frac{\int_{-x_i}^{x_i} \mu(x)n(x)dx}{\int_{-x_i}^{x_i} n(x)dx} \tag{2.13}$$

This can be interpreted as a weighted average of the mobility, weighted by the density of the charge carries, meaning that areas without or with very few mobile charges does not contribute to the overall effective mobility and areas with a high charge carrier concentration have a large contribution. x is once again the distance from the semiconductor/oxide interface. $2x_i$ is the width of the channel. This is shown in fig. 2.2 for a fin with oxide on both sides of the semiconductor.

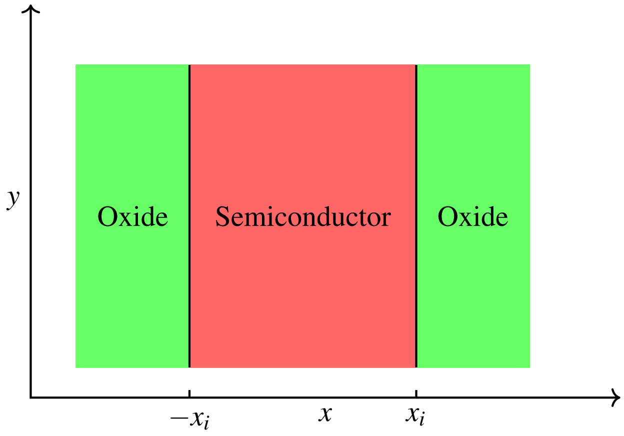


Figure 2.2: Illustration of semiconductor/oxide interface.

Since it is hard to measure mobility at a specific point another way of calculating the effective channel mobility is used from the definition of mobility given by equation 2.1. In an actual device the resistivity can be related to the drain conductance g_d and the dimensions L_g and L_w (length and width) of the channel. $q \cdot n$ is the charge, Q_n, inside the channel and the final formula for the effective mobility then becomes [2]:

$$\mu_e = \frac{g_d L_g}{L_w Q_n} \tag{2.14}$$

2.1.3 Generation and Recombination

Electrons in the conduction band can jump back into the valence band. When they do so they release energy, either in the form of light or as heat into the semiconductor. There are different ways this process can occur. One way would be that an electron jumps directly from the conduction band to the valence band. This is called band-to-band recombination. Another way would be if the electron uses a trap state inside the band gap to reach the valence band. Such trap states could come from unwanted defects, but it could also be from the doping atoms as they introduce new states in the band gap. This type of recombination is called Shockley-Read-Hall (SRH) recombination or trap-assisted recombination. All recombination processes can be reversed, meaning an electron can be excited from the valence band to the conduction band if it receives energy, either from light or from the heat of the crystal. This reverse processes is called generation.

Another very important generation process is the so called impact ionization. When charge carries travel through the material they cannot travel very far before hitting an atom and losing energy. If however a very high electric field is applied the carriers can have so much energy that when they collide with an atom they excite more carriers. These new carriers will also feel the strong electric field and will therefore accelerate and excite more carriers with the same process. This leads to an avalanche effect and a rapid increase in charge carrier concentration which in turn means a large current starts flowing. Therefore this is also known as avalanche breakdown. For a transistor this process is unwanted and limits how much voltage the transistor can block. By increasing the size of the transistor this effect can be mitigated as the voltage is applied over a larger distance, thus decreasing the maximum electric field.

Shockley-Read-Hall Recombination

At equilibrium the concentration of electrons n and of the holes p follow the law of mass action meaning that:

$$n \cdot p = n_i^2 \tag{2.15}$$

where n_i is the intrinsic carrier concentration, or how many free electrons and holes the undoped semiconductor has. If however, the semiconductor is forced out of equilibrium, generation (or recombination) will occur to try to move it back to this equilibrium. To characterize the rate at which this occurs, the concept of carrier lifetime τ_n and τ_p is useful. They describe how long a carrier can exist in the conduction band (or valence band for the holes) before recombining. Putting this together and adding an extra term n_1 and p_1 given by

$$n_1 = n_i \exp\left(\frac{E_{trap}}{kT}\right) \tag{2.16}$$

and

$$p_1 = n_i \exp\left(\frac{-E_{trap}}{kT}\right) \tag{2.17}$$

to account for the trap states in the band gap, yields an overall expression for the SRH recombination as [8]:

$$R_{net}^{SRH} = \frac{np - n_i^2}{\tau_p(n + n_1) + \tau_n(p + p_1)}. \tag{2.18}$$

Impact Ionization

The generation rate due to impact ionization is dependant on the current and the electric field. A general description of impact ionization can be given as [8]:

$$G_{ii} = \frac{1}{q}\left(\alpha_n J_n + \alpha_p J_p\right) \tag{2.19}$$

where J_n and J_p are the electron and hole current density and α_n and α_p are the field dependant, so-called, *impact ionization coefficients*. The impact ionization coefficient is the number of electron-hole pairs excited over 1 cm [4]. One model to describe the impact ionization coefficients is the van Overstreaten - de Man model which states that [9, 10]:

$$\alpha(E) = \gamma a \exp\left(-\frac{\gamma b}{E}\right) \tag{2.20}$$

with a and b as fitting parameters and γ being a temperature dependant parameter.

