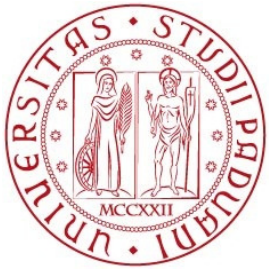


UNIVERSITY OF PADUA
Faculty of Engineering – Department of Information Engineering

Simulations of High Mobility AlGaN/GaN Field Effect Transistors

Mobility and Quantum Effects

Omenetto Leonardo
Academic Year 2011/2012

University Supervisor: Prof. Meneghesso Gaudenzio
Infineon Supervisor: Dr. Curatola Gilberto

Table of contents

Acknowledgements

I would like to thank my supervisors, my professor Gaudenzio Meneghesso that gave me the opportunity to to do my thesis in Infineon Technologies Austria, and Gilberto Curatola, my Infineon supervisor, that has helped me in my internship, guiding me in this “new” field, the Power Gallium Nitride devices; I have learned from him several things regarding Gallium Nitride and the Sentaurus simulator.

I cannot forget to mention professor Giovanni Verzellesi, that introduced me to the simulation software, giving me several aids.

Many thanks to Valentina for her patience, for the support she gave me to begin this experience and for all the period I was in Villach and for the help she gave me with some precious hints for my thesis.

Thanks to my parents for their support during my internship and for the help they gave me in some moments.

Thanks to all my friends in Villach, particularly to Tommaso that helped me several times, and to Manuel, Mariana, Javier, Vinicius, Garazi, Francesco, Fabio, David, Vicent and many others. Thanks for this great experience.

Introduction

1 Introduction

During the past few years Si-based devices have reached their limits regarding sizes, maximum frequencies and maximum output powers. Nowadays the devices need higher and higher performance and smaller dimensions.

In the last decades enormous progress has been made in the development of III-Nitride-based devices. Despite there is not the same knowledge of the silicon based structures, this new technology has shown a lot of improvement. It exploits the heterostructure between a ternary and a binary compound (for example AlGaN/GaN or AlGaAs/GaAs) which creates a quantum well, at the interface, full of electrons with high mobility.

Thanks to high breakdown field, high electron saturation velocity and high density of carriers in the channel, these high electron mobility transistors (HEMTs) have a high Johnson´s figure of merit, a property that characterizes the tendency to work at high power and high frequency.

One of the main properties of Gallium Nitride is the high direct energy gap that allows to apply high electric field and to work at higher temperature than traditional transistors. Another important characteristic, present even in the ternary material AlGaN, are the polarizations of this compound. The spontaneous one is due to the intrinsic dipole and the piezoelectric one is due to the possible strains of the materials.

Finally, when these two materials are grown one above the other, it forms a two-dimensional electron gas (2DEG) at the interface, present at the equilibrium too, that increases the mobility in the channel.

Nevertheless, AlGaN/GaN HEMTs have some weaknesses, due to the presence of traps and degradation that nowadays are still poorly understood. In fact, these traps, localized on the surface, in the barrier layer and in the buffer layer, cause a worsening of a lot of parameters. For this reason this technology need more development to improve reliability.

To limit the production costs', the productors are trying to grow the GaN-based devices on Silicon substrate, one of the most common and cheap semiconductors. As will be explained in one of the next chapters, the lattice mismatch between Si and GaN leads to the generation of defects and dislocations at interfaces. Besides, in case of growth of GaN onto GaN, the quality would be excellent, but costs would be too high to bid against other substrates.

This work will be focused on GaN-based structures, particularly in their mobility and in the quantum effects in the channel. These topics are very important. Mobility is a critical parameter in the channel of AlGaN/GaN HEMT and all the commercial simulators have been calibrated for Si-based devices. Quantum effects must be analyzed very well because the dimensions of the channel of these devices have the same order of magnitude of the one taken in account by the Quantum Mechanics.

1.1 Outline of the thesis

In the second chapter are presented the wide band gap materials, specially the binary and ternary compounds, with particular attention to Gallium Nitride, the base of High Electron Mobility Transistors used for this thesis.

The third chapter introduces the High Electron Mobility Transistors, especially the GaN-based HEMTs, which exploit the AlGaN/GaN heterostructure, by showing the polarization in these materials, the two-dimensional gas of electrons at the interface and the important negative role of the traps.

After this introduction to materials, in the fourth chapter, starts one of the main parts of this thesis, which is focused on the mobility in these transistors. In the first part there is an introduction to the concept of mobility, its dependences and the mechanisms that limit it, and then the implementation of a model for GaN-based devices in the Sentaurus simulator, a simulation tool developed by Synopsys.

Finally, in the fifth chapter, are presented the quantum effects in the AlGaN/GaN HEMT. By using the Sentaurus simulator, will be shown the different behaviour of the available models. After this, will be studied the confinement of electrons in the channel and several solutions to improve the confinement.

2 Wide band gap materials

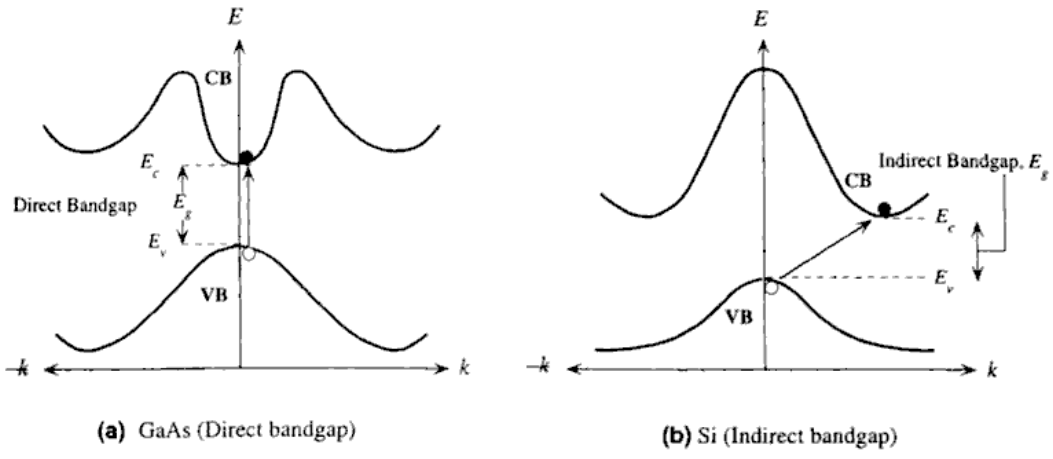
Since the end of the 80´s these materials have obtained a lot of success in optoelectronic and electronic applications. Semiconductors with a band gap greater than 1.7 eV are considered wide band gap materials.The high energy gap allows them to operate at high temperature and to tolerate high applied voltages. This intrinsic feature also brings electronic transitions in the visible range, therefore facilitates the radiative phenomena, and hence important for light-emitting devices such as LED and LASER. It´s important to mention that solid state lighting could reduce the amount of energy required to provide lighting as compared with incandescent lights, which are associated with a light output of less than 20 lumens per watt. The efficiency of LEDs is on the order of 160 lumens per watt.

Wide band gap semiconductors (WBG) can also be used in RF signal processing. Silicon-based power transistors have reached their limits of operating frequency, power and current density and breakdown voltage. These compounds are excellent substitutes and they can be used in high-temperature and power switching applications.

The magnitude of the coulombic potential, the size of atoms and their electronegativities are three factors that determine this high band gap. Small atoms and strong electronegative atomic bonds are associated with wide band gaps.

Elements high on the periodic table are more likely to be wide band gap materials. With regard to III-V compounds; nitrides are the most used because they have the largest band gaps. Band gaps can often be engineered by alloying and Vegard's Law states that there is a linear relation between lattice constant and composition of a solid solution at constant temperature.

The position of the conduction band minima versus the valence band maxima, in the band diagram, determine whether a band gap is direct or indirect. Most of wide band gap materials are associated with a direct band gap, with SiC and GaP as exceptions. This is fundamental for optoelectronic, because it allows the radiative recombinations to occur.



(a) GaAs (Direct bandgap) **(b)** Si (Indirect bandgap)

Figure 2.1: Difference between direct and indirect band gap

WBG semiconductors have a high breakdown voltage. This is due to a larger electric field required to generate carriers through impact mechanism. The drift velocity reaches a peak at an intermediate electric field and undergoes a small drop at higher fields.

The most common WBG structures are wurtzite and zincblende. Wurtzite phases allow spontaneous polarization in the (0001) direction. A result of the spontaneous polarization and piezoelectricity is that the polar surfaces of the materials are associated with higher sheet carrier density than the bulk. The polar face produces a strong electric field, which creates high interface charge densities.

Property	**GaN**	**AlN**	**InN**	**SiC**	**Si**	**GaAs**
Band gap (E_g) [eV]	3.44	6.2	0.79	3.26	1.12	1.43
Electric breakdown field (E_c) [MV/cm]	3.0	1.4-1.8	3.0	0.3	0.4	0.5
Saturated electron velocity (v_{sat}) [x10^7cm/s]	2.5	1.7	4.5	2.0	1.0	1.0
Electron mobility (μ_n) [cm^2/Vs]	900 (bulk) 2000 (2DEG)	135	3200	700	300	8500
Electron effective mass (m_c) [m_0]	0.22	0.4	0.11	0.2	1.18	0.63
Hole effective mass (m_v) [m_0]	0.8	3.53	0.27	1.0	0.55	0.52
Lattice constant [Å]	3.175	3.111	3.533	3.073	5.431	5.653

Table 2.1: Most important properties of wide band gap materials

2.1 Binary materials

[1] The elements Al-In-Ga-N are the base for the binary compound. The class of compounds will be take in account in this work are the Nitrides: gallium nitride (GaN), aluminium nitride (AlN) and indium nitride (InN) .

There are three common crystal structures for III-Nitrides (III-N):

- Wurtzite

- Zincblende

- Rock salt

GaN, AlN and InN have a wurtzite structure at room temperature, while in BN usually prevails the cubic structure. It is also possible to find GaN and InN with a zincblende structure, only for thin films, while it is pretty impossible to find AlN in zincblende phase becuase it is unstable. Last phase, rock salt, is not important for electronic purposes.

2.1.1 GaN

Initially the development of GaN was slowed by some intrinsic limits: difficulty to find appropriate substrate, the presence of a lot of defects during growth and difficulty to find a p-type dopant; nowadays is still hard, or almost impossible, to dope GaN with acceptors. Thank to new techniques of growth, such as molecular beam epitaxy (MBE) and metal organic chemical vapor deposition (MOCVD), and new improvements, GaN-based is slowly becoming more mature and some GaN products have already appeared in the market, especially for RF applications.

[1] Nowadays Gallium Nitride is the basic material of III-N material class. If an application require fast carrier transport and high breakdown voltage in a device this is the first choice.

2.1.1.1 Lattice structure

GaN is a binary compound formed by one atom of gallium (III group, Z = 31) and one atom of nitrogen (V group, Z = 7).

As written in the "Binary materials" paragraph, GaN can be found usually in wurtzite structure, the most common in nature and the most used in electronic devices. Zincblende structure is created by epitaxial growth of thin films in (011) plane using as substrate Si, SiC, MgO and GaAs. Rock salt phase realization requires very high pressure.

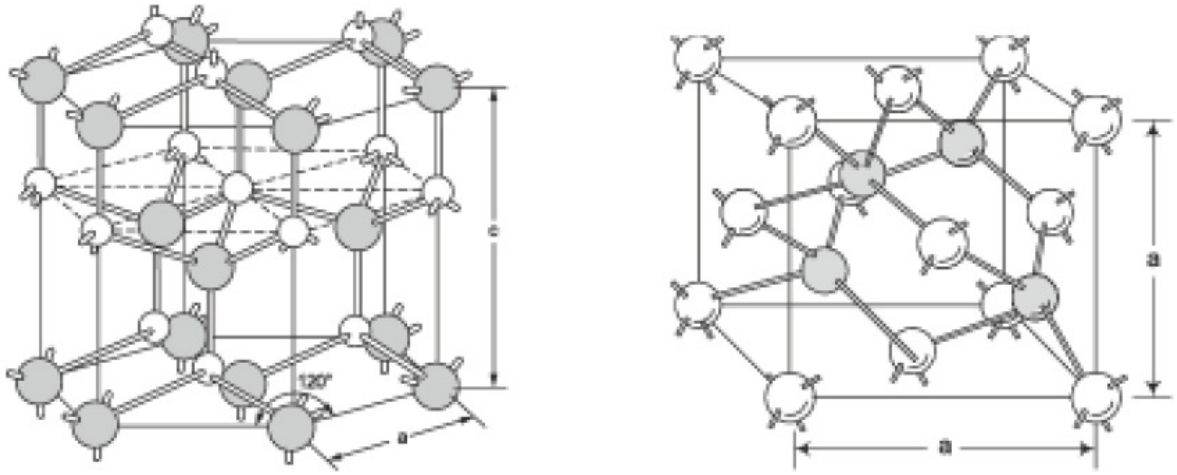


Figure 2.2: Lattice structures of GaN: wurtzite(left) and zinc blende (right)

Structure of cristalline lattice is very important to determine the property of the material. For example energy gap is inversely proportional at lattice constant.

2.1.1.2 Properties

The first chemical experiments, performed with GaN-based materials, have shown two important properties: high stability and high hardness. These peculiarities, that remain unchanged with increasing of temperature, are due to the tetrahedral structure of the material; in fact every atom has strong ionic ties with the four neighbors.

It is significant to mention the thermal expansion coefficent (TEC), that define the variation of lattice parameters with changing of temperature, and it's very important for growth. This value depends on intrinsic material factors, such as stoichiometry, defects and concentration of charge carriers. In fact there is a dispersion of this value when GaN is grown on different substrate due to high manufacture temperature.

As discussed in the former paragraphs, GaN is used in high power and high frequency applications, therefore it´s useful to introduce the Johnson´s figure of merit, which characterizes these devices performance:

$$F = \max(P_{out}, f_T) \qquad (2.1)$$

Where P_{out} is the power density and f_T is the cut-off frequency of the device.

For the output power, Johnson has used maximum output voltage multiplied for maximum output current, with dependence on breakdown electric field and velocity saturation

$$V_{MAX} = \frac{\varepsilon_s E_{bk}^2}{2qn} \qquad I_{MAX} = qndv_{SAT} \qquad (2.2)$$

$$P_{out} = \frac{1}{2}\varepsilon_s E_{bk}^2 dv_{SAT} \qquad (2.3)$$

And for the cut off frequency he has used a formula with dependence on velocity saturation and channel length

$$f_T = \frac{v_{SAT}}{2\pi L_g} \qquad (2.4)$$

In the next table it is possible to note the great value for GaN, which is two orders of magnitude higher than Si or GaAs

	Si	**GaAs**	**InP**	**SiC**	**GaN**
JFOM	1	7.8	14.7	400	625

Table 2.2: Johnson figure of merit for several materials

2.1.1.3 Polarization

All the III-V compounds have an intrinsic polarization. In particular, GaN can have two polarizations: spontaneous and piezoelectric. The first one is due to the different electronegativity between the atoms of Gallium and the atoms of Nitrogen: this difference causes an electric dipole that characterizes these materials. The second one is not intrinsic, in fact it´s due to mechanical lattice stress during growth above a different substrate.

In the heterostructures, such as AlGaN/GaN, the total spontaneous polarization is given by the difference between the individual spontaneous polarizations of two materials. Then, due to the lattice constant mismatch, the strain causes the piezoelectric polarization. This phenomenon is determinated not only by the lattice mismatch, but even by a thermal mismatch, induced by values of TEC, that are different in the two compounds.

In a subsequent paragraph this topic will be deepened.

2.1.1.4 Energy gap and band structure

The energy gap is very important for GaN. First, it's very high, therefore it's possible to apply very high voltage to the devices, and second, it's direct, and this is very significant for optoelectronic devices.

	Si	GaAs	InP	SiC	GaN
Band Gap[eV]	1.12	1.43	1.34	3.26	3.44

Table 2.3: Energy band gap for several materials

This high direct band gap characterizes both the wurtzite structure and the zincblende structure. In GaN the conduction band minima is in correspondence of the valence band maxima; therefore an electronic transition requires an energy exchange, but not a momentum exchange.

In III-N the value of band gap is inversely proportional to lattice constant, as it is possible to see in Figure 2.3.

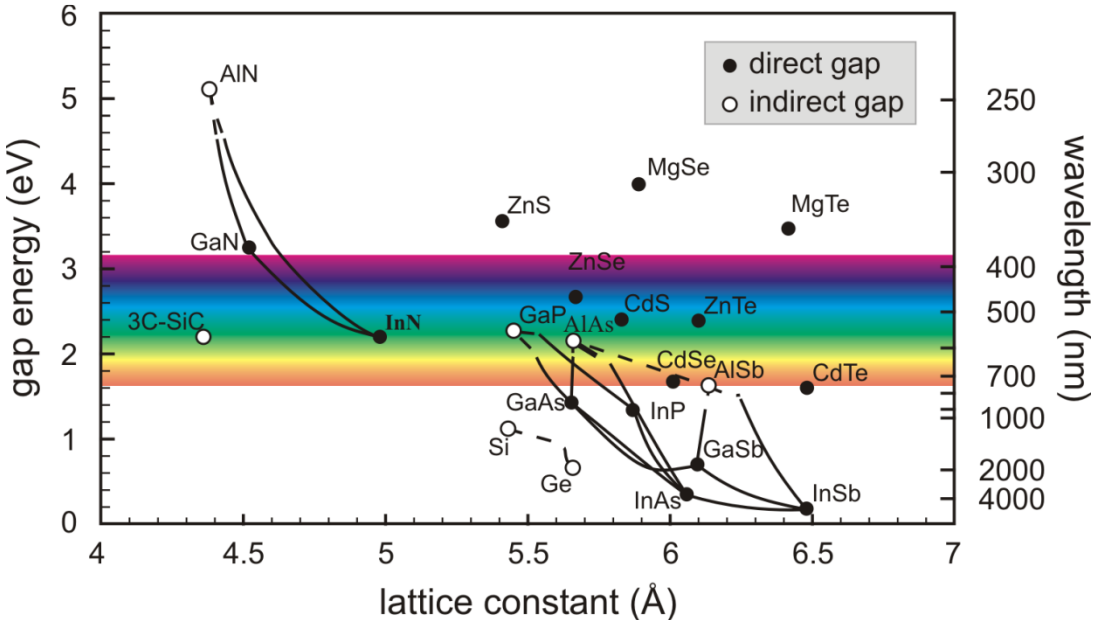


Figure 2.3: Energy band gap for several materials as function of lattice constant

It is important to note the high value of the AlN, that is used in the GaN-based HEMT, compared with GaN or InN.

The high value of energy gap is associated with a high electric breakdown field, in fact it is necessary a high applied voltage to activate the impact ionization between the bands [2]. This great value allows to reduce the sizes of the devices and even the distance between drain and gate. Finally it allows to work at high temperature because it´s necessary a high thermal energy to cause problems to the device.

The Figure 2.4 shows the GaN´s bands for the wurtzite structure and for the zincblende structure.

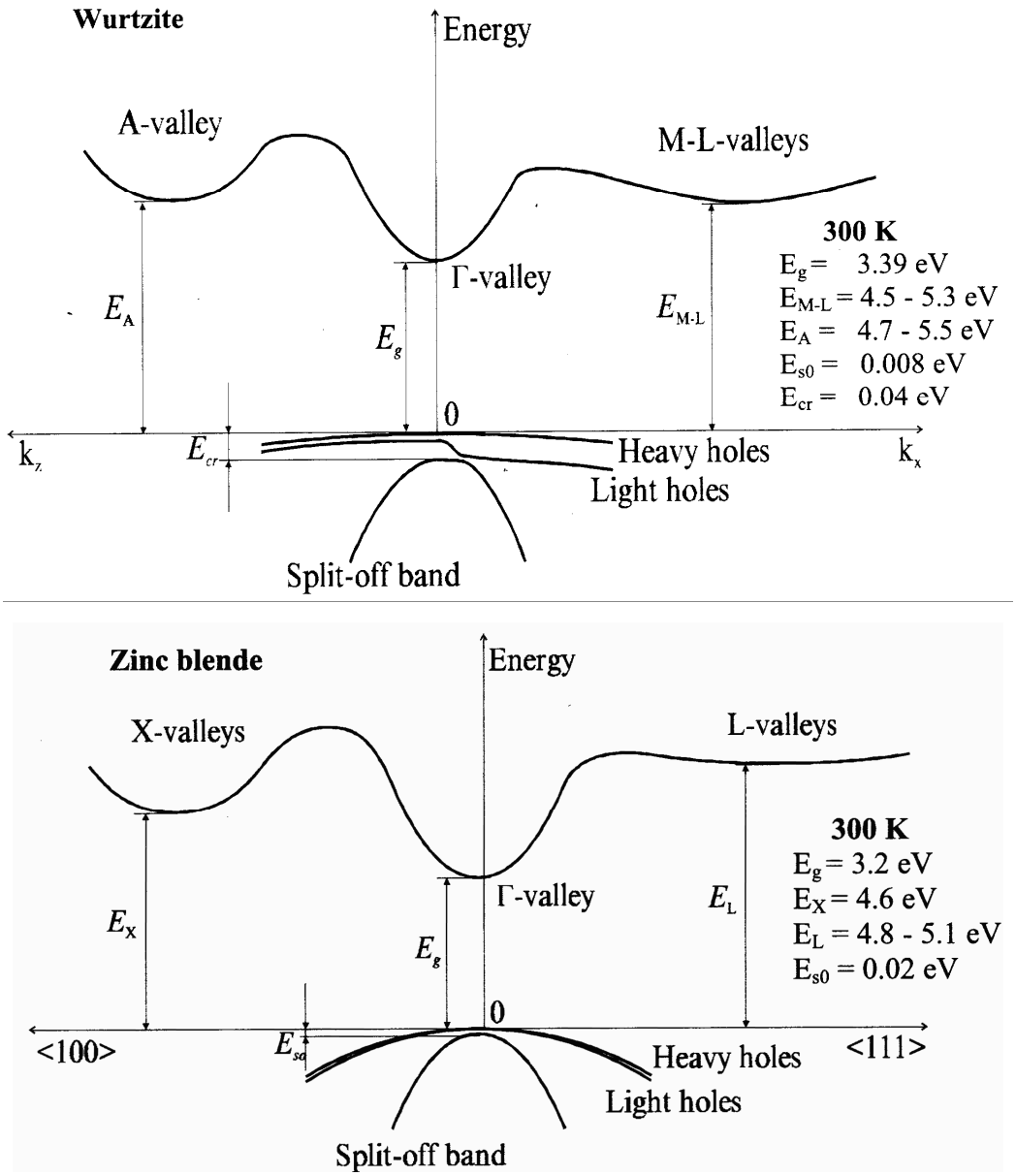


Figure 2.4: Energy bands for Gallium Nitride: wurtzite (up) and zinc blende (down)

As explained above, it is possible to see the direct band gap for both the structures, in fact the conduction band minima has the same momentum of the valence band maxima.

For valence band there are three different curves, due to the strain of the lattice and to the high electric field, and they are defined HH (Heavy holes), LH (light holes) and split-off band. For conduction band it is shown the lowest energy valley, the lambda-valley, and the closest valleys, that have higher energy.

2.1.1.5 Substrates

For many years the problem to find a suitable substrate for GaN has slowed its development. With the III-V compounds, to grow a thin layer onto a substrate of the same material (homoepitaxy) is too exspensive [3]; to solve this problem, usually the growth is made by heteroepitaxy, that is a kind of epitaxy performed with materials that are different from each other. For this reason it is very important to have substrates with lattice constant and TEC similar to GaN.

The correct lattice constant is fundamental to avoid mechanical stress at the interface and consequently to avoid defects and dislocations during growth; even the correct TEC is important, because heteroepitaxy requires high manufacturing temperature.

The main characteristics of a substrate are [1]:

- Price and price per area

- Lattice mismatch relative to the material

- Thermal conductivity and coefficient of thermal expansion

- Maximum electrical isolation

- Availability with respect to the diameter

- Crystal quality and residual defect density

- Surface properties and residual defect density

- Wafer warp and wafer bowing

- Mechanical and chemical properties with respect to thinning and viahole etching

First attempts were made by using, as substrate, silicon (Si), zinc oxide (ZnO), magnesium oxide (MgO) and compounds based on silicon carbide, such as silicon on polySiC (SopSic) and silicon carbide on polySiC (SiCopSiC); nowadays the most common and suitable substrates are sapphire (Al2O3) and silicon carbide (SiC).

First, the lattice mismatch between SiC and GaN is 3.5% and the thermal mismatch is 3.2 %, that are lower than sapphire ones; thank to these parameters there is a low generation of defects. Second, SiC has a high electric and thermal conductivity, therefore it´s possible to grow very thin films on substrate and to work with at high power devices [4].

Recently, due to high costs of production of SiC, manufacturers have tried to develop other cheaper materials, such as SopSiC and SiCopSiC, obtaining good results.

Sapphire is often used as substrate because of its crystalline structure with hexagonal simmetry; a lot of applications use this material for this reason, because is similiar to GaN.

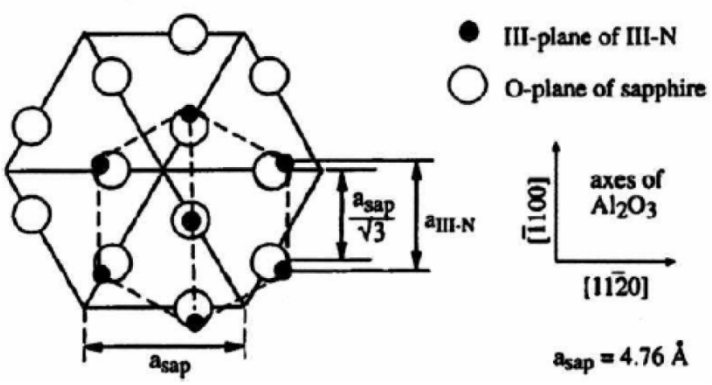


Figure 2.5: Comparison between Sapphire and Gallium Nitride(wurtzite)

Despite this material has quite high values of mismatches, lattice and thermal, it is stable at high temparature and it is possible to obtain a good Johnson´s figure of merit. Furthermore it is an insulator and it is transparent to radiation, from ultraviolet to microwave [3].

This material, in spite of the high generation of defects that degradates the devices and limits mobility, is still widely used because growth processes are good developed and therefore the manufacturing costs are lower than SiC substrates. A solution, that keeps alive sapphire substrates, is to introduce a layer, the nucleation layer, between the compounds aiming to limit the lattice mismatch.

With MOCVD and MBE, that will be explain in next paragraph, it is possible to build very thin layers of material, such as the nucleation layer, made in AlN, or an AlN spacer between GaN and AlGaN to limit alloy scattering.

Finally, in Table 2.4, are shown some important parameters for several substrates.

Parameters	**GaN**	**Al_2O_3**	**SiC**	**Si (111)**
Lattice constant [Å]	3.19	4.75	3.08	5.43
Lattice mismatch [%]	0	49	-3.5	70
Thermal expansion coefficient [$10^{-6}K^{-1}$]	5.6	-7.5	4.46	3.59
Thermal mismatch [%]	0	33.9	-20.3	-35.8
Therm. conductivity [W/cmK]	1.3	0.5	5	1.5

Table 2.4: Important parameters to grow GaN onto several materials

2.1.1.5 Techniques of growth

Nowadays exist many techniques to grow GaN on sapphire substrate and on silicon carbide substrate. The target is to obtain a heterojunction with the minimum number of defects. This paragraph will summarize the most used growth processes.

The MOCVD (Metal Organic Chemical Vapor Deposition), also known as MOVPE (Metallic Organic Vapour Phase Epitaxy) exploits a chemical reaction at high temperature; precursors of the material, at vapour phase, are deposited onto a substrate and then, at high temperatures (500°C-1100°C) [1], they react till to form a semiconductor layer. This technique is widely used for III-V devices, especially for optoelectronic applications, because of its precision and its growth velocity (about 2 µm/h) [1].

In Figure 2.6 it is shown an example of MOCVD.

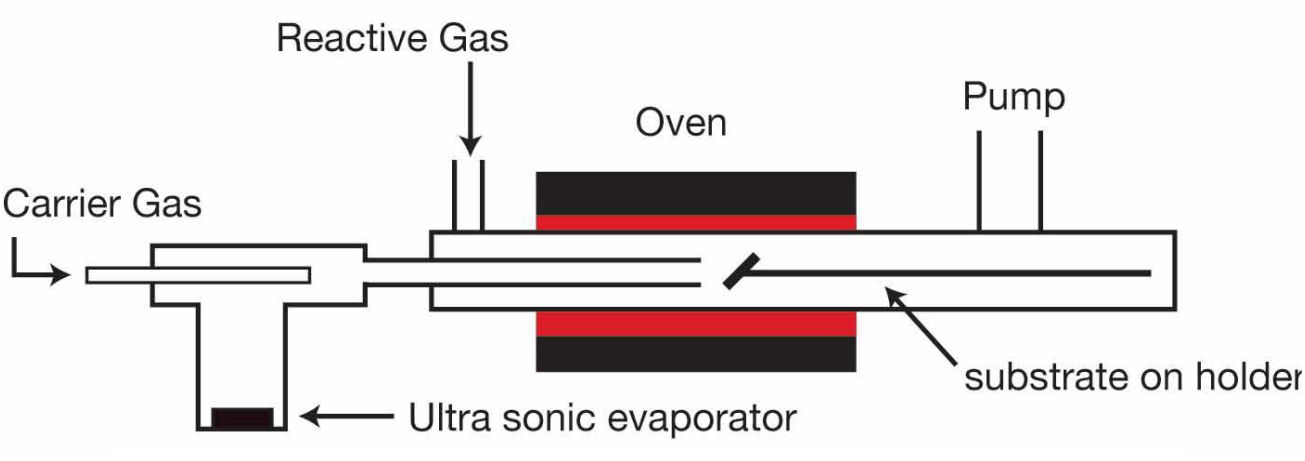


Figure 2.6: Simplified operation of the MOCVD

In III-N materials, in particular for GaN, the main reaction that controls the growth on the substrate is

$$Ga(CH_3)_3(v) + NH_3 \rightarrow GaN(s) + 3CH_4(v)(l) \qquad (2.5)$$

In this process the substrate is placed on a rotating support of graphite, in a reaction chamber, and for each material there are different growth temperatures. This reaction takes place not in a vacuum, but from the gas phase at moderate pressures (2 to 100 kPa).

The weaknesses of this technique are the low concentration of electrons that it can provide to the final structure and the difficulty to build very thin layer of some materials, such as InGaN; nevertheless exists a variation of this technique, the two-flow MOCVD, that solves some problem of MOCVD using a second flow of gases to improve the efficient of the main flow.

In Table 2.5 are summarized the precursors of the most used materials.

Materials	**Precursors (phase)**
Aluminium	Trimethylaluminium (liquid)
	Triethyaluminium (liquid)
Gallium	Trimehylgallium (liquid)
	Triethygallium (liquid)
Indium	Trimethylindium (solid)
	Di-isopropylmethylindium (liquid)
	Ethyldimethylindium (liquid)
Germanium	Isobutylgermane (liquid)
	Tetramethylgermane (liquid)
Nitrogen	Phenyl hydrazine (liquid)
	Dimethylhydrazine (liquid)
	Ammonia (gas)
Phosphorus	Phosphine (gas)
	Bisphonoethane (liquid)

Table 2.5: Precursors for MOCVD for the most used materials

The second technique widely used is the MBE (Molecular Beam Epitaxy). It exploits ionized beams to deposit the materials on a rotating substrate [4]. This epitaxy requires high vacuum condition and this prevents the use of high concentration of doping. The main advantages are the very high precision, due to the shutters that control the surface emitting and receiving molecular beams, and the semplicity to determinate a certain polarity [4]. It is important to say that this technique is very slow and, because of that, it is not used in power industrial applications, but only for tests in the university labs.

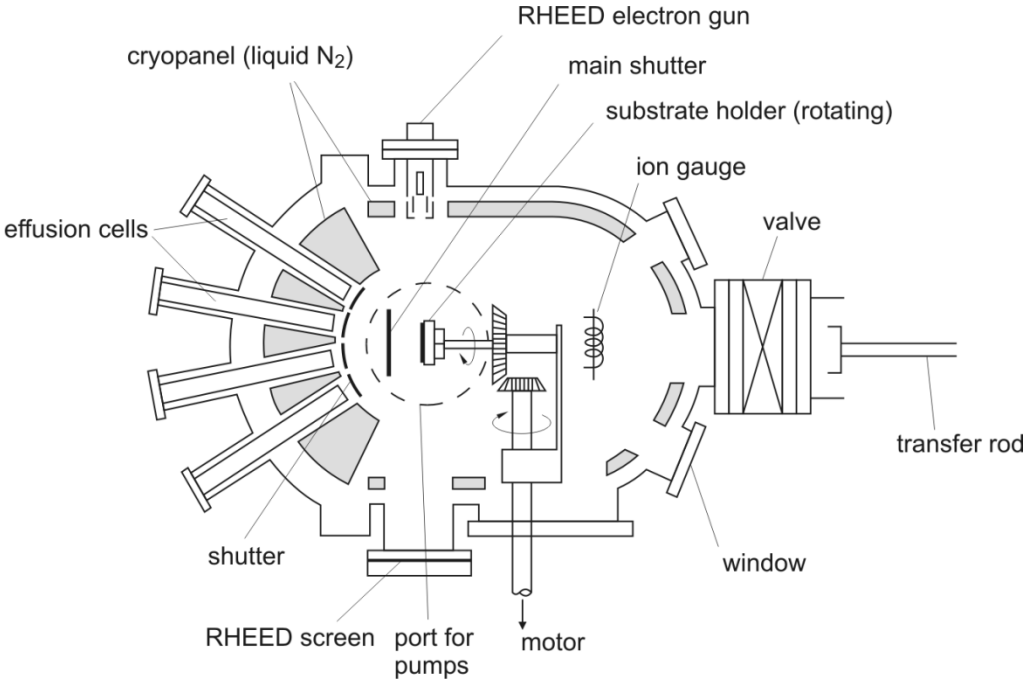


Figure 2.7: Simplified operation of the MBE

The molecular beams are created, using thermal evaporation, in the effusion cells. In these cells, for the process, is required a plasma source that is inside the Knudsen´s cells, that have a pressure of 10-6 Torr. In the former figure is possible to see the RHEED (Reflection High Energy Electron Diffraction) gun and screen. This measurement characterize the surface of the material.

The weakness of MBE, due to the low manufacturing temperature, is that it is slower than other techniques, such as MOCVD or HPVE (only 0.5-1 umh-1) [1], but it allows to grow thicker layer of material.

Finally it´s important to mention that it´s necessary to grow a buffer layer over the substrate, tipically Al, with other techniques, such as MOVPE, because only MBE, to grow GaN, determines low performance[1].

A further technique is the HVPE (Hydride Vapour Phase Epitaxy). It allows to obtain relatively high performance while maintaining low costs; in particular it can grow thick layers in short times (0.3-0.5 um min-1) [1].

It needs a furnace with three different zones: source zone, preparation zone and deposition zone. To obtain GaN it exploits a reaction between GaCl e NH_3 as well as shown in next chemical relationship

$$GaCl(g) + NH_3(g) \rightarrow GaN(s) + HCl(g) + H_2(g) \qquad (2.6)$$

This process is used only for homoepitaxy, because heteroepitaxy creates a lot of defects.

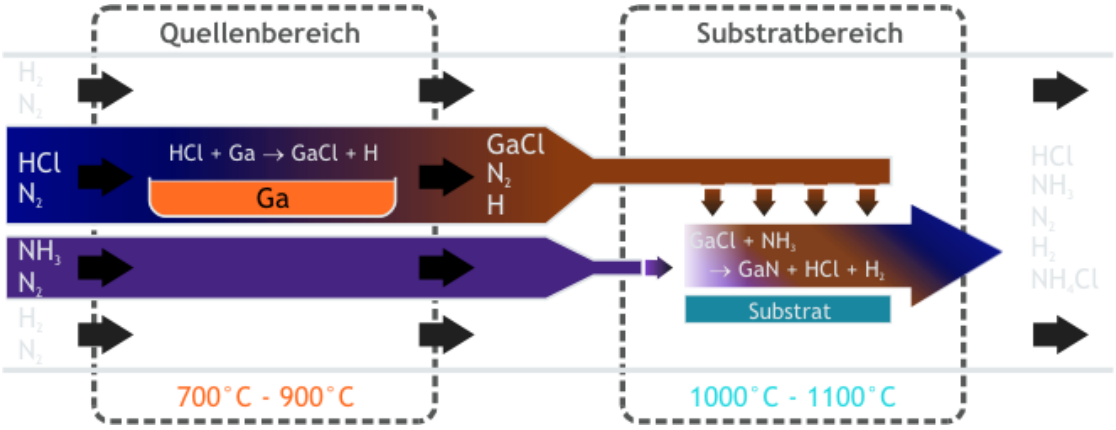


Figure 2.8: Simplified operation of the HVPE

Finally, another technique is the ELO (Epitaxial Lateral Overgrowth) that reduces the density of dislocations in the material. Thanks to this peculiarity it is usually used for optoelectronic devices, particularly for LASER (life time about 10000 hours [5]).

2.1.1.7 Doping

Although doping does not seem so important for GaN-based HEMT, thanks to the efficiency of heterostructure, for some improvements it can give a significant contribution. It can be very important to create p-n junctions, hence to support avalanche phenomena, and it could also facilitate the creation of normally-off transistors. With a good development, could be possible to build NMOS and PMOS as now happen with silicon.

N-type is mainly causes by Nitrogen vacancies inside material, Gallium interstitials or Oxygen incorporation. Hence it is unintentional and it depends on the growth technique. For an intentional doping the element used is the Silicon. In GaN, its actiovation energy is 12-18meV; in AlGaN it strongly depends on Aluminium mole fraction, thus it´s possible to find completely different values. [6]

Silicon doping can serve as additional carrier source, e.g. when it puts on the top of the AlGaN layer a Si-doped GaN sheet in such a way to provide more electrons in the 2DEG.

[7], [8] P-Doping is used when high buffer resistivity wants to be achieved, for example when it wants to have both a complete channel pinch-off and an efficient off-state high voltage blocking. Indeed for empty the channel a sort of compensation will help and at the same time it will ensure a more difficult path for the electrons that are moving from the source to the drain. This latter behaviour is called punch-through and it is one of the factors that limit the break-down voltage in the HEMT devices. The reason for the compensation is that the P-doping captures some electrons, preventing their movement and therefore the participation at the current flow. The most used element is the Magnesium that has an activation energy of 160-200meV [9].

2.1.2 AlN

[1] One of the most important binary materials in the III-N family, after GaN, is the AlN; usually it can be found as its ternary compuond AlGaN, e.g. in barriers heterostructures. Its wurtzite phase, the most common, has a direct energy band gap of 6.2eV, such as an insulator, giving it potential application for deep ultraviolet optoelectronics.

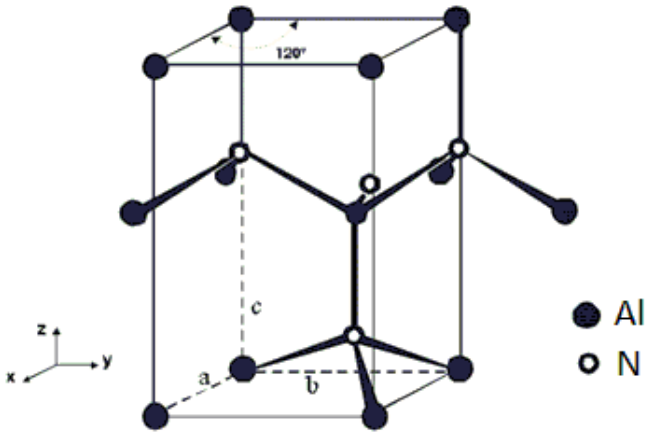


Figure 2.9: Lattice structure of AlN

[1] The high band gap of this material allows the band gap of AlGaN to be modified in a broad band from the value of GaN to that of AlN. With InN, instead of GaN, the band gap range of InAlN material can be even wider.

The electron transport in wurtzite AlN has been investigated by Monte Carlo (MC) simulations. The electronic characteristics are the large band gap and the relatively high effective electron mass of 0.48. These values lead a low field mobility of 135cm^2/Vs at room temperature, a very high critical field of 450kV cm-1 and a saturation velocity of 1.4x$10^7$$cm^{-3}$.

AlN is used as nucleation layer to start the growth on SiC or on Sapphire substrates. The thermal expansion of AlN is similar to GaN, moreover the intrinsic thermal conductivity is very high, even better than that of the other semiconductors, apart from the one of BN, SiC and diamond.

[7] AlN nucleation layer has another function, it limits the breakdown voltage (BV) of the GaN buffer layer. It has been shown that BVs of the fabricated devices strongly depended on the thickness of the AlN nucleation layer. These characteristics make AlN a potentially attractive substrate material.

[10] AlN is also used as a spacer at the interface in AlGaN/GaN HEMT; a thin layer, 2-3nm, can limit the influence of alloy scattering at the interface and, moreover, it increases the band discontinuity improving the confinement of electrons.

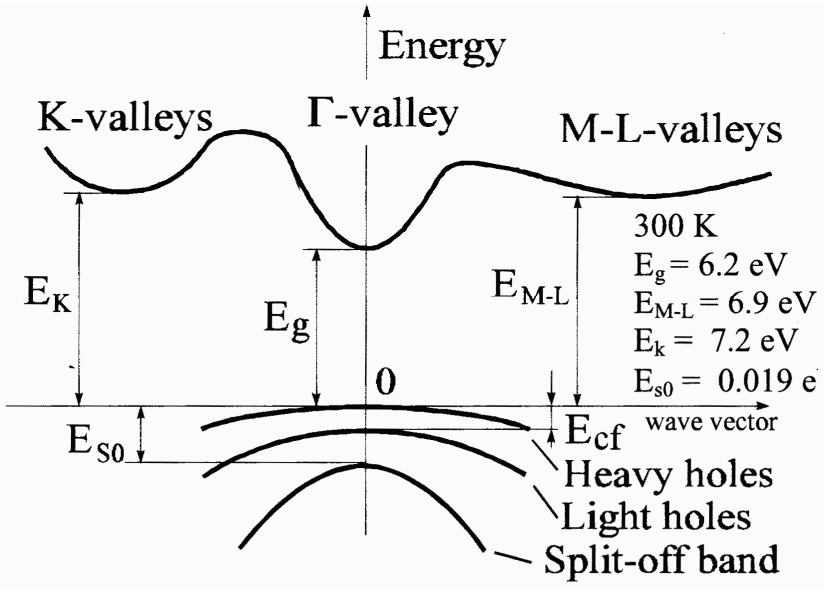


Figure 2.10: Energy bands of AlN

2.1.3 InN

[1] InN and its compounds $In_xAl_{1-x}N$ and $In_xGa_{1-x}N$ are not yet widely used in electronic devices. The indium content is low to achieve the lattice matching to GaN buffer layer. The MOCVD growth technique of InN is complicated caused by the high temperature required and MBE method causes high amount of defects in the structure, due to the nitrogen vacancies. The MBE growth is under development and allows improved material quality and thus the use of the full range of material composition in the material $In_xGa_{1-x}N$. High quality AlN has been grown by MBE and a bulk electron mobility of 3.570cm^2/Vs at 300K is obtained. The thermal expansion coefficient and the lattice constant suggest the growth on sapphire substrate.

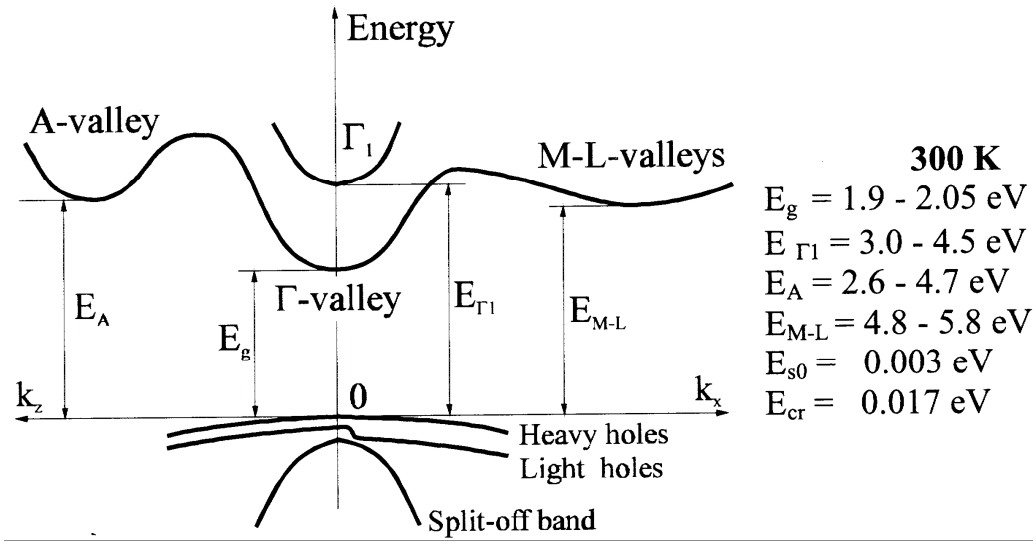


Figure 2.11: Energy bands of InN

Due to the new research on samples with improved material quality, the band gap and the optical functions of InN are reconsidered. This fact had a dramatic impact on the calculations of the transport properties. Further MC calculations on the wurtzite material give a carrier velocity of up to 4.2x10^7 cm/s at a critical field of 52-65KV cm^{-1}. These properties are promising, however, compared to GaAs or InGaAs materials, they are

not really surprising when considering the low effective mass, the low band gap and maximum electron velocity. InN has a very low band gap energy of 0.77eV, different respect the level found in the beginning, 1.89eV . This discrepancy is explained by the existence of oxy-nitrides witch have a larger band gap. Optically it means that a very broad range of wave length is available in the III-N compound materials, ranging from deep ultraviolet to red region.

2.1.4 Material Limitations

[1] One of the greatest problems that affects the device performances is the presence of electron trapping centers inside the materials, that leads to several limitations.

In the Table 2.6 are shown some important parameters for the most common compounds.

Materials	**Breakdown field [MV/cm]**	**Energy gap [eV]**	**Dielectric constant**
Si	0.3	1.12	11.9
GaAs	0.4	1.43	12.5
InP	0.45	1.34	12.4
GaN (Wz)	3.3	3.44	9.5
AlN (Wz)	8.4	6.2	8.5
InN (Wz)	1.2	0.7	15.3
6H - SiC	3.8	2.86	10
Diamond	5	5.6	5.5

Table 2.6: Breakdown field, Energy gap and Dielectric constant for several compounds

It is possible to note that the critical fields increase with the increasing of band gap. The WBG materials have breakdown fields higher than those of conventional semiconductors.

To conclude the electrical limits it is mandatory to mention the impact ionization and the hot electrons, that will be discussed later.

In addition to their electrical limits III-N semiconductors devices are subject to string thermal self-heating. In Table 2.7 there are the thermal properties of binary materials.

Materials	**K_{300} ($Wm^{-1}K^{-1}$)**	**α**	**C_{300} ($JK^{-1}kg^{-1}$)**
Si	148	-1.35	711
GaAs	54	-1.25	322
InP	68	-1.4	410
GaN (Wz)	130	-0.43	491

AlN (Wz)	285	-1.57	748
InN (Wz)	45		325
6H - SiC	390	-1.5	715
Sapphire	42		750

Table 2.7: Thermal properties for several compounds

The heat capacity of the binary semiconductors, given for constant pressure, is lower than of the substrate material. These parameters must be taken in account during growth of heterostructures.

2.2 Ternary III-N materials

[1] The existence of ternary, and even quaternary materials, is very important to other WBG. The possibility of growing AlGaN, InGaN and InAlN in heterostructures with the III-N binaries allows significant improvements. This has tremendous impact on the electronic and optoelectronic application of the materials.

The materials parameters of the quantity *P* are combined by quadratic interpolation in the following two approches:

$$P_{A_xB_{1-x}N} = P_A \cdot x + P_B \cdot (1-x) + C_{P,AB} \cdot x \cdot (1-x) \qquad (2.7)$$

In the second approach, the former formula can be written as:

$$P_{A_xB_{1-x}N} = a + b \cdot x + c \cdot x^2 \qquad (2.8)$$

That can be directly correlated to the binary materials. Sometimes this is extended to a third-order polynomial.

All the relevant property will be discussed in the following sections using the aforementioned formulae.

2.2.1 Aluminium Gallium Nitride ($Al_xGa_{1-x}N$)

[1] $Al_xGa_{1-x}N$ is the most important ternary compound, as the lattice mismatch relative to GaN can be controlled for nearly all material compositions. The material properties of AlGaN ($P_{AxB1-xN}$) can be derived starting from those of GaN (P_{BN}) and AlN (P_{AN}) following the Vegard's law, shown below:

$$P_{A_xB_{1-x}N} = P_{AN} \cdot x + P_{BN} \cdot (1-x) \qquad (2.9)$$

By using this, it can obtain information about energy gap, dielectric constant, lattice constant and other important electronics values. Varying the aluminium mole fraction (*x*) the band gap and other physical parameters can be tailored as needed from x=0 of GaN to x=1 of AlN, allowing for example the realization of a more efficient barriers for power electronic and optoelectronic applications.

2.2.2 Indium Gallium Nitride ($In_xGa_{1-x}N$) and Indium Aluminium Nitride ($In_xAl_{1-x}N$)

[1] The importance of InN and its ternary compounds is due to the smaller band gap relative to GaN, allowing for a broader variety of layers for band gap engineering also into the visible optical range in optoelectronic devices. High quality $In_xGa_{1-x}N$ layers were recently grown by MBE, mostly on sapphire substrate. $In_xAl_{1-x}N$ is lattice-matched to GaN for x = 0.17, which has recently grown attention to this material for HEMT device applications. Good quality films, especially for high In contents, have not been realized with MOCVD growth. MBE has instead allowed this goal.

3 High Electron Mobility Transistor

The High Electron Mobility Transistors are field effect transistors characterized by a great value of mobility, which make them very suitable for high power and high frequency devices. Nowadays, compared to other technologies, e.g. MOSFET and MESFET, HEMTs have the highest cut-off frequency and for a lot of applications they are considered the best solution.

The first transistors tested in the 80´s were made by AlGaAs/GaAs heterostructures and they showed incredible results for electrons transport, just think about the room temperature mobility of 8000 cm^2/Vs, till the value of 100000 cm^2/Vs at liquid nitrogen temperature [11]. In those years HEMTs were used for low-noise amplifiers in radio telescopes. At the end of 80´s the cut-off frequencies were 150-200 GHz, really unreachable for other transistors.

The decreasing of gate length and the reliability issues have slowed the development of these devices; nevertheless, the hopeful properties have pushed the productors to invest money in the research.

In the last years, the increasing request of devices that can sustain higher powers and higher frequencies, has validated the choice to develop these transistors.

3.1 Heterostructure

In Figure 3.1 is proposed the heterostructure of an AlGaN/GaN HEMT.

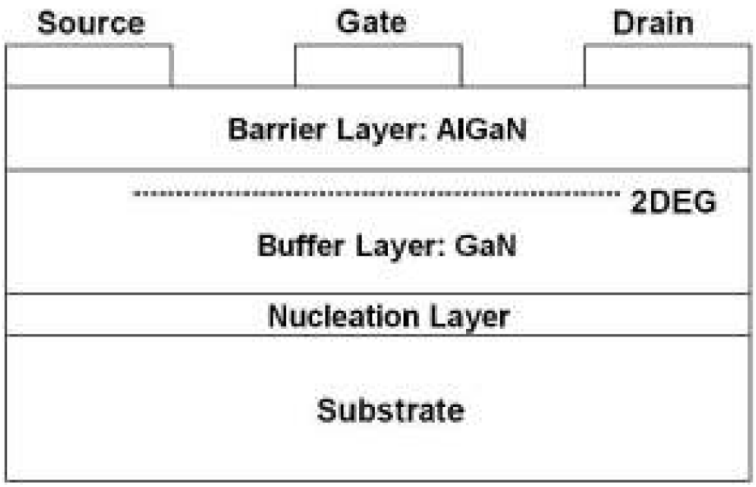


Figure 3.1:AlGaN/GaN heterostructure

The substrate is made of silicon, as all the devices simulated for this thesis. This is important for power applications and, as mentioned before, to limit the production costs.

Onto this, a nucleation layer is grown, tipically AlN, that allows to reduce the lattice and thermal mismatches. This is very important to reduce compressive or tensile strain and defects or dislocations.

Above the nucleation layer is built a layer of GaN, called buffer layer; tipically its thickness depends on the maximum applicable voltage at the transistor. Higher is the voltage, higher will be the thickness: it can range between 1 µm and 4 or 5 µm. The GaN is undoped to avoid further scattering mechanisms that could limit mobility.

Over all there is an AlGaN layer deposited on buffer, called barrier layer. It must have a minumun thickness for the formation of 2DEG and for power applications it is generally smaller than 30 nm. In this layer it is possible to introduce dopant donors to increase the electrons in the heterojunction.

At the interface AlGaN/GaN there is the channel, called 2DEG, that will deepened later.

The contacts are ohmics for source and drain, instead, for gate, there is a schottky contact (or an isolated gate where a dielectric isolation is interposed between the AlGaN barrier and the metal electrode); even this topic will be discussed later.

HEMTs exploit the confinement of electrons in a triangular quantum well, as shown in Figure 3.2. For this reason it is necessary to have two materials with a great difference of energy gap, but at the same time similar lattice constant and same crystallographic orientation.

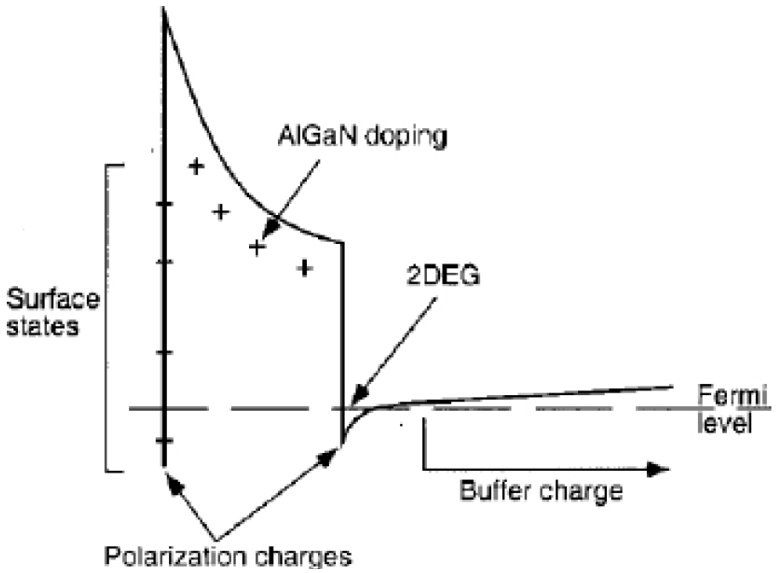


Figure 3.2: Conduction band of AlGaN/GaN HEMT

The discontinuity, between GaN and AlGaN, can confine the electrons; they can move only in y and z direction, not in x direction. Their bidimensional properties create a bidimensional gas, defined 2DEG.

Finally it is important to remember that AlGaN/GaN heterojunctions have replaced AlGaAs/GaAs heterojunctions thank to the properties shown below, such as higher energy band gap and better Johnson´s figure of merit, and also because GaN allows higher concentration of electrons. However, there is another important characteristic, which AlGaAs/GaAs transistors have not: the piezoelectric polarization. This polarization has a great influence in the property of the compound and is also important for the generation of the bidimensional gas. In next paragraph it will be deepened.

3.2 Polarization and charge density

As previously introduced, GaN and AlN have a strong spontaneous polarization (P_{SP}). This is due to the high value of electronegativity of nitrogen atoms and the low values for Ga and Al; this difference creates an intrinsic dipole, as shown below.