2.1.4 Breakdown

From eq. (2.19) it is clear that if the impact ionization coefficients are large enough even a small current can give rise to a large amount of electron hole pairs. If there is an electric field applied then these carriers will start moving and increase the current which in turn will lead to more generation and so on. When this occurs it is no longer possible to control the current in the device and breakdown occurs. A useful concept when talking about breakdown is the *multiplication factor*, $M(x)$, which gives the total number of electron-hole pairs generated from a single electron-hole generated at position x. $M(x)$ is related to the impact ionization coefficient and under the assumption that $\alpha = \alpha_n = \alpha_p$ can be expressed as [4]:

$$M(x) = \frac{1}{1 - \int_0^W \alpha dx} \tag{2.21}$$

where W is the width of the drift region. Using the multiplication factor avalanche breakdown can be defined to occur when $M(x)$ tends to infinity or when:

$$\int_0^W \alpha dx = 1 \tag{2.22}$$

This means that α needs to be kept small in order to avoid breakdown. α has a dependence on the electric field, as seen in eq. (2.20) meaning that there is some *critical electric field*, E_C, at which breakdown will occur. This field strength is material dependant and for GaN it is $E_c = 3.3\,\mathrm{MV\,cm^{-1}}$ [11]. Using this the width of the depletion region can be related to the critical electric field and the desired blocking voltage V_{BV}:

$$W_D = \frac{2V_{BV}}{E_C} \tag{2.23}$$

Another important characteristic of the drift region is the doping concentration. When a voltage is applied over the drift region the mobile charges there are forced out of the drift region, thereby depleting it of charge carriers and leaving the fixed charges behind. Normally there is an equal amount of fixed and mobile charges meaning that there is no net charge. However when the drift region becomes depleted there is an excess of fixed charges causing a net charge. This charge will then generate an electric field opposing the externally applied electric field and thereby prevent current from flowing. If the doping concentration is high in the drift region then many charges can be removed over a small distance and then the depletion region becomes very small. However from eq. (2.23) it is clear that in order to achieve a specific blocking voltage the length of the drift region has to be sufficiently large. This means that a limit on the doping level can be made as well and is given by [4]:

$$N_D = \frac{\varepsilon_S E_C^2}{2qV_{BV}} \tag{2.24}$$

Since both the length of the depletion region and its doping concentration can be related to both the breakdown voltage and the resistance of the drift region a relation between the two can be made. For larger power devices the on-state resistance is dominated by the resistance of the drift region and is therefore referred to as *ideal specific on-resistance of the drift region*. Using eqs. (2.1), (2.23) and (2.24) an overall expression can be derived as:

$$R_{on-ideal} = \frac{4V_{BV}^2}{\varepsilon_s \mu_n E_C^3} \tag{2.25}$$

The relation between the square of the breakdown voltage and the on state resistance ($\varepsilon_s \mu_n E_c^3/4$) is referred to as *Baliga's figure of merit* or BFOM:

$$\text{BFOM} = \frac{\varepsilon_s \mu_n E_c^3}{4} = \frac{V_{BV}^2}{R_{on}} \tag{2.26}$$

In theory the BFOM should remain constant for different designs, however since the breakdown voltage and R_{on} are design dependant the BFOM will vary for different desins. It can then be used as a measure of how close to the ideal limit a particular design is.

2.2 Device Structure

The structure simulated was first presented in [12] and further improved in [2]. It is a vertical FinFET made in GaN and the overall structure is shown in fig. 2.3. When a negative bias is applied to the gate the fin becomes depleted turning the transistor off. By selecting a metal with the right work function, this behaviour can be achieved at 0 V gate voltage and the device is then normally off. For this device Tungsten was used with a work function of about 4.5 eV. The thick drift layer will support the voltages while the device is off and therefore extending the device in the vertical direction should improve breakdown voltage. The thin fin helps improve electrostatic control and to further improve this a high-κ dielectric is used. In this case HfO_2 was used which has a dielectric constant of roughly 20. The n^+ regions has a doping concentration of $1 \cdot 10^{19}\,\mathrm{cm}^{-3}$ and the n region has a doping of $4 \cdot 10^{16}\,\mathrm{cm}^{-3}$. An Al_2O_3 spacer is added to move the gate contact away from the drift layer and therefore decrease the capacitance between the gate and the source. This also has the effect of creating an “ungated” part of the fin, where the gate contact is far away from the fin. This part will therefore not be affected by any applied voltage on the gate.

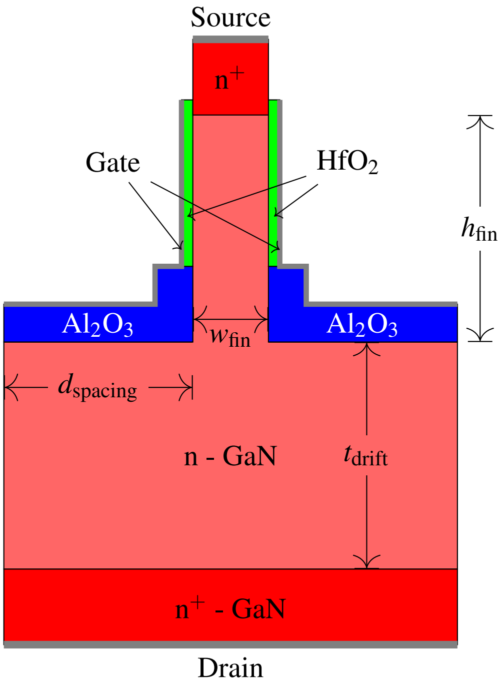


Figure 2.3: Device structure of a vertical FinFET.