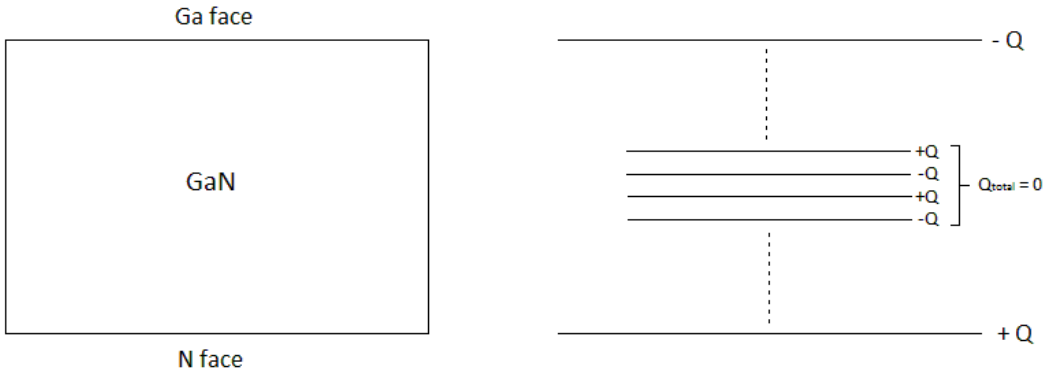


Figure 3.3: Polarization charges in GaN

Gallium nitride wurtzite structure, due to this internal dipole, presents two kinds of surface: Ga-face and N-face. By depending of the orientation of growth it can be shown one of them, as in Figure 3.4:

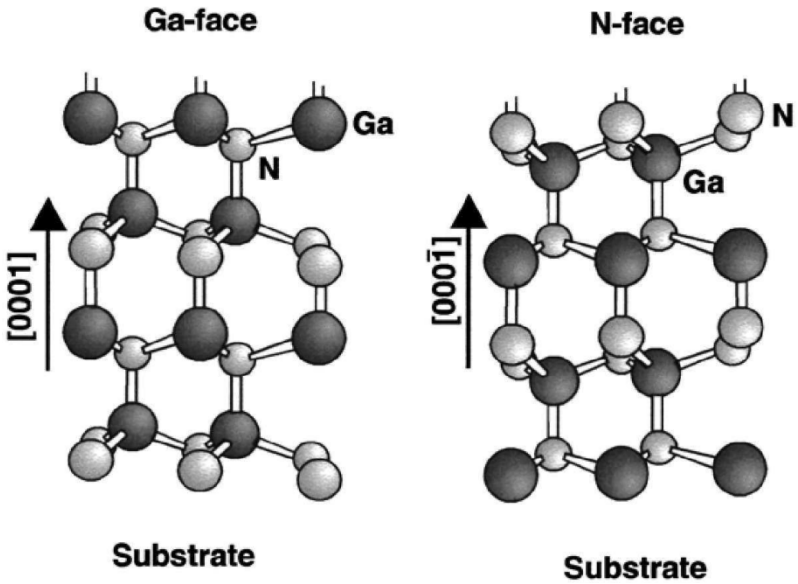


Figure 3.4: stick-ball representation of wurtzite GaN: Ga-face (left) and N-face (right)

The exposed surface is also due to the growth technique: MOCVD usually produces Ga-face structures, instead MBE produces N-face structures.

A piezoelectric polarization (P_{PZ}) appears on GaN, AlN or AlGaN when the material is strained, e.g. when AlGaN is grown onto GaN.

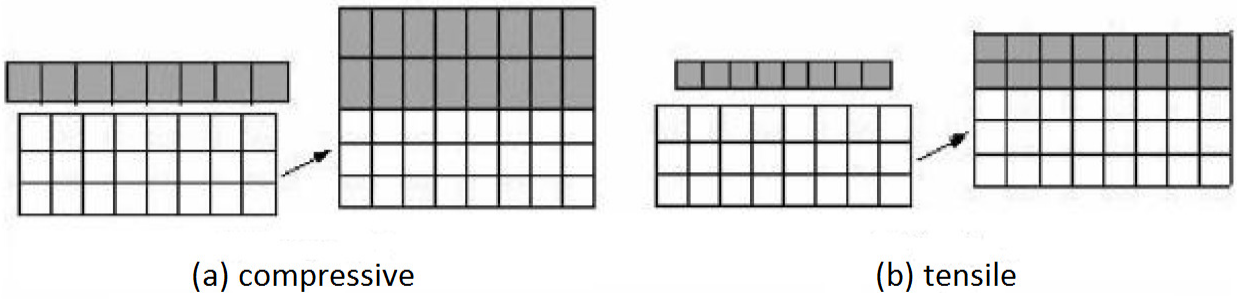


Figure 3.5: Strain in AlGaN/GaN HEMT: compressive (left) and tensile (right)

By using the work of Ambacher et al. [13] is possible to calculate the total polarization.

The spontaneous polarization along the c-axis of the wurtzite crystal is $\boldsymbol{P_{SP}} = P_{SP}\,\boldsymbol{z}$. The piezoelectric polarization can be calculated with piezoelectric coefficient e_{33} and e_{13} as

$$P_{PZ} = e_{33}\epsilon_z + e_{13}\left(\epsilon_x + \epsilon_y\right) \qquad (3.1)$$

where ϵ_z is the strain along the c-axis, ϵ_x and ϵ_y indicate the in-plane strain:

$$\epsilon_z = \frac{c - c_0}{c_0} \qquad \epsilon_x = \epsilon_y = \frac{a - a_0}{a_0} \qquad (3.2)$$

The relation between the lattice constants of the hexagonal *GaN* is given to

$$\frac{c - c_0}{c_0} = -2\frac{C_{13}}{C_{33}}\frac{a - a_0}{a_0} \qquad (3.3)$$

where C_{13} and C_{33} are elastic constants.

By using the former expressions the amount of the piezoelectric polarization in the direction of the c-axis can be determined by

$$P_{PZ} = 2\frac{a - a_0}{a_0}\left(e_{31} - e_{33}\frac{C_{13}}{C_{33}}\right) \qquad (3.4)$$

All of the previous parameters are collected below.

Wurtzite	**AlN**	**GaN**	**InN**
P_{SP} [C/m^2]	1.46	0.73	0.97
e_{33} [C/m^2]	1.55	1	
e_{31} [C/m^2]	-0.6	-0.49	-0.57
e_{15} [C/m^2]	-0.48	-0.3	
ε_{11}	9	9.5	
ε_{33}	10.7	10.4	

Table 3.1: Spontaneous polarization, piezoelectric and dielectric constants of AlN, GaN and InN

Since $[e_{31} - e_{33}(C_{13}/C_{33})] < 0$ for AlGaN over the whole range of compositions, the piezoelectric polarization is negative for tensile and positive for compressive strained barriers, respectively. The spontaneous polarization for GaN and AlN was found to be negative, meaning that for Ga(Al)-face heterostructures the spontaneous polarization is pointing towards the substrate, as depicted in Figure 3.6.

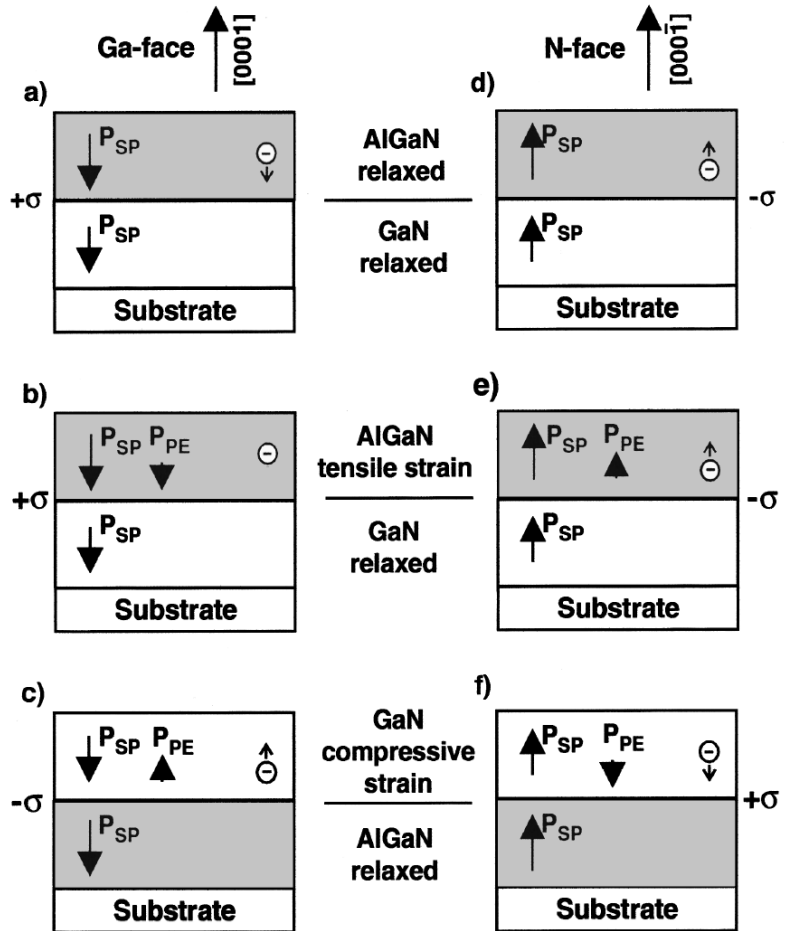


Figure 3.6: Polarization induced, sheet charge density and directions of the polarizations in Ga- and N-face strained and relaxed AlGaN/GaN heterostructures

As a consequence, the alignment of the piezoelectric and spontaneous polarization is parallel in the case of tensile strain, and anti-parallel in the case of compressively strained top layers. If the polarity flips over from Ga-face to N-face material, the piezoelectric, as well as the spontaneous polarization changes its sign.

In Figure 3.7 is shown the polarization dependence on aluminium mole fraction in AlGaN. By increasing the Al-content of the barrier, the piezoelectric and spontaneous polarization of AlGaN will increase.

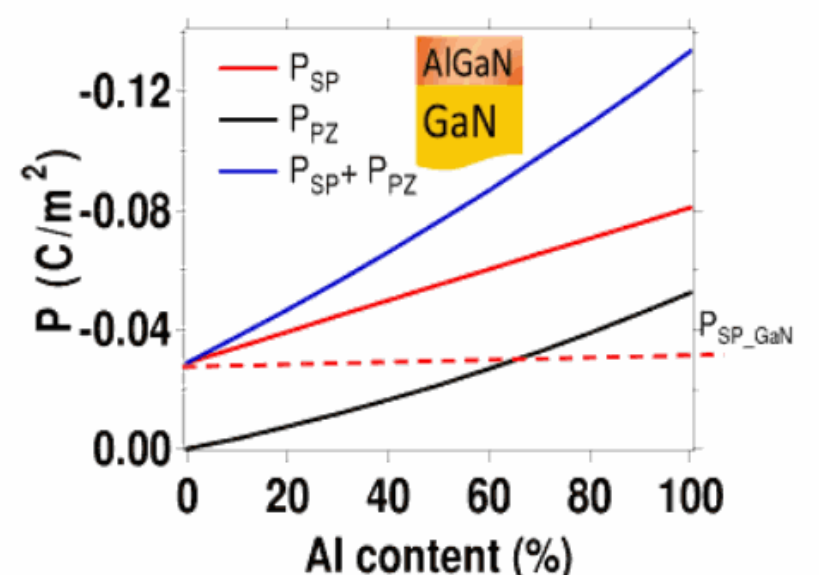


Figure 3.7: Spontaneous (P_{SP}) and piezoelectric (P_{PZ}) polarization in AlGaN grown onto GaN in relation with its Al content

Associated with a gradient of polarization in space is a polarization induced charge density given by $\rho_P = \nabla P$. In analogy, at an abrupt interface of a top/bottom layer (AlGaN/ GaN or GaN/AlGaN) heterostructure the polarization can decrease or increase within a bilayer, causing a polarization sheet charge density defined by

$$\sigma = P(top) - P(bottom) = \{P_{SP}(top) + P_{PZ}(top)\} - \{P_{SP}(bottom) + P_{PZ}(bottom)\}$$

Although, variations in composition, surface roughness or strain distribution, will alter the local distribution of polarization induced sheet charge density. However, the total sheet charge, which is associated with the change of polarization across the interface region will be very nearly equal to that present at an abrupt interface. If the polarization induced sheet charge density is positive (+σ), free electrons will tend to compensate the polarization induced charge. These electrons will form a 2DEG with a certain sheet carrier concentration (n_s), assuming that the AlGaN/GaN band offset is reasonably high and that the interface roughness is low. A negative sheet charge density (-σ) will cause an accumulation of holes at the interface.

In Figure 3.8 it is shown the polarization induced charge at the interface in function of aluminium mole fraction.

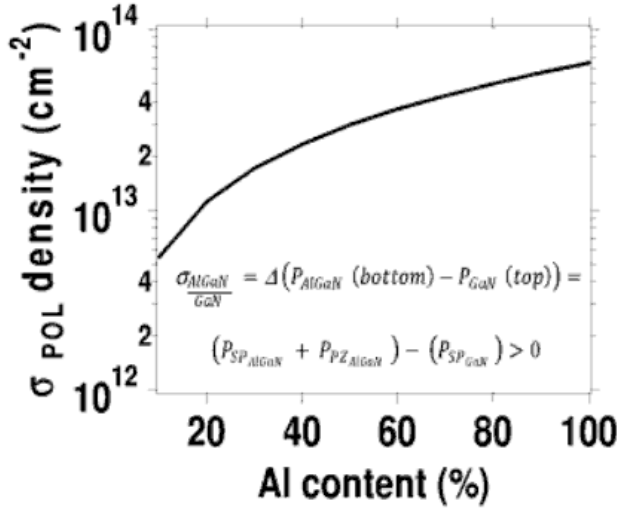


Figure 3.8: Polarization induced charge at the AlGaN/GaN interfaceas function of Al mole fraction in AlGaN

3.3 2DEG

[14] The positive polarization induced charges at the interface of AlGaN/GaN, they attract free electrons, and then it forms, spontaneously, a two-dimensional electron gas (2DEG) between the two materials.

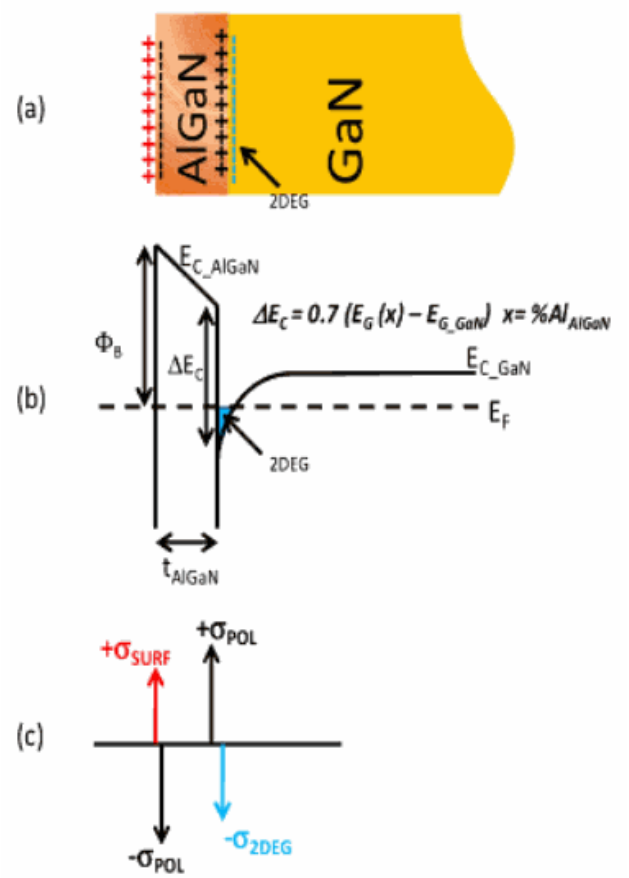


Figure 3.9: Schematic illustration of sign of the charges in the heterostructure (a), band diagram and 2DEG (b) and the magnitude of the charges (c)

In contrast to the polarization charges that are fixed at the interface, the electrons on the 2DEG are free to move. Moreover, electrons are confined in a quantum well, as schematically illustrate in last Figure 3.9(b), and they can move only in two direction, z and y, not in the x direction.

[14] Despite the positive polarization induced charges at the AlGaN/GaN interface might be responsible for the tendency of electrons to be collected at the heterointerface, these polarization charges cannot be the source of electrons; therefore, this has to be found elsewhere.

In an AlGaN/GaN heterostructure is important to explain very well the origin of the 2DEG. First, in absence of an external applied electric field all the charges must be compensated and the total space charge, in the structure, must be zero: in fact it is, by their nature, an electric dipole, as shown above in the former paragraph. Second, the 2DEG cannot be due to electrons generated thermally in the GaN buffer, since the buffer charges must be negative, otherwise the electrons in 2DEG cannot be confined at the interface. Moreover, in a well design epitaxial layer, the buffer charges should be as small as possible, in such a way they can be neglected.

Eventually, by neglecting the buffer charges and by cancelling the contribution of spontaneous and piezoelectric polarization induced charges (since they are dipoles) is possible to obtain the former charge balance equation that guarantees the neutral charge of the structure:

$$\sigma_{SURF} + \sigma_{2DEG} = 0 \qquad (3.5)$$

Where it is assumed, that an undoped AlGaN barrier layer is used; otherwise a contribution from the ionized AlGaN doped charges must be taken in account.

In the last formula, the total carriers present in the 2DEG must be equal to the number of positive charges present at the AlGaN surface. Therefore, it was suggested that the source of electrons can be found on ionized donor-like states present at the AlGaN surface, consequently, any 2DEG electrons are due to donor-like surface states. This is defined the surface donor model.

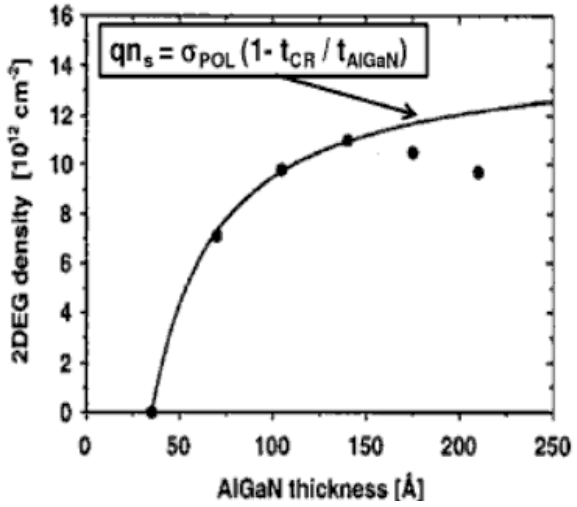


Figure 3.10: 2DEG density as a function of AlGaN thickness (at room temperature and Al = 34%)

To have the formation of the 2DEG is necessary to grow an AlGaN barrier beyond a critical thickness, as shown after.

Assuming the presence of donor-like surface state (i.e. neutral when occupied and positive when empty) at the energy E_D below the conduction band (Figure 3.11(a)), the surface charge and, consequently, the 2DEG charges, depends on the occupancy of this state and thus on its energy relative to the Fermi level (E_F). If the state is sufficiently deep, it lies below the E_F, and there is not 2DEG since $\sigma_{SURF} + \sigma_{2DEG} = 0$ (Figure 3.11(a)). At the critical thickness the donor energy reaches the Fermi level, therefore, electrons are then transferred from occupied surface states to empty conduction band states at the interface, creating the 2DEG, thus leaving behind positive surface charges.

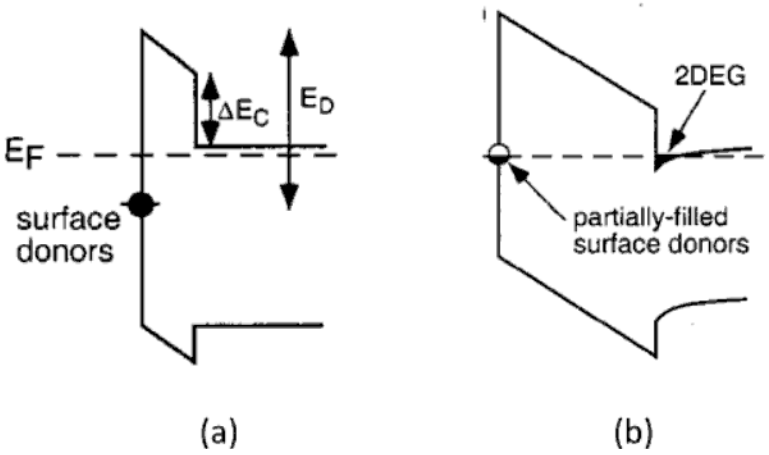


Figure 3.11: Simplified evolution of band diagram when the AlGaN thickness is increased

Furthermore, is also possible to obtain the critical barrier thickness at which the 2DEG starts to form [14], by analytical calculation as shown below

$$t_{cr} = \varepsilon \frac{(E_D - \Delta E_C)}{q\sigma_{POL}} \qquad (3.6)$$

Where ε is the AlGaN relative dielectric constant, σ_{POL} is the polarization induced charge at the interface, ΔEc is the AlGaN/GaN band discontinuity and E_D is the energy level of the surface donors.

Then, from last equation, is possible to deduce that the larger the polarization induced charge density, the smaller is the critical barrier thickness for the 2DEG formation, as shown in Figure 3.12.

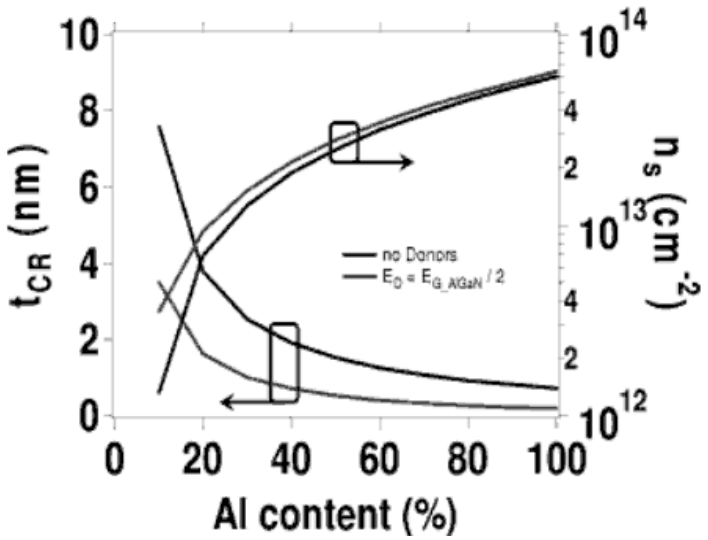


Figure 3.12: Minimum AlGaN thickness to form the 2DEG in function of Al content (left scale) and the resulting 2DEG density for a 10nm AlGaN layer (right scale)

Finally, it is also possible to calculate the 2DEG carriers density based on the surface donor model as follows:

$$qn_s = \sigma_{POL}\left(1 - \frac{t_{CR}}{t_{AlGaN}}\right) \; for\; t_{AlGaN} > t_{CR} \qquad (3.7)$$

Where t_{AlGaN} is the thickness.

From Figure 3.10 is possible to note, that this model correctly predicts that the 2DEG density rapidly increases once the critical barrier thickness is exceeded and then gradually saturates at σ_{POL} /q for t>>t_{cr}.

From the whole discussion presented above it is clear that the 2DEG density is strongly influenced by the thickness and the Al content of the AlGaN barrier. Ideally, a higher 2DEG density is obtained by growing a thicker AlGaN barrier with a higher Al content (to the extreme case of pure AlN). This would also result in a larger bandgap discontinuity at the interface, therefore, providing a better carrier confinement.

Unfortunately, another critical thickness exists beyond which the tensile strained AlGaN barrier layer starts to relax, i.e. by cracking; this phenomenon lowers the piezoelctric components and consequently decreases the 2DEG concentration.

Furthermore, this critical thickness is a function of the Al content: the higher the Al mole fraction, the lower the critical thickness (Figure 3.11). This means that a tradeoff between these values has to be found.

Conventionally, the Al content is from 15 to 35% and the AlGaN thickness is from 10 to 30 nm. Nevertheless, this still results in a very large 2DEG density close to 10^{13} cm^2, as calculated before. Such a high carrier density is also routinely measured in a lot of devices, confirming experimentally the predicted values.

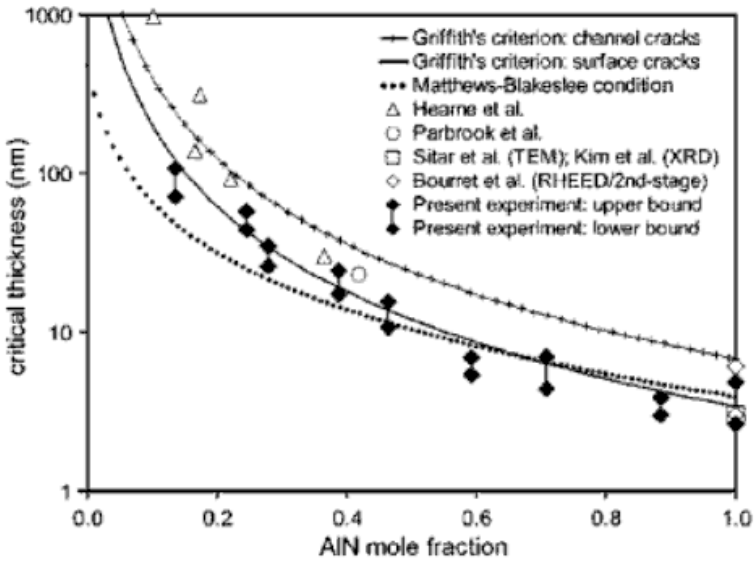


Figure 3.11: Comparison between measured (data points) and calculated (lines) critical thickness

3.4 Contacts

In HEMTs, tipically, there are two kinds of contact:

- *Ohmic contact*

- *Schottky contact*

The first is used for source and drain. It is realized by a high doping of the barrier layer, hence the Schottky barrier can be neglected because the electrons can pass it by tunneling effect[12].

The current-voltage (I-V) characteristic is linear and symmetric. This contact is made by low work function materials, e.g. Ti. Actually, at the state of the art is preferred, due to conductivity reasons, to use Ti-based alloys, such as Ti/Al/Ni/Au, Ti/Al/Ti/Au and Ti/Al/Nb/Au [12].

It is used for its low resistance (< 10^{-6} Ωcm^2) and because it is thermally stable, i.e. does not degrade at elevated temperature or react with oxygen.

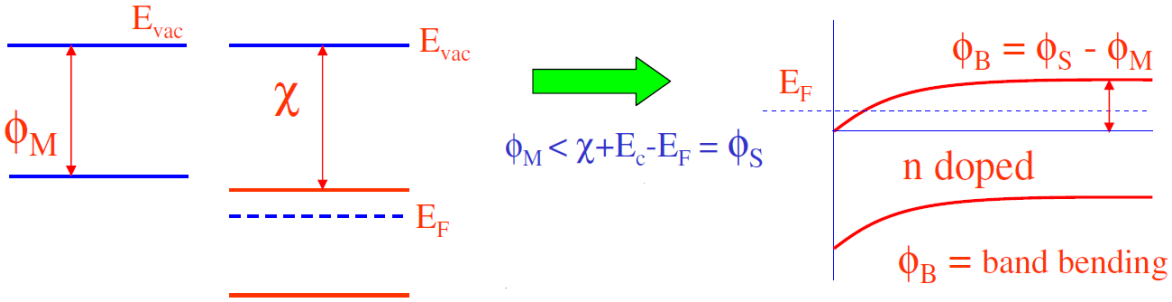


Figure 3.22: Ohmic contact

Where Φ_M is the work function of metal, χ is the electron affinity and Φ_B is the barrier height.

Usually for compound semiconductors it is difficult to realize an ohmic contact by band alignment due to surface states and Fermi pinning.

The Schottky contact is a rectifying contact, made by the deposition of metal onto a semiconductor (barrier layer). This is usually used for gate, because it limits the leakage with high work funciton materials, such as Ni (5,15eV), Pl (5,15eV) and Pd (5.65eV). The contact can be improved by introducing Au or Al to increase the conductivity and by introducing Ti to reduce the thermal mismatch between different materials.

This contact are formed when the doping is not very high (< ~5x10^{18} cm^{-3}) and when the metal work function is greater (lower) than the n-type (p-type) semiconductor work function.

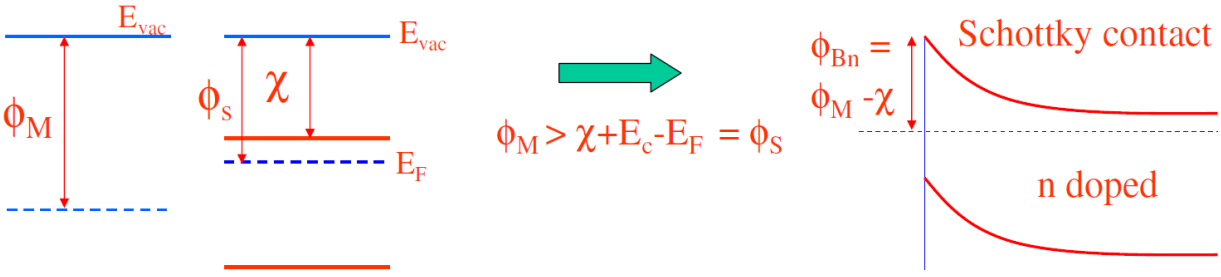


Figure 3.23: Schottky contact

Where Φ_M is the work function of metal, Φ_S is the work function of the semiconductor, χ is the electron affinity and Φ_B is the barrier height.