Chapter 3

Simulation

Different simulations were created to simulate the different behaviors of the transistor and to calibrate the physics model. All simulations were done using Sentaurus TCAD software.

3.1 Standalone Fin Simulation

The first simulation created was the standalone fin simulation. This was used to calibrate the interface mobility models described in section 2.1.2. The effective mobility of the fin was calibrated to match the measured effective mobility in [2]. The fin was 1 µm long and 100 nm wide. The left and right contact have a heavily doped region with a doping concentration of $n^+ = 1 \cdot 10^{19}\,\mathrm{cm}^{-3}$ and the fin is lightly doped with $n = 4 \cdot 10^{16}\,\mathrm{cm}^{-3}$. The gate runs above and below the fin and has a work function of 4.8 eV. The full structure is shown in fig. 3.1.

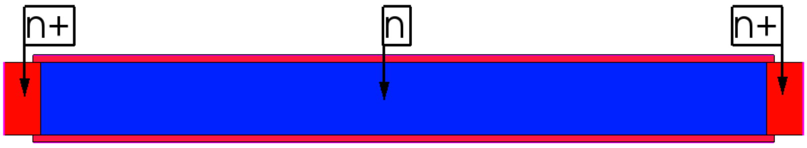


Figure 3.1: Structure of fin simulation.

In order to model the mobility dependence on the gate voltage accurately, the Enhanced Lombardi Model was used together with Coulomb scattering. Since no measurements on temperature had been done the simulation was done at 300 K which meant that the k in eqs. (2.8) and (2.11) could be ignored during calibration. In addition there were trapped charges along the GaN/HfO$_2$ interface with a concentration of $1 \cdot 10^{12}\,\mathrm{cm}^{-2}$. SRH recombination and doping dependant mobility was also used. For the non field dependant mobility the Aurora model was used and parameters were set to get a mobility of $450\,\mathrm{cm}^2/(\mathrm{Vs})$ at a doping concentration of $4 \cdot 10^{16}\,\mathrm{cm}^{-3}$.

Using the built in optimizer tool in Sentaurus the parameters were updated to reduce the difference between the simulated effective mobility of the channel and the measured mobility. To save time only some parameters were updated by the optimizer. The three parameters with the greatest impact on the error were identified to be c_{trans}, δ and B. Both the default value for η and the values given in [13] followed the relation $\eta = 1 \times 10^{16}\delta$. This was kept for the optimized value of δ as well. Some additional parameters were still updated, since in the process of investigating which parameters had the greatest impact, random values of all parameters were tested. By doing so, better values for these parameters were found.

Table 3.1: Calibrated parameters values. Sources for the initial values are given.

Parameter	Unit	Initial value	Optimized Value
Acoustic Phonon Scattering [14]			
B	cm/s	$1 \cdot 10^6$	$10 \cdot 10^6$
C	$\mathrm{Kcm/s(V/cm)^{-3/2}}$	$3 \cdot 10^6$	$1 \cdot 10^6$
λ	1	0.03	0.01
Surface Roughness Scattering [13]			
δ	$\mathrm{cm^2/Vs}$	$1.6 \cdot 10^{14}$	$4.737 \cdot 10^{14}$
η	$\mathrm{V^2/cms}$	$1.6 \cdot 10^{30}$	$4.737 \cdot 10^{30}$
c_{trans}	$\mathrm{cm^{-3}}$	$1 \cdot 10^{15}$	$9 \cdot 10^{14}$

The updated parameters are given in table 3.1 as well as their initial values and the resulting effective mobility is shown in fig. 3.2. The parameters not mentioned in table 3.1 were kept at their default values in Sentaurus. Of note is that the optimizer was only run on the mobility values for $V_g > 0.75\,\mathrm{V}$ as the measured mobility for low gate voltages was not reliable.

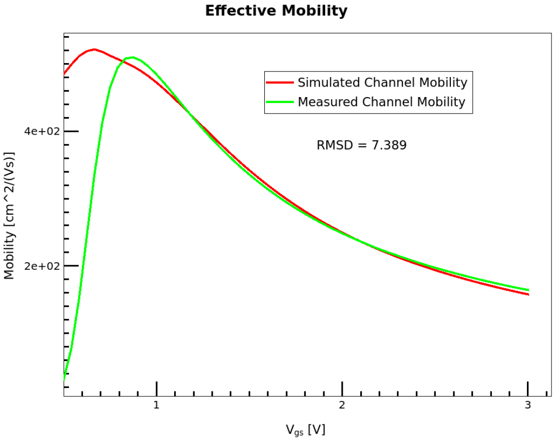


Figure 3.2: Effective mobility of fin after calibration.

3.2 Full Device Simulations

Using the calibrated mobility models the full device could be simulated. This device has a drift layer of 1 μm, a fin width of 100 nm and a fin length of 1 μm. The models used were the same as for the fin simulation, with the addition of a series resistance on the source contact of $5 \cdot 10^{-6}\,\Omega\mathrm{cm}^2$. For the drain voltage sweep and the breakdown simulation mobility saturation using the Extended Canali model was added and impact ionization was also part of the breakdown simulation. All simulations used periodic boundary conditions. The first simulation was a V_g sweep.

3.2.1 V_g Sweep

The result of the V_g sweep is shown in fig. 3.3. The gate voltage was swept between $-3\,\mathrm{V}$ to $3\,\mathrm{V}$ while the drain voltage was kept constant at $50\,\mathrm{mV}$. The values are similar to the once shown in [2], however there the current is normalized to the fin area and not the device area. The threshold voltage is also very close, it was measured to be $V_t = 0.58\,\mathrm{V}$ and simulated as $0.39\,\mathrm{V}$. The discrepancy could be explained by differing work function for the gate in the simulation compared to the actual Tungsten metal.

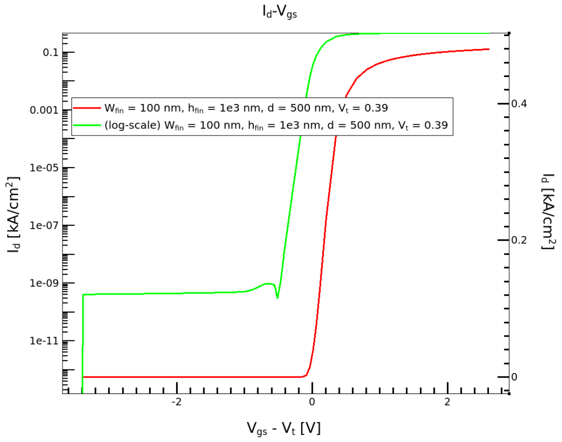


Figure 3.3: Gate voltage sweep. Current is normalized to the area of the drift layer. Sweep was performed at $V_{ds} = 50\,\mathrm{mV}$.

3.2.2 V_d Sweep

The drain voltage sweep was performed at multiple different gate voltages. The threshold voltage (V_t) calculated from the V_g-sweep was added to the gate voltage to account for the different threshold voltages for different devices. Higher gate voltage increased the saturation current. The results are shown in fig. 3.4. From the drain voltage sweep the on-state resistance was extracted from the point of greatest slope.

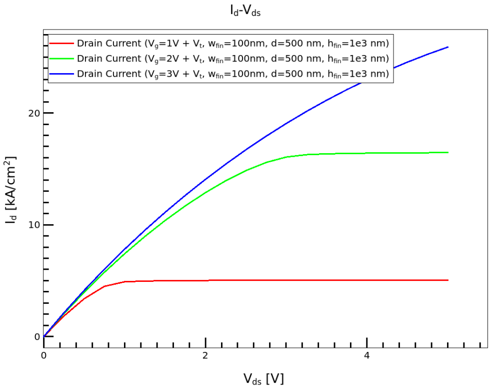


Figure 3.4: Drain voltage sweep at different gate voltages. Current normalized to device area. $V_t = 0.39\,\mathrm{V}$.

3.2.3 Breakdown Simulation

The final simulation was the breakdown simulation. From the V_g-sweep the gate voltages at which a current of $1 \cdot 10^{-10}\,\mathrm{A\,\mu m^{-1}}$ flows through the devices was determined. This was then used to bias the breakdown simulation and give a leakage current of about the same size. With a larger negative gate bias the leakage current was dominated by a displacement current, leading to unrealistic behaviour and convergence issues. The result is shown in fig. 3.5. A large contact resistance of $1 \cdot 10^{11}\,\Omega\,\mathrm{\mu m^{-1}}$ was also added in series with the drain. With a current of $1 \cdot 10^{-10}\,\mathrm{A\,\mu m^{-1}}$ this amounts to a voltage drop of $10\,\mathrm{V}$ across the resistor. This resistor helped the simulation to converge once breakdown started occurring in the simulation since at that point the current rises rapidly with very small changes of voltage. Since the simulation steps the voltages, most of the increase in voltage will fall over the resistor and therefore the increase in voltage over the transistor is limited. The plot shown in fig. 3.5 shows the voltages over the transistor, not including the voltage over the series resistor. This explains why the voltages drops as the current increases.

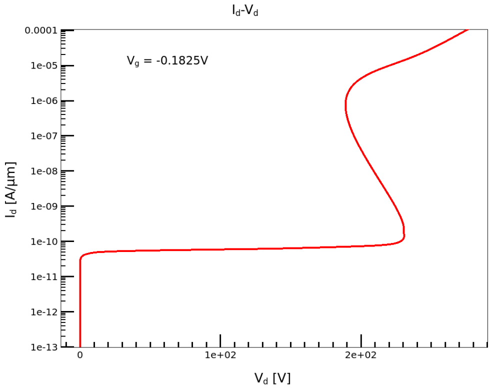


Figure 3.5: Breakdown simulation with $w_{\text{fin}} = 200\,\mathrm{nm}$.

3.3 Scaling

With the working simulations the effect of scaling was investigated. The width of the fin w_{fin}, length of the fin h_{fin}, distance between the fin and the edge d_{spacing} and the thickness of the drift layer t_{drift} was varied. All combinations of the values given in table 3.2 were simulated. The simulated dimensions are illustrated in fig. 3.6. The results for the 8 µm drift layer are shown in figs. 3.7 to 3.9 and the rest of the results are given in appendix A.