The conduction mechanisms of this contact are:

- Thermionic emission: electrons can pass over the barrier and the probability of direct tunnelling is very low (valide for low doping: < ~10^{17} cm^{-3})

- Thermionic-field emission: electrons, by using the thermal energy, can pass through the upper thin part of the barrier by tunnelling (valide for medium doping: ~10^{17} – ~10^{18} cm^{-3})
- Field emission: this is a direct tunnelling, the depletion region is very narrow (almost ohmic contact, valide for heavy doping > ~10^{18} cm^{-3})
- Leakage current: there is a high probability of defect-assisted tunnelling and simple conduction (can happend in poor material/interface quality, due to dislocations)

3.5 Degradation phenomena

3.5.1 Traps

[15] The structures built using wide band semiconductors suffer from trapping and detrapping of 2DEG electrons both inside the layer structure and at the semiconductor surface. These trapping effects give rise to the formation of quasi-static charge distributions that cause the current-voltage (I-V) characteristics at microwave frequencies to be considerably lower than under direct-current (DC) conditions. Consequently the microwave output power capability of the devices is significantly lower than expected from the DC output (I-V) characteristics and the chosen operation class.

Whereas the occurrence of traps heavily affects the device performances, it's very important understanding the traps spatial and energy locations. The presence of traps is partially due to the heteroepitaxy growth technique and for this reason traps could be generated both within the layer and in the heterointerface. The former type is due to the not still perfect material quality obtained by the growth process. The latter type instead is due to the difference in the lattice constant of the two semiconductors, meaning that the material grown on top produces defects because it has to adapt its lattice parameter at the one of the lower material (Figure 3.12).

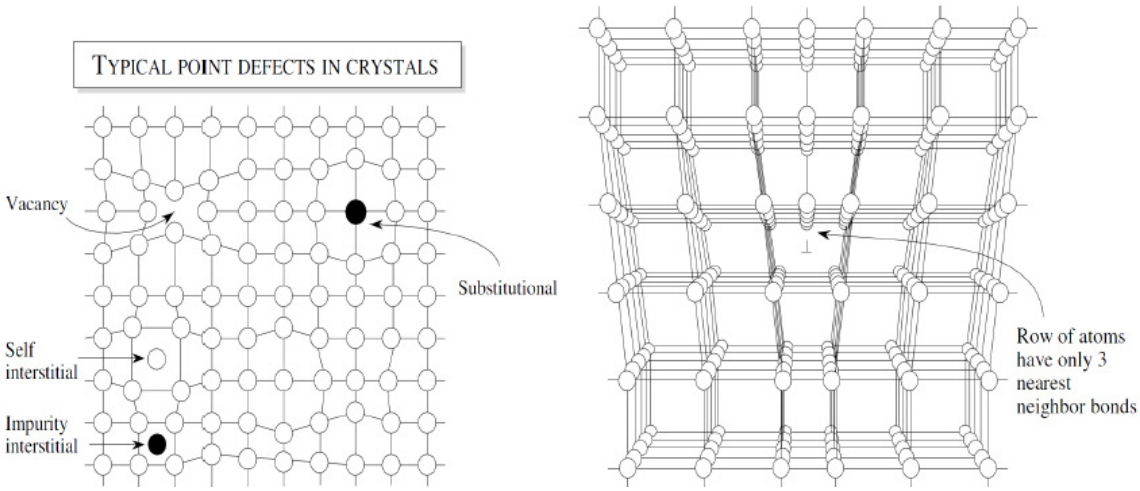


Figure 3.12: Main defects in a crystal (left) and volumetric dislocation (right)

The main consequence of these traps is the current collapse; other effects are the gate-lag and drain-lag, due to a virtual gate close to the real one, that lead to the phenomenon of dispersion frequency.

The traps in a material can be in the surface, *surface traps*, in the barrier layer, *barrier traps*, and in the buffer, *buffer layer.* For power application, the main effect due to them is are a R_{DSon} degradation, a shift of the threshold voltage and the dynamic issues of current collapse.

These traps are shown below.

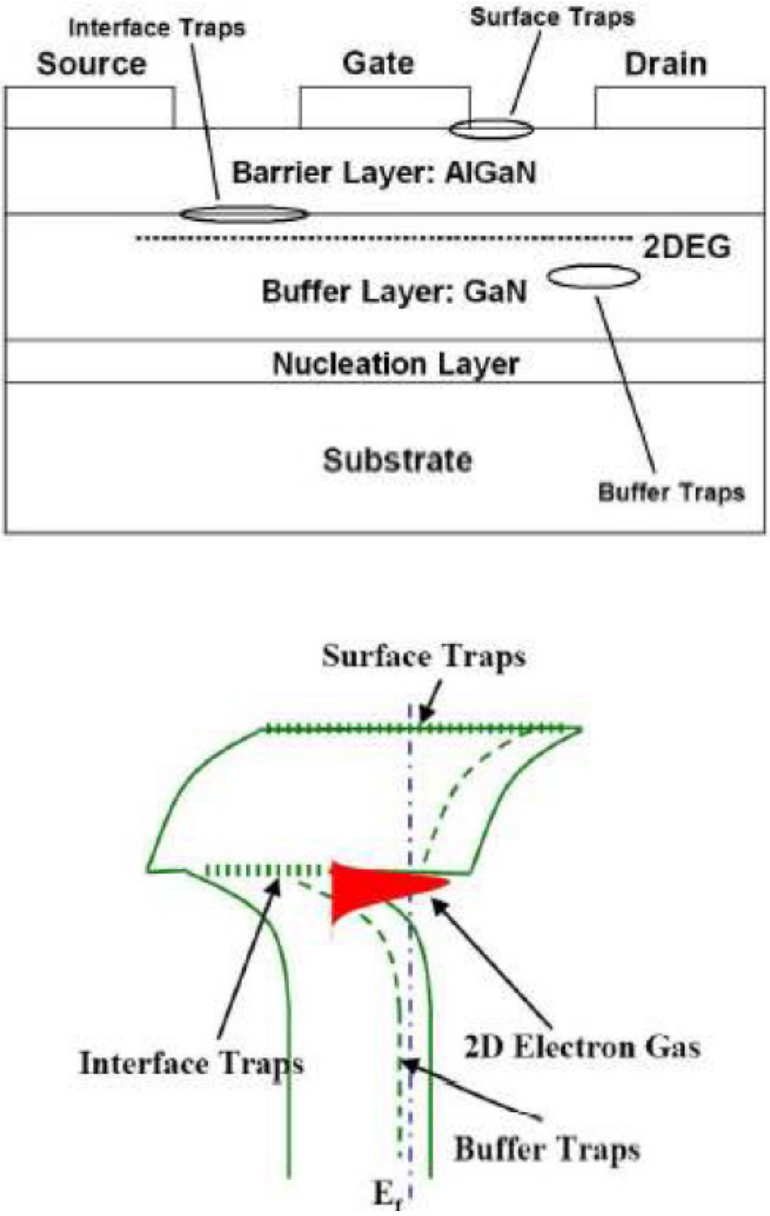


Figure 3.13: Traps in AlGaN/GaN HEMT: in the structure (up) in the band diagram (down)

3.5.1.1 Current collapse

This phenomenon appears when devices are stressed with high current drain-source. In this condition, electrons are captured in buffer layer. Here, the donor-like traps, situated at mid band gap, work as recombination states; as higher are their concentration as lower are the electrons concentration in the channel, hence the drain current decreases.

As shown in Figure 3.14, at low voltages traps have the greatest influence; instead, for high voltages, electrons can be released.

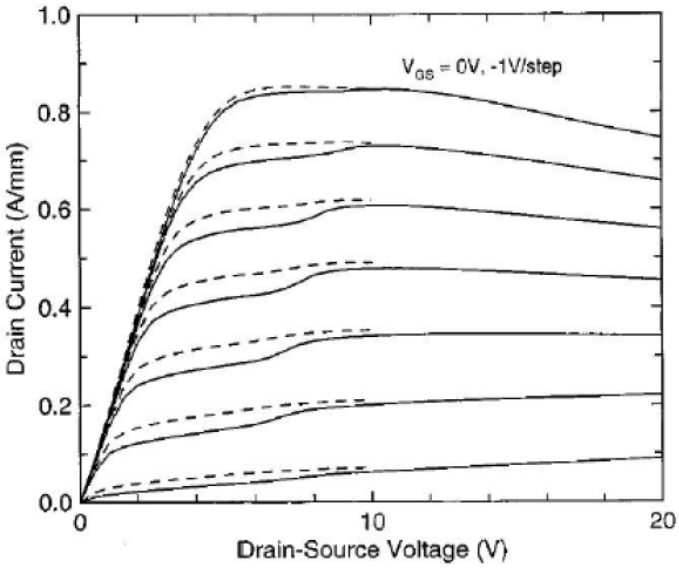


Figure 3.14: Id-Vd characteristic fora n AlGaN/GaN HEMT: bias voltages 20V (solid lines) and 10V (dashed lines)

It is possible to see that this effect is present below 8V and it is greater at high bias voltages.

3.5.1.2 Dispersion frequency

The main effects of this dispersion are a decrease of current and an increase knee voltage. In practice, these lead to a degradation of high frequency performance.

It is mainly due to the traps in the AlGaN barrier and on the surface. In Figure 3.15 is depicted this phenomenon.

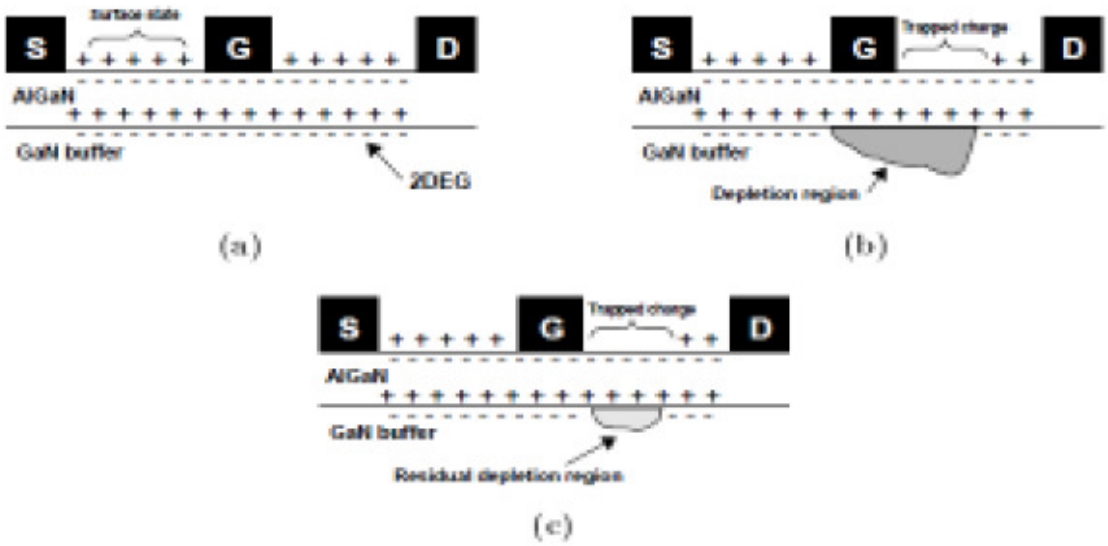


Figure 3.15: Dispersion frequency

electrons and leading to extension of the gate depletion region. Hence, the effect of surface negative charge is to act like a negatively biased metal gate as shown in Figure 3.16.

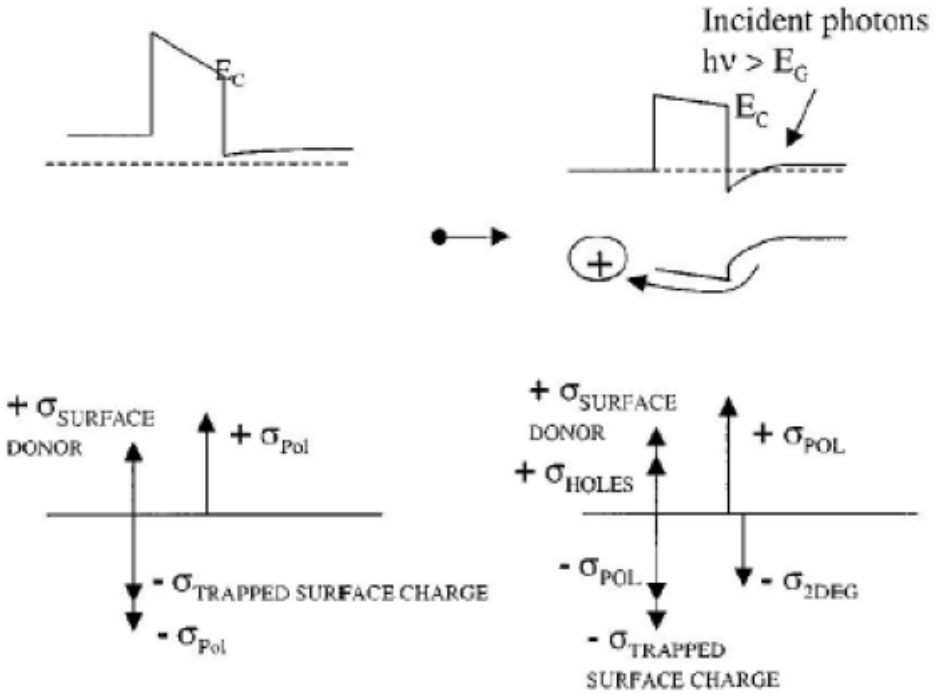


Figure 3.16: Charges in the HEMT and effects due to a interaction with a phonon

Hence, now there are two gates on the surface, between the source and drain, connected in series. The potential on the second gate is controlled by the total amount of trapped charge in the gate drain access region. This second gate is referred to as the virtual gate.

3.5.1.3 Solution to reduce these degradation effects

The most common solutions are:

- *cap layers*

- *field-plate*

The first is used to reduce the dispersion. The cap layers are made in GaN or AlGaN; in this way is possible to increase the distance between surface and channel, therefore, the influence of the traps is reduced. However R_{DSon} increases, lead to a worsening of performances. In Figure 3.17 is depicted an example.

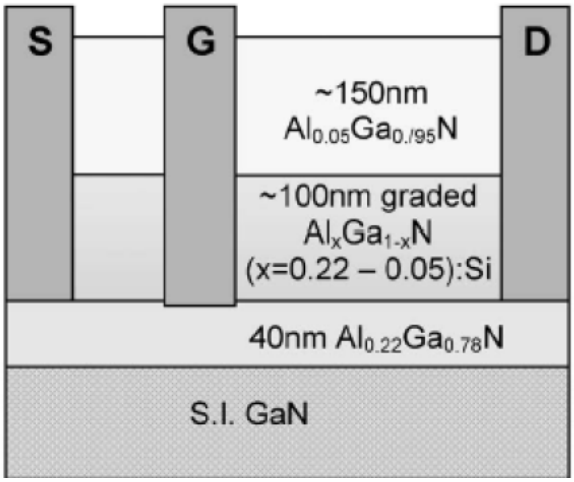


Figure 3.17: Structure with deep recessed gate

The field-plate (FP) helps to reduce the gate leakage and to increase the electric breakdown field. This technique can be different in dependence of where it is connected: gate connected FP and source connected FP. In Figure 3.18 is shown a gate connected FP.

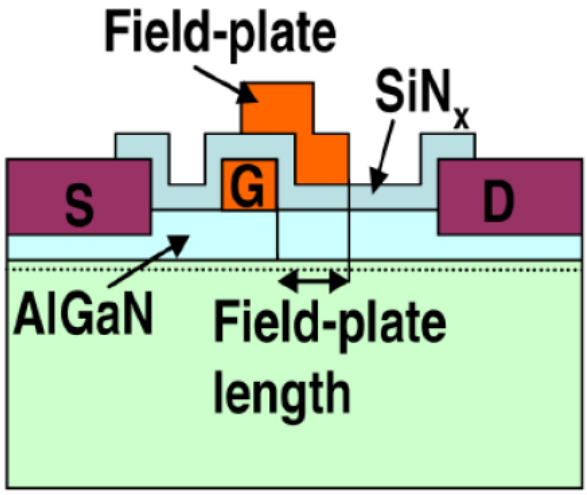


Figure 3.18: Gate connected field-plate

This FP reduces the effects of the traps by changing and reducing the electric field. However, it forms a capacitance, between FP and drain, that decrases the current gain and the the cut-off frequency.

The source connected FP, shown in next figure, can remove the high capacitance between gate and drain; nevertheless it can introduce parasite capacitance at the input of the device. However, this configuration allows to build smaller devices at parity of power.

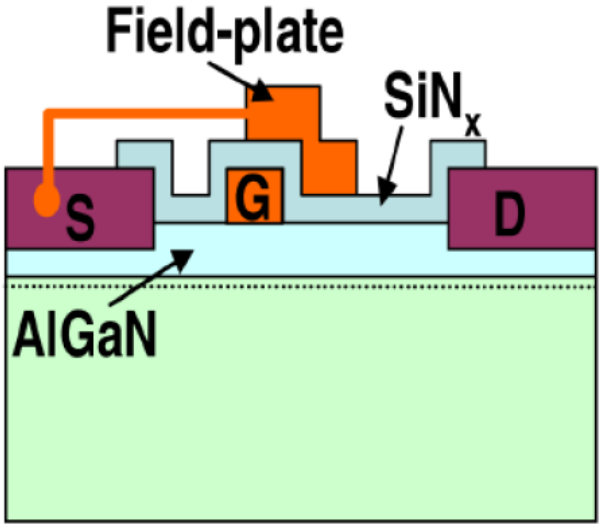


Figure 3.19: Source connected field-plate

Finally, it is important to mention the passivation. This technique aims to limit the effects of the surface traps. As shown in the previous figures, it consists in the deposition of an insulator material onto the surface of the device. It allows to compensate the positive charges and to neutralize the density of charge on the surface, hence it can eliminate the secondary depletion region in the channel.

4 Mobility

In GaN-based HEMTs the mobility is a very important and very critical parameter and an optimal evaluation of it is fundamental; in this chapter, after an introduction about the mobility and its characteristics, it has been implemented a GaN-based model, found in literature, in the Sentaurus simulator. The motivations of this work are that, at the state of the art, in the commercial simulators there are not models for GaN applications, but only models developed and calibrated for Si-based devices.

4.1 Introduction

[51] In the solid-state physics, the electron mobility characterizes the movements of the electrons through a metal or semiconductor, when they are under the effect of an electric field. For the semiconductors exists an analogous quantity for holes, called hole mobility. When is used the term carrier mobility, it is referred to both electron and hole mobilities in materials.

When an electric field E is applied across a piece of material, the electrons respond by moving in a direction with an average velocity called the drift velocity, V_D, and in next paragraph it will be deepened. The common measurement unit for electron mobility is $cm^2/V{\cdot}s$. This is different from the SI unit of mobility, $m^2/V{\cdot}s$.

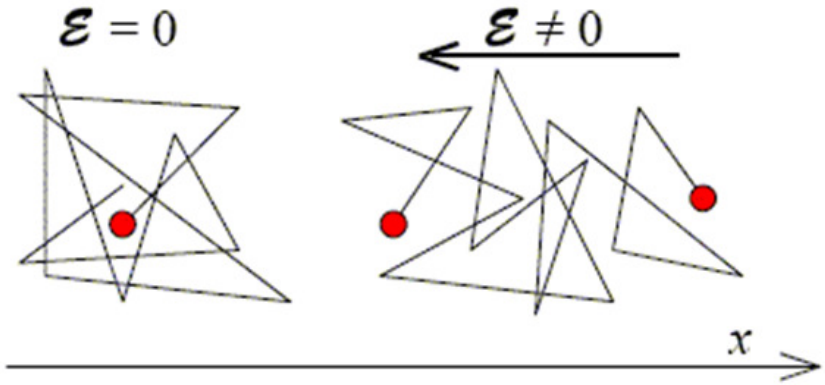


Figure 4.1: Movements of a carrier when pulled by electric field

The conductivity is another property of the materials; it is proportional to the product of mobility and carrier concentration, e.g. the same value of conductivity could come from few electrons with high mobility for each, or a lot of electrons with a low mobility. Anyway is important to note that for semiconductors, the behavior of the devices can be very different depending on which is the case that is present in the material, few electrons or a lot of electrons. Therefore mobility is a very important parameter for semiconductor materials. Almost always, higher mobility leads to better device performance, with other things equal.

It is also important to mention that the semiconductor mobility depends on the impurity inside of them (including donor and acceptor concentrations), defects, temperature and electron and hole concentrations.

It also depends on the electric field, particularly at high fields when velocity saturation occurs. In a next paragraph the mobility dependences will be deepened.

Typical low field electron mobility for Si at room temperature (300 K) is 300 $cm^2/V{\cdot}s$.

The mobility is very important, and very high, for several low-dimensional systems, such as two-dimensional electron gases (2DEG) (3,000,000 $cm^2/V{\cdot}s$ at low temperature) [17], carbon nanotubes (100,000 $cm^2/V{\cdot}s$ at room temperature) [18] and more recently, graphene (200,000 $cm^2/V{\cdot}s$ at low temperature). [19]

4.1.2 Drift velocity

[51] When there is not applied electric field, the electrons (or, in the case of semiconductors, both electrons and holes) move randomly in the material; in fact the carrier have not an overall motion in a particular direction.

However, when there is an electric field, electrons are accelerated by the electric field. If the electrons were in a vacuum, they would be accelerated without velocity limits (called ballistic transport), till the maximum possible velocity. However, in a solid material, the electrons are rapidly conditioned by the scattering mechanisms; they move with a finite average velocity, called the drift velocity.

In a semiconductor electrons and holes typically have different drift velocities for the same electric field.

Quasi-ballistic transport is possible in solids if the electrons are accelerated across a very small distance (as small as the mean free path), or for a very short time (as short as the mean free time). In these cases, drift velocity and mobility are not meaningful.

The definition is:

$$v_D = \mu E$$

E is the magnitude of the electric field applied to a material, v_D is the magnitude of the electron drift velocity (hence the electron drift speed) caused by the electric field and μ is the electron mobility. The measurement unit for drift velocity is m/s and for electric field is V/m. The hole mobility is defined by the same equation. Both electron and hole mobilities are positive by definition.

Usually, this velocity in a material is directly proportional to the electric field, which means that the electron mobility is a constant (independent of electric field). When this is not true (for example, in very large electric fields), the mobility has still a great dependence on the electric field.

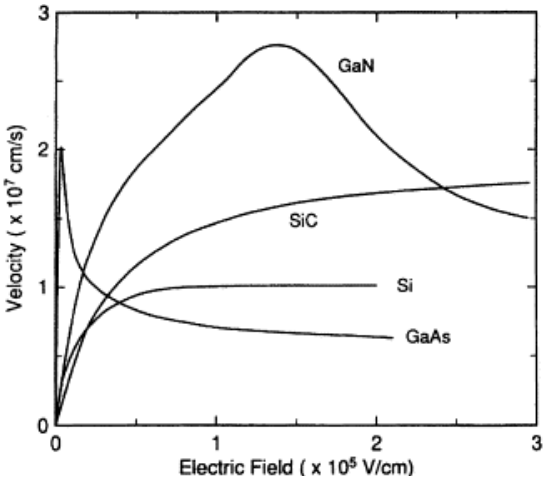


Figure 4.2: Electron drift velocity for several materials

In Figure 4.2 is shown the electron drift velocity for several materials. For GaN (for III-V in general) is important to note the portion in the plot of the velocity versus electric field with negative slope (negative differential velocity); in fact at high electric field the electrons in the conduction band minima have enough energy to jump to the upper valleys, where they have higher effective masses, therefore the drift velocity decreases. In fact another way to define the electron mobility is

$$\mu = \frac{q \cdot \tau}{m_0}$$

Where q is the elementary charge of the carrier, τ is the time between the collision and m_0 is the effective mass. In the next chapter the effective mass will be explained.

4.1.3 Conductivity

[51] Between mobility and electrical conductivity the relation is very simple, by using n as the number density of electrons and μ_e their mobility. [20]

$$\sigma = ne\mu_e \qquad (4.1)$$

Where $ne\mu_e$ is the current density and e is the elementary charge.

This formula is valid when the conductivity is due entirely to electrons. In a p-type semiconductor, the conductivity is due to holes instead, but the formula is the same. If p is the density of holes and μ_h is the hole mobility, then the conductivity is

$$\sigma = pe\mu_h \qquad (4.2)$$

If a semiconductor has both electrons and holes, the total conductivity is[1]

$$\sigma = e(n\mu_e + p\mu_h) \qquad (4.3)$$

4.1.4 Mobility dependence

[51] At low fields, the drift velocity v_D is proportional to the electric field E, so mobility μ is only a function of doping and temperature. This value of μ is called low-field mobility.

As the electric field is increased, however, the carrier velocity increases almost linearly and asymptotically towards the saturation velocity v_{sat}, that is the maximum possible value. The tipical value of v_{sat} is on the order of 1×10^7 cm/s for both electrons and holes in Si. For GaN it can arrive until $2.5\text{x}10^7$. This velocity is a characteristic of the material and it is a strong function of doping, impurity levels and temperature. It is one of the most important device properties that determine the performance, such as a transistor's ultimate limit of speed of response and frequency.

On of the most important mechanism which influences the saturation velocity is the optical phonon scattering. At high electric fields, carriers are accelerated enough to gain sufficient kinetic energy between collisions to emit an optical phonon and they do it very quickly, before being accelerated once again. The velocity that the electron reaches before emitting a phonon is:

$$\frac{m_o v_{emit}^2}{2} \approx \hbar\omega_{phonon\,(opt.)} \qquad (4.4)$$

Where $\omega_{phonon\,(opt.)}$ is the optical phonon angular frequency and m_0 is the carrier effective mass in the direction of the electric field. The saturation velocity is only one-half of v_{emit}, because the electron starts at zero velocity and accelerates up to v_{emit} in each cycle. [19]

Velocity saturation is not the only possible high-field behavior. Another can be the Gunn effect, where a sufficiently high electric field can causes intervalley electron transfer, which reduces drift velocity.

In the regime of velocity saturation (or other high-field effects), mobility is strongly influenced by the electric field. This means that mobility is a somewhat less useful concept, compared to simply discussing drift velocity directly.

[51] With increasing temperature, phonon concentration increases and causes several scattering phenomena, which will be discussed later. This lattice scattering lowers the carrier mobility more and more at higher temperature. By using theoretical calculations is possible to obtain the mobility dependence on temperature. From experiments, the values obtained for the temperature dependence of the mobility for some materials are listed in Figure 4.1. [20]

	Si	**Ge**	**GaAs**
Electrons	$\propto T^{-2.4}$	$\propto T^{-1.7}$	$\propto T^{-1.0}$
Hole	$\propto T^{-2.2}$	$\propto T^{-2.3}$	$\propto T^{-2.1}$

Table 4.1: Electron and hole mobilities dependance on temperature for several materials

[51] Finally, thenumber of charge carriers in semiconductors, i.e. electrons and holes, are controlled by the concentrations of impurity elements, such as doping concentration. This concentration of dopants has a great influence on carrier mobility.

In the next formula is presented the mobility dependence on doping concentration for Silicon. This expression is often characterized by the empirical relationship: [21]

$$\mu = \mu_0 + \frac{\mu_1}{1 + \left(N/N_{ref}\right)^{\alpha}} \qquad (4.5)$$

Where N is the doping concentration (either N_D or N_A) and N_{ref} and α are fitting parameters.

The previous expression is used sometime to obtain prediction of the doping dependence for different materials.