Table 3.2: Dimensions simulated.

w_{fin}[nm]	h_{fin}[nm]	d_{spacing}[nm]	t_{drift}[µm]
10	300	160	1
30	500	300	5
50	800	500	8
100	1000	800	–
200	–	–	–

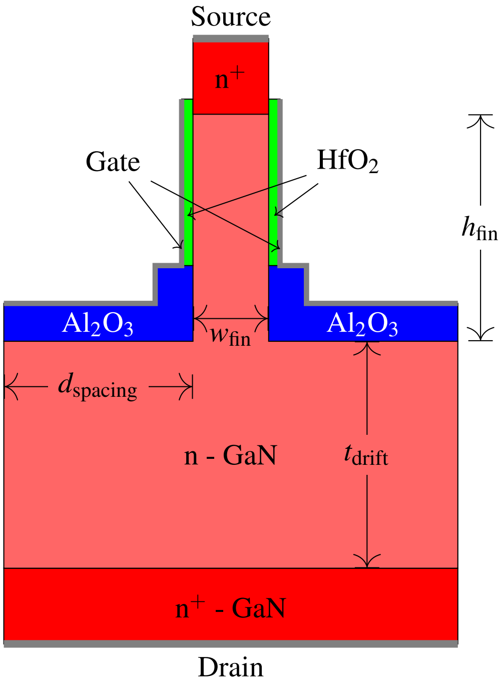
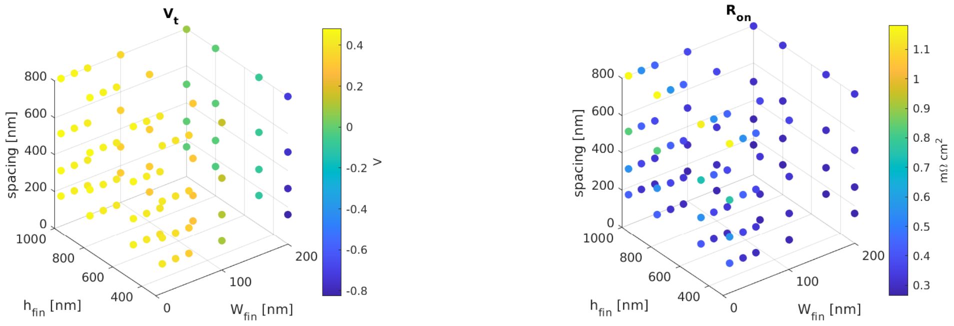


Figure 3.6: Examined dimensions of the device.

For all drift layers the width of the fin has the largest impact on the threshold voltage. This is to be expected as a narrower channel improves the electrostatic control from the gate which would improve the built in potential and raise the threshold voltage. The height of the fin plays some role, however this is mostly in the devices with a wide fin.

This shows that for wider fins a long enough channel is needed to counteract the loss of electrostatic control. The width of the device and the length of the drift layer does not impact the threshold voltage. This is good as the threshold voltage should be determined from the channel and the rest of the device should not impact it. The on-state resistance also behaves as expected. A longer drift layer increases the resistance and so does the length of the fin. Since the resistance is normalized to the width of the device ($2d_{\text{spacing}} + w_{\text{fin}}$), an increase in resistance can be seen for the wider devices.



(a) Threshold voltage

(b) On-state resistance

Figure 3.7: Threshold voltage and on-state resistance of 8 μm drift layer device.

The saturation current increases when d_{spacing} decreases. Since the current is normalized to the area of the device, a smaller area gives a higher normalized current so this is expected. However another trend can also be seen. A wider fin increases the current. This could be due to the top contact resistance as a wider fin would give a smaller contact resistance. It could also be due to the n$^+$ region under the contact is also adding a substantial resistance or the part of the fin where the Al_2O_3 separates the gate from the fin. In either case it is most likely due to an increase in the resistance of the fin, limiting the saturation current at narrow widths. The height of the fin does not impact the saturation current which indicate that the part of the fin with the HfO_2 does not limit the current. If the resistance of this "gated" part would be substantial then a decrease in current should be seen for longer fins.

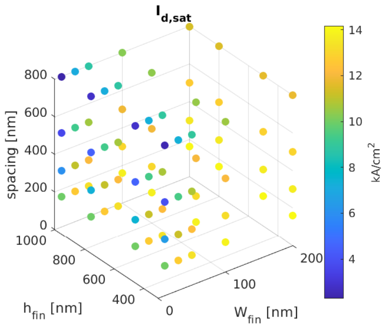
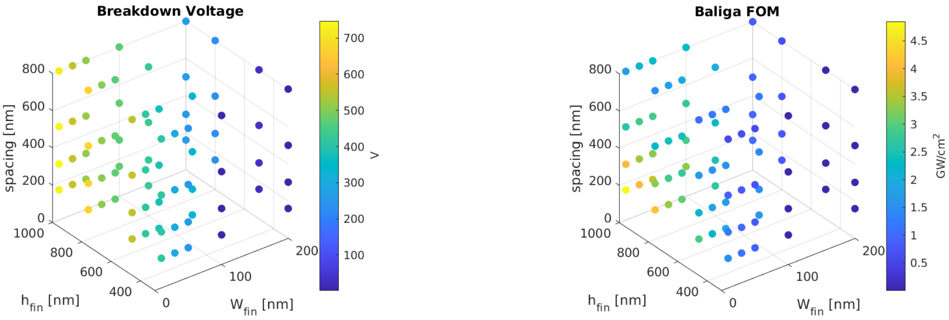


Figure 3.8: Saturation current of 8 μm drift layer device.