4.1.5 Scattering Mechanisms

The main factor determining drift velocity (different than effective mass, which will be discussed in next chapter) is scattering time, i.e., the time between two collisions of the carriers, when they are accelerated by an electric field, and this is leads to different direction and/or energy. The main causes of scattering in a typical semiconductor are ionized impurity scattering and acoustic phonon scattering (also called lattice scattering). In some cases, e.g. high electric field, heterojunctions with different compounds, etc., other sources of scattering could be important, such as alloy scattering, optical phonon scattering, surface scattering and defect scattering. [22]

[51] An elastic scattering is when the energy is (almost) conserved during the different collisions. Some elastic scattering processes are scattering from acoustic phonons, impurity scattering, piezoelectric scattering, etc. In acoustic phonon scattering, electrons move from state k to k', by emitting or absorbing a phonon. This phenomenon leads to lattice vibrations which cause small shifts in energy bands. The additional potential causing the scattering process is generated by the deviations of bands due to these small transitions from frozen lattice positions. [22]

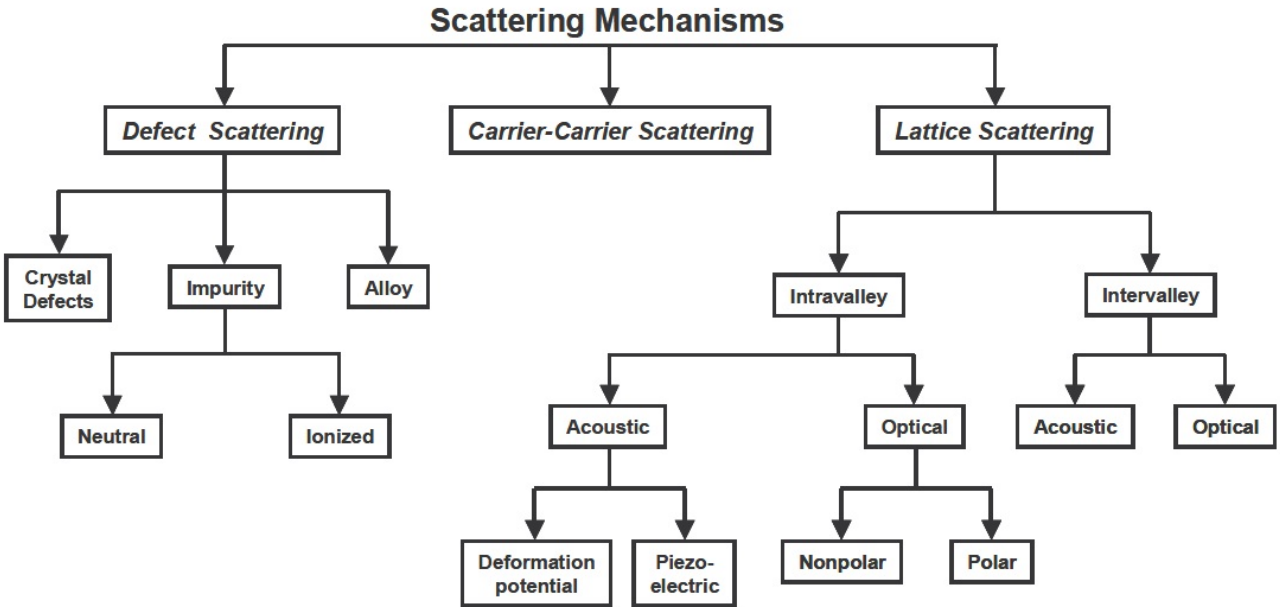


Figure 4.3: Scattering Mechanisms

If a semiconductor is doped with donors and/or acceptors, these are typically ionized, and hence they introduce charges in the material. The Coulombic forces can deflect an electron, or a hole, that is moving toward the ionized impurity. This is known as *ionized impurity scattering*. The amount of deflection depends on the speed of the carrier and its proximity respect the ion. If a semiconductor is heavly doped, the probability, in a given time, that a carrier will collide with an impurity is higher, and it means that the time between collisions is shorter and the mobility is lower. When determining the strength of these interactions

due to the long-range nature of the Coulomb potential, other impurities and free carriers cause the range of interaction with the carriers to reduce significantly compared to bare Coulomb interaction.

At the interface between the semiconductors, due to the crystal defects, disorders, dislocation and dandling bonds occurred during the growth, the problems, due to scattering mechanisms, increase. These defects leads to the formation of charge trapping centers inside the material and in the surfaces . Scattering happens because after trapping a charge, the defect becomes charged and therefore starts interacting with free carriers. If scattered carriers are in the inversion layer at the interface, the reduced dimensionality of the carriers makes the case differ from the case of bulk impurity scattering as carriers move only in two dimensions. Interfacial roughness also causes short-range scattering limiting the mobility of quasi two-dimensional electrons at the interface. [22]

[51] Another important scattering mechanism which limits mobility is the *lattice (phonon) scattering*. At any temperature above absolute zero, the vibrating atoms create pressure (acoustic) waves in the crystal, which are termed phonons. The phonons can be considered to be particles, like electrons. As mentioned before, a phonon can collide with an electron (or hole) and scatter it. At higher temperature, there are more phonons, therefore this scattering mechanism increases and leads to a reduction of the mobility.

This scattering can be: [23]

- *intravalley*: after the collision the electron is still in the same valley of the conduction band; it happens when the phonons have low energy

- *intervalley*: due to the high energy of the phonons, the electron has enough energy to jump to other valleys.

The *deformation potential scattering* is induced by the lattice deformation associated to lattice vibrations. The lattice deformation induces a variation of the band structure and hence of the potential energy felt by electrons. The main effect is the variation of the lattice constants, instead the effective masses are less influenced.

In ionic, or partially ionic, semiconductors, i.e. all compound semiconductors, exist a further scattering mechanism related to the interaction between electrons and lattice vibrations; the deformation perturbs the dipole moment of the atoms, causing a scattering due to polar phonons (acoustics or optics). This is called *piezoelectric scattering*. The mechanism due to acoustic phonons has influence only at low temperature, while the other one is dominant at room temperature.

[51] Another scattering mechanism influenced by the normal electric field, like phonon scattering, is the *surface roughness scattering*. It is caused by defects or dislocations at the interface and it has a great influence to limiting the mobility in 2DEG. From high-resolution transmission electron micrographs, it has been determined that the interface is not abrupt on the atomic level, but after the growth of the layers, is very difficult to obtain a very smooth surface. These defects are random and cause fluctuations of the energy levels at the interface, which then causes scattering. [22]

[51] In compound (alloy) semiconductors, e.g. AlGaN or InGaN, the scattering caused by the perturbation of crystal potential due to the random positioning of substituting atom species in a relevant sublattice is known as *alloy scattering*. This can only happen in ternary or higher alloys, because in their crystal structures some atoms are replaced randomly by new atoms of one of the three elements. Generally, this phenomenon is quite weak but in certain materials or circumstances, it can become dominant effect limiting the mobility of the device. In bulk materials, interface scattering is usually ignored. [22]

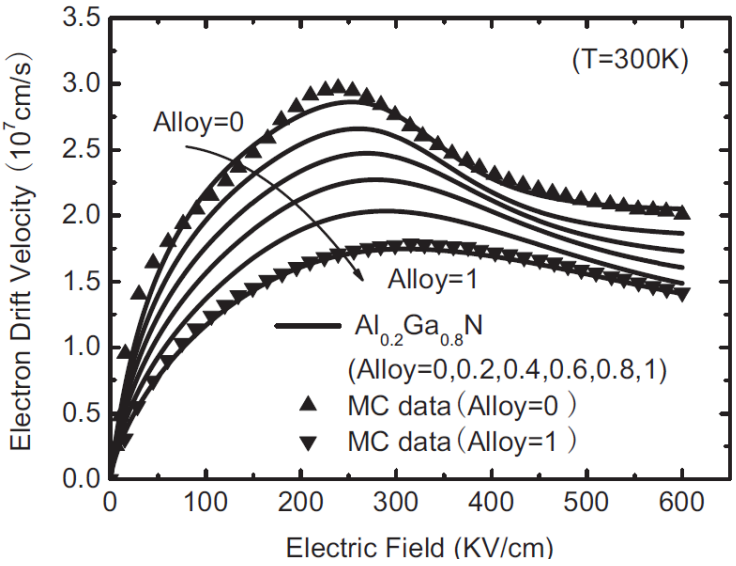


Figure 4.4: Degradation due to alloy scattering. MC data from [Farahmand]

In the former figure is depicted the effects of the degradation due to alloy scattering in a AlGaN/GaN HEMT.

[51] Finally is important to mention the scattering between the carriers, the *electron-electron scattering*. Due to the Pauli exclusion principle, electrons can be considered as non-interacting if their density does not exceed the value 10^{17} cm^{-3} or electric field value 10^3 V/cm. However, significantly above these limits electron-electron scattering starts to dominate. Long range and nonlinearity of the Coulomb potential governing interactions between electrons make these interactions difficult to predict and to evaluate. [22]

[51] All of these models, particularly their scattering time, have a relation with the mobility. It is assumed that after each collision, the carrier moves randomly, so it has zero average velocity. After that, it accelerates uniformly in the electric field, until it scatters again. The resulting average drift mobility is:[25]

$$\mu = \frac{q}{m^*}\bar{\tau} \qquad (4.6)$$

Where q is the elementary charge, m^* is the carrier effective mass and τ is the average scattering time.

If the effective mass is anisotropic (direction-dependent), m* is the effective mass in the direction of the electric field.

Normally, in presence of several scattering mechanisms, a very good approximation to combine their influences is to use the "Matthiessen's Rule" (developed from work by Augustus Matthiessen in 1864):

$$\frac{1}{\mu} = \frac{1}{\mu_{impurities}} + \frac{1}{\mu_{lattice}} + \frac{1}{\mu_{defects}} + \cdots \qquad (4.7)$$

Where μ is the total mobility, $\mu_{impurities}$ is the mobility that the material would have if there was impurity scattering but no other source of scattering, $\mu_{lattice}$ is the mobility that the material would have if there was lattice phonon scattering but no other source of scattering, and so on for others.

Matthiessen's rule can also be stated in terms of the scattering time:

$$\frac{1}{\tau} = \frac{1}{\tau_{impurities}} + \frac{1}{\tau_{lattice}} + \frac{1}{\tau_{defects}} + \cdots \qquad (4.8)$$

where τ is the true average scattering time and $\tau_{impurities}$ is the scattering time if there was impurity scattering but no other source of scattering, etc.

[51] As mentioned above, Matthiessen's rule is an approximation and in some circumstances it is not valid. This rule is not valid if the factors affecting the mobility depend on with each other, because individual scattering probabilities cannot be summed unless they are independent of each other. The average free time of flight of a carrier and therefore the relaxation time is inversely proportional to the scattering probability. [22] For example, lattice scattering alters the average electron velocity (in the electric-field direction), which in turn alters the tendency to scatter off impurities. There are more complicated formulas that attempt to take these effects into account.

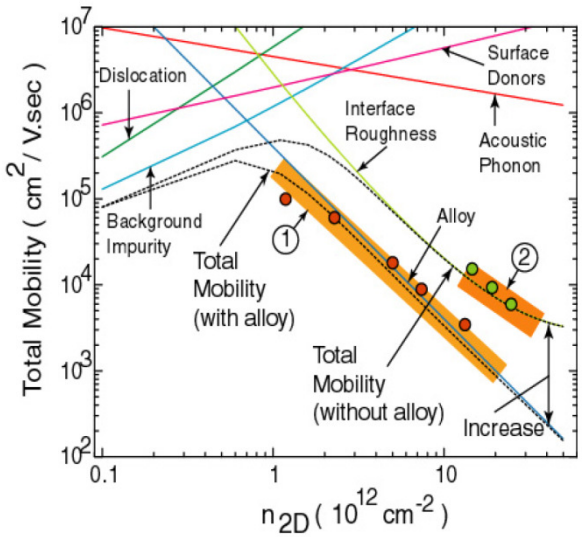


Figure 4.5: Total mobility in function of density of carriers with the contribution of the scattering mechanisms. Experiments from [23]

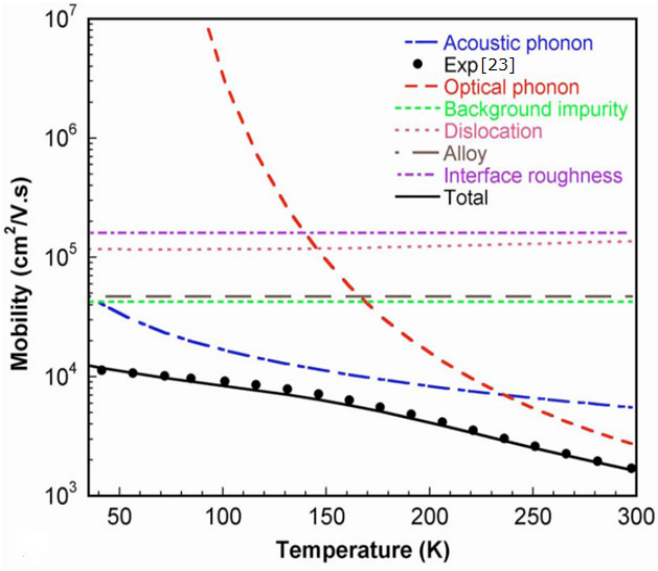


Figure 4.6: Total mobility dependence on temperature with the contribution of the scattering mechanisms. Experiments from [27]

4.2 Literature overview on GaN

4.2.1 Models

As introduced above, a reliable simulation model is very important and it determines the quality of the results. In this regard, it was necessary to find, in literature, the best model for GaN-based HEMTs, that can reproduce in the better way the experiments in the devices.

4.2.1.1 Farahmand Model

The first complete model, i.e. both expressions for low field mobility and for high field mobility, is a model proposed by Farahmand et al. [27]. This model takes in account all the mobility dependence on doping, temperature and electric field. It is important to note that it is developed and calibrated for bulk GaN.

The general expression that it presents for low field mobility (tipically used in drift-diffusion simulations) is given as

$$\mu_0 = \mu_{min}\ (T/300)^{\beta 1} + \frac{(\mu_{max} - \mu_{min})\ (T/300)^{\beta 2}}{1 + \left[\frac{N}{N_{ref}\ (T/300)^{\beta 3}}\right]^{(T/300)^{\beta 4}}} \qquad (4.9)$$

Where *T* is the temperature, *N* is the total doping density and *α, β1, β2, β3, β4, μ_{max}, μ_{min}, N_{ref}* are parameters that can be determined either from experiment or from Monte Carlo simulation.

The previous formula takes in account the mobility dependence on doping and on temperature. At room temperature the doping has the main influence.

For high field mobility this is the expression proposed

$$\mu = \frac{\mu_0 + v_{sat}\left(E^{\delta-1}/E_c^{\delta}\right)}{1 + \alpha(E/E_c)^{\gamma} + (E/E_c)^{\delta}} \qquad (4.10)$$

This formula is a function of low field mobility (μ_0) and electric field, hence it takes in account the mobility dependence on electric field. v_{sat}, *Ec*, *α*, *γ* and *δ* are determined from a least squares fit to the results of Monte Carlo simulation. Their values are extracted for both ternary compounds for the two bracketing cases of alloy scattering.

4.2.1.2 Yang Model

Several years later, Yang et al. [28] proposed a new model based on Farahmand´s model. In the old one an electron concentration- and temperature-dependent GaN low field mobility model is suitable for doping concentration of $1x10^{16}$-$1x10^{18}$ cm^{-3}.

[28] At higher doping concentrations, mobility is overestimated, but it has little influence on simulation because heavy doping is normally used in an ohmic contact region rather than an active region of a device. Anyway, the formula for low field mobility used is the Farahmand one.

For AlGaN, a reduction in low-field mobility is mainly caused by variation of electron effective mass m* with Al composition *x*. Yang proposed a coefficient $f_Z(x)$ to represent parabolic dependence of mobility on Al composition, as given by

$$f_Z(x) = 1/(1 + ax + bx^2 + cx^3) \qquad (4.11)$$

Where parameters *a, b* and *c* depend on a doping concentration *N*. They are extracted from the MC data.

[28] By using the previous formulas is possible to obtain the low field mobility for the ternary compound as function of the Al mole fraction.

$$\mu_0(Al_xGa_{1-x}N) = f_Z(x)\,\mu_0(GaN) \qquad (4.12)$$

This model also indicates that the temperature effect of AlGaN is similar to that of GaN. Figure 4.7 shows the calculated temperature dependence of $\mu_0(Al_xGa_{1-x}N)$ by the present model, where *x* ranges from 0 to 1, doping is 1×10^{17}cm^{-3}, and the random alloy potential is neglected. Calculations using Farahmand´s model and available experimental and theoretical data are also given as comparison.

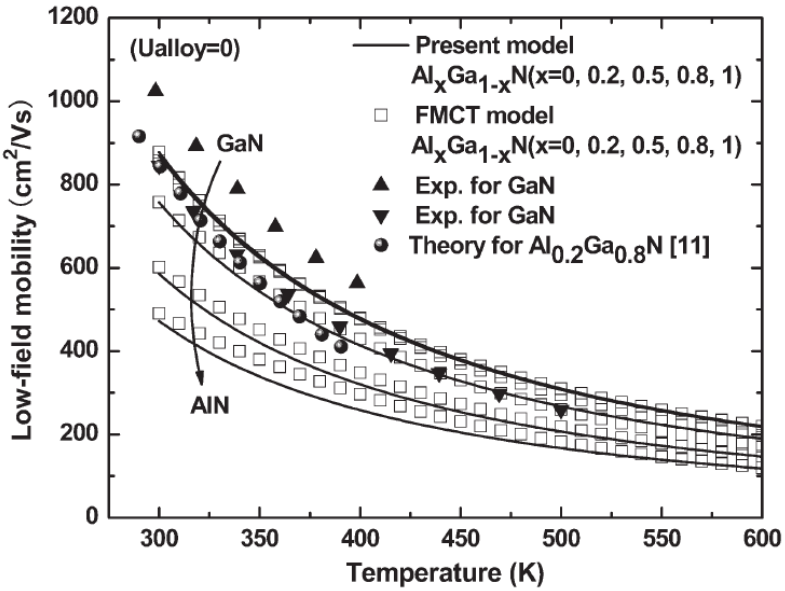


Figure 4.7: Low field mobility dependence on temperature. "Present model" is Yang model [28] and the others data are from Farahmand model or experiments made by Farahmand [27]

For high field mobility, as discussed above, the base formula is the same of Farahmand´s model. Obviously, it is a kind of a drift–diffusion model.

[28] It has been found that the model shows evident errors at high temperatures basically due to the fact that the fitting parameters v_{sat}, *Ec*, δ, γ and α are defined as constants, except for μ_0, which depends on temperature in the Farahmand mobility model. The Yang model introduces the temperature dependence for these parameters.

For the mobility model of wurtzite n-GaN doped at ~10^{17} cm^{-3}, the formulas are shown below:

$$v_{sat} = v_0[v_1 + v_2(T/300) + v_3(T/300)^2] \quad [cm/s] \qquad (4.13a)$$

$$E_c = E_0[E_1 + E_2(T/300) + E_3(T/300)^2] \quad [V/cm] \qquad (4.13b)$$

$$\delta = \delta_0[\delta_1 + \delta_2(T/300) + \delta_3(T/300)^2] \qquad (4.13c)$$

$$\gamma = \gamma_0[\gamma_1 + \gamma_2(T/300) + \gamma_3(T/300)^2] \qquad (4.13d)$$

$$\alpha = \alpha_0[\alpha_1 + \alpha_2(T/300) + \alpha_3(T/300)^2] \qquad (4.13e)$$

Where the previous fitting parameters are extracted from previous MC simulation data.

All of these parameters have been adjusted for AlGaN, by using formulas with dependence on Al mole fraction, as shown below

$$v_{sat(AlGaN)} = v_{sat}\,(1 + v_{Z1}x + v_{Z2}x^2) \quad [cm/s] \qquad (4.14a)$$

$$E_{c(AlGaN)} = E_c\,(1 + E_{Z1}x + E_{Z2}x^2) \quad [V/cm] \qquad (4.14b)$$

$$\delta_{(AlGaN)} = \delta\,(1 + \delta_{Z1}x + \delta_{Z2}x^2 + \delta_{Z3}x^3) \qquad (4.14c)$$

$$\gamma_{(AlGaN)} = \gamma\,(1 + \gamma_{Z1}x + \gamma_{Z2}x^2) \qquad (4.14d)$$

$$\alpha_{(AlGaN)} = \alpha\,(1 + \alpha_{Z1}x + \alpha_{Z2}x^2) \qquad (4.14e)$$

Where v_{sat}, Ec, δ, γ and α are determined by GaN´s formulas, respectively. The Al-composition-dependent parameters are extracted from the MC data.

The aforementioned AlGaN mobility model does not include the influence of a random alloy scattering. As shown in the *Scattering mechanisms* paragraph, this is a serious phenomenon of degradation in GaN.

[28] Indeed, it cannot be completely ignored because of the limit in crystal quality of AlGaN. The influence of random alloy and related scatterings on mobility is represented by a random alloy potential U_{alloy} in which a conduction band offset between the binaries ΔEc is defined as maximum U_{alloy}. This new model introduces a modification factor f_{alloy} to describe the influence of random alloy potential on AlGaN low-field mobility as follows:

$$f_{alloy}(p) = 1 - p(f_1 x - f_1 x^2) \qquad (4.15)$$

Where p is defined as $(U_{alloy}/\Delta Ec)^2$ that yields the values of 0–1 to indicate the percentage of maximum U_{alloy}. Thus, the improved AlGaN low-field mobility including the random alloy potential is given as

$$\mu_{0(alloy)} = f_{alloy}(p)\,\mu(Al_xGa_{1-x}N) \qquad (4.16)$$

With this model is possible to calculate the AlGaN low field mobility at room temperature; as depicted below, the alloy scattering degrades considerably both the mobility dependence on temperature and on Al content.

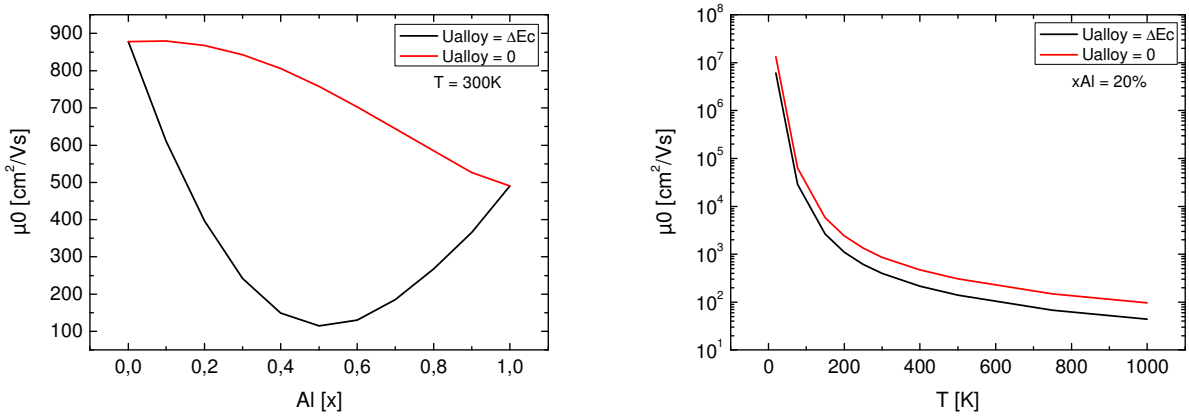


Figure 4.8: Low field mobility dependence on the Al mole fraction and the temperature; by taking in account the alloy scattering (black curves) or not (red curves)

[28] Finally this model also introduces the effect of U_{alloy} into AlGaN high-field mobility by modifying the parameters calculated before.

Now it is possible to calculate also the total mobility, taking in account the alloy scattering; as shown below, for the mobility dependence on temperature, Al mole fraction and electric field, this scattering has a great influnce and it degrades considerably all the characteristics.

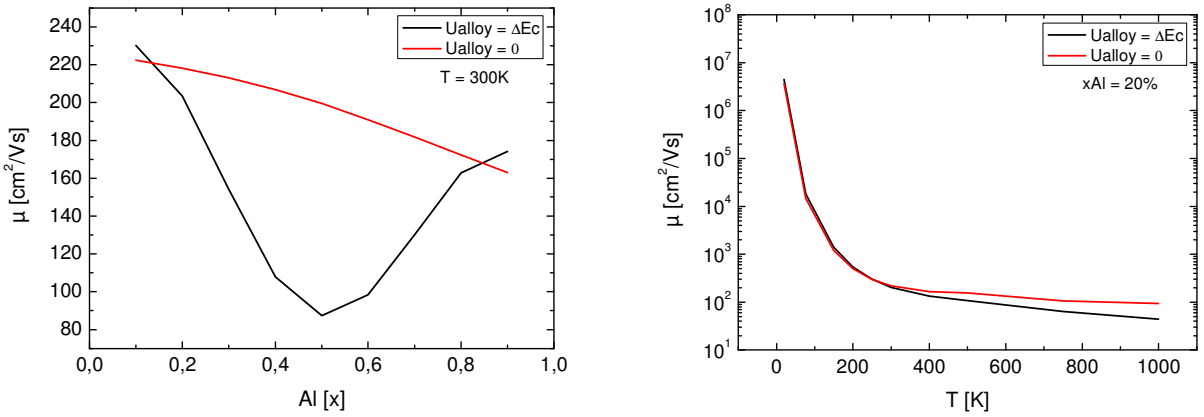


Figure 4.9: Total mobility dependence on the Al mole fraction and the temperature; by taking in account the alloy scattering (black curves) or not (red curves)

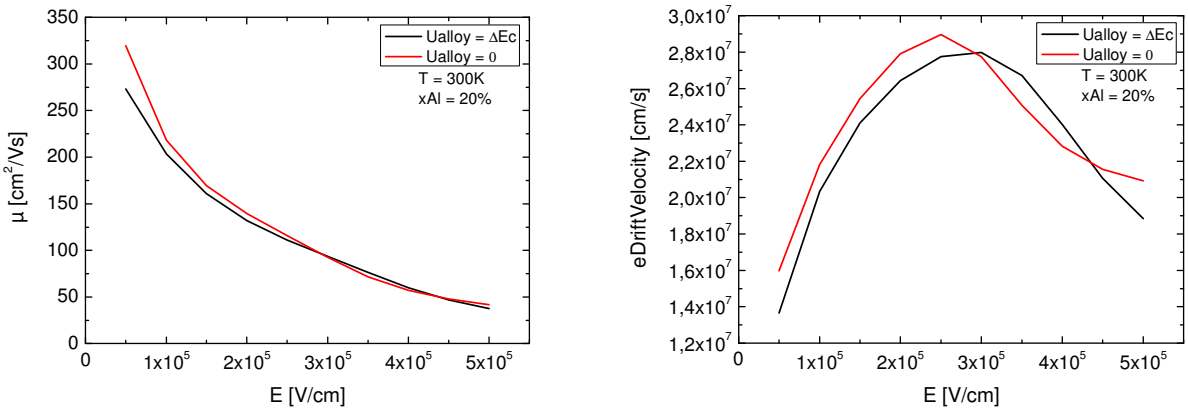


Figure 4.10: Total mobility and electron drift velocity dependences on the electric field; by taking in account the alloy scattering (black curves) or not (red curves)

Finally it is proposed a comparison between this new model and the previous model of Farahmand. In Figures 4.11 and 4.12 is shown the mobility dependence on temperature and Al mole fraction, and the dependence on electric field, all for bulk GaN.

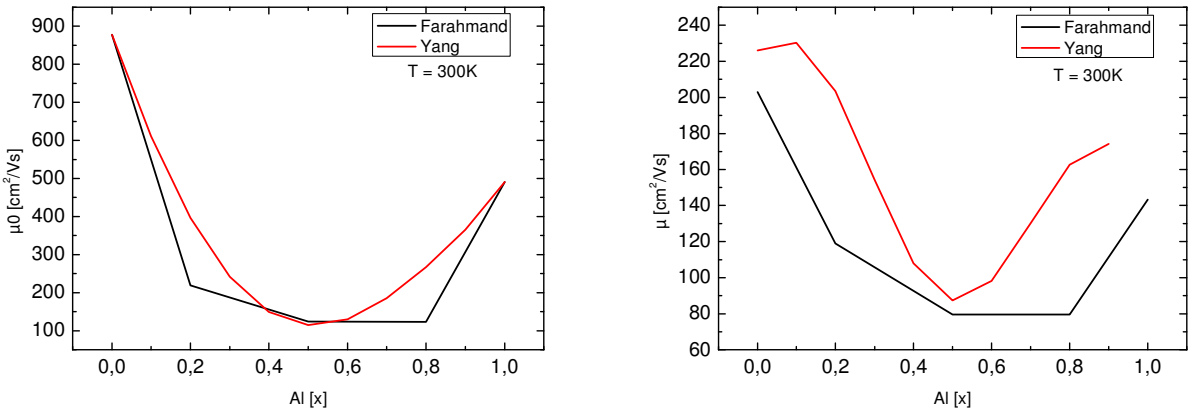


Figure 4.11: Comparison between the models presented in this paragraph, for low field mobility and total mobility dependences on Al content. Farahmand model (black curves) and Yang model (red curves)

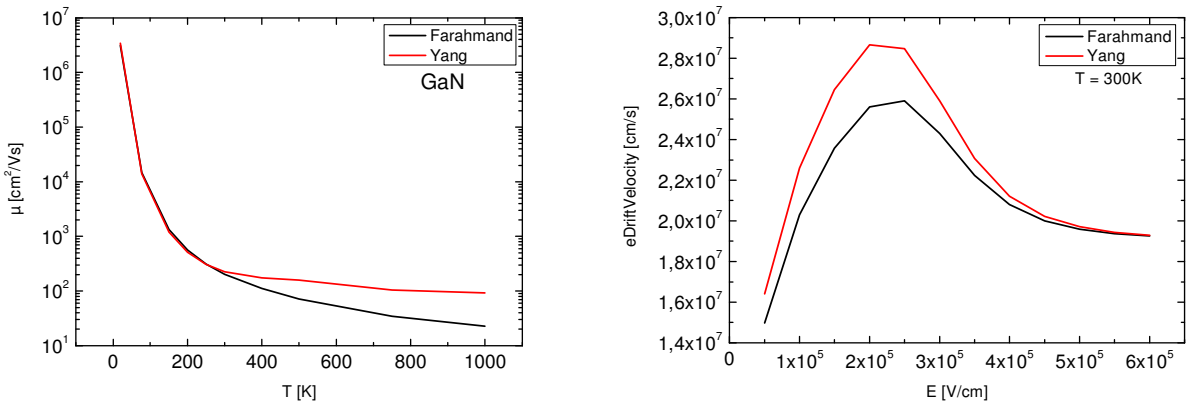


Figure 4.12: Comparison between the models presented in this paragraph, for total mobility dependence on temperature and for electron drift velocity on electric field. Farahmand model (black curves) and Yang model (red curves)

4.2.2 Experiments

To validate all of these calculations about this model, it is necessary to compared them with experiments in GaN-based devices.