The breakdown improves for narrower and longer fins. This could be because the narrower fins offer greater electrostatic control which improves the band bending providing a greater barrier to overcome before breakdown. The longer fins would also provide a longer barrier which could improve the breakdown. However more likely is that part of the fins also becomes depleted meaning that the fin adds to the length of the depletion region. This would improve the breakdown voltage. With a depletion layer of 2 μm a breakdown voltage (using equation eq. (2.23)) of 330 V is expected. This is very close to the breakdown voltage of the 1 μm drift layer with a 1000 nm long fin seen in fig. A.1.

The optimal Baliga's figure of merit for both the 5 μm and 8 μm drift layer occur at $h_{\text{fin}} = 1000\,\text{nm}$, $d_{\text{spacing}} = 160\,\text{nm}$ and $w_{\text{fin}} = 10\,\text{nm}$. For the 1 μm device the result is similar, however the fin was slightly wider with $w_{\text{fin}} = 100\,\text{nm}$.



(a) Breakdown Voltage **(b)** Baliga's Figure of Merit

Figure 3.9: 8 μm drift layer.

Chapter 4

Discussion

It is clear that the models used for the mobility of the fin describe the physics of the device well. Even though these models were originally observed in silicon they still work well, with modified parameters, for GaN. The error can probably be improved if more parameters are optimized at the same time, however it would take longer to run. Something else worth considering is that the measured mobility is based on a lot of estimates of e.g. the contact resistance, the resistance of the drift layer etc. Given this error it might not be worth improving the model further as there is still an uncertainty in the reference mobility.

In spite of only calibrating the low field mobility, the simulation result was fairly close to the actual device. The simulation could be further improved by calibrating more models. This would however require more measurements. One example of how the V_d-sweep could be improved could be with measurements of the drift velocity at different field strengths in order to calibrate the Extended Canali Model for the mobility degradation at high electric fields.

The simulated breakdown was a lot higher than the measured breakdown. This was because in the physical device the breakdown happened due to gate leakage and all current flowed between the drain and the gate. An attempt was made to run the breakdown simulation with Fowler-Nordheim tunneling which is a type of field dependant tunneling that allows current to flow through the gate oxide at high electric fields. However in Sentaurus this extra gate current is added as an additional boundary condition on the Al_2O_3/GaN interface, which causes the convergence of the simulation to suffer. This is usually counteracted with a finer mesh at the interface and this was also tested, however a finer mesh lead to a slower simulation which made it unfeasible to simulate gate leakage. Another problem with the breakdown simulation was that if a large negative gate bias was applied then the leakage current was dominated by a displacement current which is not very physical. In order to create a more reasonable leakage current the gate voltages was raised slightly. In the physical device there was a significant leakage current even when the gate was biased to $-3\,\mathrm{V}$. This could be due to traps at the HfO_2/GaN interface limiting the amount the bands can bend and therefore preventing the transistor from turning off completely. If that is the case, then applying a larger bias would be equivalent to limiting how much the bands bend. However it still shows that not all the physics is completely captured in the models used and maybe more models are needed to describe the device accurately. If so then improving the models used could also improve convergence when running the simulation with gate leakage. Usually a more accurate simulation leads to better convergence.

Even though the breakdown did not match the measured results it was still a reasonable value. Using equation eq. (2.23) an estimate for the breakdown can be made. For a 1 µm, 5 µm and 8 µm a breakdown voltage in GaN of 165 V, 825 V and 1320 V is expected. The best 1 µm drift layer device had a breakdown voltage of around 300 V. This could indicate that a significant part of the fin is also depleted which raises the breakdown voltage. Another conclusion that can be made is that the doping concentration of the drift layer can probably be raised as the breakdown voltage exceeds the expected value. This would improve the on-state resistance and in turn improve Baliga's figure of merit. For both the 5 µm and 8 µm device the breakdown voltage is the same which is below the expected value. This indicates that the doping concentration is too high in the drift layer. Using eq. (2.24) an estimate for the correct doping concentration can also be made based on the estimated breakdown voltage. GaN has a relative permittivity of roughly 8.5 [15] which gives a doping concentration of $3.1 \cdot 10^{16}\,\mathrm{cm}^{-3}$ and $1.9 \cdot 10^{16}\,\mathrm{cm}^{-3}$ for a 5 µm and 8 µm drift layer respectively. Decreasing the doping concentration would raise the on-state resistance.

Bibliography

[1] José Millán. “A review of WBG power semiconductor devices”. In: *Proceedings of the International Semiconductor Conference, CAS*. Vol. 1. 2012. DOI: `10.1109/SMICND.2012.6400696`.

[2] P Gribisch et al. “Capacitance and Mobility Evaluation for Normally-Off Fully-Vertical GaN FinFETs”. In: *IEEE Transactions on Electron Devices* 70.8 (2023), pp. 4101–4107. ISSN: 1557-9646 VO - 70. DOI: `10.1109/TED.2023.3287820`.

[3] S. M. Sze and M.K. Lee. *Physics of Semiconductor Devices: Physics and Technology*. 2012.

[4] B.Jayant Baliga. *Fundamentals of power semiconductor devices*. Cham, Switzerland: Springer, 2019. ISBN: 3030067653.

[5] Narain D. Arora, John R. Hauser, and David J. Roulston. “Electron and Hole Mobilities in Silicon as a Function of Concentration and Temperature”. In: *IEEE Transactions on Electron Devices* 29.2 (1982). ISSN: 15579646. DOI: `10.1109/T-ED.1982.20698`.

[6] C. Canali et al. “Electron and Hole Drift Velocity Measurements in Silicon and Their Empirical Relation to Electric Field and Temperature”. In: *IEEE Transactions on Electron Devices* 22.11 (1975). ISSN: 15579646. DOI: `10.1109/T-ED.1975.18267`.

[7] Claudio Lombardi et al. “A Physically Based Mobility Model for Numerical Simulation of Nonplanar Devices”. In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 7.11 (1988). ISSN: 19374151. DOI: `10.1109/43.9186`.

[8] Synopsis. “Sentaurus Device User Guide”. In: *Simulation* December (2022).

[9] R. Van Overstraeten and H. De Man. “Measurement of the ionization rates in diffused silicon p-n junctions”. In: *Solid State Electronics* 13.5 (1970). ISSN: 00381101. DOI: `10.1016/0038-1101(70)90139-5`.

[10] A. G. Chynoweth. “Ionization rates for electrons and holes in silicon”. In: *Physical Review* 109.5 (1958). ISSN: 0031899X. DOI: `10.1103/PhysRev.109.1537`.

[11] Hao Wu et al. “Breakdown voltage improvement of enhancement mode AlGaN/GaN HEMT by a novel step-etched GaN buffer structure”. In: *Results in Physics* 29 (2021). ISSN: 22113797. DOI: `10.1016/j.rinp.2021.104768`.

[12] Philipp Gribisch et al. “Tuning of Quasi-Vertical GaN FinFETs Fabricated on SiC Substrates”. In: *IEEE Transactions on Electron Devices* 70.5 (2023). ISSN: 15579646. DOI: `10.1109/TED.2023.3263154`.

[13] Viktoryia Uhnevionak et al. “Comprehensive Study of the Electron Scattering Mechanisms in 4H-SiC MOSFETs”. In: *IEEE Transactions on Electron Devices* 62.8 (2015). ISSN: 00189383. DOI: `10.1109/TED.2015.2447216`.

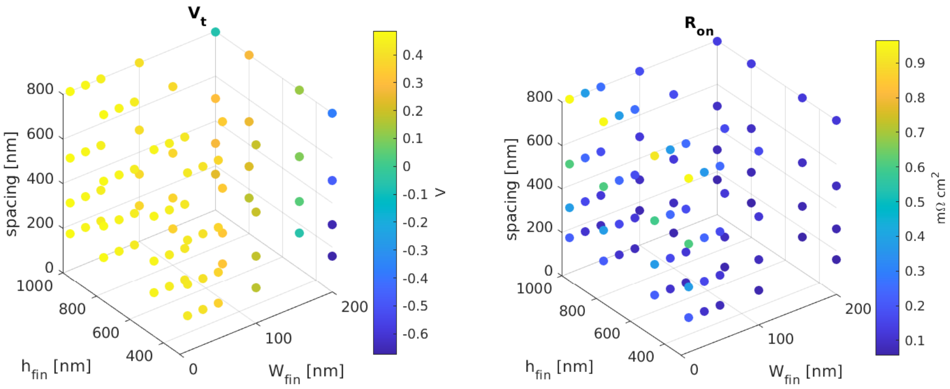
[14] A. Pérez-Tomás et al. “Field-effect mobility temperature modeling of 4H-SiC metal-oxide- semiconductor transistors”. In: *Journal of Applied Physics* 100.11 (2006). ISSN: 00218979. DOI: `10.1063/1.2395597`.

[15] S. Yu Davydov. “Estimates of the spontaneous polarization and permittivities of AlN, GaN, InN, and SiC crystals”. In: *Physics of the Solid State* 51.6 (2009). ISSN: 10637834. DOI: `10.1134/S1063783409060249`.