Initially, the idea was to make measurements on our devices, but due to high leakage current in the measured devices and process instability, we have decided to make use of published experimental data already available in journal papers.

It has been done an extensive literature search, focused on obtaining experimental data for bulk mobility, where the previous models are calibrated, and for 2DEG mobility, where is very important to evaluate this parameter; to obtain the most trusted values, the research has been expanded to the Monte Carlo simulations. After the comparison between experiments, MC data and indicative informations from Infineon devices, it was possible to extrapolate the graphs shown below.

The first experimental data are for bulk GaN; particularly for low field mobility, total mobility and electron drift velocity. For the total mobility, in Figure 4.13a, the experiments are provided with different doping concentrations.

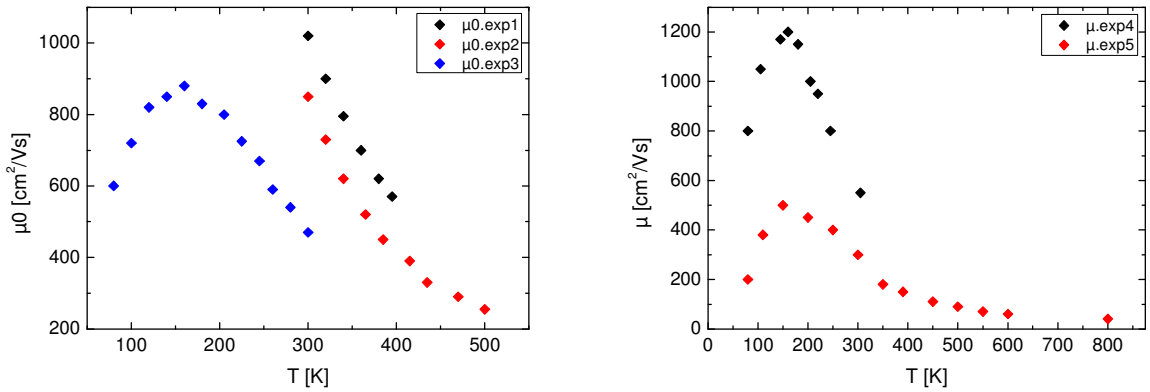


Figure 4.13a: Low field mobility and total mobility experiments for bulk GaN. Exp 1) → [29]; Exp 2) → [30]; Exp 3) → [31]; Exp 4) → [32]; Exp 5) → [33]

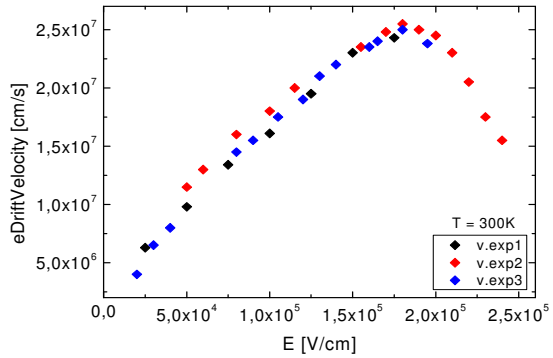


Figure 4.13b: electron drift velocity experiments for bulk GaN. Exp 1) → [34]; Exp 2) → [35]; Exp 3) → [36]

For the mobility in 2DEG several data were found, both for the dependences on temperature and on Al mole fraction; in Figure 4.14, we have collected the experiments more trusted in the range of temperature and Al content most used for power applications.

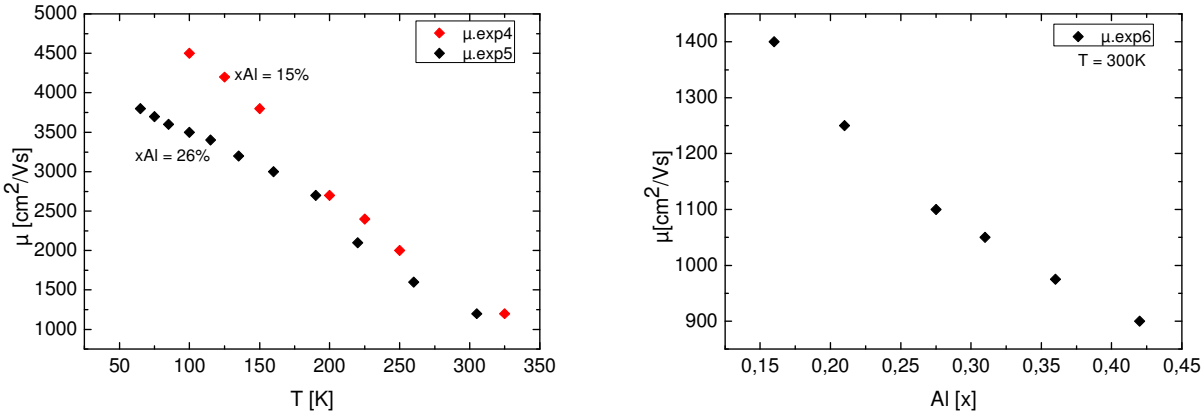


Figure 4.14: Total mobility in the 2DEG. Exp 4) → [37]; Exp 5) → [38]; Exp 6) → [39]

Finally has been done a comparison between the experiments and the analytical calculations of the Yang`s model. In next figures are shown low field mobility, total mobility and electron drift velocity. It is possible to see that this model can reproduce well all the mobility dependence in the range useful for power applications. The areas, where the model is far from the experiments, are those where the device rarely will work, so we considered the model reliable. The new model was tested with different sets of parameter values, to obtain the best trade-off between experiments and MC data.

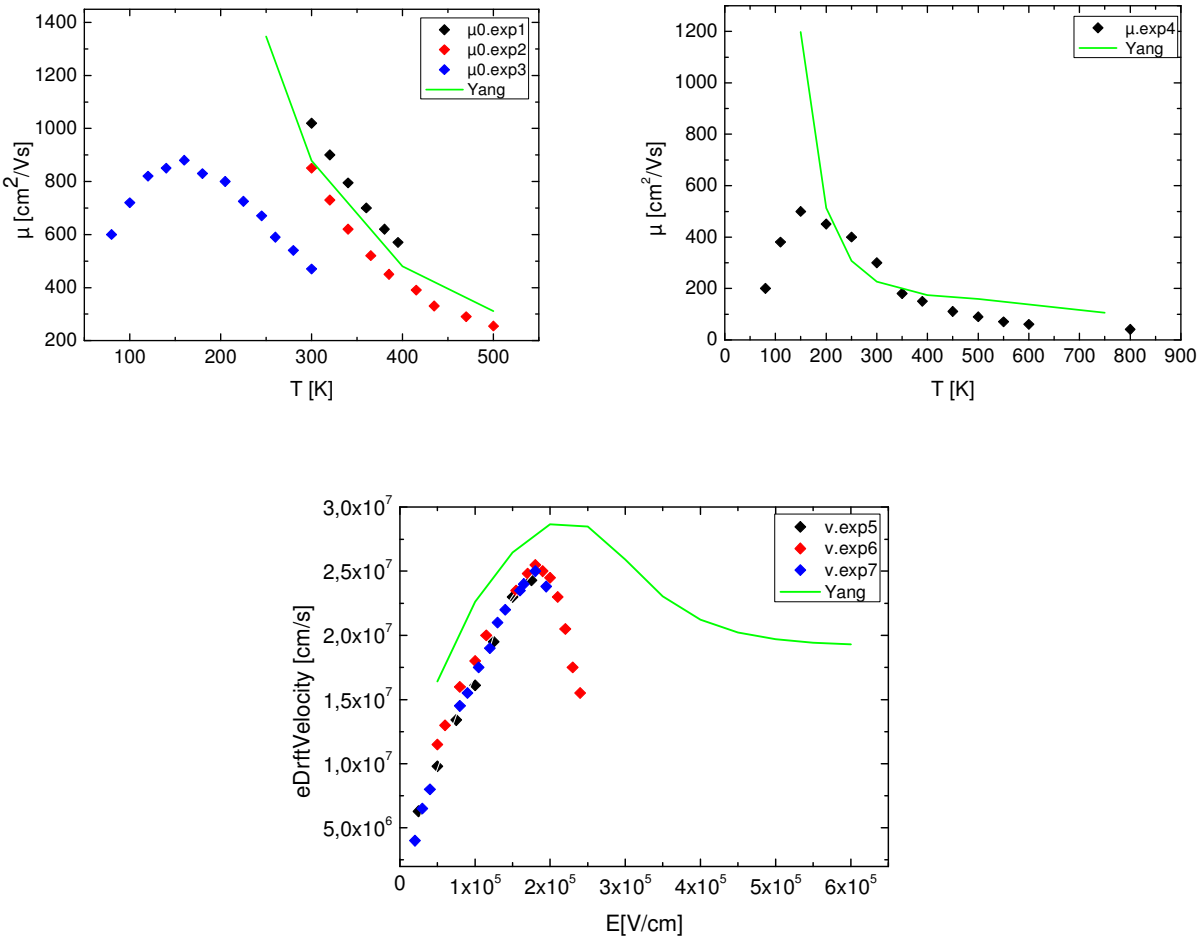


Figure 4.15: Comparison between experiments and Yang model. Exp 1) → [29]; Exp 2) → [30]; Exp 3) → [31]; Exp 4) → [33]; Exp 5) → [34]; Exp 6) → [35]; Exp 7) → [36]

4.3 Mobility models in simulator

In this paragraph are presented the mobility models available in Sentaurus simulator. In next paragraphs, these models will be tested to implement the Yang model in the simulator. All of them are calibrated for Si-based devices and only one, a model for high field mobility, is calibrated for III-V devices.

4.3.1 Low field mobility

These models take in account the mobility dependence on doping and on temperature at low electric field, where it haven´t got a great influence.

The default model used by Sentaurus Device to simulate doping-dependent mobility in silicon was proposed by Masetti et al. [40]

$$\mu_{dop} = \mu_{min1}\exp\left(-\frac{P_c}{N_{A,0}+N_{D,0}}\right) + \frac{\mu_{const}-\mu_{min2}}{1+((N_{A,0}+N_{D,0})/C_r)^{\alpha}} - \frac{\mu_1}{1+(C_s/(N_{A,0}+N_{D,0}))^{\beta}}$$

The second model is the Arora model [41]

$$\mu_{dop} = \mu_{min} + \frac{\mu_d}{1+((N_{A,0}+N_{D,0})/N_0)^{A^*}}$$

And last is the model proposed by the University of Bologna [42]

$$\mu_{dop}(T) = \mu_0(T) + \frac{\mu_L(T)-\mu_0(T)}{1+\left(\frac{N_{D,0}}{C_{r1}(T)}\right)^{\alpha}+\left(\frac{N_{A,0}}{C_{r2}(T)}\right)^{\beta}} - \frac{\mu_1(N_{D,0},N_{A,0},T)}{1+\left(\frac{N_{D,0}}{C_{s1}(T)}+\frac{N_{A,0}}{C_{s2}(T)}\right)^{-2}}$$

4.3.2 High field mobility: Transverse electric field

In Sentaurus simulator it is possible to use separately the mobility dependence on transverse and normal electric field. For transverse electric field, as presented in next models, it´s possible to see the fundamental role of the scattering mechanisms, in this case phonon scattering, surface roughness scattering and Coulomb scattering.

All of these models use the Mathiessen´s rule to combine the scattering contributions. The first model is the Enhanced Lombardi [43]

$$\frac{1}{\mu} = \frac{1}{\mu_b} + \frac{D}{\mu_{ac}} + \frac{D}{\mu_{sr}}$$

With the contribution of, in order, bulk mobility, phonon scattering and surface roughness scattering.

The second is the inversion and accumulation layer model [44], it is similar to the Lucent (Darwish) model but contains additional terms that account for two-dimensional Coulomb impurity scattering in the inversion and accumulation regions of a MOSFET.

$$\frac{1}{\mu} = \frac{1}{\mu_{ph}} + \frac{D}{\mu_{sr}} + \frac{1}{\mu_C}$$

With the contribution of, in order, phonon scattering, surface roughness scattering and Coulomb scattering.

The last is the University of Bologna surface mobility model [42], it is calibrated for temperature of 25°C-648°C, and it has the contribution of Coulomb scattering, phonon scattering and surface roughness scattering.

$$\frac{1}{\mu} = \frac{1}{\mu_{bsc}} + \frac{D}{\mu_{ac}} + \frac{D}{\mu_{sr}}$$

4.3.3 High field mobility: Parallel electric field

All of these models are a function of the low field mobility and of the parallel electric field.

The first, the Extended Canali model is originated from the Caughey-Thomas formula [45], but has temperature-dependent parameters, which were fitted up to 430 K by Canali et al. [46]:

$$\mu(F) = \frac{(\alpha+1)\mu_{low}}{\alpha + \left[1 + \left(\frac{(\alpha+1)\mu_{low}F_{hfs}}{v_{sat}}\right)^{\beta}\right]^{1/\beta}}$$

The second is the only model calibrated for III-V-based devices, that has a negative diffential mobility for high driving fields. This effect is caused by a transfer of electrons into an energetically higher side valley with a much larger effective mass. Sentaurus Device includes a transferred electron model for the description of this effect, as given by [47]:

$$\mu = \frac{\mu_{low} + \left(\frac{v_{sat}}{F_{hfs}}\right)\left(\frac{F_{hfs}}{E_0}\right)^4}{1 + \left(\frac{F_{hfs}}{E_0}\right)^4}$$

Finally the Meinerzhagen-Engl model [48], that requires hydrodynamic simulations

$$\mu = \frac{\mu_{low}}{\left[1 + \left(\mu_{low}\frac{3k}{2q}\frac{(T_c - T)}{\tau_{e,c}v_{sat}^2}\right)^{\beta}\right]^{1/\beta}}$$

4.4 Model implementation in Sentaurus

By starting from the models and the experiments it has been attempted to implement a model in the Sentaurus simulator based on GaN experimental data. All the models in the simulator have been tested.

For low field mobility the best available Synopsys model is the Arora model. It has an excellent correspondence with the Yang model, as shown below

$$\mu_0 = \mu_{min} + \frac{\mu_d}{1 + \left[\frac{N_D + N_A}{N_0}\right]^{A^*}} \qquad (4.17)$$

$$\mu_{min} = A_{min}\,(T/300)^{\alpha_m} \quad \mu_d = A_d\,(T/300)^{\alpha_d} \quad N_0 = A_N\,(T/300)^{\alpha_N} \quad A^* = A_a\,(T/300)^{\alpha_a} \qquad (4.17a)$$

The parameters inside these formulas are editable, and by using analytical calculations, we tested different sets of them, to obtain the best match with the experiments.

With this model for low field mobility, the simulations have been very slow and the values at low temperature are too high, as despicted below.

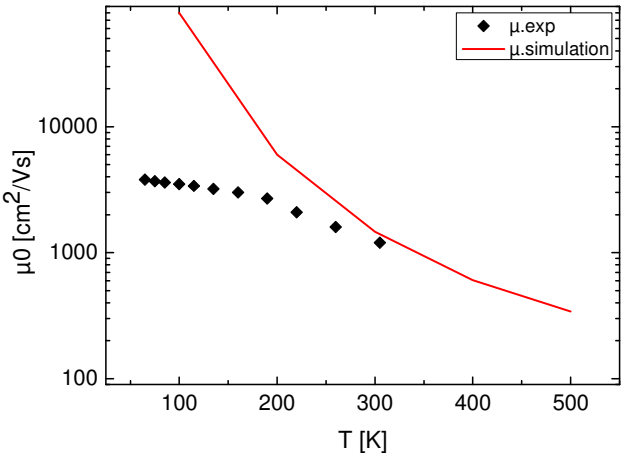


Figure 4.16: Comparison between experiments [38] and Arora model, for low field mobility

After this attempt, other models were tested; a good trade-off was found with the Constant Mobility model, the basic low field model, that can reproduce quite well the mobility dependence on temperature, as depicted in Figure 4.17

$$\mu_{const} = \mu_L \left(\frac{T}{300K}\right)^{-\zeta} \qquad (4.18)$$

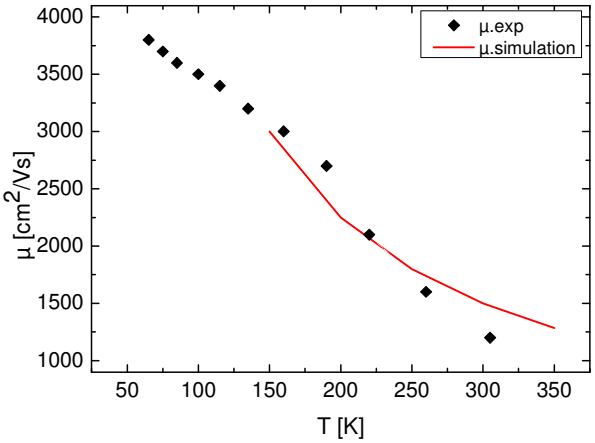


Figure 4.17: Comparison between experiments [38] and Constant Mobility model, for low field mobility

For high field mobility, the first attempt was made by using the only III-V model available, the Transferred Electron model, that shows a good similarity with the Yang model; the only difference is the denominator. However, it has been seen by analytical calculations, that the missing term is really small value if it is compared with other terms.

$$\mu = \frac{\mu_0 + \frac{v_{sat}}{E}\left(\frac{E}{E_0}\right)^4}{1 + \left(\frac{E}{E_0}\right)^4} \qquad (4.19)$$

The main difference between this model and Yang's one, are the denominator´s exponents, in fact, here they are constants; therefore, the only way to fit this model to the Yang's one, is to modify the parameter E_0.

For the next simulations presented, it was used the Constant mobility model for low field mobility and the Transferred Electron model for the high field mobility; in this way, it is possible to reproduce accurately both the mobility dependences on temperature and on Al mole fraction, as shown in Figure 4.18.

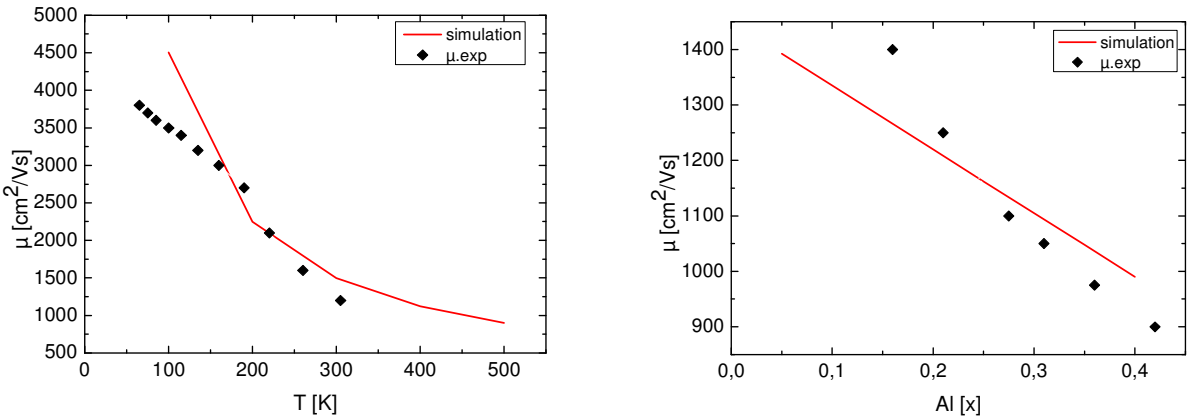


Figure 4.18: Comparison between experiments [38],[39] and total mobility with Transferred Electron model, for the dependences on temperature and Al mole fraction

However, this combination of models is not able to reproduce the electron drift velocity dependence on the electric field.

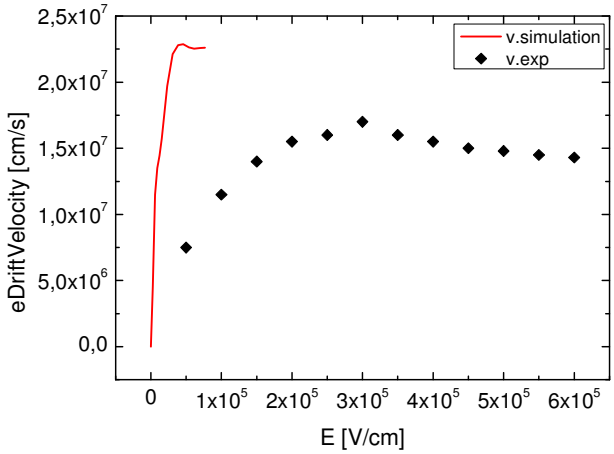


Figure 4.19: Comparison between experiments [27] and electron drift velocity with Transferred Electron model, as function of the electric field

This model leads to unstable simulations and long computation times; for these reasons we decided to not use this combination of models for next calculations and to looking for new solutions.

It overestimates the electric field, therefore, the currents are very high and it is impossible to reproduce the I-V characteristics.

In Figure 4.20 is possible to see the current and the mobility (black curves) compared with a standard simulation used for GaN-based devices (red curves), with Masetti model for low field mobility and Extended Canali model for high field mobility. The standard simulation is optimazed for GaN-based devices and it is fitted to data obtained from Infineon transistors, however, it is not rigorously based on experiments.

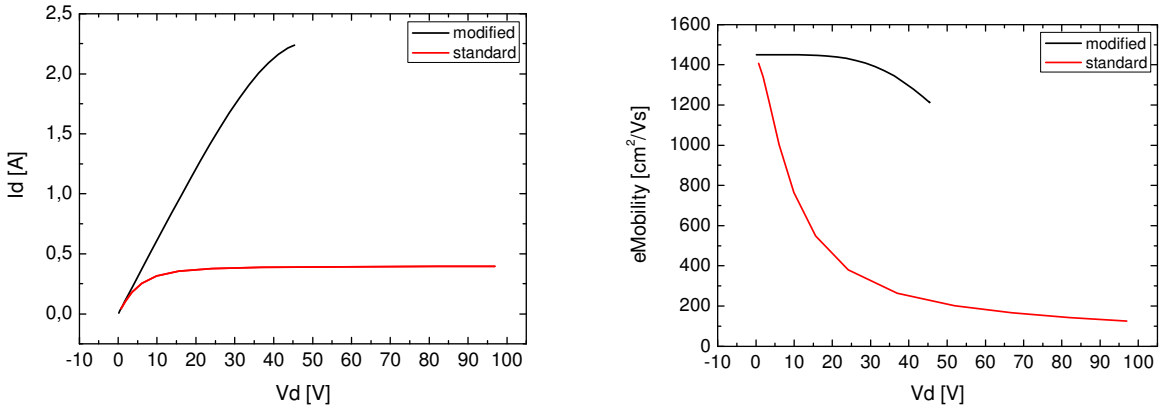


Figure 4.20: Drain current and electron mobility for different models; Constant Mobility and Trasferred electron (black curves), Masetti and Extended Canali (red curves)

It has been also tested an ad-hoc mobility model via C++ interface. This model is completely editable and is possible to implement all the formulas of the Yang model; nevertheless, it was affected by several convergence issues in the simulations and the computing times were very long. Therefore, we decided to use different models.

Finally, the models for Si-based devices have been tested; before, by using analytical calculations and editing the modifiable parameters, the best match was found with the Canali Extended model. This high field mobility model can reproduce carefully both the mobility dependence on temperature and on Al mole fraction, as shown below.

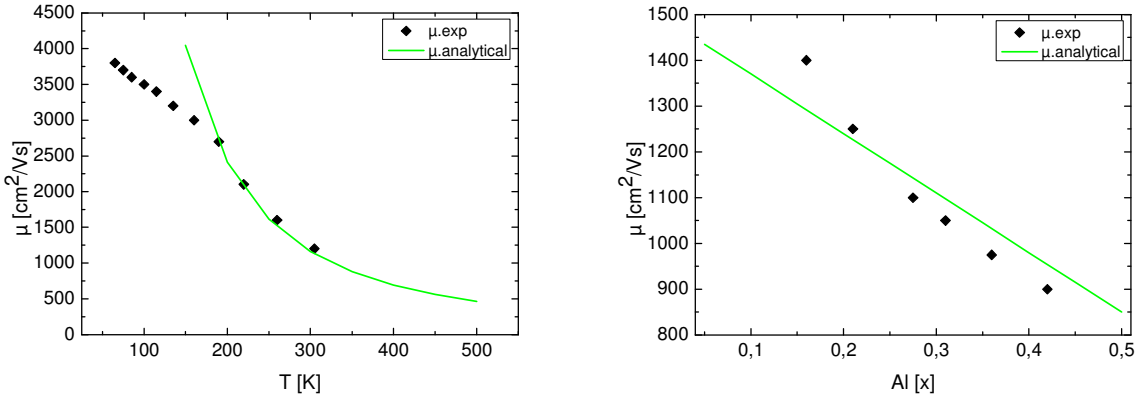


Figure 4.21: Comparison between experiments [38], [39] and total mobility with Extended Canali model, for the dependences on temperature and Al mole fraction

It can reproduce quite well also the electron drift velocity dependence on electric field. Although it provides a good approximation, it cannot reproduce the typical peak of GaN-based devices, because it is a Si-based model, as shown in Figure 4.22.

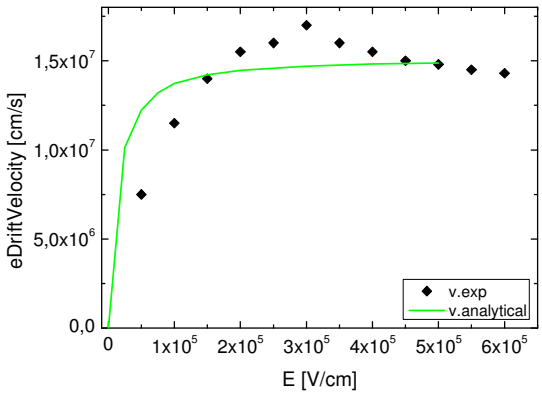


Figure 4.22: Comparison between experiments [27] and electron drift velocity with Extended Canali model, as function of the electric field

After these attempts, it was possible to implement the best solution and to simulate the devices. In next figures are shown the mobility dependences on temperature and on Al mole fraction. For the mobility dependences, particularly for temperature one, we fitted the model only in the operating zone where the devices will work.

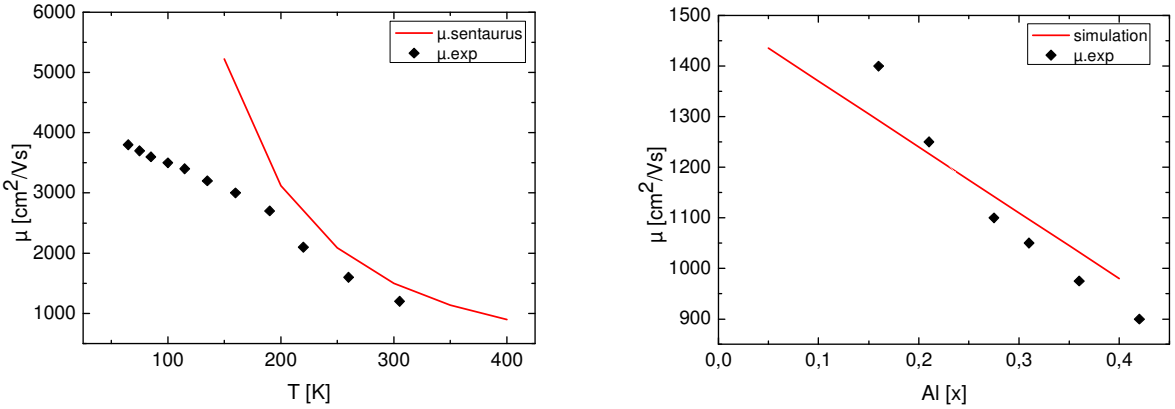


Figure 4.23a: Comparison between experiments [38] and [39], and simulations made by the Constant Mobility model and Extended Canali model, for mobility dependences

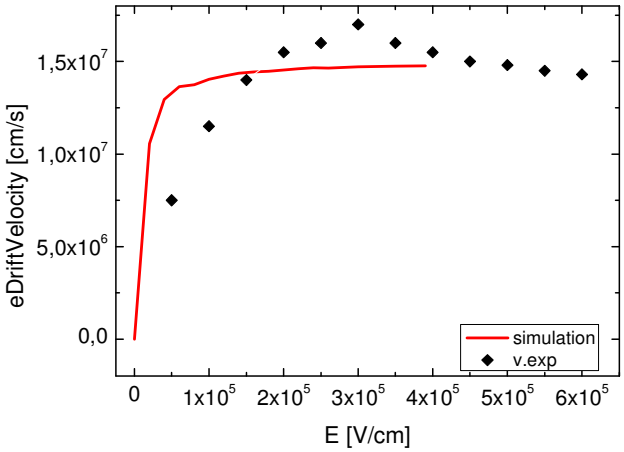


Figure 4.23b: Comparison between experiments [27], and simulations made by the Constant Mobility model and Extended Canali model, for electron drift velocity

Finally, are proposed the drain current and the electron mobility as function of drain voltage for this model. It is important to remember, that the model called standard, as mentioned before, is fitted to measurements on GaN-based devices producted in Infineon. The R_{DSon}, that is the first linear part of the curve, is unchanged, and the difference in saturation drain current, the last part of the curve, is limited, about 10%.

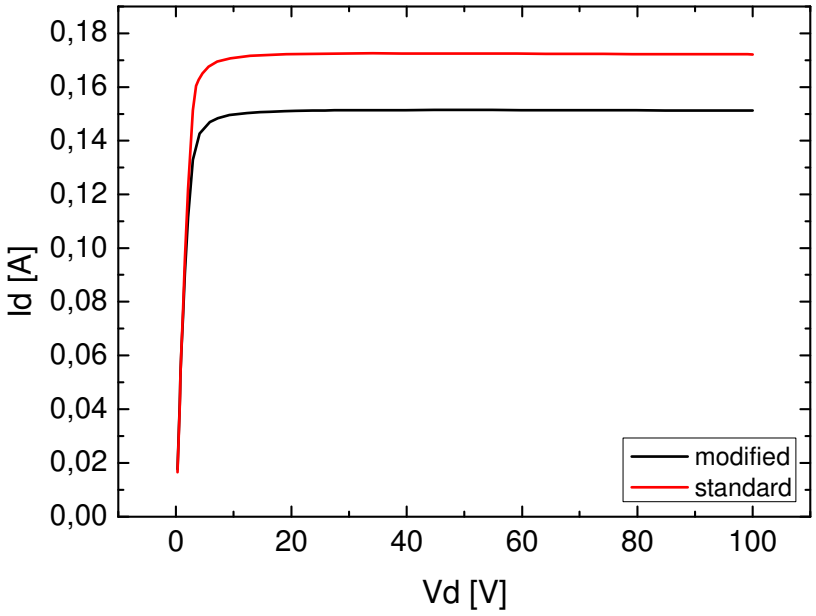


Figure 4.24: Drain current as function of drain voltage; modified model (black curves) and standard model (red curves)

For the electron mobility, it is important to note the difference between this model and a constant mobility model. The improvement is considerable, in fact, if the model were constant, the blue line in Figure 4.25, the current would be highly overestimated.

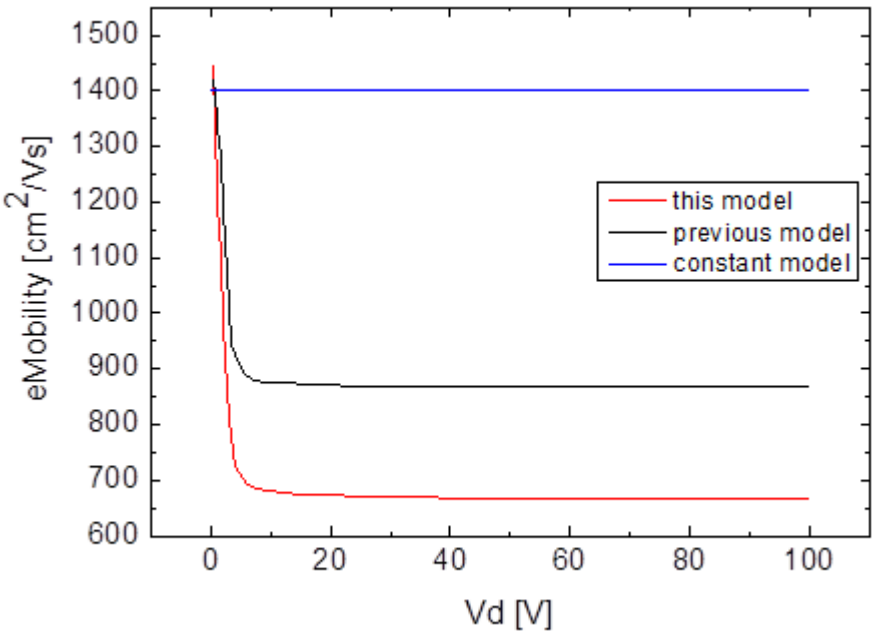


Figure 4.25: Electron mobility as function of drain voltage; modified model (black curves) and standard model (red curves); constant mobility model (blue curve)

4.5 Summary

After a brief introduction about mobility in HEMTs, it has been presented the wide work of research in literature, focused on experiments, Monte Carlo data and models for the mobility in GaN-based devices.

In the second part, we used a model introduced by Yang et al. [28], and it has been presented the implementation of it in the simulator, by using the models available.

To implement the new model, it was sought to find a good trade-off between accuracy and computing time, and to reproduce carefully the experimental data in the useful range for power applications, for the mobility dependences; in fact, in the industrial environment, are important both the precision and the fast simulations to test thousands of devices.

5 Quantum effects

[49] At the end of nineteenth century, the physics was divided in two branches: the mechanics of Newton and the electromagnetism of Maxwell. Actually these theories could not explain several phenomena. Subatomic or thermodynamic experiments were not understandable with the classical physics. The differences between theory and experiments were highlighted when the results were not unchanged with different inertial systems. Several new theories were introduced:

- *Relativity theory*: to solved the incompatibility between the first two theories for different inertial systems
- *Statistical mechanics*: for systems formed by many particles
- *Quantum mechanics*: a generalization of classical physics, to study the matter at atomic and subatomic level

This chapter will be focused on the last topic, in particularly about the quantum effects inside the channel of the HEMTs. It is very important to evaluate these effects, because we are talking about nanometric sizes of the quantum well; despite the power devices have lateral sizes of 20-30 µm, the barriers have a thickness of 20-30 nm, therefore, differences of few nanometers in the peak of electron density in the channel, lead to remarkable consequences.

5.1 Introduction

[49] According to classical physics, an atom could be considered in equilibrium between the electrostatic force and the gravitational force. If it were true, the atomic model would be the same as our solar system, with precise orbits. Moreover, for the conservation of energy, an electron under the centripetal acceleration, should lose energy in the form of electric charge, and it should collapse against the nucleus.

In the *quantum mechanics*, the concept of orbit does not exist; in fact it is replaced by the orbitals, zones with a certain density of probability to find or not an electron. This electron, if excited, can jump in other levels, that are quantized, i.e. they are defined, discrete and multiples of a fundamental level.

A further innovation was the quantization of the light: by affecting a metal with a electromagnetic field, it is possible to extract electrons; this is the *photoelectrical effect*, and it was discovered and formalized by Einsten. In a metal, there is a cutoff level of energy, Fermi level, beyond which the electrons can not remain. However, by providing enough energy to overstep the forbidden level, with a beam of light, it is possible to extract them. The quantity of electron extracted depends on the frequency of the beam.

$$E = \hbar\omega \qquad (5.1)$$

Where E is the energy of a photon, i.e. a quantum of light, and $\hbar$ is the reduced Planck constant.

The key concept for quantum mechanics is the *wave-particle duality*: all particles have both wave and particle proprieties; this duality addresses the inability of classical physics to fully describe the behavior of quantum-scale objects.

[50] This is reflected by the concept of matter waves or *de Broglie waves*, that he proposed in 1924; the de Broglie relations show that the wavelength is inversely proportional to the momentum of a particle and is also called de Broglie wavelength. He deduced also that the frequency is proportional to the total energy of the particle.

$$E = h\upsilon = \hbar\omega \quad \omega = 2\pi\upsilon \qquad (5.2)$$
$$p = h/\lambda \quad k = 2\pi/\lambda \qquad (5.3)$$

Therefore, for each particle, with momentum *p* and energy *E*, is associated a wave, with frequency *v* and wavevector *k,* and viceversa, for each photon is possible to associate an energy and a momentum.

This theory was confirmed several years later by the *Davisson-Germer experiment*. By using an electron gun, they fired a beam of electrons on a slab of nickel with two little slots and, beyond the piece, they observed interference phenomena, constructive and destructive, typical of electromagnetic waves that cause diffraction phenomenon; hence they demonstrated that an electron, i.e. a particle, has a behavior as a wave.

Another concept, that characterizes the probabilistic nature of the quantum mechanics, is the *uncertainty principle of Heisenberg*. It says that it is impossible to know concurrently the momentum and the position of a particle with certainty. Generally, the principle can be applied at any couple of related variables, such as space-momentum or time-energy. It establishes the minimum error that you make by measuring concurrently two variables:

$$\Delta x \cdot \Delta p \geq \hbar \qquad \Delta t \cdot \Delta E \geq \hbar \qquad (5.4)$$

Where $\Delta x, \Delta p, \Delta t, \Delta E$ are the uncertainties of position, momentum, energy and time.

This principle introduces an intrinsic error in the instantaneous determination of position and velocity, accordingly they can be expressed only by statistical terms; therefore, it is only possible to define the probability that a particle could pass in a given position.

The quantum mechanics is a mathematical formulation of the physics. It replaces the variables, such as position, momentum, energy, etc., with mathematicl operators, such as gradient, derivative, etc., but without changes the static variables that are intrinsic in the particle, such as mass, electric charge, etc. With these premises, at a particle is associated a wavefunction that has as domain position and time. Generally the position is described by the three coordinates x, y and z in a tridimensional system. In special cases, such as confinements in one or two dimensions, the system can become at one or two dimensions.

5.1.1 Wavefunction

The wavefunction $\Psi(x, y, z, t)$ is linked to the probability to find a particle in the spatial coordinate (x, y, z) at the instant t. Tipically, its values are complex numbers

This is an abstract mathematical function that is fundamental for the quantum mechanics. Ψ is a full description of all the information that can be measured of a particle, such as position, momentum, energy etc.

In Figure 5.1 is depicted a simplified quantum well with the wavefunctions for each descrete level of energy.

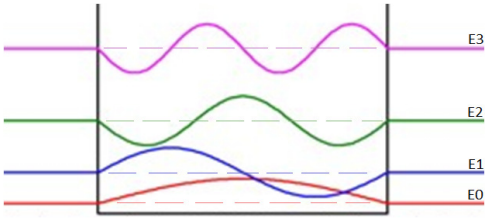


Figure 5.1: Simplified quantum well with wavefunctions

The Schrödinger equation, that will be explained in next paragraph, describes the evolution of the wavefunctions over the time. The wavefunction has a behaviour like a water wave, because the Schrödinger equation is mathematically a type of wave equation and this gives rise to the wave-particle duality.

5.1.2 Schrödinger equation

The classical mechanics imposes that the energy in a closed and isolated system is constant. Hence, with E the total energy, E_K the kinetics energy and U the potential energy, is $E = E_K + U$. By using, for simplicity, a one-dimensional system, several operators will be introduced:

$$E \to i\hbar\frac{\partial\Psi}{\partial t} \qquad E_k \to -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} \qquad U(x) \to U(x)\Psi \qquad (5.5)$$

With these changes is possible to obtain the differential equation called *Schrödinger equation*:

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + U(\mathrm{x})\Psi \qquad (5.6)$$

And it can be rewritten in a compact way as

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi \qquad where\ H = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + U(x) \qquad (5.7)$$

H is the Hamiltonian operator. This equation is the fulcrum of the quantum mechanics and it is the conservation of energy in quantum terms. The previous expression is the time-indipendent equation. By using the next ones is possible to describe how the quantum state of a physical system changes with time; it was formulated in 1925 by the Austrian physicist Erwin Schrödinger.

For a particle, such as an electron, is possible to calculate its wave function by using the following formula:

$$i\hbar\frac{\partial\Phi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Phi + [E_{c0}(\boldsymbol{r}) + U_{lt}(\boldsymbol{r})]\,\Phi(\boldsymbol{r},t) \qquad (5.8)$$

Where m is the electron mass, as it were in the vacuum, E_{c0} is the potentials, built-it or applied, and U_{lt} is the periodic electrostatical potential due to atoms in the lattice.

By introducing the Block function u_k and by assuming E_{c0} to be known, is possible to do a great simplification in the previous expression with the effective mass equation:

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2}\nabla[(m_0)^{-1}\nabla\Psi] + E_{c0}(\boldsymbol{r})\Psi(\boldsymbol{r},t) \qquad (5.9)$$

Where m_0 is the effective mass, which will be explained later.

Ψ is the eigenfunction for each descrete level of energy inside the quantum well. To obtain the total wavefunction is sufficient to multiplied Ψ with the Block function

$$\Phi(\boldsymbol{r}, t) = \Psi(\boldsymbol{r}, t) \cdot u_{k=k_{min}} \qquad (5.10)$$

5.1.3 Semiclassical Approch and Quantum Approch

First, it is important to remember that the HEMTs are normally-on devices, and the 2DEG is only due to the polarization charges, therefore, it is formed already at the equilibrium; instead in a conventional Si-based MOSFETs, which are normally-off, to have the quantum well and, consequentely, the quantum effects, it is necessary to apply a positive voltage at the Gate, and then it is possible to confine electrons in the oxide/silicon inteface. Therefore, in any condition, is very important to have a good knowledge of the situation inside the channel.

The Semiclassical approach is obtained only by solving the electrons transport equation of Boltzmann. Instead, the Quantum approach takes in account the quantum effects following the Schrödinger equation. In Figure 5.2 is possible to note the differences of the models, in a MOSFET.

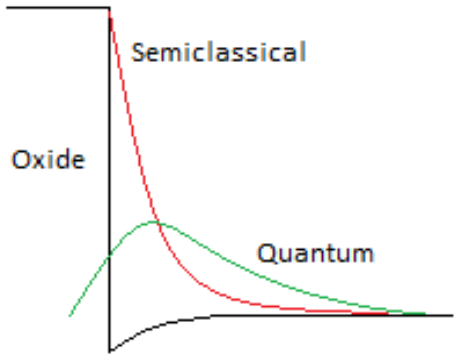


Figure 5.2: Comparison of the electron densities in the channel for Semiclassical (red) and Quantum (green) models

The Semiclassical model (red curve) shows a peak of density at the interface oxide/semiconductor and then a rapid drop of the electrons concentration after the channel. For this approach, there are not electrons inside the oxide. The first important difference between these models, is that the Quantum approach (green curve) admits a non-zero probability to find electrons in the oxide; the second difference is that the peak of electron density is not at the interface, but few nanometers below, in the Silicon. This leads to a great difference when they are applied in a small quantum well, e.g. tens of nanometers, as is shown below.

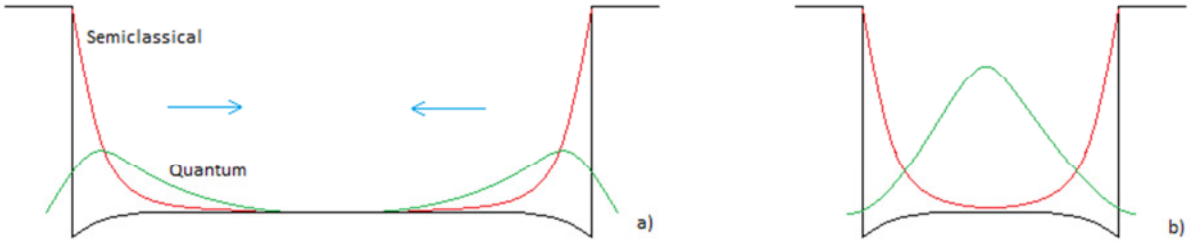


Figure 5.3: Electron density in a small quatum well

In Figure 5.3 a), which is a large quantum well, the situation is the same of Figure 5.2 in both sides of the well. However, when the distance between the barriers is decreased as in Figure 5.3 b), the electron densities are completely different. In fact for the Quantum approach the peak of electrons concentration is in the middle of the well, not at the interfaces as in Semiclassical approach. This means that the equivalent oxide thicknesses are different, as shown in Figure 5.4.

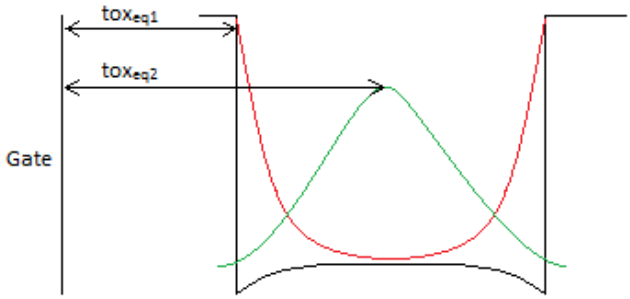


Figure 5.4: Equivalent oxide thickness for Semiclassical (red) and Quantum (green) approches

By using the expression for capacitance

$$C = \frac{\varepsilon_{ox}}{t_{ox_{eq}}} \qquad (5.10)$$

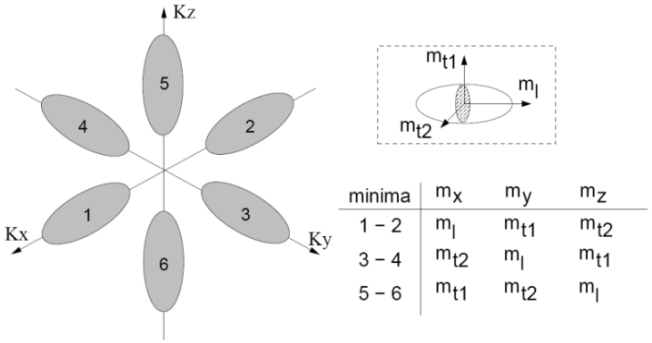
It is simple to understand that in the two approaches the capacitance will have different values, because of the different position of the charges, and also the current, which is linked to the peak of electron density, will be very different.

5.1.4 Effective mass

In a solid material the forces between atoms affect the movement of the free electrons and their behaviour can not be described with Newton´s law. Therefore, the concept of effective mass must be introduced to explain the movements of electrons; indeed carriers, in a crystal, show a behavior as if they were particles with a mass dependence in their direction of travel. The effective mass is calculated as if the electron were in the vacuum, but taking in account the effect of energy. It can be negative or different due to circumstances.

In the semiconductors, each minima has three different effective masses, one for each direction, therefore, one is longitudinal and two are transversals.

Silicon has six equivalent valleys in the conduction band and its minima are ellipsoids, because the longitudinal mass is $0.98m_0$ and the transversals are $0.19m_0$. In Figure 5.5 is shown a simplified model of the minima.



minima	m_x	m_y	m_z
1 – 2	m_l	m_{t1}	m_{t2}
3 – 4	m_{t2}	m_l	m_{t1}
5 – 6	m_{t1}	m_{t2}	m_l

Figure 5.5: Minima of the Silicon

From this concept was born the *stress engineering,* which exploits the strain of the materials to modify the probability of occupation of the minima. For example, with a confinement in z direction, by using a strain, it is possible to move the minima 1-2 and 3-4 to higher levels of energy, hence their probability to have electrons decrease. Consequently the minima 5-6 have higher probability because they move to lower energies; in this way, the gap between first and second level (E_0 and E_1) is higher, limiting also the intervalley scattering. Minima 5-6, which are the most occupied of electrons, in the direction of current flow in the channel, have the lowest transversal effective mass, as summarized in Figure 5.5; by using Formula (4.6), it means that carriers will experience the highest mobility when being confined only in minima 5-6. Higher mobility means higher current capability and, in general, better device performance and figures of merit.

Gallium Nitride has only one equivalent valley in the conduction band and the effective masses are $0.18m_0$ and $0.2m_0$, longitudinal and transversal, respectively; hence the minima is almost a sphere.

The most important difference between Si and GaN, is that in GaN, the valleys A- and M-L- have energies just above the minima (Γ-valley).

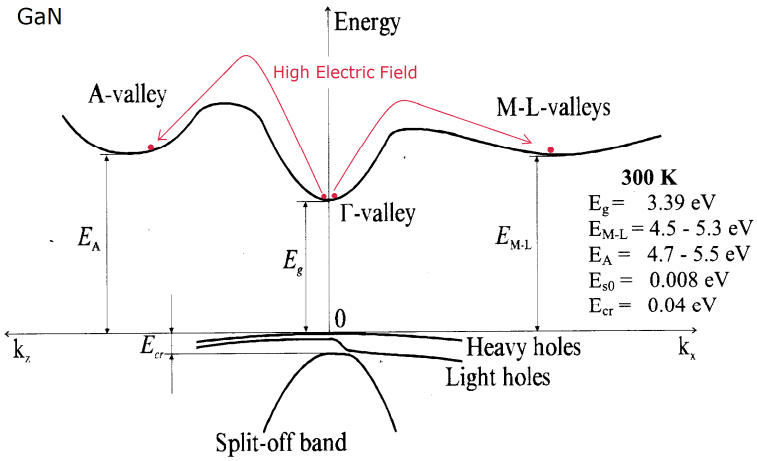


Figure 5.6: Band diagram of GaN

At high applied electric fields, electrons in the minima have enough energy to jump to the higher valley; in these valleys the electron effective masses are higher, about $0.36m_0$, therefore the mobility, following (4.6), decreases, as depicted in next figure, showing a peak at $2x10^5$ V/cm.

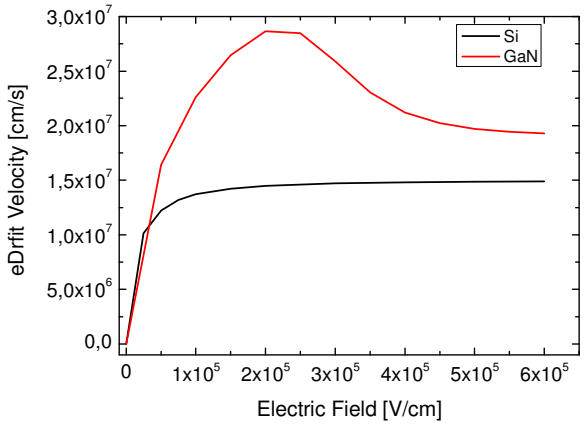


Figure 5.7: Electron drift velocity for Si and GaN

5.2 Simulations of Quantum models

In Sentaurus simulator is possible to activate different models to take in account quantum effects. The models are:

- Van Dort model
- Modified Local-Density Approximation model (MLDA)
- Density Gradient
- Schrödinger model

The first two models are robust and fast during simulations, but they are not calibrated for quantum wells, particularly the first is suitable only for bulk MOSFET.

Density Gradient model is slower than the first models and it introduces several semplifications in the calculation of the Schrödinger equation; however, it can reproduce accurately the quantum effects in the channel.

The previous models do not solve the Schrödinger equation, in fact they are empirical approaches which are able to reproduce the quantum effects and the charge density of devices.

Schrödinger model is the slowest, particularly because it uses a very sophisticated set of equations. This is the most precise method to compute the quantum effects in the channel, but its greatest limit is that it works only at the equilibrium.

Instead of other models which introduce only correcting factors, in this case, the simulator solve Poisson equation, continuity equations for electrons and hole, and the Schrödinger equation.

In Figure 5.8 is shown a comparison between Semiclassical model, Density Gradient model and Schrödinger model, done at the equilibrium in an AlGaN/GaN HEMT, with an oxide layer under the Gate contact to limit the current leakage.

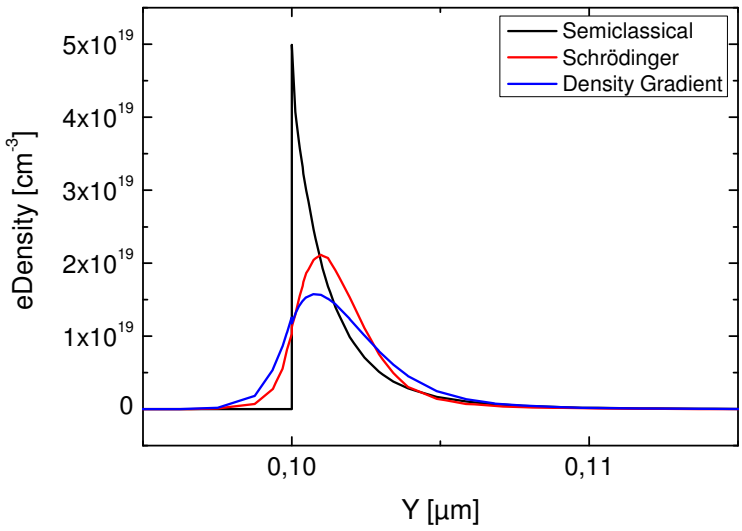


Figure 5.8: Comparison of electron density for Semiclassical model, Schrödinger model and Density Gradient model

In Figure 5.8, as explained before, is possible to see the peak of electrons at the interface of the Semiclassical model and the non-zero probability to find electrons in the barrier of the Quantum models.

Regarding the differences in the characteristic Id-Vd (Figure 5.9), we were expecting an higher difference in currents, but, finally, it seems to be more important to have a good and reliable mobility model to fit it to the experiments. If the quantum effects are taken in account, it is possible to obtain a ~5% plus of accuracy.

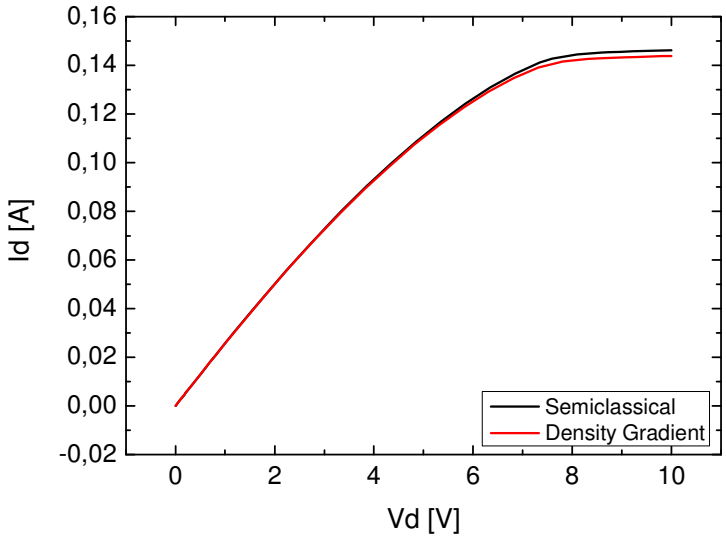


Figure 5.9: Comparison of Id-Vd between Semiclassical and Density Gradient

It is important to say, that the oxide layer hides the differences between the two approaches, limiting the differences in the I-V characteristics. In fact, with an oxide layer with a thickness of tens of nanometers, if the peak of electrons is moved from the interface only few nanometers, is difficult to see the changes. We tested these approaches also with a schottky gate, hence without oxide, but for power applications is necessary this layer to reduce the current leakage.

5.3 Back Barrier

During the development of AlGaN/GaN HEMTs a lot of solutions have been tried to the pursuit of improvement of the characteristic of these transistors. One of the most important targets is to reduce the current leakage and to increase the breakdown voltage, but without compromising the R_{DSon}.

To reach this target, one of the possible solutions is to insert in the transistor a back barrier below the channel, trying to put a high discontinuity band before the buffer layer, which can confine the electrons. This method, as will be shown later, it is capable to cut the “tail” of electrons in the buffer layer, limiting leakage current and, accordingly, it increases the breakdown voltage. Despite it reduces the amount of electrons in the channel, slightly increasing the R_{DSon}, the improvements are considerable.