Appendix A

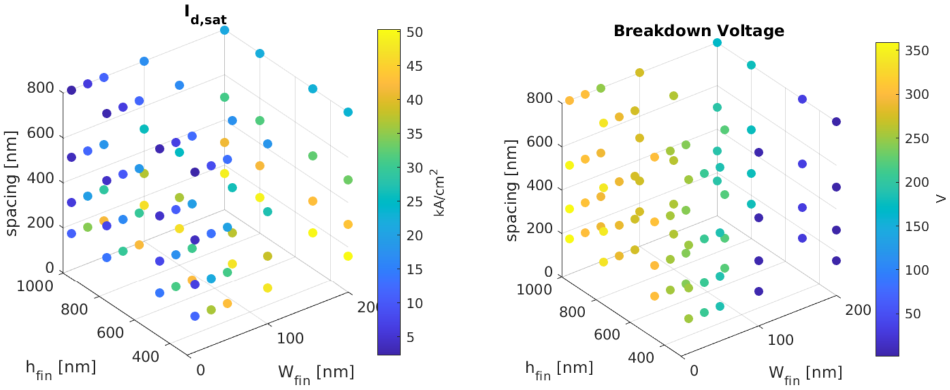
Scaling Data

A.1 $1\,\mu m$ drift layer



(a) Threshold voltage

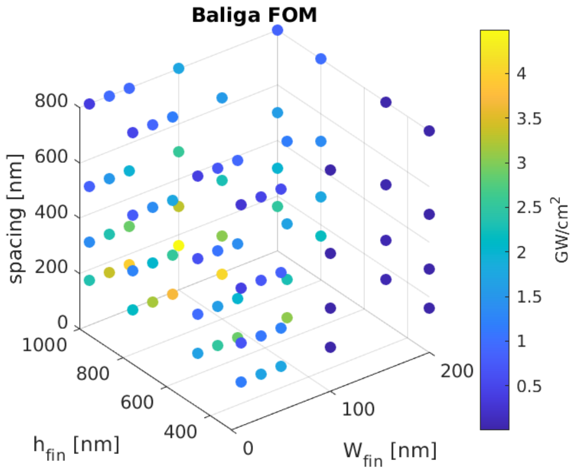
(b) On-state resistance



(c) Saturation current

(d) Breakdown Voltage

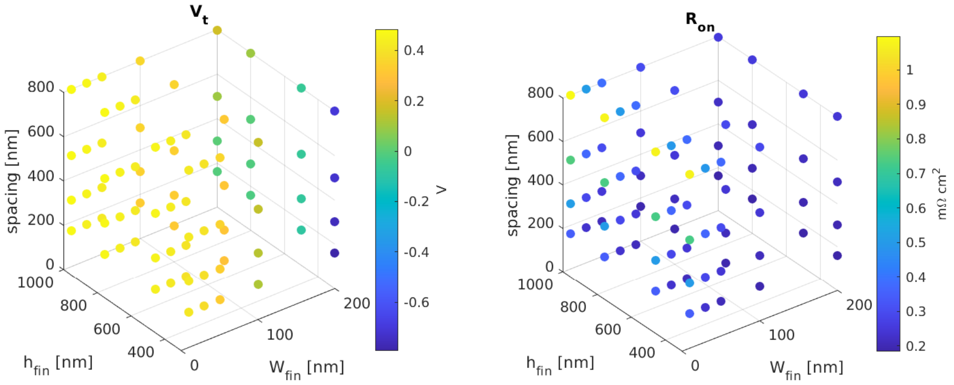
Figure A.1: $1\,\mu m$ drift layer.



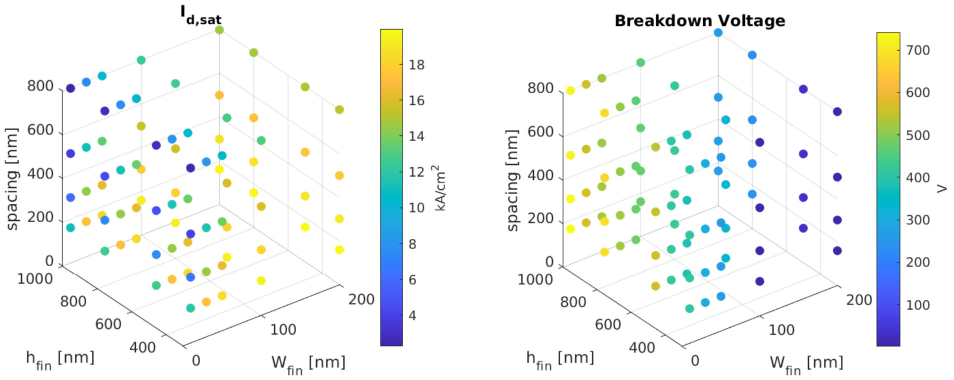
(e) Baliga's Figure of Merit

Figure A.1: $1\,\mu m$ drift layer (cont).

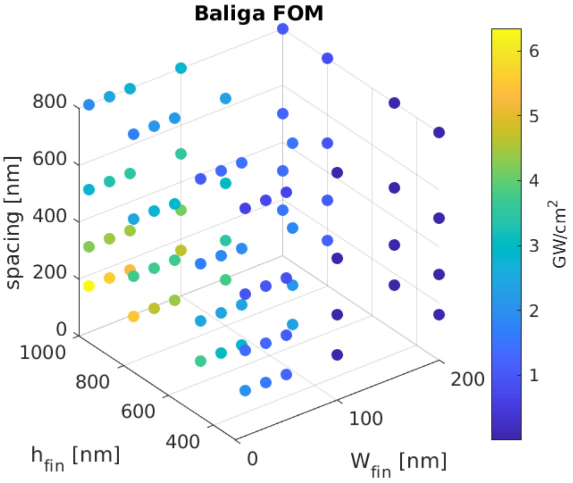
A.2 5 μm drift layer



(a) Threshold voltage **(b)** On-state resistance



(c) Saturation current **(d)** Breakdown Voltage



(e) Baliga's Figure of Merit

Figure A.2: 5 μm drift layer.

Printed by Tryckeriet i E-huset, Lund 2024

LUND UNIVERSITY

Series of Master's theses
Department of Electrical and Information Technology
LU/LTH-EIT 2024-983
http://www.eit.lth.se