In this chapter are presented a couple of solutions, the most used:

- Back barrier in InGaN
- AlGaN buffer

In the next paragraphs will be shown a comparison between these structures and a conventional HEMT, to show all the properties, the improvements and the weaknesses. All of the structures have, under the gate contact, an insulator with a thickness of 30 nm, an AlGaN barrier of 20-30 nm with Al mole fraction of 20-30 %, and a total thickness from the barrier to the substrate of 1.7 µm.

5.3.1 InGaN back barrier

This solution is made by introducing a thin layer of InGaN below the channel as shown in Figure 5.10.

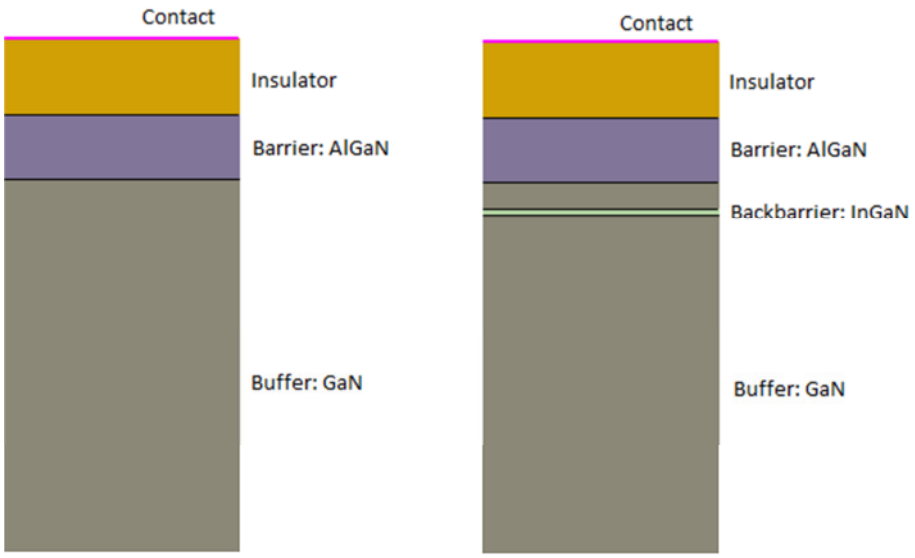


Figure 5.10: Structures of conventional HEMT (left) and HEMT with back barrier in InGaN (right)

In this case the back barrier is inserted at 11 nm from the channel, it has a thickness of 2 nm with In mole fraction of 10%. The resulting conduction band is shown below.

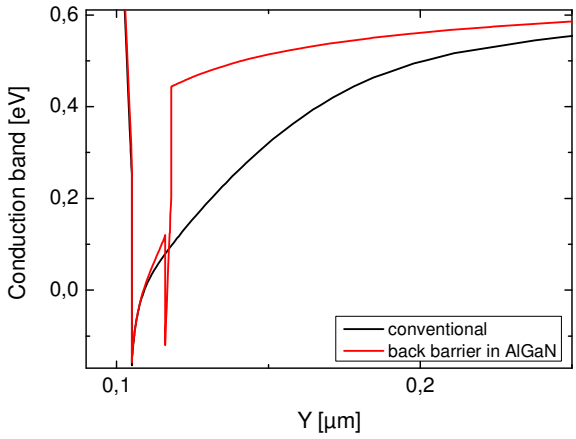


Figure 5.11: Conduction bands of conventional HEMT (black) and HEMT with back barrier in InGaN (red)

It is important to note the confinement introduced by the back barrier, which eliminates the "tail of the electrons" present in the conventional HEMT. In Figure 5.12 are shown the electron densities calculated with and without taking in account the quatum effects. In the InGaN layer, there is a second peak, similar to a secondary 2DEG, which is due to polarization charges introduced in the simulations; this is a side effect, that is not possible to remove.

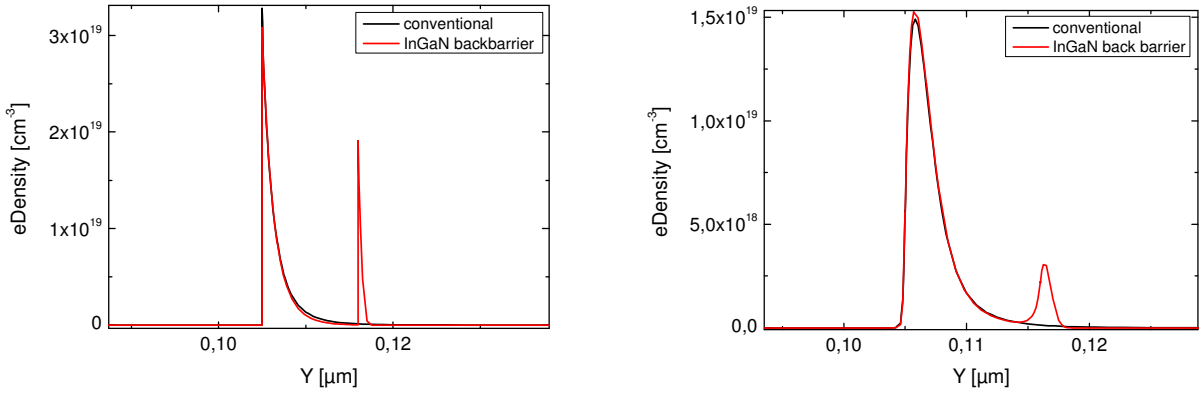


Figure 5.12: Electron densities of conventional HEMT (black) and HEMT with back barrier in InGaN (red), without (left) and with (right) quantum effects

In Figure 5.12, the main peak of electron density, evaluated with the quantum models, decrease from $3x10^{19}$ to $1.5x10^{19}$ electrons per cm^{-3}, as explained above.

This second peak, also present in literature, is really high and leads to a total concentration of electrons higher for the HEMT with back barrier (Figure 5.13 and Table 5.1), instead of a lower one. Consequentely the threshold is shifted to lower voltages, more than we expected. After a lot of attempts, it has been found that maybe the problem is the polarization charges of the InGaN and it has been opened a case with Synopsys to try to solve the issue.

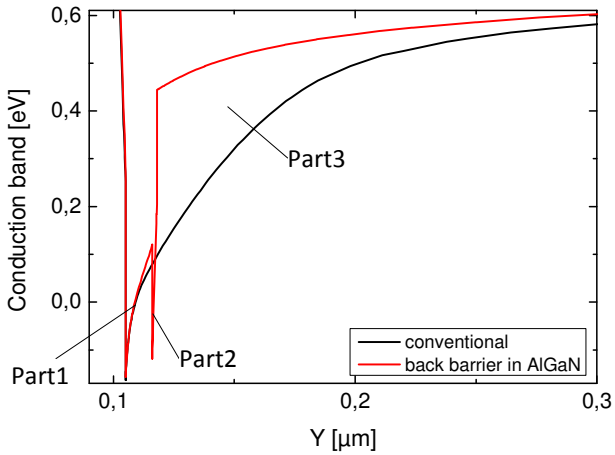


Figure 5.13: Conduction bands of conventional HEMT (black) and HEMT with back barrier in InGaN (red)

	part1	**part2**	**part3**	**total**
$n_{s(conventional)}$ [cm^{-2}]	4,4533e12	1,8306e10	2,0538e10	4,4922e12
$n_{s(backbarrier)}$ [cm^{-2}]	4,2349e12	7,0795e11	1,0716e05	4,9429e12

Table 5.1: Electrons concentration in the different parts of the conduction band

In the first part the values are similar, in the third part is possible to see that the back barrier "cut" the tail of electrons and, in the second part, the problem due to polarization charges leads to a concentration too high of electrons. According to Synopsys, we are still trying to understand the problem; we have already analysed the effects of the polarization charges, of band gap and of electron affinity, unfortunately without results.

Finally, the electrons concentration is shown as function of In content and distance from channel of back barrier.

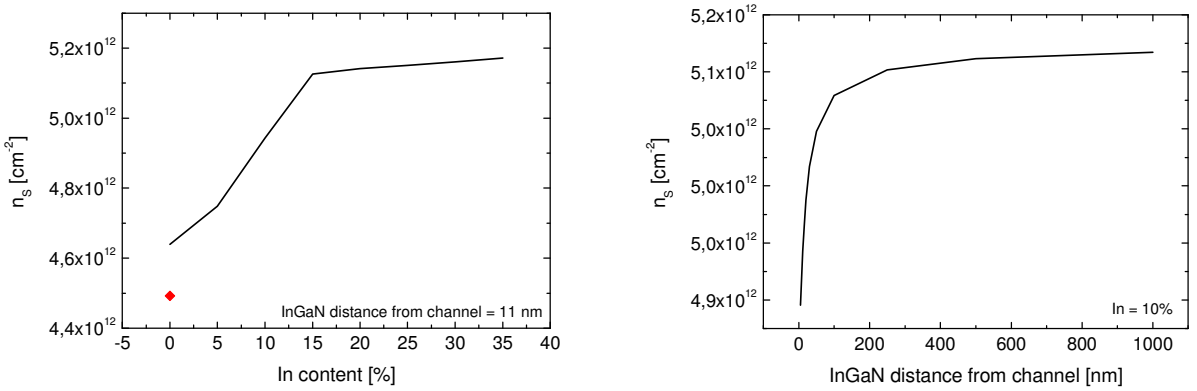


Figure 5.14: Electron density as function of In content (left) and distance of back barrier from channel (right). The red dot is the electron density for the conventional HEMT

For the aforementioned problem, for In content of 0% the concentration is not the same as conventional HEMT and for large distance from channel, where the back barrier should not have influence, the concentration is still not the same as a conventional HEMT.

5.3.2 AlGaN buffer

The second solution presented is a HEMT with a buffer in AlGaN, as despicted below.

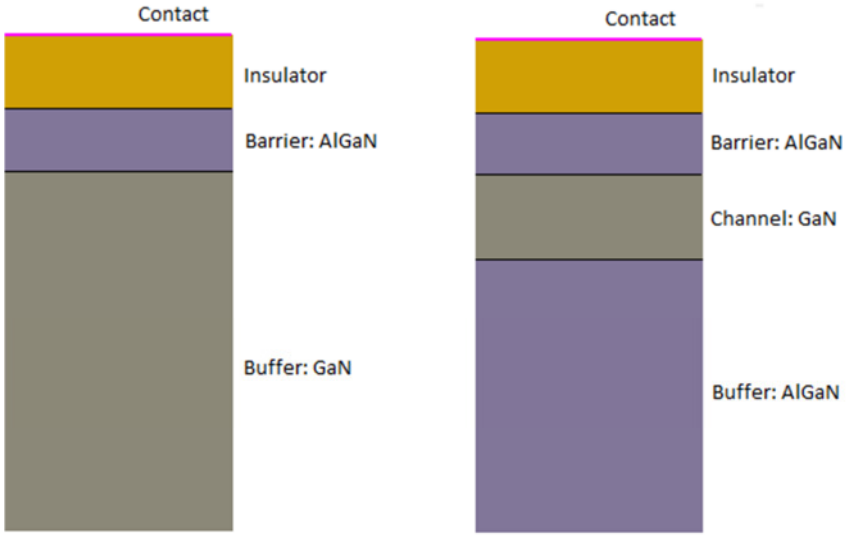


Figure 5.15: Structures of conventional HEMT (left) and HEMT with AlGaN buffer (right)

This structure presents a channel in GaN of 35 nm, which has a compressive strain due to the growth onto a relaxed AlGaN layer, and a buffer made in AlGaN, that works as back barrier, with an Al mole fraction of 5%. In Figure 5.16 are shown the conduction bands.

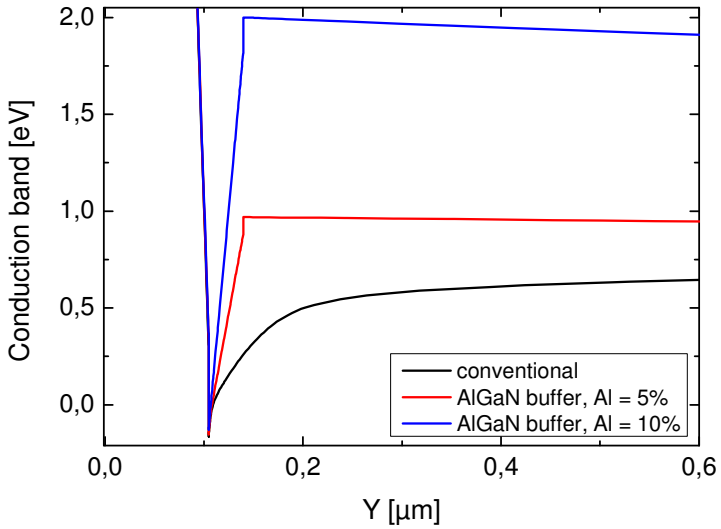


Figure 5.16: Conduction bands of conventional HEMT (black) and HEMT with AlGaN buffer: Al content 5% (red) and Al content 10% (blue)

Even in this structure is possible to note the excellent confinement introduced by the back barrier. In previous picture are presented two different Al mole fraction in the buffer layer; in next simulation will be used a buffer where Al is 5%, which guarantees a good confinement and a good concentration of electrons in the channel. In Figure 5.17 are shown the electron densities without and with the quantum effects.

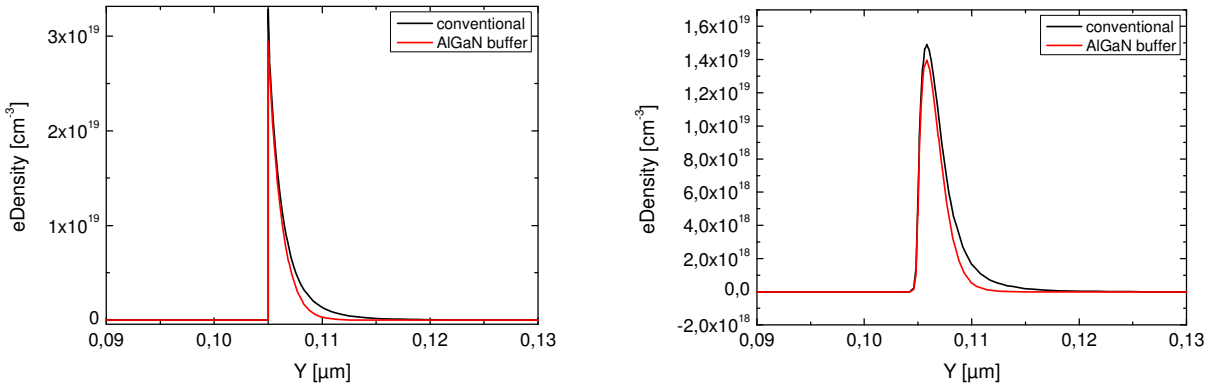


Figure 5.17: Electron density for conventional HEMT (black) and HEMT with AlGaN buffer (red), without quantum effects (left) and with quantum effects (right)

	No quantum effects	**With quantum effects**
$n_{s(conventional)}$ [cm^{-2}]	4,4922e12	4,4195e12
$n_{s(AlGaN\ buffer)}$ [cm^{-2}]	3,5195e12	3,4441e12

Table 5.2: Electrons concentration for conventional HEMT and HEMT with back barrier

With this solution even the capacitances have been simulated, to compare them with the conventional ones, as depicted in Figure 5.18.

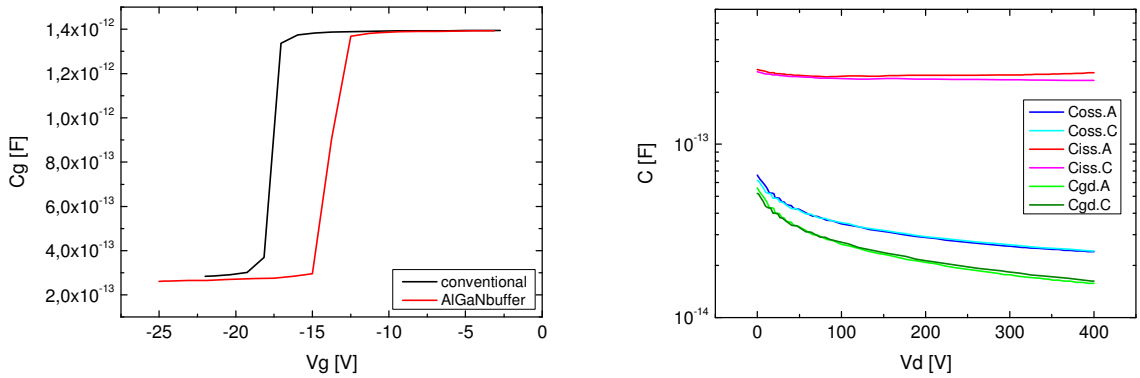


Figure 5.18: Capacitances for conventional HEMT and HEMT with AlGaN buffer; Cg (left) and Coss, Ciss and Cgd (right)

In the previous figure, the capacitances of gate, as function of gate voltage, are similar, except for the shift toward right of the structure with AlGaN buffer, due to the shift of threshold introduced by the back barrier that decrease the amount of electrons in channel.

In the right part of figure are shown other capacitances as function of drain voltage: Coss, which is the output capacitance, i.e. Cds and Cdg, Ciss, which is the input capacitance, i.e. Cgd and Cgs, and Cdg, the capacitance from drain to gate; all of these, respect of the conventional transistor, are almost unchanged.

This solution gives good results: it reduces the leakage current, it moves slightly the threshold voltage and it reduces the electron concentration. This is the trade-off that we have to optimize: increase the breakdown voltage and reduce the leakage, seeking to not damage the R_{DSon}.

5.4 Summary

In this chapter we discussed about the quantum effects in HEMTs, starting from the quantum mechanics, its characteristics and its fundamental rule, the Schrödinger equation.

After we tested the models available in Sentaurus simulator; it seems that a reliable and good mobility model is very important and it is possible to increase lightly the accuracy with a good quantum model. It is important to note that with an oxide layer above the barrier, it is possible to reduce leakage but the influence of quantum effects are lower.

Finally, we applied the quantum models in two AlGaN/GaN HEMTs with a back barrier. Despite the first structure tested, with a back barrier in InGaN, was not completely studied because of the problem explained before; the second has given good results regarding the confinement of electrons and it is capable to reduce leakage and, accordingly, to increase breakdown voltage. However, the less amount of electrons in the channel leads to a R_{DSon} higher than a conventional transitor; anyway, these changes keep unchanged the figures of merit linked to the capacitances.

Bibliography

Bibliography

[1] Rüdiger Q., *Gallium Nitride Electronics,* Springer, 2008

[2] Mishra, U.K., Singh, J., *Semiconductor Device Physics and Design,* Springer, 2008

[3] Liu L, Edgar H., *Substrates for gallium nitride epitaxy,*Report: a Review Journal, 2002

[4] Morkoc, H., *Handbook of Nitride Semiconductors and Devices,* Wiley-VCH, 2008

[5] Nakamura, S., Senoh, M., Nagahama, S.,Iwasa, N., Yamanda, T., Matsushita, T., Kiyoku, H., Sugimoto, Y., Kozaki, T., Umemoto, H., Sano, M. and Chocho, K., Applied Physics Letters, 1998

[6] Götz et al., *Activation energies of Si donors in GaN*, Appl. Phys. Lett., 1996

[7] Y. Chul Choi, M. Pophristic, H.-Y. Cha, B. Peres, M. G. Spencer, and L. F. Eastman, *The effect of a Fe-doped GaN buffer on offstate breakdown characteristics in AlGaN/GaN HEMTs on Si substrate*, Electron Devices, IEEE Transactions on, 53(12), 2926 –2931, dec. 2006

[8] Y.C. Choi, L.F. Eastman, and M. Pophristic, *Effects of a Fe-doped GaN buffer in AlGaN/GaN power HEMTs on Si substrate,* Solid-State Device Research Conference, 2006

[9] Kipshidze et al., *Mg and O codoping in p-type GaN and AlGaN (0<x<0.08)*, Appl. Phys. Lett., 2002

[10] Teke et al., *The effect of AlN interlayer thicknesses on scattering processes in lattice-matched AlInN/GaN two-dimensional electron gas heterostructures*, New Journal of Physics, 2009

[11] Mishra U.K., Nguyen D., Larson L.E., *Ultra-High-speed Modulation-Doped Field-Effect Transistors: A - torial Review,* 1992

[12] Mishra U.K., Parikh P., WU Y.F., *AlGaN/GaN HEMTs—An Overview of Device Operation and Applications,* 2002

[13] Ambacher O. et al., *Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaN/GaN heterostructures,* Journal of Applied Physics, 1999

[14] Marcon D., *Reliability study of power gallium nitride based transistors,* 2011

[15] M.C.J.C.M. Krmer, *Gallium Nitride-based Microwave High-Power Heterostructure Field-Effect Transistors, design, technology, and characterization*. 2006

[16] *Electrical properties of Silicon,* Ioffe Institute Database

[17] Harris J. J., Foxon C. T., Barnham K. W. J., Lacklison D. E., Hewett J., White C., *Two-dimensional electron gas structures with mobilities in excess of* $3\times10^6\ cm^2\ V^{-1}\ s^{-1}$, Journal of Applied Physics, 61: 1219, 1987

[18] Dürkop T., Getty S. A., Cobas E., Fuhrer M. S., *Extraordinary Mobility in Semiconducting Carbon Nanotubes*, Nano Letters 4: 35, 2004

[19] Bolotin, K; Sikes K, Jiang Z, Klima M. Fudenberg G., Hone J., Kim P., Stormer H., *Ultrahigh electron mobility in suspended graphene*, Solid State Communications, 2008

[20] B.Van Zeghbroeck, *Chapter 2: Semiconductor Fundamentals,* Online textbook

[21] B. L. Anderson and R. L. Anderson, *Fundamentals of Semiconductor Devices*, Mc Graw Hill, 2005

[22] Ferry David K., *Semiconductor transport,* London: Taylor & Francis, 2000

[23] Verzellesi G., Slides of the course of "Device simulations"

[24] Peter Y. Yu, Manuel Cardona, *Fundamentals of Semiconductors: Physics and Materials Properties*, Springer, 2010

[25] Parish G., et al, *AlGaN/AlN/GaN High Electron Mobility Transistors with Improved Carrier Transport,* IEEE, 2005

[26] Tülek R., et al, *Comparison of the transport properties of high quality AlGaN/AlN/GaN and AlInN/AlN/GaN two-dimensional electron gas heterostructures,* Journal of Applied Physics, 2009

[27] Farahmand et al., *Monte Carlo simulation of electron transport in the III-Nitride wurtzite phase materials system: binaries and ternaries*, IEEE trans. Electron devices, vol.48, no.3, pp.535-512, 2001

[28] Yang et al., *Improved negative differential mobility model of GaN and AlGaN for terahertz gunn diode*, IEEE trans. Electron devices, 2011

[29] D. C. Look et al*., Defect donor and acceptor in GaN,* Phys. Rev. Lett.,vol. 79, no. 12, pp. 2273–2276, Sep. 1997.

[30] W. Götz et al., *Hall effect analysis of GaN films grown by hydride vapor phase epitaxy,* Appl. Phys. Lett., vol. 72, no. 10, pp. 1214–1216, Mar. 1998

[31] A. F. M. Anwar et al*., Temperature Dependent Transport Properties in GaN, AlGaN and InGaN Semiconductors,* IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 48, NO. 3, MARCH 2001

[32] A. F. M. Anwar et al*., Temperature Dependent Transport Properties in GaN,* $Al_xGa_{1-x}N$*, and* $In_xGa_{1-x}N$ *Semiconductors,* IEEE, 2001

[33] Ilegems, M., Montgomery, H.C., J. Phys. Chem. Solids 34, 1972

[34] J. Mateos et al., *Monte Carlo simulation of AlGaN/GaN heterostructures,* IEEE, 2007

[35] J. M. Barker et al., *Bulk GaN and AlGaN/GaN heterostructure drift velocity measurements and comparison to theoretical models,* J. Appl. Phys., vol. 97, no. 6, p. 063 705, Mar. 2005

[36] Bertazzi et al., *Theory of high field carrier transport and impact ionization in wurtzite GaN. Part I: A full band Monte Carlo model,* J. Appl. Phys., 106, 2009

[37] Oded Katz et al., *Characteristics of InAlN–GaN High-ElectronMobility Field-Effect Transistor* , IEEE transaction on electron devices, vol.52, 2005

[38] Makoto Miyoshi et al., *Nanostructural characterization and two-dimensional electron-gas properties in high-mobility AlGaN/AlN/GaN heterostructures grown on epitaxial AlN/sapphire templates*, J. Appl. Phys 98, 2005

[39] Makoto Miyoshi, *Structural characterization of strained AlGaN layers in different Al content AlGaN/GaN heterostructures and its effect on two-dimensional electron transport properties ,* J. Vac. Sci. Technol., Jul/Aug 2005

[40] G. Masetti, M. Severi, and S. Solmi, *Modeling of Carrier Mobility Against Carrier Concentration in Arsenic-, Phosphorus-, and Boron-Doped Silicon*, IEEE Transactions on Electron Devices, vol. ED-30, no. 7, pp. 764, 1983.

[41] N. D. Arora, J. R. Hauser, and D. J. Roulston, *Electron and Hole Mobilities in Silicon as a Function of Concentration and Temperature*, IEEE Transactions on Electron Devices, vol. ED-29, no. 2, pp. 292, 1982.

[42] S. Reggiani et al., *A Unified Analytical Model for Bulk and Surface Mobility in Si n- and p-Channel MOSFET´s*,in Proceedings of the 29th European Solid-State Device Research Conference (ESSDERC), Leuven, Belgium, pp. 240, September 1999.

[43] C. Lombardi et al., *A Physically Based Mobility Model for Numerical Simulation of Nonplanar Devices*, IEEE Transactions on Computer-Aided Design, vol. 7, no. 11, pp. 1164, 1988.

[44] S. A. Mujtaba, *Advanced Mobility Models for Design and Simulation of Deep Submicrometer MOSFETs*, Ph.D. thesis, Stanford University, Stanford, CA, USA, December 1995.

[45] D. M. Caughey and R. E. Thomas, *Mobilities in Silicon Empirically Related to Doping and Field*, Proceedings of the IEEE, vol. 55, no. 12, pp. 2192, 1967.

[46] C. Canali et al., *Electron and Hole Drift Velocity Measurements in Silicon and Their Empirical Relation to Electric Field and Temperature*, IEEE Transactions on Electron Devices, vol. ED-22, no. 11, pp. 1045, 1975.

[47] J. Barnes, R. J. Lomax, and G. I. Haddad, *Finite-Element Simulation of GaAs MESFET´s with Lateral Doping Profiles and Submicron Gates*, IEEE Transactions on Electron Devices, vol. ED-23, no. 9, pp. 1042, 1976.

[48] B. Meinerzhagen and W. L. Engl, *The Influence of the Thermal Equilibrium Approximation on the Accuracy of Classical Two-Dimensional Numerical Modeling of Silicon Submicrometer MOS Transistors*, IEEE Transactions on Electron Devices, vol. 35, no. 5, pp. 689, 1988.

[49] Tebaldi A., Fisica dei Materiali per l´Elettronica, 2007

[50] Cester A., Corso di Nanoelettronica – Parte 1

[51] Electron Mobility, Wikipedia

Abbreviations

2DEG	Two-Dimensional Electron Gas
BV	Breakdown Voltage
ELO	Epitaxial Lateral Overgrowth
FP	Field Plate
HEMT	High Electron Mobility Transistor
HPVE	Hydride Vapour Phase Epitaxy
III-N	Compounds made by one element from group III and Nitrogen
III-V	Compounds mabe by one element from group III and one from group V
I-V	Current-Voltage characteristics
LASER	Light Amplification by Stimulated Emission of Radiation
LED	Light Emitting Diode
MBE	Molecular Beam Epitaxy
MESFET	Metal-Semiconductor Field Effect Transistor
MOCVD	Metal-Organic Chemical Vapor Deposition
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
R_{DSon}	Resistance between Drain and Source with the device on
RF	Radio Frequency
RHEED	Reflection High Energy Electron Diffraction
TEC	Thermal Expansion Coefficient
WBG	Wide Band Gap

Chemical formulas

GaN	Gallium Nitride
InN	Indium Nitride
AlN	Aluminium Nitride
BN	Boron Nitride
Si	Silicon
GaAs	Gallium Arsenide
InP	Indium Phosphide
SiC	Silicon Carbide
InGaN	Indium Gallium Nitride
ZnO	Zinc Oxide
MgO	Magnesium Oxide
Al_2O_3	Sapphire